Enhancing the Shared Decision-Making Experience of Multiple Sclerosis Via Model-Driven Decision Support

A Dissertation
Presented in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
with a
Major in Computer Science
in the
College of Graduate Studies
University of Idaho
by

Rayan Alshamrani

Approved by:
Major Professor: Xiaogang Ma, Ph.D.
Committee Members: Terence Soule, Ph.D.; Min Xian, Ph.D.; Audrey Qiuyan Fu, Ph.D.
Department Administrator: Terence Soule, Ph.D.

May 2023

Abstract

Multiple sclerosis (MS) is a disorder that strikes the central nervous system of the human body. Due to the complexity of this disorder, healthcare sectors increasingly need shared clinical decision-making tools that provide practitioners with insightful knowledge and information about MS. These tools ought to be comprehensible by both technical and non-technical healthcare audiences. Decision support systems (DSSs) enhance decision-making by resolving real-world problems through timely practical decisions. DSSs are promising in medicine. That is why several studies reveal the increasing importance of DSSs in most medical domains. DSSs are hypothetically promising tools in the MS domain as they could provide decision-makers with needed information and critical decisions. However, the utilization of DSSs in MS practices remains an open challenge due to the disease's rarity and data scarcity. In general, three issues are needed to be addressed and examined in detail in MS research from the data and computer science perspective: (1) how to predict the type of MS in affected patients? (2) how to determine the most suitable disease-modifying therapy (DMT), specifically the DMT's administration route, according to an MS patient needs? (3) how to increase the MS's public awareness by presenting a visually reliable source of information? Model-driven DSS is proposed to address these research gaps and enrich the MS field with modern decision-making technology.

The proposed design utilizes machine learning (ML) algorithms to expedite proper solutions for the first two concerns. Therefore, this dissertation investigates several supervised ML models in predicting the type of MS in affected persons and assesses the models' effectiveness using a real-world MS dataset. A framework for training and comparing the results of various traditional and ensemble algorithms to predict MS episodes is presented. Clinical baseline data from a database of questionnaires was obtained, and then several traditional models and ensemble classifiers were trained against this dataset. To address the first research gap, random forest (RF), decision trees (DT), bagging, and gradient boosting

classifiers showed consistently promising accuracy, sensitivity, and specificity results. However, hyperparameter tuning did not yield considerable increases in any evaluation metrics.

With the availability of various DMTs for controlling MS, an accurate selection of the best DMT regimen is critical for improving the quality of life of MS patients. However, selecting the best route of administration suitable for an MS patient depends on the evaluation risks versus the efficacy of a specific DMT. Thus, one of the studies presented in this dissertation demonstrates a framework for a model-based system that utilizes ML algorithms for predicting the best route of administration for delivering DMTs to MS patients. The best-performing models were the gradient boosting and RF. This study serves as a proof-of-concept for the ML application in decision-making regarding MS DMT prescriptions.

Broadly, MS is a preference-sensitive condition. Thus, the shared decision-making regarding the diagnosis and treatment of MS is vital yet complicated. Physicians should have broad domain knowledge about MS, and patients should also know the disorder's key concepts. A knowledge graph of MS is a viable way to support communication and shared decision-making between physicians and patients. The MS Knowledge Graph (MSKG) using ontology engineering methods and semantic technologies to facilitate and address the third research gap is structured and reviewed in this dissertation. MSKG provides necessary medical terminologies related to MS and organizes them in a logically-coherent framework. Target beneficiaries of MSKG would be patients diagnosed recently or at high risk of developing MS, their first-degree relatives, and anyone interested in obtaining more knowledge about MS.

In short, the dissertation's objective is to design a DSS prototype that emulates an MS expert to accurately identify the type of MS and make appropriate DMT recommendations. The presented DSS uses knowledge representation via ontology engineering and ML to enhance decision-making. Future

work should focus on using DSS based on ML and deep learning to understand MS occurrence patterns, etiology, effects on quality of life, and correlations with other disorders. Access to MS data should be readily available, and knowledge graph applications in the MS domain can be improved through knowledge-based reasoning methods. DSS technologies can be pragmatic in the MS domain and research, and further research is necessary to fully realize their potential benefits.

Acknowledgments

I would like to express my deepest appreciation to all members of Taif University, especially my colleagues and officials at the College of Computers and Information Technology. I acknowledge the executives at the Office of Vice President for Graduate Studies and Scientific Research for their support. The presented dissertation was supported by Taif University, which granted me a full scholarship to pursue graduate degrees. I would also like to express my sincere gratitude to the faculty, staff, and officials at the University of Idaho who have been a vital part of my academic journey, especially those affiliated with the Department of Computer Science. I would like to sincerely thank Sean Sullivan, Larry Stauffer, Ph.D., P.E., John Crepeau, Ph.D., P.E., Michele Mattoon, Jennifer Neelon, Jeremy Tamsen, and members of the Office of Information Technology for their effort in assisting the entire Vandal family.

I acknowledge Hollie Schmidt and David Gwynne of the Accelerated Cure Project for Multiple Sclerosis for their valuable involvement while obtaining the study materials. I sincerely thank Jamil Chaudri, Ph.D., and John Biros, Ph.D. of Marshall University, for imparting their knowledge and expertise at the beginning of my academic pursuits. My sincerest appreciation to Sarita Said-Said, M.D. of Palouse Neurology and Lucas McCarthy, M.D. of Virginia Mason Hospital for their endless support and valuable contribution.

My genuine acknowledgment to my colleagues, IDEA lab past and present fellows, and friends; thank you for the camaraderie, discussions, and debates and for making the learning experience more enjoyable and memorable. I sincerely thank Husam Samkari, Ph.D., Waled Alshhrani, D.D.S, Ayman Almajnoony, Ph.D., Mazen Alyobi, Ph.D., Amruta Kale, Ph.D., and Bassim Alshedokhy M.S. for encouraging and motivating me during the critical stage of my journey.

To Terry Wahls, M.D. and Deborah McGuinness Ph.D., thank you for inspiring and motivating me to "*defying all the odds*" and to be "*very open about my*" MS "*journey*".

To my mentor and role model Dr. Xiaogang (Marshall) Ma, I would not have made it this far without your collective supervision, support, guidance, motivation, and encouragement. Your dedication has played a significant role in shaping me into the person I am today. Your wisdom and expertise have been instrumental in sharpening my knowledge and improving my research skills. I am also grateful for your commitment to fostering a supportive and stimulating academic environment. I feel fortunate to have had the opportunity to work with you. Finally, thank you, Dr. Ma for your unwavering belief in my abilities, which has given me the confidence and motivation to pursue my academic and professional goals.

Dedication

All praise is to Allah. The one and only Almighty God, thank you for bestowing your blessings upon my family and me.

My beloved country the Kingdom of Saudi Arabia, thank you for this opportunity that made me achieve my goals and made my dreams come true.

My dear family members and in-laws, many thanks to each and every one of you, especially my mother-in-law Nadia Rais, Aunt Amani Hadhrawi, Aunt Maryam Hadhrawi, and Grandmother Salha Alwasem for believing in me.

My deceased grandfather General Fared Hadhrawi, I wish you were here now. You will always be remembered and missed. This work is dedicated in memory of you.

To Yara, Raneem, and Anan, the successes you achieved in your lives have been boosting my confidence to do more and to work harder. You are the source of my strength and pride.

My father Brigadier General Mohammed Alshamrani, you worked day and night to provide for us, and you pushed us very hard to pursue higher education. You gave up your dreams to make ours. I will always look up to you humbly.

Behind every great man is a great woman, and that is you, my lovely wife Ohood. You stood up by my side since day one. Every day that goes by with you is a blessing.

My mother Mrs. Norah Hadhrawi, you dropped out of school and devoted your life to us. I vowed to earn as many degrees as possible for you. Mom, this is a starting point, and I will never stop.

Seba and Wafi, I was truly gifted with you. You gave me hope in the worst circumstances that befell me. I love you my kids to the moon and back.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACP ..... Accelerated Cure Project for MS

AI ..... Artificial Intelligence

AMS3 ..... Alemtuzumab in MS Safety Systems

ANN ..... Artificial Neural Networks

AUC ..... Area Under ROC Curve

CAD ..... Computer-Assisted Decision

CDSS ..... Clinical Decision Support System

CID ..... Chronic Inflammatory Disease

CIS ..... Clinically Isolated Syndrome

CV ..... Cross Validation

DE ..... Differential Evolution

DF ..... Deformation Field

DMA ..... Disease-Modifying Agent

DMT ..... Disease-Modifying Therapy

DSS ..... Decision Support System

DT ..... Decision Tree

DTI ..... Diffusion Tensor Imaging

DWT ..... Discrete Wavelet Transform

EA ..... Evolutionary Algorithm

EDSS ..... Expanded Disability Status Scale

EMR ..... Electronic Medical records

FN ..... False Negative

FP ..... False Positive

HAT ..... Home Automated Tele-management

IRB ..... Human Research Protections

KNN ..... K-Nearest-Neighbor

LASSO ..... Least Absolute Shrinkage and Selection Operator Regression

LDA ..... Linear Discriminant Analysis

LOOCV ..... Leave-One-Out CV

LR ..... Logistic Regression

LS-SVM ..... Least-Squares Support Vector Machine

MHO ..... Mental Health Ontology

ML ..... Machine Learning

MLR ..... Multiple Logistic Regression

MRI ..... Magnetic Resonance Imaging

MS ..... Multiple Sclerosis

MSKG ..... MS Knowledge Graph

NB ..... naive Bayes

ND ..... Neurological Disease Ontology

NSWM ..... Non-Specific White Matter

OWL2 ..... Web Ontology Language

PCA ..... Principal Component Analysis

PMS ..... Progressive MS

PPMS ..... Primary Progressive MS

PRMS ..... Progressive Relapsing MS

RF ..... Random Forest

RFE ..... Recursive Feature Elimination

RIS ..... Radiologically Isolated Syndrome

ROBOKOP ..... Reasoning Over Biomedical Objects linked in Knowledge Oriented Pathways

RRMS ..... Relapsing Remitting MS

rsfMRI ..... Resting-State Functional MRI

SCD ..... Sickle Cell Disease

SGD ..... Stochastic Gradient Descent

SLCMSR ..... Sylvia Lawry Centre for Multiple Sclerosis Research

SMSreg ..... Swedish MS Registry

SPIN ..... SPecific Tests Rule IN

SPMS ..... Secondary Progressive MS

SVM ..... Support Vector Machine

SVM-rbf ..... SVM with a Radial Basis Function Kernel

SWE ..... Stationary Wavelet Entropy

TN ..... True Negative

TP ..... True Positive

Turtle ..... Terse RDF Triple Language

XGBoost ..... Extreme Gradient Boosting

Statement of Contribution

Chapter 2: I, myself, conceptualized the research idea and design. I developed the methodology. Myself, Ashrf Althbiti, and Dr. Xiaogang Ma verified the structure of the study and the research outputs. I conducted the research and investigation process with a help from Yara Alshamrani. I wrote and revised the first draft. Myself, Fatimah Alkomah, and Dr. Xiaogang Ma reviewed and edited the manuscript. Dr. Xiaogang Ma supervised the study.

Chapter 3: I, myself, conceptualized the research idea and design. I developed the methodology. Myself, Amruta Kale, and Dr. Xiaogang Ma verified the structure of the experiment and the research outputs. I conducted the research and investigation process, then provisioned the study materials with Dr. Sarita Said-Said. I wrote and revised the first draft. Myself and Dr. Xiaogang Ma reviewed and edited the manuscript. Dr. Xiaogang Ma supervised the study.

Chapter 4: I, myself, conceptualized the research idea and design. I developed the methodology. Myself, Mazen Alyobi, Ayman Almajnoony, and Dr. Xiaogang Ma verified the structure of the experiment and the research outputs. I conducted the research and investigation process, then provisioned the study materials with Sarita Said-Said. I wrote and revised the first draft. Myself and Dr. Xiaogang Ma reviewed and edited the manuscript. Dr. Xiaogang Ma supervised the study.

Chapter 5: Myself and Dr. Xiaogang Ma conceptualized the research idea and design. I developed the methodology. Myself, Ashrf Althbiti, Dr. Xiaogang Ma, and Dr. Sarita Said-Said verified the structure of the experiment and the research outputs. I conducted the research and investigation process, then provisioned the study materials. I wrote and revised the first draft. Dr. Xiaogang Ma and Dr. Sarita Said-Said reviewed and edited the manuscript. I visualized the methodology and the result. Dr. Xiaogang Ma supervised the study.

Chapter 1 Introduction

Background

Multiple sclerosis (MS) is a chronic, neurological, and the most common inflammatory demyelinating disorder that stimulates the immune system to attack the central nervous system of the human body (Sospedra and Martin, 2005; Lassmann, Brück and Lucchinetti, 2007; Chiaravalloti and DeLuca, 2008; Milo and Kahana, 2010). Several genetic and environmental factors like vitamin D deficiency, Epstein-Barr virus, ambient ultraviolet radiation, obesity, and smoking are the potential confirmed or yet to be confirmed causes that trigger MS and its progression (Trapp *et al.*, 1998; Ascherio and Munger, 2016; Reich, Lucchinetti and Calabresi, 2018; deAndrés-Galiana *et al.*, 2019). However, the exact causes of MS are still unknown. MS patients may experience several symptoms independently or concurrently during the course of the disease such as sensory, visual, motor, cognitive, and cerebellar disorders (Miri Ashtiani *et al.*, 2018). Although, these symptoms and their impacts vary among MS patients (Pinto, 2018). MS affects several millions of people around the globe, especially young adults (Ho, Ghosh and Unnikrishnan, 2013). Generally, men are less likely to develop MS than women (Kragt *et al.*, 2009). MS is manifested as one of the following progressions: clinically isolated syndrome (CIS), relapsing remitting (RRMS), progressive MS (PMS) that refers to the combination of secondary progressive MS (SPMS) and primary progressive MS (PPMS), progressive relapsing MS (PRMS), and radiologically isolated syndrome (RIS) (Okuda *et al.*, 2009; Granberg *et al.*, 2013; Lublin *et al.*, 2014; Mahad, Trapp and Lassmann, 2015; D'Amico *et al.*, 2016; Efendi, 2016; deAndrés-Galiana *et al.*, 2019).

The MS diagnostic procedure requires neurological examinations such as magnetic resonance imaging (MRI), lumbar punctures, and blood tests to confirm MS cases (Ghasemi, 2017). In addition, neurologists prescribe existing MS therapies in order to control the symptoms and the progression of the disease, as it turns out that this disorder cannot be cured or prevented (Karaca *et al.*, 2017). Thus, neither the treatment nor the diagnosis of MS is easy. This is because MS shares several clinical features

with other diseases, and there is no consensus approach to diagnose MS (Ho, Ghosh and Unnikrishnan, 2013). In fact, the decision-making in treating and diagnosing MS is critical and relies heavily on the neurologist's experience and judgment. Therefore, the quality of decisions for that matter remains doubtful due to the presence of uncertainties associated with MS. Notably, MS is a preference-sensitive condition, meaning that both the physician and the patient participate in the decision-making process, i.e., shared decision-making (Colligan, Metzler and Tiryaki, 2017). Hence, this would impose a great responsibility upon the contributors in the decision-making process as they must have complete knowledge about the current state of the condition and the potential risks and benefits of all possible options in order to achieve the optimal decision. Comprehensively, it would be beneficial to have easy-to-use automated solutions that propose several optimal alternatives to make the shared decision-making easier for all participants in this process.

A decision support system (DSS) is a computer-based system devoted to people concerned with decision-making to enable them to solve real-world problems via worthy decisions (Eom and Kim, 2006). DSSs are largely accepted in modern commercial businesses and have accomplished significant success (Kawamoto *et al.*, 2005). DSSs in the medical realm are very promising, especially for enhancing decision-making. Recent studies have demonstrated the increasing importance of DSSs in medicine, i.e., clinical decision support systems (CDSS), to help intended decision-makers to nominate the right decision among several alternatives as often as possible (Pota, Esposito and Pietro, 2012). Healthcare organizations increasingly need DSSs that are understandable by non-technical audiences such as healthcare providers. DSSs aid clinical decision-making by giving practitioners insight and information about their patients to generate suitable assessments or recommendations (Power, 2002). In the MS domain, DSSs are just as important as their use in other healthcare areas. DSSs technologies are potentially favorable tools in the MS domain. For instance, DSSs could provide decision-makers with useful information (e.g., alerts, warnings, or predictions) about MS cases (Fraccaro *et al.*, 2015). The benefits of using DSSs in the MS domain include: enabling access to neurologists, enhancing

clinical documentation and prescription processes, escalating diagnosis accuracy, minimizing time loss and healthcare expense, enhancing diagnostic predictions, maximizing the quality of patients' lives and care provided, and improving the quality of decisions (Rothman, Leonard and Vigoda, 2012; Arani *et al.*, 2018). With that said, applying DSSs in the MS realm would be phenomenal, especially for enhancing the shared decision-making process.

However, with all these benefits, the utilization of DSSs in the MS field is not encouraging. To the best of the contributor's knowledge, the investigation of the DSSs in the MS domain is still insufficient for two reasons. First, the number of published articles on the presented topic is minimal. Hence, there is no clear intention or motivation to pursue this research domain path. Second, the number of existing systems used in the MS's daily clinical practices is exiguous. Indeed, the acceptance of DSSs within the MS domain remains inadequate. From the computer and data science perspectives, this condition is not getting the proper attention compared to other incurable diseases due to its uncommonness and data scarcity. Therefore, a model-driven DSS is presented to address this concern and enrich the MS domain with the most appropriate yet modern technologies. Generally speaking, a model-driven DSS is a type of DSS that uses complex and quantitative models that provide a simplified and straightforward knowledge representation to decision-makers (Power and Sharda, 2007). Two characteristics distinguish Model-driven DSSs: (1) a model in a model-driven DSS is made accessible to experts with no technical background, and (2) DSSs of this type are reusable in equivalent decision situations (Power and Sharda, 2007). Undoubtedly, the development of DSSs using different methodologies for improving the decision-making process in the MS's clinical practice experience, with a special focus on applying model-driven approaches used to construct model-driven DSSs, should be considered in this type of research.

Research Motivation and Problem Statement

MS is thought to be caused by environmental factors affecting people who are genetically susceptible

(Pinto, 2018). This disease is not easily detected due to its similarity to other diseases, in addition to the fact that the pattern of symptoms is ambiguous and that the duration of their appearance is not fixed and indeterminate. Furthermore, MS specialists usually detect MS activities by comparing initial diagnosis and follow-ups reports. Nevertheless, these procedures require intensive knowledge and experience, given the inconclusiveness of this medical condition due to the lack of consensus diagnostic procedures. To aid MS specialists, early diagnosis and/or sooner prediction of the type of MS would gain momentum in the MS domain. This is because the prognosis of several MS manifestations and their severe progression could take time (Ma, 2018). Therefore, determining the progression level in MS patients is decisive in undertaking the necessary and immediate medical interventions.

MS manifestation varies among patients. Therefore, different patients require tailored treatment plans as each plan is associated with a patient's conditions and symptoms. The variety of MS disease-modifying therapies (DMTs) offers a range of potential benefits such as disease course modification, attacks number and severity reduction, relapses prevention, disability accrual, and symptoms control, to name a few (Ma, 2018; Eskyte *et al.*, 2019). On the other hand, these different DMTs may also tolerate life-threatening risks. Therefore, when choosing the appropriate treatment solution, decision-making has become more complex and challenging for physicians and patients as they decide what tradeoffs, consequences, and long-term benefits they are willing and ready to make (Rahn *et al.*, 2015, 2018; Eskyte *et al.*, 2019). Unfortunately, MS patients rely mostly on information that they find online to self-educate themselves about the available DMT options (Langhorne, Thomas and Kolaczkowski, 2013). Thereby, using a reliable decision-making tool would assist both parties of the shared decision-making process, a clinician and a patient, to select the most suitable treatment regimen, explicitly selecting the best routes of administration for the prescribed DMTs.

Comprehensive information about MS, together with the availability and accessibility of credible resources and services, are scarce globally for MS patients and those who are eager to know more about

this condition (Browne *et al.*, 2014). MS patients increasingly need reliable information about this disorder, its symptoms, and ways to control and manage it, to name a few (Meca-Lallana *et al.*, 2017). Thus, providing MS patients, especially newly diagnosed individuals as well as people close to them who are interested in knowing more about MS, with an accessible and easy-to-use source of information would reinforce patient empowerment in informed decisions associated with MS. Hence, visualizing MS's fundamental information in the form of a knowledge graph using ontology engineering would be a genuine endeavor. This would promote having an available open-access source of information and elevate public awareness about MS.

Researchers nowadays are applying several technologies, as standalone approaches or as the basis of several DSSs, to support decision-making in the MS medical routine. However, the use of DSSs in the MS field is limited due to the condition's rarity, even though decision-making is a critical task in MS practices. Ultimately, the existing MS DSSs in use or even the proposed ones in research are implemented for several specific purposes inconsistent with the research motivation of this dissertation. Primarily, most of these DSSs detect and classify the MS lesions by recognizing them when analyzing the affected patients' MRIs. Likewise, several DSSs are utilized to distinguish individuals with MS by examining cerebrospinal fluid and blood findings. Other DSSs are developed for predicting the MS's short- and long-term prognosis, disability rate, and treatment effects. In addition, several DSSs are used for decision-making associated with organizing MS patients' important care routines.

Arguably, there are several reasons why DSSs simulating MS specialists may not exist. For instance, to develop such a DSS, sufficient data must be available to train the system. In the case of MS, there may need to be more data available to build an accurate DSS that impersonates a specialist (Alshamrani *et al.*, 2020). Moreover, MS is a complex and unpredictable disease with a wide range of symptoms and treatment options (Goldenberg, 2012; Pérez del Palomar *et al.*, 2019; Saccà *et al.*, 2019). Thus, it may be challenging to develop a DSS that can accurately reflect the expertise of a specialist in this

field. In addition, MS specialists use a combination of their training, experience, and intuition to make clinical decisions (Alshamrani *et al.*, 2020). Replicating such proficiency in a DSS, which typically relies on rules and algorithms, is complex. Also, there may be concerns about the liability and trust associated with using a DSS in clinical decision-making, especially in a complex and unpredictable disease like MS. That being said, there are DSSs that provide guidance for specific aspects of MS care, which can be useful tools for healthcare providers. However, these systems are not designed to replace the expertise of a specialist and are typically used in conjunction with clinical judgment. Particularly, no DSSs support decisions about determining MS type, prescribing suitable DMTs, and providing inclusive information about MS.

Study Objectives

To address the aforementioned challenges, this Ph.D. dissertation aims to apply supervised machine learning (ML) and knowledge graph to design a DSS suitable for enhancing the MS's periodical decision-making. The designed model-driven DSS is mainly comprehensible for (1) indicating the type of MS in affected patients, (2) assisting the MS's shared decision-makers in deciding the best treatment solution for individuals affected by MS, and (3) representing comprehensive information about MS in an easy visualized manner.

Research Questions

This work answers the following research questions:

1) How to practically discover the MS's stage in patients diagnosed with MS?
2) Would it be possible to assist an MS patient and a provider in deciding on the best MS DMT among alternatives?
3) How to discover and observe the risk factors associated and mostly correlated with MS types?
4) What features that are mostly correlated when determining the most suitable treatment plan?
5) How would the ML algorithms be used, and which performs better in the above matters?

6) Could ontology engineering be used to form a suitable data source and knowledge representation used by those who need brief and initial information about MS?

Hypotheses

To address the research questions mentioned above, the following hypotheses are considered to facilitate the investigation process of this study:

1) Supervised ML models are prime for predicting the progression level that affects persons with MS, in addition to predicting the best DMT route of administration for MS patients. (research questions 1-5)
2) One and only one supervised ML model among all can outperform the rest. (research questions 1-5)
3) Knowledge representation using ontology engineering would increase the MS's public awareness due to its ability to visualize the fundamental and essential facts about MS. (research question 6)

Based on the empirical experiments conducted in this dissertation, proofing and confirming the feasibility of the abovementioned claims would be emphasized to promote decisions associated with MS.

Scientific Contribution

The potential scientific contribution of the dissertation focuses on the following topics:

- The essentiality of predicting the MS type after confirming the diagnosis.
- Supporting the decisions made by MS patients and specialists regarding formalizing treatment solutions.
- Finding and then applying several supervised ML algorithms to be used in the MS research domain.

- Reinforcing the MS's public awareness by employing ontology engineering as a knowledge representation technique to form a knowledge graph associated with MS.

Overall, this study and the prospective experiments provide a route map for the MS's shared decision-makers regarding developing a DSS prototype that would be used to secure the dissertation's aims.

Organization of the Dissertation

The dissertation consists of six chapters, four of which (Chapter 2 – Chapter 5) contain primary contributions toward addressing the research questions mentioned above. The following paragraphs briefly outline these chapters and highlight their objectives.

Chapter 2: Literature Review. This chapter reviews the state-of-the-art DSSs in MS research, specially model-driven DSSs. This chapter observes the utilization of knowledge representation, ontology engineering, and ML approaches in DSSs proposed in MS research to simplify the complex process for decision-makers. This chapter comprehensively summarizes recent studies highlighting the importance of DSSs in the MS realm and addresses the future work of applying DSS technologies in the MS field.

Chapter 3: Examination of Supervised Machine Learning Classifiers and Ensemble Learning for Predicting the Type of Multiple Sclerosis. This chapter investigates several supervised ML models in predicting the type of MS in affected patients and assesses the models' effectiveness using a real-world MS dataset. A framework for training and comparing the results of various traditional and ensemble algorithms to predict MS types is presented in this chapter.

Chapter 4: The Application of Predictive Machine Learning Models to Support Shared Decision-Making in Selecting the Best DMT Route of Administration. This chapter proposed a design for a model-based system that utilizes ML algorithms to predict, using a real-world MS dataset, the best

route of administration for delivering DMTs to MS patients. The proposed framework in this chapter trains and compares the results of various traditional and ensemble models.

Chapter 5: Using Knowledge Graph to Improve Informed Multiple Sclerosis Diagnosis and Treatment Decisions. This chapter represents a blueprint of an MS knowledge graph structured using ontology engineering methods and semantic technologies. In this chapter, the proposed knowledge graph provides necessary medical terminologies related to MS and organizes them in a logically-coherent framework. Target beneficiaries of the proposed knowledge graph would be patients diagnosed recently or at high risk of developing MS, their first-degree relatives, and anyone interested in obtaining more knowledge about MS.

Chapter 6: Conclusion, Limitations, and Future Recommendations. This chapter briefly revisits the above chapters, demonstrates the overall significance and novelty of the findings, underlines the limitations, and provides recommendations for future research.

Chapter 2 Literature Review

This chapter is retrieved from:

Alshamrani, R., Althbiti, A., Alshamrani, Y., Alkomah, F., & Ma, X. (2020). Model-Driven Decision Making in Multiple Sclerosis Research: Existing Works and Latest Trends. Patterns, 1(8), 100121. https://doi.org/10.1016/j.patter.2020.100121

Introduction

Multiple sclerosis (MS) is a chronic, neurological, and the most common inflammatory disorder that stimulates the immune system to attack the central nervous system of the human body (Kasper *et al.*, 2012; Arani *et al.*, 2018; Mowry *et al.*, 2018). Genetic and environmental factors such as vitamin D deficiency, Epstein-Barr virus, ambient ultraviolet radiation, obesity, and smoking are the potential confirmed or yet to be confirmed causes of MS and its progression (Trapp *et al.*, 1998; Leray *et al.*, 2016; Reich, Lucchinetti and Calabresi, 2018; deAndrés-Galiana *et al.*, 2019; Dobson and Giovannoni, 2019; McGinley, Goldschmidt and Rae-Grant, 2021). However, the exact causes of MS are still unknown. MS patients may experience several symptoms independently or concurrently during the course of the disease such as sensory, visual, motor, cognitive, and cerebellar disorders (Miri Ashtiani *et al.*, 2018). MS affects several millions of people around the globe, especially young adults (Ho, Ghosh and Unnikrishnan, 2013). In general, men are less likely to develop MS compared to women (Kragt *et al.*, 2009). Four medical terms represent MS in terms of the progression level: relapsing-remitting MS (RRMS), secondary-progressive MS (SPMS), primary-progressive MS (PPMS), and progressive-relapsing MS (PRMS) (deAndrés-Galiana *et al.*, 2019).

The MS diagnostic procedure requires neurological examinations such as magnetic resonance imaging (MRI), lumbar punctures, and blood tests to confirm MS cases (Ghasemi, 2017). In addition, neurologists prescribe existing therapies in order to control the symptoms and the progression of MS as it turns out that this disorder cannot be cured or prevented (Karaca *et al.*, 2017). Thus, neither the treatment nor the diagnosis of MS is easy. This is because MS shares several clinical features with other

diseases and has no consensus approach in MS diagnosis (Ho, Ghosh and Unnikrishnan, 2013). In fact, the decision-making in both treatment and diagnosis of MS is critical and relies heavily on the experience and the judgment of the neurologist. Therefore, the quality of decisions for that matter remains doubtful due to the presence of uncertainties associated with MS. Moreover, MS is a preference-sensitive condition, so both the physician and the patient participate in the decision-making process, i.e., shared decision-making (Colligan, Metzler and Tiryaki, 2017). Hence, this would impose a great responsibility upon the contributors in the decision-making process as they must have full knowledge about the current state of the condition and the potential risks and benefits of all possible options in order to achieve the optimal decision. Comprehensively, it would be beneficial to have easy-to-use automated solutions that could propose several optimal alternatives to make the shared decision-making easier for all participants in this process.

Decision support systems (DSSs) are computer-based systems devoted to people who are concerned with decision-making so they could solve real-world problems via worthy decisions (Eom and Kim, 2006). DSSs are largely accepted in modern commercial businesses and have been accomplishing significant successes (Kawamoto *et al.*, 2005). DSSs in the medical realm are very promising especially for enhancing the decision-making process. Recent studies demonstrate the increasing importance of DSSs in medicine, i.e., clinical decision support systems (CDSS), for helping intended decision-makers to nominate the right decision among several alternatives as often as possible (Pota, Esposito and Pietro, 2012). DSSs' technologies are potentially favorable tools in the MS domain. For instance, DSSs could provide decision-makers with useful information (e.g., alerts, warnings, or predictions) about MS cases (Fraccaro *et al.*, 2015). The benefits of using DSSs in MS domain include: enabling access to neurologists, enhancing clinical documentation and prescription processes, escalating diagnosis accuracy, minimizing time loss and healthcare expense, enhancing diagnostic predictions, maximizing the quality of patients' lives and care provided, and improving the quality of decisions (Rothman, Leonard and Vigoda, 2012; Arani *et al.*, 2018). However, with all these benefits, the utilization of DSSs

in the MS field is not encouraging. To the best of our knowledge, aside from this review, the investigation of the DSSs in the MS domain is still insufficient because there is no published review on the presented topic.

The goal of this chapter is to analyze the state-of-the-art DSSs in MS research. This work aims to answer the following questions: (1) What DSSs are currently used in the MS domain? (2) What are their key fundamental methodologies? (3) What are they used for? and (4) What are the current most promising technologies associated with decision-making in the MS domain? Answering these questions would (1) demonstrate the importance of adopting DSSs in MS research and (2) show the extent of technologies, mostly correlated with DSSs, adopted for decision-making purposes in the scope of MS.

A model-driven DSS is a type of DSS that uses complex and quantitative models that provide a simplified and straightforward knowledge representation to decision-makers (Power and Sharda, 2007). Model-driven DSSs are distinguished by two characteristics: (1) a model in a model-driven DSS is made accessible to experts with no technical background, and (2) DSSs of this type are reusable in equivalent decision situations (Power and Sharda, 2007). Definitely, both the data science community and the MS community need to have a clear overview of the current state of the automated tools used to support decision-making in the MS domain.

This review is directed to data scientists, especially those who are interested in complex modeling approaches within the health informatics field. This review demonstrates the current DSS technologies within the domain of MS research, so data scientists can glance over the recent trends and the potential future research paths in order to enrich this research field with new automated decision-making technologies. Surely, MS community members and MS specialists will find useful information about the most recent technologies that could help them in their daily clinical practices.

The remainder of this chapter is organized as follows: Section 2 explains the methodology conducted to acquire the needed resources. Section 3 describes the results in accordance with the analysis methodology. Section 4 discusses remarkable findings and trends.

Literature Search Methodology

All relevant articles used to carry out this literature review were collected from six database sources: Google Scholar, DBLP Computer Science Bibliography, Web of Science, PubMed, ACM Digital Library, and IEEE Xplore Digital Library. The searching strategy utilized Multiple Sclerosis, Decision Support, Ontology, Semantic Web, Machine Learning, and Knowledge Graph as the search phrases, where the first phrase (i.e., MS) is combined with each of the latter ones (i.e., technical terms) with the logical operator "AND" to form a searching keyword. Each keyword was used to retrieve articles that observe current trends of MS research with the help of the technical term noted in the keyword. This is done in order to determine if there are correlations in MS research between the retrieved articles and the use of DSSs or the decision-making process in the MS domain. Articles published in English between 2007 and 2019 were collected during September and December 2019 and were screened out. Narrowing the number of the selected articles took two steps. First, the abstract (along with the title) of each article was analyzed to compose a subset of articles corresponding to the searching keywords. Afterward, the full text of each article in the composed subset was resolved to report the most interesting articles that cope with the inclusion and exclusion criteria applied to the searching strategy of this literature review.

Articles reviewed in this work comply with two inclusion criteria. The first criterion ensures that the machine-driven model applied in an article has a detailed description of its functionality. The second criterion justifies that the whole purpose of an article being analyzed is for diagnosing, treating, classifying, or predicting MS specifically. Research papers that have pure medical knowledge about

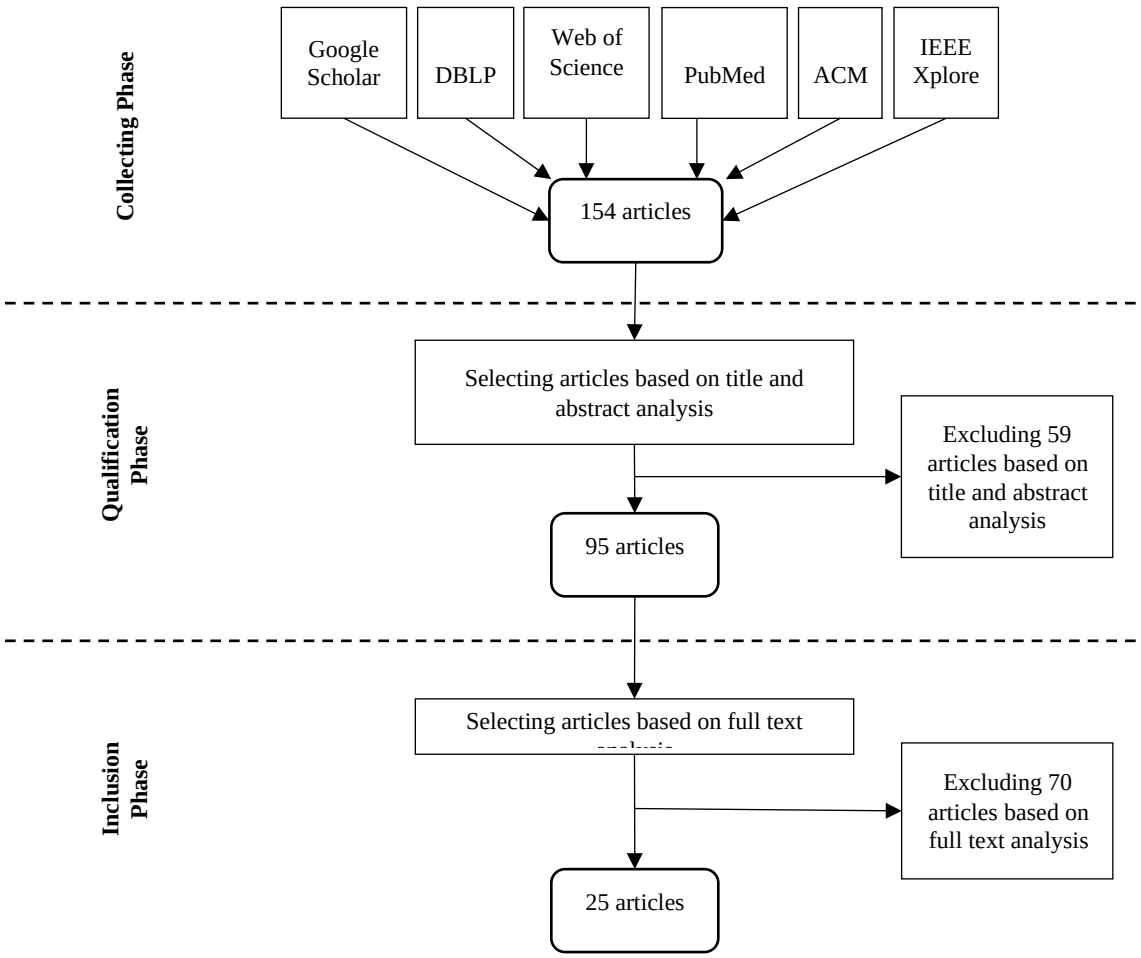


Figure 2.1: Sources and steps in the literature selection process

MS and others that employ solid mathematical theories with no automated models or running computer systems are therefore excluded. Likewise, papers that are dedicated to general neurology, except the ones that have MS as an example or as a case study, are excluded as well. The acquired data from each article consists of the author, the main intention of the paper, dataset information, system in use (if applicable), applied model and algorithms (i.e., research methodology), outcomes and remarks, and evaluation approaches and results (if reported). Figure 2.1 briefly illustrates the literature selection process.

Literature Analysis and Results

Using the searching strategy presented earlier, a total number of 154 articles were retrieved from several

Table 2.1: Summary of the analyzed MS DSSs articles

Author(s)	Year	Objective				Methodology basis	System in use
		Classification	Diagnosis	Prediction	Treatment		
Esposito and De Pietro	2011	✓				Knowledge-based	
De Falco et al.	2016	✓				DE	
Siddiqui et al.	2015	✓				DWT, PCA, and LS-SVM	
Esposito et al.	2011	✓				Knowledge-based	
Linder et al.	2009		✓			MLR and ANN	CAD tool
Pourakbari et al.	2014		✓			Image processing	
Dogan and Duru	2011		✓			SVM and k-means	
Almasi and Moradkhani	2015		✓			Case-based reasoning and rule-based reasoning	
Veloso	2013			✓		Agent-based simulation model	
SLCMSR et al.	2007			✓		OLAP tool uses matching algorithm	Individual Risk Profile project
Finkelstein et al.	2011				✓		HAT DSS
Veloso	2014				✓	Agent-based simulation model	
Hillert and Stawiarz	2015				✓		SMSreg
Reddel et al.	2019				✓		AMS3 CDSS

electronic databases (see Figure 2.1). After scanning through the articles' abstracts, the list was narrowed down from 154 to 95 articles. After the full-text analysis phase, 25 among the 95 articles were selected based on the inclusion and the exclusion criteria. The overall objective of each of the selected research papers was determined for categorizing them into subgroups. It should be noted that one article could belong to more than one category, but it was categorized under the subgroup that was the most appropriate given the goal of that article. The following subsections outline the main relevant methods and provide examples of how technologies related to DSSs and decision-making processes are implemented for the MS domain.

The Role of DSSs in MS Quest

Healthcare organizations are increasingly in need of DSSs, namely Clinical DSSs (CDSSs), that are

understandable by non-technical audiences such as healthcare providers. CDSSs aid clinical decision-making by providing practitioners with insight knowledge and information about their patients for generating suitable assessments or recommendations (Power, 2002). In the MS domain, the use of DSSs, or CDSSs, is just as important as their use in other healthcare areas. Indeed, the existing DSSs technologies tied with MS, as explained next and summarized in Table 2.1, are for the sake of classifying, diagnosing, predicting, or treating MS.

Classification tasks have motivated a number of experts to implement DSSs particularly to be used in the MS domain. Esposito and De Pietro (Esposito and De Pietro, 2011) developed an ontology-based fuzzy DSS to assist neurologists in classifying MS lesions, i.e., white matter lesion. They performed their study on a dataset that contained brain MRIs of 120 patients between 20 and 63 years with clinically definite MS. The methodology of this DSS relied on a knowledge-based mechanism that integrated ontologies (to elucidate the structure of the knowledge semantically and to formulate clear outputs) and fuzzy logic (to comprehend the dataset's uncertainty and fuzziness) as knowledge representation techniques in order to embed an expert's high-level medical knowledge into the DSS. The DSS combined the obtained knowledge, in terms of fuzzy rules and ontologies, with Fuzzy Inference Ruled by Else-Action, i.e., the FIRE method. As a matter of fact, this methodology comprised three stages: knowledge elicitation, knowledge representation, and knowledge reasoning respectively. Thereby, this DSS was able to classify white matter lesions and to obtain measures of their volumes. The authors argued that their proposed DSS provided better outcomes for patients with large lesions compared to patients with small lesions. They supported their argument by evaluating the performance of this DSS using the area under ROC curve (AUC) and Similarity Measures. The result of the AUC was ranged between 0.82 - 0.87. Using different thresholds (0.25, 0.50, and 0.75), the similarity measures got the following results: 0.72 - 0.97 for the similarity index, 0.67 - 0.97 for the overlap fraction, and 0.01- 0.37 the extra fraction.

De Falco et al. (De Falco, Scafuri and Tarantino, 2016) proposed a DSS that utilized Differential Evolution (DE), an evolutionary algorithm (EA), for automating the classification of potential MS lesions. This work used brain MRIs of 120 confirmed cases of MS. The methodology of this DSS extracted explicit knowledge, a set of explicit IF-THEN rules (i.e., classification rules), from the data in use. Furthermore, the methodology proceeded by finding separately the best set of rules for each class. The best set of rules at the end of the evolution emerged, so the classification here was all about searching for the optimal specification among others. As a result, this study reported a set of rules obtained in the 12th run for fold 3 as the best set of rules for the MS cases. The authors of this research used accuracy (81.21% over the training set, 85.92% over the testing set, and 81.68% over the whole dataset), sensitivity (82.06% over the training set, 88.18% over the testing set, and 82.64% over the whole dataset), specificity (78.69% over the training set, 80.25% over the testing set, and 78.87 over the whole dataset), and AUC (79.00% over the training set, 80.72% over the testing set, and 79.19% over the whole dataset) as an evaluation strategy to back their effort.

Siddiqui et al. (Siddiqui, Reza and Kanesan, 2015) established a design of an intelligent medical DSS for classifying brain MRIs as normal or abnormal. The primary motivation behind this design was to introduce a generalized DSS that can operate efficiently and effectively on various brain MRI datasets associated with neurological disorders. The researchers here ran their DSS against two datasets which consisted of T1-weighted and T2-weighted brain MRIs of 340 right-handed patients diagnosed with major brain disorders including brains affected by MS. In addition, the datasets also covered the patients' demographics and clinical details. This DSS took the advantage of the discrete wavelet transform (DWT), the principal component analysis (PCA), and the least-squares support vector machine (LS-SVM) approaches to secure the goal of this study. The methodology of this work started with utilizing DWT in the feature extraction phase. Then, PCA performed feature reduction. The last step was to train the LS-SVM classifier by using the extracted reduced features. The authors claimed that their DSS classified the human brain as healthy or diseased with promising accuracy. Accordingly,

their experiment yielded better results and outperformed other classifiers regarding sensitivity (100%), specificity (100%), and accuracy (100%). Moreover, the result proved that the DSS has a notable generalization ability.

Esposito et al. (Esposito, De Falco and De Pietro, 2011) implemented an evolutionary-fuzzy DSS to support neurologists by recognizing MS lesions in order to evaluate the health status of individuals affected by MS. They conducted their experiment on the same dataset presented in (Esposito and De Pietro, 2011; De Falco, Scafuri and Tarantino, 2016). Essentially, the methodology of this work composed knowledge representation, knowledge reasoning, and knowledge tuning respectively. Knowledge representation interpreted and encoded the required medical knowledge of experienced clinicians in terms of linguistic variables, linguistic values, and IF-THEN rules. Knowledge reasoning specified a fuzzy inference technique that fitted the structure of the knowledge used for medical inferences. Knowledge tuning adopted DE to tune the knowledge through membership functions optimization for each linguistic variable involved in the rules ultimately to achieve the highest correct classification rate. This system obtained the best outcomes exceeding several classification algorithms compared to it in the study's literature. To aid this finding, the authors of this study evaluated and compared their DSS's accuracy, sensitivity, and specificity with several classification techniques, namely machine learning (ML) algorithms. The study recorded the average results over the 10 folds (accuracy on the training set was 89.10%, accuracy on the testing set was 88.79%, sensitivity was 0.88, and specificity was 0.88) and the results for the best fold in terms of the highest percentage of accuracy on the testing set (accuracy on the training set was 88.71%, accuracy on the testing set was 92.93%, sensitivity was 0.96, and specificity was 0.84). Moreover, the study reported the 10-fold classification accuracy of the proposed system (88.79) and compared it with other classifiers.

Easing the diagnostic procedure has encouraged several researchers to model DSSs. Linder et al. (Linder *et al.*, 2009) discussed proof of principle study by demonstrating the use of a computer-assisted

decision (CAD) support that aimed to diagnose MS patients. The idea here revolved around the ability to distinguish between 73 MS patients, 22 patients of other chronic inflammatory diseases (CIDs) of the central nervous system, and 12 psychiatric patients (control group) in terms of cerebrospinal fluid and blood findings (i.e., standard laboratory findings). In other words, this CAD facilitated MS diagnosis by discovering any significant differences between MS patients and the other two groups of patients (MS vs. CID and MS vs. control group) based on the major parameters of the standard laboratory findings. To obtain the desired results, the authors here made the use of univariate and multivariate analyses using multiple logistic regression (MLR) and artificial neural networks (ANN). MLR categorized patients based on their characteristics while ANN performed feature selection on all parameters of the standard laboratory findings specified in the study. As a result, CAD was able to differentiate between MS patients and the control group. In comparison, CAD lacked the ability to deliver meaningful decision support when differentiating MS and CID patients since it did not disclose common parameters. Sensitivity, specificity, and accuracy assessed the performance of CAD as an eligible DSS. Noteworthy, the study evaluated the parameter sets MLR2 and MLR5 (MLRs with two and five parameters respectively) and the ANN. The ANN was able to perform with 84.9% sensitivity, 54.5% specificity, and 77.9% accuracy when differentiating MS and CID patients. Similarly, The MLR2 and the MLR5 recorded respectively 94.5% sensitivity, 22.7% specificity, and 77.9% accuracy (for MS vs. CID). Furthermore, The ANN distinguished MS and CID patients with 95.9% sensitivity, 66.7% specificity, and 91.8% accuracy. Likewise, The MLR2/MLR5 had 94.5%/95.9% sensitivity, 75.0%/83.8% specificity, and 91.8%/94.1% accuracy (for MS vs. control group).

Pourakbari et al. (Pourakbari *et al.*, 2014) designed a DSS suitable for diagnosing MS as early as possible. The study investigated that the analysis of postural impairment, a type of quantitative movement disorders, was valid for detecting MS even in its early stages for the cause of managing the disease before its severity progression. This study recorded the movement signals of 14 MS patients in the early stages (able to walk without an assistive tool) with an age range of 21-53 years. Also, the

medical examination of this work documented the postural behaviors of 20 healthy subjects with an age range of 20-60 years in order to compare their results with MS patients. This DSS used image processing algorithms to calculate the postural oscillations rate as spatial signals. By obtaining proper features (via statistical analysis) from these signals, this method separated control subjects from patients. However, the authors of this work did not evaluate the performance of their DSS.

Dogan and Duru (Dogan and Duru, 2019) created a DSS for physicians by using image processing, supervised, and unsupervised ML algorithms in order to detect lesions for diagnosing MS. Further, the study compared the functionality of two types of ML tasks (supervised and unsupervised) concerning the objective. The presented techniques analyzed a dataset, collected by (Loizou *et al.*, 2011), containing brain MRIs of 38 MS and clinically isolated syndrome (CIS) patients with average age equal to 34.1. For the methodology, the linear support vector machine (SVM) was the supervised ML algorithm in use while k-means (with k = 4) acted as the unsupervised ML algorithm. The outcomes were acceptable and promising especially for SVM regarding the segmentation process. This is because k-means relied on objective assignment compared to SVM, which benefited from spatial coordinates of data. Calculating the result accuracy for both ML algorithms was the only evaluation mechanism presented in this work (70.24% for k-means and 91.04% for SVM).

Almasi and Moradkhani (Almasi *et al.*, no date) depicted a DSS prototype that aimed to minimize the time required to diagnose MS with the help of artificial intelligence (AI) appropriate techniques. In this work, the authors adopted two AI methods: case-based reasoning and rule-based reasoning. This work appeared to be limited as it lacked the experimental data and the design evaluation that support the researchers' arguments. Predicting the course of MS has pulled the attention of several researchers. For instance, Tintore et al. (Tintore *et al.*, 2015) presented a notable paper that addressed MS prediction by analyzing the most common demographic, clinical, radiological, and biological features that have a strong correlation with the prognosis of MS. This study used a multivariate approach in the experiment

and successfully stated that demographic characteristics, oligoclonal bands presence, and brain MRIs are considered as the impact prognostic factors, ordered from the lowest to highest impact factors respectively. However, this study did not use DSS technology.

Several studies elaborate on the application of DSSs to predict optimally the probability of MS occurrence and progression. Veloso (Veloso, 2013) demonstrated the use of an agent-based simulation model that aimed to aid healthcare providers with a simulation tool. This tool was able to predict long term disability and treatment effects on individuals affected by MS. For testing the methodology, a model that populated 100 virtual agents simulating patients with RRMS was created. Despite that, the validation task of this tool used real data from a group of 50 patients diagnosed with RRMS that lasted for at least 10 years. The dataset used was selected from a total of 173 patients. The author concluded by claiming that this simulation model can be used in everyday clinical practice for monitoring the disability episodes as it might scale for an individual with RRMS over 30 years. Aside from this, the tool was able to perform the treatment effect assessment over the same timeframe. Because of the lack of real medical data for experimental use, evaluating the performance of the model was not presented.

The Sylvia Lawry Centre for Multiple Sclerosis Research (SLCMSR) et al. (the Sylvia Lawry Centre for Multiple Sclerosis Research *et al.*, 2007) presented the "Individual Risk Profile" project that aimed to accurately predict short, mid, and long-range prognosis of MS during the lifetime of an individual affected by any type of this disorder. This project consists of an OLAP-tool that uses a comprehensive database (contained data of 20,000 patients pooled from the academic and corporate sources), that is available to practitioners experienced in MS with an interest in clinical trials for decision-making purposes. To conduct this study, the researchers here derived only the data of 1059 patients from the comprehensive database. The OLAP tool applied a matching algorithm as a strategy to match the patient of interest with all similar patients retrieved from the database. By doing so, the tool was able to predict the course of MS during the life span of the patient of interest by determining the disease course of all

patients in the database that are similar to the patient of interest. The developers of this project argued that this work has potential advantages compared to purely model-based predictors. However, they also discussed that this tool needs improvements as it presented several limitations, so they did not use any evaluation metrics to assess the performance of this tool.

To aid MS patients with treatment decisions, numerous pieces of research have manifested the use of DSSs for this essence. Finkelstein et al. (Finkelstein, Wood and Shan, 2011) discussed the blueprint of the Home Automated Tele-management (HAT) DSS for MS patients. This system took the benefits of the current technologies to provide MS patients with the most convenient therapy and exercise plans during the rehabilitation phase. Furthermore, The HAT system adopted Wagner's model of chronic disease care (Wagner, Austin and Korff, 1996). The designers of this DSS stated that HAT was a successful pilot DSS with promising outcomes for MS patients. The system would enhance the quality of life and the awareness of MS patients by minimizing frequent doctor visits, allowing patients to self-observe their health frequently, educating them on this condition, and guiding them through exercise routines needed during the rehabilitation procedures. To this end, this work observed a standalone system that was tested and evaluated based on end-user opinions. Therefore, the study lacked prior data acquisition, algorithm modeling, and performance evaluation.

Veloso (Veloso, 2014) proposed a web-based computer prognostic simulation model that addressed the needs to start/modify treatment plans, the transformation likelihood of a patient with CIS to definite MS, the long-term prognosis of MS, and the level of disability associated with MS progression. This simulation model applied, reformulated, and extended the simulation model presented in (Veloso, 2013) with the usage of distinct algorithms. This simulation model used two sets of data. At first, the author obtained a dataset from reference studies dealing with the natural history of MS for experimenting on the proposed model. The researcher then used a dataset of 50 patients, who had been living with RRMS for at least 10 years, for validating the simulation process and its result. The study

conductor proclaimed that this model answered the patients' fundamental questions about their current state with MS at various evolutionary stages during the disease course. Lastly, this work had no evidence regarding the performance evaluation of the model.

Hillert and Stawiarz (Hillert and Stawiarz, 2015) presented a review article that demonstrated The Swedish MS registry (SMSreg). SMSreg was developed as a web-based system to help all departments of neurology across Sweden. This system functioned as a decision support tool. Plus, it provided practitioners with patient information needed at clinical visits. SMSreg included data on 14,500 patients and recruited the data of 1000 new MS patients reaching coverage of almost 80% throughout the country. As a decision support tool, SMSreg was valuable as it summarized the information needed to make decisions concerning DMTs. In addition, this tool offered the ability to make decisions by comparing similar patients together. Nevertheless, there was not much to say about the framework of the methodology of this system nor about whether it got evaluated.

Reddel et al. (Reddel *et al.*, 2019) explained the idea of Alemtuzumab in MS Safety Systems (AMS3). AMS3 was developed as a CDSS to determine and organize the MS patients' important care routines such as identifying risks associated with alemtuzumab therapy, scheduling periodic tests, sending reminders when needed, and analyzing test results, just to name a few. The study used a dataset that included a total of 10 patients with active RRMS receiving alemtuzumab treatment. The authors described the system's overall architecture without describing in more detail the structure of the methodology, so this CDSS was evaluated based on its acceptance within the healthcare community. The designers of AMS3 argued that this CDSS accomplished the expected result.

Ultimately, based on the above-mentioned literature analysis, the use of DSSs in the MS field is quite limited due to the rarity of this medical condition. However, decision-making is a critical task in MS practice. Researchers nowadays are applying several technologies, alone or together with DSSs, to

support the process of decision-making in the MS medical routine. To keep up with this matter, the scope of this review was expanded to cover a specific number of technologies associated with decision-making in MS studies. As noticed, knowledge-based and ML approaches often form the basics of the DSSs' methodologies analyzed before. Therefore, the next subsections outline the decision-making process in MS care using ontologies, the knowledge-based approach of interest, and ML algorithms.

The Uses of Ontology in MS Research

The semantic web is an extension of the current World Wide Web that gives the information well-defined meaning, so the contents of the web become both machine-readable and human-readable (BERNERS-LEE, HENDLER and LASSILA, 2001). To make this possible, a key component of the semantic web, i.e., ontology, would offer a structured representation of the semantics that is relevant to one or several knowledge domains (Sánchez *et al.*, 2015). An ontology is an explicit and formal specification of the shared conceptualization of a domain by means of classes, instances, properties, and semantic relationships (Ma *et al.*, 2014).

Capturing and personalizing knowledge (e.g., knowledge about chronic disorders) in formal, simple, powerful, and incremental ways and then applying appropriate reasoning processes to the personalized captured knowledge would be a remarkable finding in the biomedical domain research (Riaño *et al.*, 2012). Such an idea would be phenomenal in biomedicine as it reinforces sharing and reusing medical knowledge among health practitioners for decision-making purposes. This applies to MS as it is a chronic disorder. Existing publications about ontologies that serve the process of decision-making related to MS are recapped next.

Hadzic et al. (Hadzic, Chen and Dillon, 2008) created a Mental Health Ontology (MHO) for deriving knowledge that aimed to prevent, diagnose, and treat and control mental disorders using data mining algorithms for exposing the patterns in mental health data. MHO consisted of subontologies

representing mental disorder types, factors causing a specific type, and treatments suitable for a certain type. According to the findings of this work, MS was a physical factor affecting mental health because it may result in mood disorders.

Alfano et al. (Alfano *et al.*, 2007) developed an ontology and a rule-based system that can automatically measure the load of the brain lesions (especially those that cannot be assessed visually) of MS patients. This was useful in terms of monitoring responses to treatments and studying the level of progression during the course of MS. This work utilized ontology as a knowledge representation model while the rule-based system acted as a reasoner to infer a new set of knowledge. To test the reasoning process, the proposed method used the MRI data of an MS patient. The authors stated that their approach showed greater sensibility as it recognized more lesions compared to an algorithmic procedure.

Jensen et a. (Jensen *et al.*, 2013) originated the Neurological Disease Ontology (ND) that aimed to formally structure common and accurate representation of a variety of neurological diseases with more specifics for MS and Alzheimer's disease for clinical applicability and research purposes. The ultimate goal of ND was to represent each disease along with its genesis (genetically or environmentally), etiology, symptoms, syndromes, progression levels, diagnostic criteria, treatments, and relationships with other neurological disorders as a means to maximizing the potential reasoning capability. Up to the date of this work, this ontology contained nearly 450 classes in addition to over 700 classes imported from external ontologies.

Malhotra et al. (Malhotra *et al.*, 2015) proposed the MS Ontology specified for clinical and translational research related to MS. The MS Ontology used a conceptual hierarchy to represent medical knowledge specific to MS, which were retrieved from scientific literature, database sources, and electronic medical records (EMR). Moreover, this ontology identified a huge amount of data that define the associations between risk factors, molecules, therapies, and several other diseases aiming to improve the societies'

understanding of MS. The authors argued that MS Ontology could acquire a wide range of MS concepts. They supported their argument by claiming that MS Ontology obtained knowledge more accurately compared to PubMed advance searches.

All of these ontologies formulate knowledge bases that represent MS's domain knowledge. Several reasoning and inference engines can make good use of the experts' knowledge captured in knowledge bases to automate the decision-making process, i.e., knowledge-based systems. This would enhance decisively the MS's medical practice and possibly equalize or even outperform the medical justification of qualified MS specialists.

The Utilization of Machine Learning in MS Studies

ML is an AI major discipline that draws attention towards its ability to learn patterns from input data using an increasing variety of algorithms (supervised and unsupervised) dedicated to automating the observation process which overcomes real-world challenges (Wottschel *et al.*, 2015). Supervised ML algorithms train models using determinated prior information (i.e., input data) representing class labels in order to automatically classify new objects or data not seen before (Mateos-Pérez *et al.*, 2018). In contrast to supervised algorithms, unsupervised ML algorithms do not require prior information about the class labels as they train models to discover hidden structures and patterns (i.e., determine class labels) from unlabeled target variables (Lim, Tucker and Kumara, 2017) ML gains momentum in the medical realm for mining and analyzing large collections of medical data (Sweeney *et al.*, 2014). In fact, healthcare organizations, as most public and private organizations, have begun to apply ML as a central phase for analyzing medical knowledge for decision-making purposes (Tuggener *et al.*, 2019).

Notably, MS experts have adopted ML techniques mainly to distinguish MS from other pathologies and to investigate crucial characteristics of MS during its course (Zurita *et al.*, 2018). The vague patterns of MS (e.g., in terms of etiology, progression, clinical presentation, and response to drug therapies)

elevate ML algorithms as the optimal set of tools that automate the recognition of patterns and regularities in MS data. An overview of articles that describe the use of the ML algorithms that yield better decision-making regarding classification, diagnosis, and detection of MS are demonstrated next. In order to keep things simple, it should be noted that this overview is limited to articles of interest published during the last five years.

The success of ML has given the opportunity to pioneer algorithms able to provide a better classification of MS. Zurita et al. (Zurita *et al.*, 2018) developed SVM classifiers able to recognize brain areas (affected by MS) that may assist to better diagnose potential cases of RRMS. This experiment used a dataset containing diffusion tensor imaging (DTI) and resting-state functional MRI (rsfMRI) data of 107 RRMS patients and 50 control subjects. Using 12 well-constructed rsfMRI and DTI based linear SVM learners, the authors here stated that these classifiers reliably discriminated (was able to avoid bias and overfitting) between RRMS patients and control subjects with accuracies up to 89%.

Lopez et al. (Lopez *et al.*, 2018) utilized an unsupervised ML algorithm to cluster MS patients based on their genomic similarity and potentially discover valuable differences among these clusters. This algorithm clustered instances of a dataset that contained DNA samples from 191 MS patients. The methodology of this work employed an agglomerative hierarchical clustering algorithm with multiple linkage methods in order to identify underlying cluster structures with the help of the majority vote approach. In addition, the methodology employed a Silhouette index as an internal validity metric to select the appropriate number of clusters. The outcomes of this study revealed that the methodology presented here was able to identify patient clusters genetically without specifying the number of clusters in advance or indicating any prior input parameter. According to the authors, this methodology outperformed others found in the study's literature regarding overfitting, as it had a significant Rand index greater than the benchmarked methods.

Ion-Mărgineanu et al. (Ion-Mărgineanu *et al.*, 2017) built multiple binary classifiers to automatically differentiate between patients with different MS clinical forms. Namely, the researchers performed nine binary classification tasks for different combinations of MS types. This work used clinical data, lesion loads, and metabolic features of 87 MS patients and 18 individuals served as healthy control subjects. The idea of this work was to compare the outcomes of the linear discriminant analysis (LDA), SVM with a radial basis function kernel (SVM-rbf), and random forest (RF) models. The results of this work showed that SVM-rbf, trained on clinical data and lesion loads, was the best classifier for distinguishing CIS from RRMS (F1-score = 71) or RRMS plus SPMS (F1-score = 72). Nevertheless, LDA, trained with clinical data, performed better when discriminating RRMS from PPMS (F1-score = 85) or SPMS (F1-score = 84).

Wang et al. (Wang *et al.*, 2018) aimed for segmenting MS lesions and non-specific white matter (NSWM) lesions separately based on their shape and spatial location features by adopting a spherical harmonics descriptor using an ML pipeline. To perform the experiment, the authors obtained two datasets. The first dataset contained 234 MS lesions and 190 NSWM lesions. The second dataset included 160 MS lesions and 119 NSWM lesions labeled by location. The authors trained three different ML models: logistic regression (LR), SVM, and extreme gradient boosting (XGBoost). The authors continued by arguing that the proposed pipeline successfully classified MS and NSWM lesions with good accuracy (70.52% - 87.97% for the logistic regression, 70.29% - 74.89% for the SVM, and 85.58% - 90.43% for the XGBoost) and AUC (83.76% - 95.42% for the logistic regression, 70.49% - 87.01% for the SVM, and 93.45% - 96.43% for the XGBoost).

Automating the process of differentiating stable from potentially evolving MS cases is a research topic highly in demand. Salem et al. (Salem *et al.*, 2018) integrated intensity and deformation-based approaches for automatically detecting new T2-w lesions. The study used a dataset consisted of images from 60 different patients with the CIS or early relapsing MS with 36 of them with confirmed MS cases

due to the appearance of new T2-w lesions in their scans. This work used a deformation-subtraction based logistic regression model, i.e., a logistic regression algorithm that adopts the deformation field (DF) aspect, to detect new T2-w legions inside the white matter region. The authors declared that there was a significant difference in the model's performance when including DF as it turned out that it improved the model's accuracy. In fact, the combination of DF and logistic regression helped to boost the performance when detecting new T2-w lesions. To uphold this finding, the researchers compared their model with two state-of-the-art approaches and three variants of their model with fewer features. Their full model outperformed all the other models and had the best values for all the evaluation measures (sensitivity (74.30±28.70), specificity (11.86±18.40), and dice similarity coefficient (0.77±0.23 for detection and 0.56±0.23 for segmentation)) except when detecting very small lesions.

Zhang et al. (Zhang *et al.*, 2016) proposed a study that compared the performance of three ML algorithms with the intention to detect MS in the brain by using stationary wavelet entropy (SWE). The authors conducted their experiment on a dataset that included brain images of 38 MS patients and several healthy control subjects (the population of healthy control is not specified). The authors applied three ML algorithms: decision tree (DT), k-nearest-neighbor (KNN), and SVM. The experiment recorded KNN as the best performer in terms of specificity (99.32%), precision (99.09%), and accuracy (97.94%), while the SVM performed the best in sensitivity only (97.34%). In contrast, the evaluation results of DT were the worst in all of the four measures. Thus, KNN yielded the best classification performance among the three algorithms in this detection process.

McGinnis et al. (McGinnis *et al.*, 2017) proposed a technique for estimating walking speed using a wearable device. The researchers of this proposal used accelerometers worn in several body locations to characterize gait speeds. To compare their mobility capabilities, this work recruited 30 subjects diagnosed with MS and 7 healthy controls. The authors utilized SVM models to measure walking speed features indicated from the wearable accelerometer. Additionally, the authors analyzed the correlation

between speed estimation accuracy and device location combinations. They clarified that placing additional accelerometers in proximal locations would improve the accuracy of estimating the gait speed. The authors concluded their observation by claiming that there was a high correlation between the ground truth and estimated speeds during comfortable walking tests.

To better understand patterns that may underlie cardinal factors of MS, several recent studies strive to infuse ML into MS research. In light of this fact, ML algorithms are adopted excessively to automatize MS's practice. The abovementioned studies show how ML benefits MS research. It is important to mention that ML studies conducted for MS prediction (e.g., predicting MS progression), treatment, and diagnosis are beyond the scope of this review because they will form the baseline of future work.

Discussion

In the preceding sections, MS studies were reviewed and analyzed about their usage of DSSs, ontology, and ML. Each of these disciplines, along with their underlying technologies, has certain benefits and drawbacks yielding the applicability of automating MS diagnosis, detection, treatment prescription, classification, and prediction. By virtue of its nature as a preference-sensitive condition, MS specialists and patients participate more intensively in MS decisions, particularly regarding diagnosis and treatment. Generally speaking, the MS diagnostic procedure imposes all MS specialists to obey the guidance of McDonald Criteria (McDonald *et al.*, 2001) in addition to performing the Expanded Disability Status Scale (EDSS) (Kurtzke, 1983) and clinical examinations (e.g., MRI and lumbar puncture) to confirm MS cases. MS specialists usually detect MS activities by comparing initial diagnosis reports with follow-ups reports. Nevertheless, these procedures require intensive knowledge and experience given the inconclusiveness in this medical condition due to the lack of consensus diagnostic procedures. Furthermore, the variety of MS drug therapies offer a range of potential benefits, but they may also tolerate life-threatening risks. In like manner, predicting and classifying MS during

and before the course of the disease are very challenging due to the ambiguity in terms of MS's progression and occurrence pattern. To fulfill the need for automated systems that could help to overcome these gaps, several pieces of research, stated in the previous section, manifest the use of modern technologies, e.g., DSSs, for this essence with remarkable findings and high-performance metrics.

Knowledge acquisition and representation of an expert is critical in developing reliable knowledge-based DSSs used for MS clinical routines. For instance, studies in (Esposito and De Pietro, 2011; Esposito, De Falco and De Pietro, 2011) demonstrated the use of fuzzy logic to handle the uncertainty in MS. This, in turn, would accurize solid knowledge representation to perform more rational knowledge reasoning that is able to make valid inferences. Significantly, the work presented in (Esposito and De Pietro, 2011) integrated two knowledge representation approaches, namely ontology in addition to fuzzy logic. The importance of applying ontology here was to provide a shared understanding of the MS domain, i.e., semantic interoperability. This mixture led to knowledge elicitation, knowledge representation, and knowledge reasoning and inference with reduced uncertainty. Similarly, the methodology proposed in (Esposito, De Falco and De Pietro, 2011) combined fuzzy logic with DE to represent, reason, and tune knowledge. DE optimizes a complex problem by improving a candidate solution iteratively, so it finds the best set of rules that guarantees the best set of knowledge. This combination can obtain the most optimal result because it leverages MS uncertainty using the best set of knowledge and rules. It is worth to mention that the researchers in (De Falco, Scafuri and Tarantino, 2016) formulated a DSS using DE that obtained explicit knowledge through an optimal set of rules. In order to find this set of rules, each class has its own rules that are used to recognize the class's instances. At the end of the DE evolution, the optimal set of rules, that is used for classifying instances to their corresponding classes, emerged to form the best classification specification. Similarly, the study conducted in (Esposito, De Falco and De Pietro, 2011) used the same DE mechanism demonstrated in (De Falco, Scafuri and Tarantino, 2016) in the knowledge tuning

phase. In fact, DE was applied to achieve the highest correct classification rate by tuning the knowledge via membership functions optimization for each linguistic variable involved in the rules. Remarkably, these pieces of research performed well in classifying MS lesions with significant evaluation metrics results.

Two studies that used simulation models were demonstrated in this review. The study presented in 28 (Veloso, 2013) used an agent-based simulation model that exhibited the disability and the treatment effects prediction. This model virtually populated 100 agents that simulated patients with RRMS. The beauty of this work is its ability to abbreviate 30 years of monitoring and observing the quality-of-life of RRMS patients. By the same token, the work presented in (Veloso, 2014) extended the previous methodology with a web presence and the utilization of distinct algorithms. This model's functionality was extended as it considered other forms of MS (CIS and SPMS), suggested to start/modify treatment plans, and evaluated medical prognosis in the long-term. After all, using such simulation systems in MS clinical practices would allow clinicians to prompt potentially several rapid and appropriate medical interventions before any complications arise in the medical status of actual patients.

Certainly, DSSs need to be implemented in such a way to highly simulate MS specialists. To optimistically reach this objective, the underlying structure of DSSs should be able to learn new patterns by observing subsets of data to produce reliable decisions without human intervention. This goal would be possible with the help of ML algorithms. The CAD tool presented in (Linder *et al.*, 2009) used ANN and LR. This work showed promising results. The LR model was able to discriminate MS patients based on their features while the ANN selected features that are correlated mostly with the MS. Likewise, the study in (Dogan and Duru, 2019) compared SVM with k-means in the lesion detection task. This comparison was in favor of the SVM, although both algorithms performed the segmentation process acceptably with promising results. However, both studies need improvements in terms of applying the most suitable ML model by comparing the results of different ML algorithms. In a word,

it should be noted that each ML algorithm performs differently depending on the problem and the dataset in use, so comparing the performance of several ML algorithms should be sufficient when adopting ML as the ground solution.

Commonly, the SVM algorithms have been used extensively in several studies presented in this review whether as a DSS basis or as standalone models. The studies demonstrated in (Siddiqui, Reza and Kanesan, 2015; Zhang *et al.*, 2016; Ion-Mărgineanu *et al.*, 2017; McGinnis *et al.*, 2017; Wang *et al.*, 2018; Zurita *et al.*, 2018; Dogan and Duru, 2019) relied entirely or partially on SVMs despite the fact that they are used for mutual or different objectives. For instance, the linear SVM was applied in the DSS of (Dogan and Duru, 2019) and in the experiment of (Zurita *et al.*, 2018) for diagnosing MS. On the contrary, the study established in (Ion-Mărgineanu *et al.*, 2017) used the SVM with rbf kernel to segregate patients with different MS clinical forms. Furthermore, the standard SVM model was used as a part of an ML pipeline structured in (Wang *et al.*, 2018) and as an individual model in (Zhang *et al.*, 2016) for detecting MS. Moreover, the study conducted in (Siddiqui, Reza and Kanesan, 2015) utilized the LS-SVM classifier to be able to classify brain MRIs as normal or abnormal. In addition, the study proposed in (McGinnis *et al.*, 2017) used the SVM model to measure walking speed features in a wearable device, which is used to characterize gait speeds to compare MS patients' mobilities.

The comparison between all of the SVMs presented in this review seems imbalanced for two reasons. First, they are used for different objectives and with different datasets. Second, they are used partially as one of the essential tools of a DSS paradigm or entirely as an independent method. However, the performances of these SVMs can be slightly explored despite the above contradictions. To emphasize this, the LS-SVM was the best SVM among the others because it guarantees higher evaluation rates with minimum computation time and complexity even when it is running against huge datasets. The LS-SVM is an enhanced, a reformulated, and an upgraded version of the classical SVM. LS-SVM ensures more accuracy by using least-squares to modify and correct the classifier's behavior in order

to minimize the errors i.e., cost function. As stated in (Siddiqui, Reza and Kanesan, 2015), the performance of the model used in the study with the LS-SVM classifier that used rbf kernel exceeded all other models especially those that applied standard SVM with different kernel values. Yet, exploring and comparing the performance of several sets of parameters within LS-SVM is something that needs to be considered in future works that adopt LS-SVM in order to have the most optimal model performance.

It could be inferred from this review that the number of studies conducted to address the usage of DSSs in the MS field is quite limited. Indeed, the acceptance of DSSs within the MS domain remains limited. This condition is not getting proper attention compared to other incurable diseases, such as Alzheimer's disease due to its uncommonness and data scarcity. Nonetheless, the direction of the future work should incline toward applying DSS technologies, and potentially knowledge graph techniques, that are able to understand MS progression and occurrence patterns. Additional work should also adopt these machine-based approaches to emphasize MS etiology and MS’s long-terms effects on the quality-of-life of the affected individuals. Moreover, the correlation between MS and other disorders (especially chronic neurological and autoimmune diseases) should be investigated. In spite of this, MS's intended researchers require an extensive amount of data, but access to them is very restricted. To overcome this issue, the FAIR principle for MS data (Peeters, 2018) should be considered in the near future. Considering these recommendations would enhance the MS’s clinical practice experience.

Chapter 3 Examination of Supervised Machine Learning Classifiers and Ensemble Learning for Predicting the Type of Multiple Sclerosis

Introduction

Multiple sclerosis (MS) is the most common inflammatory demyelinating autoimmune disorder that chronically affects the central nervous system (Sospedra and Martin, 2005; Lassmann, Brück and Lucchinetti, 2007; Chiaravalloti and DeLuca, 2008; Milo and Kahana, 2010). Globally, MS affects millions of middle-aged and young adults especially females, who are more likely to develop MS than males (World Health Organization and Multiple Sclerosis International Federation, 2008; Kragt *et al.*, 2009; Shanmuganathan *et al.*, 2020). MS can cause multiple independent cognitive, motor, and neuropsychiatric symptoms (Chiaravalloti and DeLuca, 2008; Brownlee *et al.*, 2017). The known reported symptoms of MS include abnormal sensations (paresthesia or tingling), gait impairment and spasticity, weakness, pelvic-organ dysfunctions (bladder, bowel, and sexual dysfunctions), vertigo, fatigue, anxiety, frequent falls, sleep disorder, heat/cold intolerance, visual impairment, pain, Lhermitte's sign, migraines, and depression, to name a few (Milo and Kahana, 2010). As a result, the quality of life regarding physical functioning, emotional health, and social life may be significantly affected or disrupted (Col *et al.*, 2019; McGinley, Goldschmidt and Rae-Grant, 2021).

MS is considered a complex disease because the precise causes of MS are still unknown. Although studies show that genetic and environmental factors such as vitamin D deficiency, Epstein-Barr virus, radiation, obesity, and smoking might potentially cause MS (Trapp *et al.*, 1998; Leray *et al.*, 2016; Reich, Lucchinetti and Calabresi, 2018; Dobson and Giovannoni, 2019; McGinley, Goldschmidt and Rae-Grant, 2021), vaccines, traumatic and stressful events, and allergies have not been classified as risk factors (Leray *et al.*, 2016). MS patients usually get the disease in different episodes, ranging from mild to extremely aggressive forms: MS usually begins as a clinically isolated syndrome (CIS), progresses to relapsing remitting MS (RRMS), progresses to secondary progressive MS (SPMS), and

then advances to primary progressive MS (PPMS) or progressive relapsing MS (PRMS) progressions in some rare cases (Lublin *et al.*, 2014; Mahad, Trapp and Lassmann, 2015; Efendi, 2016; deAndrés-Galiana *et al.*, 2019; Seccia *et al.*, 2020). In addition, Okuda et al. (Okuda *et al.*, 2009) defined a recent terminology describing patients without typical MS symptoms as a radiologically isolated syndrome (RIS). Another study suggested that the risk of disease development and progression is determined by factors related to immune function and activation (Mahad, Trapp and Lassmann, 2015). All things considered, it is believed that there is a lack of awareness and education about MS among healthcare practitioners, patients, and the general public (Rieckmann *et al.*, 2013). As stated in (Confavreux and Vukusic, 2006), "There may be much to be learned on this topic from detailed scrutiny of the natural history of the disease."

Since there is no sufficient single clinical feature or diagnostic test to diagnose MS, diagnostic criteria have included a combination of clinical and paraclinical practices (McDonald *et al.*, 2001). Typically, MS specialists detect MS activities by comparing initial diagnosis reports with follow-up reports, which requires extensive knowledge and experience (Alshamrani *et al.*, 2020). Nevertheless, the severity and progression of MS over time in any affected patients cannot yet be determined since the course of MS has high variations and is unpredictable (Pérez del Palomar *et al.*, 2019; Saccà *et al.*, 2019). This unpredictability would cause MS patients and specialists to have frustration and anxiety while trying to come up with useful shared decisions (Yperman *et al.*, 2020). Even for experienced MS specialists, establishing an accurate clinical course description for MS is difficult, but it is critical for decision-making (Fiorini *et al.*, 2015). Hence, MS course identification takes a long time for MS specialists and patients as they must endure stressful examinations to confirm MS cases and progression levels (Fiorini *et al.*, 2015). Given the clinical heterogeneity of this disorder, it is critical to determine the progression level of MS in affected patients (Seccia *et al.*, 2020). Therefore, predicting the type of MS would be extremely useful in making suitable decisions, such as tailoring therapy to a patient's specific needs (Seccia *et al.*, 2020). However, it is difficult to predict the course of MS given demographic and baseline

clinical data, and there are currently no proven prediction techniques to determine the progression of MS (Zhao *et al.*, 2017; Seccia *et al.*, 2020). Moreover, predicting MS disease course is challenging as MS progression and occurrence patterns are ambiguous (Alshamrani *et al.*, 2020).

To fulfill the need to overcome the gaps highlighted above, an automated system that simulates MS specialists and learns new patterns by observing fitted data without human intervention is needed. This could be done with the utilization of machine learning (ML) algorithms. For mining and analyzing vast volumes of medical data, ML is gaining attention in the medical field (Sweeney *et al.*, 2014). Healthcare organizations are starting to use ML as the primary method for assessing medical knowledge for decision-making (Tuggener *et al.*, 2019). Due to the ambiguous nature of MS (including its etiology, progression, clinical manifestation, and response to treatment), ML algorithms are the best tools for automatically identifying patterns and regularities in MS data (Alshamrani *et al.*, 2020). Therefore, the main aim of this chapter is to explore different supervised ML models in predicting the progression levels (MS types) in individuals affected by MS and comparing the performance of each model against a real-world MS dataset. A framework that compares the outcomes of different traditional and ensemble models to address the proposed research objective is presented.

The remainder of this chapter is organized as follows: Section 2 briefly presents the related work of ML algorithms in the MS domain, specifically in predicting MS progression. Section 3 explains the proposed methodology. Section 4 describes the results. Section 5 discusses remarkable outcomes and trends. Finally, Section 6 concludes the study.

Related Work

In this section, the main focus is on previous work demonstrating the application of different supervised ML models in predicting the MS's progression level. Note that the following emphases of research

papers that proposed different ML paradigms were excluded: image processing classifiers such as the ones used for determining MS lesions, ML models implemented for differentiating MS patients from healthy control, ML predictors that estimate the short- or long-term effects or disabilities caused by MS, the application of ML in research related to MS genetics, and the use of unsupervised ML algorithms in MS domain. According to the literature, several studies used ML techniques to classify MS patients based on MS type. For instance, Taschler et al. (Taschler *et al.*, no date) worked on classifying MS types using support vector machine (SVM). The researchers used data containing information about quantitative features derived from magnetic resonance imaging (MRI) images in addition to demographic and clinical data. The proposed model achieved 56.0% accuracy on the entire dataset, while yielded 39.8% accuracy on the demographic and clinical data without MRI images. Likewise, Ion-Margineanu et al. (Ion-Mărgineanu *et al.*, 2017) attempted to categorize MS patients according to the form of the disease based on MRI data, demographics and clinical data, and metabolic variables. This task was done by comparing linear discriminant analysis, SVM, and random forest (RF) classification results. For CIS vs. RRMS and CIS vs. RRMS + SPMS, respectively, the use of the SVM against clinical data and MRI information achieved 71.00-72.00% accuracy.

Moreover, Karaca and Hayta (Karaca and Hayta, 2015) presented a framework to identify the type of MS (specifically RRMS, SPMS, and PPMS) by comparing MRI data and the Expanded Disability Status Scale (EDSS) score of MS patients. Artificial neural networks algorithm, fitted with MRI data only, was able to determine the type of MS with 98.90% accuracy. The same model was able to predict with 99.90% accuracy when trained with MRI and EDSS data. Furthermore, the work proposed by Ekşi et al. (Ekşi *et al.*, 2020) outlined a methodology that discriminates between RRMS and SPMS, mainly through a binary classification. The central ML model utilized in this study was SVM, which was used to predict the MS type using magnetic resonance spectroscopy data. The outcomes of this study showed that RRMS and SPMS patients were differentiated with 83.33% accuracy. Aside from the above research, Fiorini et al. (Fiorini *et al.*, 2015) demonstrated an ML pipeline for analyzing clinical

questionnaires to detect the progression of MS, particularly RRMS vs. other progressive and benign forms. The dataset contains information about patients' mobility, fatigue, cognitive performance, emotional status, bladder continence, and quality of life. The authors explored the performance of several linear classifiers on the described binary classification problem. It was found that logistic regression (LR) outperformed the other classifiers on the entire feature set (77.28%). In comparison, ordinary least squares method was the best approach when a subset of the features was selected (78.32%) compared to regularized least squares, linear SVM, and k-nearest neighbor (KNN).

From a data science perspective, this research area is not yet enriched despite the previously proposed methodologies being solid contributions to the MS domain. To date, ML is still not fully utilized in clinical practices due to the limited confidence of MS specialists in predictive technologies and the insufficient collaboration between data and computer scientists and MS specialists (Obermeyer and Lee, 2017; Lynch and Liston, 2018; Wiens *et al.*, 2019; Seccia *et al.*, 2020). To overcome this limitation and to address this gap, the following demonstration of the proposed framework would potentially outperform the aforementioned related work. In short, the proposed work will broadly classify different MS types and benefit from using sufficient and more extensive real-world MS data. With this in mind, the presented framework leverages the functionality and performance of different traditional supervised classifiers and ensemble learners to develop a satisfactory ML ground solution.

Materials and Methodology

Data Description

The dataset obtained for this study represents the iConquerMS™ initiative founded by the Accelerated Cure Project for MS (ACP), a nonprofit organization dedicated to developing and sharing resources needed for MS research. ACP started recruiting participants since the time around 2010. The data is a clinical baseline data collected from a questionnaire-based repository. This collection is driven by MS

patients who participated in this initiative by submitting questionaries that contain information related to their MS cases. The repository keeps documents that record participants' unidentifiable demographics, medical history, MS history and symptoms, quality of life, wellness and dietary habits, overall physical activities and health self-scales, and current and past DMTs. The participation inclusion and exclusion criteria are found here https://www.iconquerms.org/swms-inclusion-requirements. The number of participants in the raw dataset is 8329, and each participant's clinical record contains up to 247 attributes. Note that this is the actual number of applied features after initially removing identifiable variables and attributes with higher degrees of freedom (many categorical values) to ensure that the dimensionality of the data is not increased.

Ethics

The legitimate use of the dataset is for research purposes and was authorized by the Office of Research Assurances of the University of Idaho through an approved Human Research Protections (IRB) application (Authorization 21-235, dated February 1st, 2022). Please refer to Appendix F for more information about the IRB.

Data Preprocessing Phase

Three important steps were taken into account to perform practical feature engineering. At first, features were assessed based on their importance in determining MS types in two rounds. Among all features presented in the dataset, the first round of feature filtering was done from a neurological perspective after consulting an MS specialist. The total number of features selected after the first round, in addition to the target variable, was 214. Then, an automated feature selection approach was used to narrow the number of features further. To perform the second round of feature selection accurately, we used a subset of the filtered dataset that includes records with no missing values. This was a key step to ensure the automated feature selection tool would work as expected. For this quest, selecting features was done

1. Train a base ML estimator on the entire dataset.
2. Rank the importance of each feature based on the contribution to the performance of the model
3. Select the top-ranked features based on a specified threshold
4. Create a subset of dataset with only the selected features.
5. Train a new ML model on the new dataset with only the selected features

Figure 3.1: SelectFromModel algorithm

based on importance weights using SelectFromModel, a scikit-learn algorithm described in Figure 3.1, after training its base RF estimator against the completed subset.

Since any real-world raw dataset is overwhelmed by records with missing values (Seccia *et al.*, 2020), it was critical to handle incomplete records before using ML models by applying missing value imputation. Hence, missing value imputation was performed using a model-based imputation method to impute missing records after the dataset's dimensionality got reduced based on the selected features. Model-based imputation techniques that use ML algorithms have been outperforming other statistical imputation techniques (Cevallos Valdiviezo and Van Aelst, 2015; Pan *et al.*, 2015; Silva-Ramírez, Pino-Mejías and López-Coello, 2015; Conroy *et al.*, 2016; Raja and Thangavel, 2020; Bertsimas, Pawlowski and Zhuo, no date). The imputation here was done using KNN imputer with hyperparameter tuning. By using this imputation technique, each instance's missing values were imputed using the mean value from n_neighbors nearest neighbors. Lastly, note that the categorical independent and target variables were encoded using OrdinalEncoder and LabelEncoder, scikit-learn encoding preprocessors, respectively. These procedures were essential to make the model training/fitting and prediction faster with potentially better results and effective performance. Indeed, feature selection will eliminate several unrepresentative features, while missing values imputation will ensure that ML algorithms can effectively analyze the complete dataset (Liu *et al.*, 2020). Most recent medical research in the MS domain use RF algorithm for feature selection and KNN algorithm for missing values imputation (Suganthi and Karunakaran, 2019; Liu *et al.*, 2020; Cui, Hu and Liang, 2021; Sharma and Kaur, 2021;

Awawdeh, Faris and Hiary, 2022; Christo *et al.*, 2022; Faisal and Tutz, 2022; Kabir and Farrokhvar, 2022; Nagarajan and Dhinesh Babu, 2022; Seba and Benifa, 2022), which is why these approaches were adopted in this analysis.

It is worth mentioning that records of patients who did not specify their MS class were removed prior to feature selection and missing value imputation to avoid model biases when imputing missing values of the target feature. A total of 3442 patients have verified the current status of their MS in terms of the stage of their conditions. Thus, the classification models demonstrated next were trained with 3442 instances (2147 RRMS, 675 SPMS, 392 PPMS, 128 CIS, 8 RIS, and 92 were not sure).

Prediction Models and Evaluation

The main objective of this proposed methodology is to investigate the best traditional or ensemble classifier that categorizes patients into RRMS, SPMS, PPMS, CIS, and RIS groups against the aforementioned historical data. We consider the following traditional ML approaches: RF, SVM, KNN, LR, decision tree (DT), naive Bayes (NB), perceptron, and stochastic gradient descent (SGD). Furthermore, we examined voting, bagging, gradient boosting, AdaBoost, stacking, and extreme gradient boosting (XGBoost) ensembles. The dataset was split into training and testing subsets using five different splitting techniques. Then, we compared the performance of the applied models using accuracy, sensitivity, and specificity. We used these evaluators to measure the predictions' accuracy and value as these metrics are widely used in medical domains (Sharma and Kaur, 2021; Faisal and Tutz, 2022; Kabir and Farrokhvar, 2022; Nagarajan and Dhinesh Babu, 2022). These evaluation strategies are computed based on the results of indicators reported in the confusion matrix. Confusion matrix indicators report predictions and real occurrences as True Negative (TN), False Negative (FN), False Positive (FP), and True Positive (TP), respectively. Hence, accuracy denotes the ratio of correctly classified instances as follows:

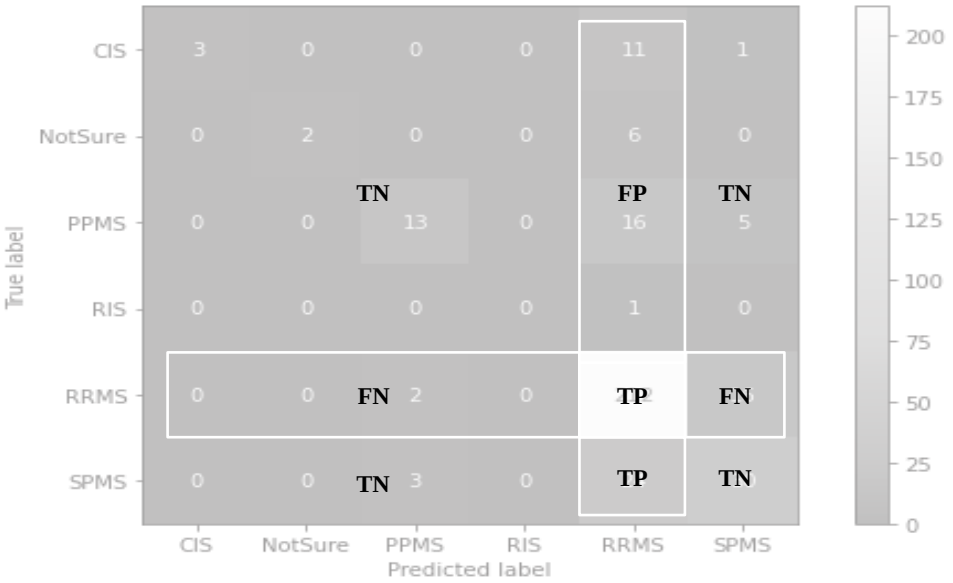


Figure 3.2: Multiclass Confusion Matrix

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

On the other hand, sensitivity tests a model's ability to determine the exact correct class for each instance. This is known as the TP rate, which is the ratio of all positive samples that were accurately predicted as positive by the classifier and calculated as follows:

$$Sensitivity = \frac{TP}{TP + FN}$$

Besides, specificity gives the TN rate, representing the ratio of all negative samples accurately predicted as negative by the classifier. Specificity checks the model's ability to not classify an observation to incorrect classes, and it is calculated as follows:

$$Specificity = \frac{TN}{TN + FP}$$

In a multiclass classification problem, the confusion matrix is usually represented as an N x N matrix, where N is the number of classes. The diagonal of the matrix represents the number of correct predictions for each class, while the off-diagonal elements represent the misclassifications. Figure 3.2 demonstrates the confusion matrix of the six classes presented in this study, where the demonstration

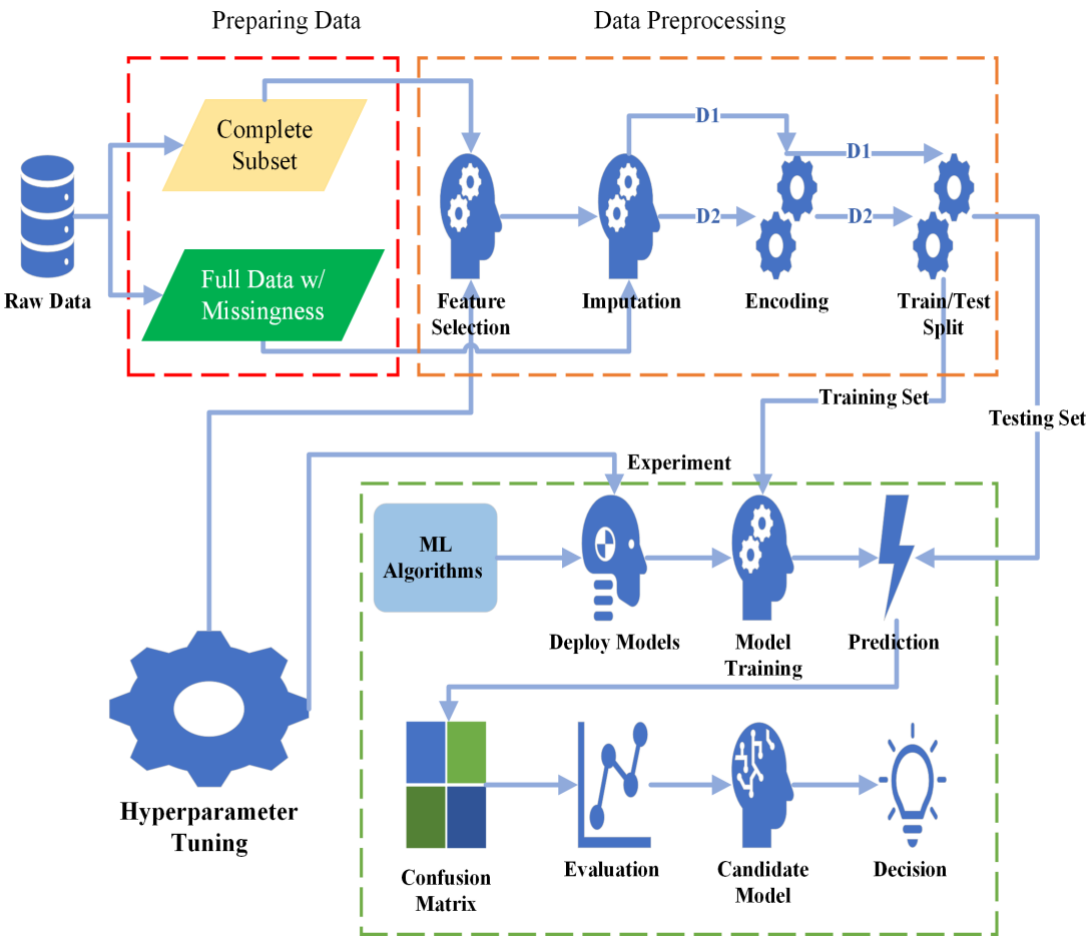


Figure 3.3: Framework design for predicting MS type

focuses on RRMS as an example for more clarification. In such a case, the above metrics are calculated as follows: Accuracy is calculated by dividing correct predictions by total predictions, while sensitivity and specificity results are recorded through macro averaging calculation. Macro averaging is a technique used in data analysis and ML to evaluate the performance of a classification model. It involves calculating the average performance of the model across multiple classes or categories. To perform macro averaging, the performance metric for each class in the dataset is calculated first. Then, the average of these metrics is calculated across all the classes. This gives an overall measure of the model's performance across all the different classes.

Comprehensively, the following demonstrates the research pipeline structured to discriminate MS patients based on the MS type. The analysis was performed using the open-source package scikit-learn on Python 3.7. First, the most important features were selected through SelectFromModel algorithm, where the base estimator was RF model with hyperparameter tuning trained against a subset of data with no missing values. Then, KNNImputer with hyperparameter tuning was used to impute the missing values of the selected features. Next, all traditional models mentioned above were trained, with default parameters' settings, using the imputed dataset and then evaluated. The top classifiers were selected for further investigation. Afterward, the ensemble learning methods mentioned before were trained with the imputed dataset using default parameters. Note that voting, bagging, AdaBoost, and stacking ensembles invoked the top traditional models as their fit method estimators. Models' training was done using three test:train rules (10:90, 20:80, and 30:70) and two cross validation (CV) methods (10-folds and leave-one-out CV "LOOCV"), respectively. All models were compared via the three evaluation metrics described above to determine the best classifier. At last, hyperparameter tuning was used to try to enhance the best model to get better evaluation metrics results. All things considered, RandomizedSearchCV, a scikit-learn method that searchers for hyperparameters, was used in steps that needed hyperparameter tuning. Figure 3.3 elucidates the overall architecture of our prediction framework.

Results

Feature Selection and Missing Value Imputation

Feature selection using SelectFromModel was done using RF, DT, and ExtraTreesClassifier as the base estimators embedded in SelectFromModel function. SelectFromModel algorithm with RF model as the base estimator (using 10:90 rule) was used to perform feature selection, which enabled the study's classifiers to perform better. As a result, the model weighted features based on their importance to the target feature. Out of the 214 features chosen based on the MS specialist's opinion, 78 variables,

including the response variable, were selected as they scored the highest weights (Appendix A lists the most important features along with their brief description, and Appendix B includes the order of the features based on their importance). Here, RF estimator was implemented with hyperparameter tuning. RandomizedSearchCV specified the following parameters as the best set of settings: the number of trees in the forest is 100, the maximum depth of a tree is 100, the minimum number of samples required to split an internal node is 10, the minimum number of samples required at a leaf node is 2, the number of features to consider when looking for the best split is specified as 'sqrt', and bootstrap is false to indicate that the entire dataset is used to build each tree. Besides, KNNImputer used the following parameters to impute missing values of the newly selected features: 35 as the number of neighboring samples to use for imputation and 'distance' as the weight function used in prediction.

Traditional Classification

The outcomes resulting from various classifiers when applying different splitting approaches are summarized in Table 3.1. By observing the results, RF model outperformed all other models in terms of accuracy in all splitting scenarios (greater than 70%), where the maximum accuracy was 75.36%. As for sensitivity scores, DT was the best in three out of five tests, with the highest recorded sensitivity at 51.06%. Besides, DT also achieved the highest specificity in three out of five experiments at 91.64%. The highest scores were obtained when splitting the data to 10:90. Aside from the above results, it is worth noting that SVM and LR came after RF in accuracy results. Similarly, NB got decent sensitivity results in the two other tests where DT performed poorly. Nevertheless, SVM got low results when it comes to sensitivity, while NB got the lowest accuracy results among all models. On the contrary, LR and DT specificity results were sufficiently comparable to RF. Thus, RF, DT, and LR were the top three classifiers nominated for further investigation.

Table 3.1: MS types prediction outcomes using traditional models.

Model	10:90			20:80			30:70			10-K CV			LOOCV		
	Accuracy	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*
KNN	64.06%	33.81%	88.16%	61.83%	28.43%	87.78%	61.96%	28.43%	87.78%	59.88%	28.11%	86.97%	63.22%	30.93%	88.10%
SVM	72.17%	31.72%	88.65%	69.81%	28.98%	88.68%	70.96%	28.98%	88.68%	69.09%	26.40%	88.08%	70.89%	29.20%	88.73%
LR	70.72%	35.37%	89.20%	69.09%	31.25%	89.25%	69.70%	31.25%	89.25%	68.45%	32.97%	89.22%	69.73%	33.38%	89.48%
DT	70.14%	51.06%	91.64%	65.31%	44.23%	89.80%	62.63%	44.23%	89.80%	59.41%	33.68%	88.50%	64.58%	38.51%	90.02%
RF	75.36%	38.01%	90.05%	73.88%	34.44%	90.40%	75.12%	34.44%	90.40%	70.92%	31.07%	88.97%	74.78%	35.17%	90.34%
NB	43.77%	28.19%	87.46%	43.83%	30.30%	87.85%	46.56%	30.30%	87.85%	44.65%	37.42%	88.01%	44.60%	38.56%	88.10%
Perceptron	54.20%	34.46%	89.14%	61.10%	32.08%	88.22%	59.44%	32.08%	88.22%	58.72%	29.14%	88.31%	59.09%	32.14%	88.54%
SGD	67.83%	36.16%	89.10%	65.31%	23.92%	86.36%	67.67%	23.92%	86.36%	62.26%	25.76%	87.07%	62.58%	32.81%	88.46%

Table 3.2: MS types prediction outcomes using ensemble models.

Model	10:90			20:80			30:70			10-K CV			LOOCV		
	Accuracy	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*	*Accuracy*	*Sensitivity*	*Specificity*
AdaBoost	75.36%	47.35%	91.47%	73.00%	42.92%	90.06%	74.54%	35.62%	90.65%	70.80%	34.51%	89.02%	74.46%	38.92%	90.33%
Bagging	75.94%	44.64%	91.24%	74.02%	36.85%	90.19%	75.51%	40.48%	90.74%	70.60%	34.55%	89.67%	74.29%	39.04%	90.78%
Stacking	74.78%	42.06%	90.83%	72.42%	36.06%	90.40%	75.61%	39.01%	91.36%	71.56%	37.35%	90.02%	74.11%	39.16%	90.87%
XGBoost	73.62%	42.67%	90.54%	71.84%	39.61%	90.19%	73.77%	38.63%	90.95%	70.28%	29.24%	88.50%	74.06%	33.46%	89.83%
Gradient	73.62%	46.98%	90.80%	72.28%	40.97%	90.81%	74.44%	42.36%	91.51%	71.67%	39.13%	90.26%	74.49%	41.76%	91.26%
Voting	75.36%	43.08%	90.61%	72.57%	36.65%	90.57%	73.18%	36.84%	90.74%	70.48%	34.00%	89.54%	74.43%	38.12%	90.90%

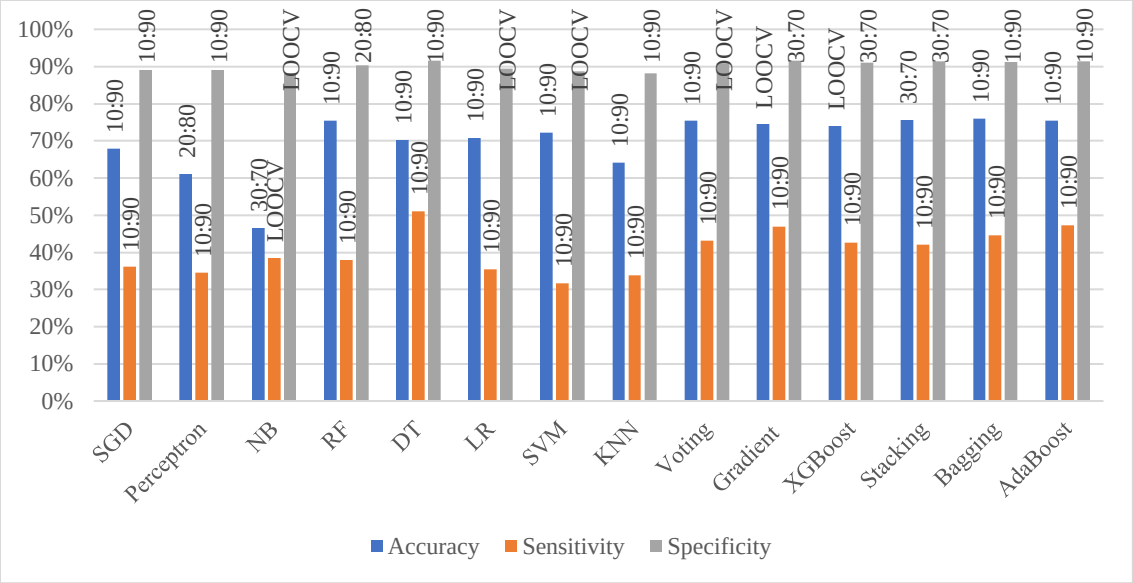


Figure 3.4: Performance evaluation for MS types prediction models

Ensemble Learning

Table 3.2 summarizes the results achieved by the ensemble learning models for all splitting scenarios. Bagging classifier, with RF as the base estimator, achieved a 75.94% accuracy rate. Also, AdaBoost classifier with DT as the base estimator scored the highest sensitivity at 47.35%. In addition, gradient boosting classifier scored the highest specificity at 91.51% when trained/tested with 30:70 rule. Again, all ensemble learning classifiers performed better when training/testing the data with 10:90 rule. All ensembles performed almost similarly regardless of the splitting techniques, as they obtained comparable results. Compared to the inclusive results of the traditional classifiers, ensembles classifiers achieved accuracy rates around 69% - 75%, sensitivity rates between 30% - 70%, and significant specificity rates estimated as 88% - 91%. All things considered, RF bagging is the best model in terms

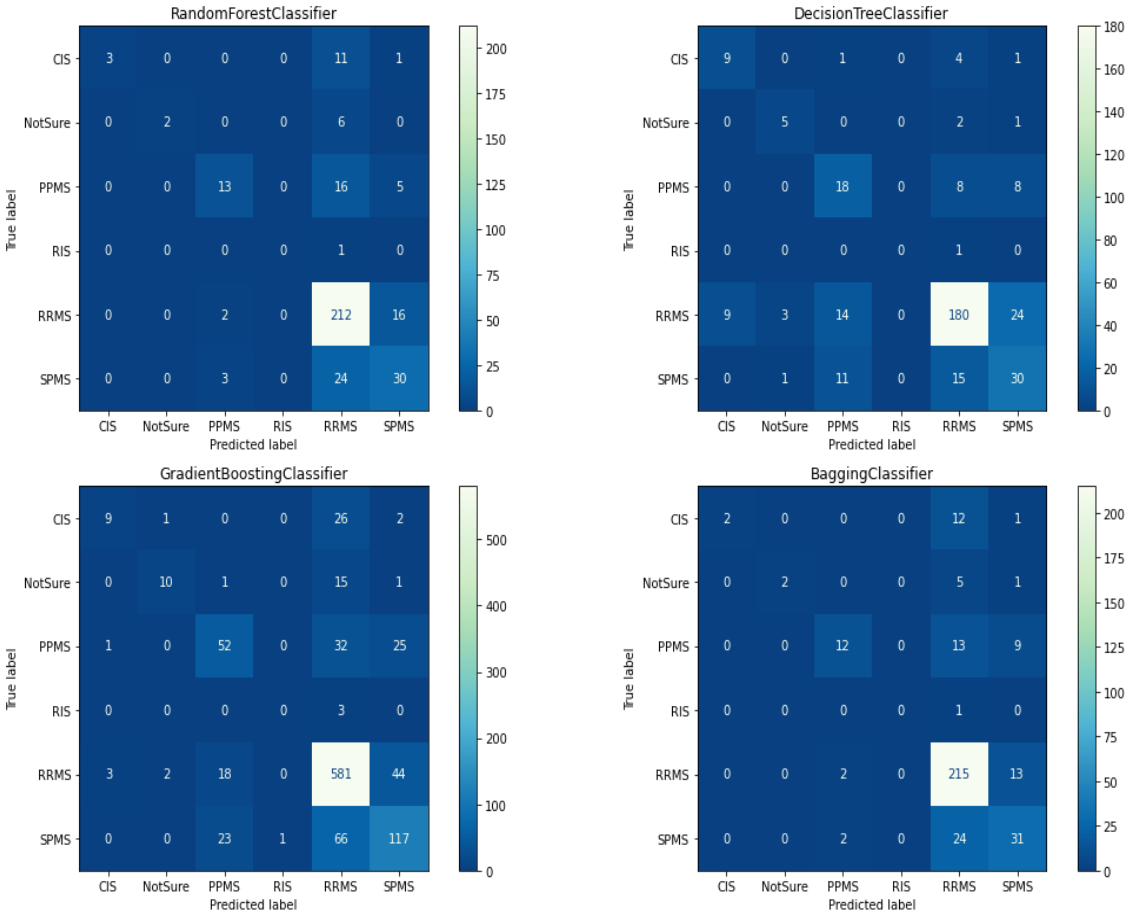


Figure 3.5: Confusion matrices for MS types top classifiers

of accuracy, while DT classifier is the top model when it comes to sensitivity and specificity. Therefore, these two models are investigated further to find better results. Figure 3.4 compares the overall results obtained after evaluating the models, and Figure 3.5 shows the confusion matrices of the top-performed models. Note that the number of observations varies between the confusion matrices illustrated in Figure 3.5 to reflect the splitting approach used in each model.

Hyperparameter Tuning

RF classifier was slightly improved after hyperparameter tuning with 75.94% accuracy, 38.15% sensitivity, and 90.26% specificity. Thus, it could be expected that the performance of bagging model could be improved after using RF with hyperparameter tuning as the base estimator of bagging model.

As for DT, hyperparameter tuning underperformed the model as it got lower accuracy (69.28%), sensitivity (28.87%), and specificity (88.98%). Then, the top bagging ensemble nominated for further analysis was then retrained using the recent hyperparameter-tuned RF as the base estimators. It was found that the accuracy and specificity of the bagging RF model were not changed, after applying RF with hyperparameter tuning, even though the sensitivity was slightly improved by 0.20%. Nevertheless, this was an insignificant improvement. On the contrary, DT with hyperparameter tuning failed to score better results or even maintain the existing outcomes in all metrics. Therefore, training the best-reported traditional models with hyperparameter tuning did not enhance the performance of the top models.

Discussion

In this chapter, we investigated several supervised ML algorithms to generate predictions on the type of MS against real-world data. Unlike (Karaca and Hayta, 2015; Ion-Mărgineanu *et al.*, 2017; Ekşi *et al.*, 2020; Taschler *et al.*, no date), who investigated ML predictions against MRI or any imaging-based data, our study focused on classifying patients based on questionnaire-based clinical data. This dataset keeps records of MS patients' demographics, medical history, quality of life, wellness, physical activities, current and past DMTs, and self-evaluated health scales. These baseline records were fitted into different prediction models with different settings. To our knowledge, the only attempt to use such clinical data in MS domain research was made by (Fiorini *et al.*, 2015), even though their dataset partially covers MS's medical features.

The analysis started by examining the prediction outcomes of the traditional models, with default parameters, based on different feature selection and missing value imputation settings. This was essential to understand the influence of feature selection and missing value imputation on the performance of the traditional models. Here, RF for feature selection and KNNImputer were set up with default parameters and hyperparameter tuning, respectively. It was found that the prediction

accuracy results of the traditional models were better when applying feature selection using RF with hyperparameters tuning. In addition, hyperparameter tuning facilitated KNNImputer in filling missing values to improve the reported accuracy of the traditional models. We first decided to perform feature selection to minimize the computational complexity (time and resources) when imputing missing values and training ML models due to the limitation in computer resources.

Eventually, we observed the prediction results after training all classifiers using the entire set of features. As anticipated, all models performed worse when considering all features. Thus, this observation verified and confirmed the importance of feature selection before fitting ML models for prediction. Besides, missing value imputation using model-based imputation methods, such as KNNImputer, was critical because ML models will not work as expected with missing values. KNNImputer has preserved the data volume and variation while sustaining data neutrality. Moreover, the best sets of parameters in all hyperparameter tuning tasks were done using RandomizedSearchCV. This random search is based on hyperparameter distributions which frequently outperforms grid search as the latter does not scale well when there are many hyperparameters to tune.

As inferred from the outcomes of the three evaluation metrics used in this study, not all traditional models have convergent performance. Several of them got shallow results, while only some performed decently. Actually, most traditional models have noticeable results disparity, especially when trained using different splitting techniques. To clarify with an example, several models performed better in one splitting strategy while did worse in another splitting setup. Also, several models were evaluated better in specific metrics while acquired low results in other metrics. For instance, SVM was one of the models that got decent accuracy, but its sensitivity and specificity could have been more encouraging. Furthermore, all results obtained after training KNN, NB, perceptron, and SGD were insufficient. Only

RF, DT, and LR showed constantly promising outcomes in different metrics using different splitting strategies. Consequently, they were selected for further investigation.

After observing the results of the ensemble classifiers, LR had no apparent effects on the performance of these models compared to RF and DT. However, better accuracy and specificity were reported when optimizing several ensembles with RF as the base estimator. On the other hand, DT has increased the sensitivity when utilized as a base estimator in several ensemble methods. Voting and stacking ensemble methods scored reasonable accuracy, sensitivity, and specificity when combining both models, but their results did not surpass the best-reported results of the top ensemble methods. What caught our attention was the fact that the results of all ensemble methods were very convergent despite the applied evaluation metric or the splitting technique. In other words, no specific ensemble models that mostly or frequently outperformed others in most of the testing scenarios. Only the top models that provided the best evaluation metrics were considered for further improvement. Unfortunately, using hyperparameter tuning did not reveal any significant improvements. Still, these remarkable findings could prove that ensemble classifiers, in general as a group, perform better compared to the group of traditional classifiers in such research studies.

This empirical study investigated the best prediction model(s) based on evaluating each model's accuracy, sensitivity, and specificity. Thus, the proposed methodology would be medically persuasive compared to the literature, as previous work only focused on evaluating models via accuracy. Furthermore, the accuracy results reported in our experiment are superior to the outcomes estimated by (Ion-Mărgineanu *et al.*, 2017; Taschler *et al.*, no date). Nevertheless, our models were within the accuracy results presented by (Fiorini *et al.*, 2015; Karaca and Hayta, 2015; Ekşi *et al.*, 2020) for several reasons. First, the dataset obtained for the presented analysis did not contain information about MRI, which is vital in decisions related to MS from a medical perspective. Also, previous work performed

predictions mainly through a binary classification, which is different from this study's objective, where the main aim was to perform multinomial classification. Nonetheless, the proposed framework made good use of a clinical baseline dataset covering more medical aspects of MS. In addition, the number of subjects recorded in the dataset was sufficient to make the discovered results rational and feasible, unlike most related research lacking enough instances to train ML models efficiently.

While the study showed satisfactory outcomes, several limitations must be outlined here to be addressed in future work. First, dataset characteristics such as variables and the number of observations are medically inadequate after removing records with missing target values. The scarce and restricted availability of more comprehensive MS data is the leading cause of this drawback. Second, the collected MS data has some degree of imbalances among MS stages (target labels), especially where RRMS was reported more frequently. Notably, the majority of samples identified people with RRMS. Hence, there is some imbalance in the classification problem. This problem made the classification work relatively composite. Future work should focus more on collecting data that represents MS types in a more balanced way, meaning that all types should have a sufficient number of observations. Likewise, future research could focus more on the primary MS types and ignore several non-progressive MS types. On the bright side, our results are potentially medically accepted because most models could be used as specific tests for ruling in patients with a specific type, i.e., SPecific tests rule IN (SPIN) (Power, Fell and Wright, 2013).

With that said, further investigation is necessary to expedite the utilization of ML in MS's clinical routines and decision-making. Special attention should be paid to implementing automated tools that facilitate the decision-making process, especially for newly diagnosed patients. For example, studying decision support systems that could detect MS lesions and determine the MS type simultaneously based on MRI readings without human interventions is recommended. Future work should concentrate on

implementing decision-support technologies that can comprehend MS progression and occurrence patterns. Additional research should use these model-based techniques to emphasize the causes of MS and its long-term effects on the quality of life of affected patients. Examining the relationships between MS and other diseases, particularly autoimmune and chronic neurological conditions, is also important. Despite that, accessing data and attaining more information associated with MS should be accelerated. These suggestions should be considered to improve MS's clinical practice.

Chapter 4 The Application of Predictive Machine Learning Models to Support Shared Decision-Making in Selecting the Best DMT Route of Administration

Introduction

Multiple sclerosis (MS) is a progressive disorder that attacks the brain and spinal cord, i.e., the central nervous system (Sospedra and Martin, 2005; Chiaravalloti and DeLuca, 2008). This disorder is the most common inflammatory disorder of the central nervous system because it is categorized as an inflammatory demyelinating disease (Lassmann, Brück and Lucchinetti, 2007; Milo and Kahana, 2010). MS causes various cognitive, motor, and neuropsychiatric symptoms that appear independently (Chiaravalloti and DeLuca, 2008; Brownlee *et al.*, 2017). These symptoms may include paresthesias or numbness, weakness or incoordination, visual dysfunction, vertigo, fatigue, bladder dysfunction, bowel dysfunction, sexual dysfunction, depression, heat intolerance, pain, Lhermitte's phenomenon, frequent falls, sleep disorder, and anxiety, to name a few (Milo and Kahana, 2010). Thus, the quality of life related to physical functioning, emotional health, and social life could be affected or disturbed (Col *et al.*, 2019; McGinley, Goldschmidt and Rae-Grant, 2021).

The exact trigger of MS is doubtful, but several genetic and environmental factors such as vitamin D deficiency, Epstein-Barr virus infection, ambient ultraviolet radiation, obesity (especially in childhood), and tobacco smoking habits are the possible reasons (confirmed or yet to be confirmed) correlated with MS, which makes it a complex disease (Trapp *et al.*, 1998; Leray *et al.*, 2016; Reich, Lucchinetti and Calabresi, 2018; Dobson and Giovannoni, 2019; McGinley, Goldschmidt and Rae-Grant, 2021). In contrast, several other nominated potentials such as vaccines, stress, traumatic events, and allergies have not been identified as risk factors (Leray *et al.*, 2016). MS patients manifest one of the following forms: clinically isolated syndrome (CIS), relapsing remitting MS (RRMS), progressive

MS (PMS) that refers to the combination of secondary progressive MS (SPMS) and primary progressive MS (PPMS), progressive relapsing MS (PRMS), and radiologically isolated syndrome (RIS) (Granberg *et al.*, 2013; Lublin *et al.*, 2014; Mahad, Trapp and Lassmann, 2015; D′Amico *et al.*, 2016; Efendi, 2016; deAndrés-Galiana *et al.*, 2019). Overall, it is believed that there is a lack of MS awareness and education among healthcare providers, patients, and the general public (Rieckmann *et al.*, 2013). A thorough examination of the disease's natural history may reveal much about the subject (Confavreux and Vukusic, 2006).

A variety of disease-modifying therapies (DMTs) for the different episodes of MS have become available, which have prompted changes in treatment procedures over the past few years (Winkelmann *et al.*, 2016). An early and accurate diagnosis of MS is critical to start an effective treatment plan especially in RRMS cases, the initial disease course, in order to maintain the quality of life of MS patients (Rieckmann *et al.*, 2013; Brownlee *et al.*, 2017; Kobelt *et al.*, 2017). MS disease management is found to be effective with the introduction and use of DMTs (Kobelt *et al.*, 2017). Around 70% of MS cases are controlled and managed with different types of DMTs (Sormani *et al.*, 2021). Existing MS DMTs reduce the relapses' frequency and slow the likelihood of developing disabilities associated with the relapses by modulating or suppressing the patient's immune system (Rieckmann *et al.*, 2013). Controlling disease activity, stopping disease progression, and, ideally, inducing the reversal of neurological impairments are the main objectives of MS DMTs (Winkelmann *et al.*, 2016). Through long-term inflammation control and the regulation of abnormal immunological responses, DMTs attenuate or mute disease activity, slow disease progression and relapse rates, and decrease disability probabilities (Winkelmann *et al.*, 2016). Based on the presented facts, DMTs are essential to ensure stability and improvement in MS patients' quality of life.

Current available DMTs vary in mode of action, risk profiles, monitoring needs, usefulness in preventing relapse and disease progression, and side effects (Arroyo *et al.*, 2017; Visser *et al.*, 2020). In addition, MS DMTs are delivered through three routes of administration: oral, injectable, and infusion (Bowen *et al.*, 2020; Visser *et al.*, 2020). Each DMT, available for prescription, offers a range of potential benefits but might tolerate life-threatening risks (Alshamrani *et al.*, 2020). Utilizing the growing number of MS-treating medications demands an in-depth understanding of treatment-associated risks, risk-reducing strategies, and procedures for monitoring and treating such adverse events (Winkelmann *et al.*, 2016). Additionally, when selecting MS DMTs, as well as throughout and after therapy, individual stratification of treatment-related risks is required (Lee Mortensen and Rasmussen, 2017). Therefore, the decisions regarding selecting the best DMT rely heavily on evaluating the risk of further MS progression and considering the risk versus efficacy of a specific DMT (Wiendl *et al.*, 2021). These tradeoffs that should be considered when choosing a specific MS DMT among a variety of options make this type of decision difficult as physicians and patients must decide what tradeoffs, consequences, and long-term benefits they are willing and ready to make.

In fact, MS is a preference-sensitive condition that requires a shared decision-making process, meaning that both the physician and the patient should participate in the decision-making process (Colligan, Metzler and Tiryaki, 2017). Thus, MS physicians and patients would have a tremendous deal of responsibility since they would need to be fully knowledgeable about the current state of MS that they are dealing with, in addition to the potential risks and benefits of all possible alternatives to achieve the optimal decisions, especially decisions about selecting the best DMTs (Alshamrani *et al.*, 2020). MS manifestation varies among patients as different patients require tailored treatment plans, as each plan is associated with a patient's conditions and symptoms. Thereby, using a reliable decision-making tool would assist both parties of the shared decision-making process, a clinician and a patient, in selecting the most suitable treatment regimen.

To address the research gap identified, an automated system that expedites the shared decision-making process without human intervention is proposed by utilizing machine learning (ML) algorithms. ML is gaining popularity in the medical field for mining and analyzing massive amounts of medical data (Sweeney *et al.*, 2014). Medical organizations have been using ML as the primary method for assessing medical knowledge for decision-making (Tuggener *et al.*, 2019). Because of the ambiguity of MS (including its etiology, progression, clinical manifestations, and response to treatment), ML algorithms are being exploited in several pieces of research in the MS realm (Alshamrani *et al.*, 2020). Therefore, the main goal of this work is to explore different supervised ML models in predicting the best route of administration for delivering DMTs to persons affected by MS and comparing the models' performance against a real-world MS dataset. A framework that compares the outcomes of different traditional and ensemble classifiers to address the proposed research objective is presented.

The remainder of this study is structured as follows: Section 2 briefly presents the related work of ML algorithms in determining the best DMTs. Section 3 explains the proposed methodology. Section 4 describes the results. Section 5 discusses the results and trends. Finally, Section 6 concludes this study.

Related Work

This section provides an overview of past research endeavors that contributed to aid decisions regarding DMTs prescriptions. Past Proceedings focusing on predicting responses to or effects of DMTs are excluded, along with other studies highlighting the predictions of MS patients' treatment adherence. Therefore, the following studies demonstrate the previous attempts to support DMTs selection decisions. Clark et al. (Clark *et al.*, 2019) investigated the most suitable medication routes of administration among the three possible alternatives for MS patients receiving DMTs. The study was conducted using data about diagnosis, clinical measures, treatment patterns, and perceptions of patient

suitability for each DMT. The total number of participants was 1978 from 5EU (UK/Germany/France/Italy/Spain) and 756 from the US. The finding of this work suggested that the populations of patients with mild forms are suited for injectable therapies, unlike populations with worsening progressions that need infusions. In addition, oral medicines exhibit great appropriateness in all populations, indicating broad use.

Talwar et al.(Talwar *et al.*, 2021) classified groups of patients based on their prescribed disease-modifying agents (DMAs) into injectable, oral, or infusion users. The analysis was done using data provided by TriNetX, the world's largest living ecosystem of real-world data, from a federated electronic medical records (EMR) network that provides information about MS patients. The total population included in the study was 12,922 representing MS patients ≥ 60 years old, where 2,455 of these patients (18.99%) have active DMA prescriptions. Multivariable logistic regression (LR) model, based on Andersen Behavioral Model, was used to assess factors associated with prescribing DMAs. Overall, multivariable LR found that groups representing patients between 60–64 years and 65-69 years were more likely to receive DMA than patients 70 years and older, with 2.39 and 1.60 adjusted odds ratios, respectively.

Li et al. (Li, Huang and Aparasu, 2022) presented a study of implementing and analyzing ML models for predicting DMA switching among MS patients. Mainly, the study focused on predicting whether an MS patient would receive a different DMA prescription than their previous prescription during the follow-up visit. The researchers trained LR, least absolute shrinkage and selection operator regression (LASSO), random forests (RF), and extreme gradient boosting (XGBoost) against TriNetX data to predict treatment switch. It was found that 16% of the 7,258 eligible MS patients with at least one DMA have changed their treatment within two years. Furthermore, RF model had the best performance with 61% accuracy, 60% recall, and 72% F1 score, even though other models got comparable results.

Due to the limited number of research on the proposed topic, the scope of the literature analysis was expanded to include a study that tried to predict suitable DMTs for patients affected by another disease. Khalaf et al. (Khalaf *et al.*, 2017) investigated the utilization of ML to indicate the effective dosage levels suitable for Sickle cell disease (SCD) patients. Various ML models were fitted with data from 1168 SCD patients acquired from the Alder Hey Children's Hospital in Liverpool, UK. Each sample in the dataset keeps records of 13 SCD's vital factors representing information associated with the disease course. The result reported in the study showed that multilayer perceptron trained using the Levenberg-Marquardt algorithm, RF, and decision tree (DT) classifiers outperformed all other models with 100% accuracy, even though that the expectations of this work were in favor of artificial neural network algorithm.

All things considered, these contributions provided prospective beneficial efforts toward implementing model-based decision support tools useful in decisions associated with determining appropriate DMT plans. However, this research domain is still inadequate. To the best of our knowledge, facilitating decisions regarding MS DMTs via automated approaches is considerably limited. To date, there has not been an actual intention to address the presented research gap as all related work demonstrated above are either preliminary conference proceedings or a corresponding study related to another disease. Therefore, the novelty of the proposed study is recognized as a genuine framework of a model-based decision support tool practical for enhancing the shared decision-making associated with selecting the best MS DMTs suited for patients' needs. Thus, the presented work is an encouraging first step toward ML applications in MS DMTs clinical specification.

Methodology

Study Material

In this study, the acquired data is a part of the iConquerMS™ initiative established by a non-profit organization called Accelerated Cure Project (ACP) for MS. Recruitment of participants for this initiative began around 2010, and the data was obtained through a questionnaire-based repository. This repository was populated by individuals diagnosed with MS who voluntarily submitted questionnaires detailing information pertaining to their medical history, MS symptoms, lifestyle habits, and DMTs. The inclusion and exclusion criteria for participation can be found at (https://www.iconquerms.org/swms-inclusion-requirements). The raw dataset consists of 8329 participants, each with a clinical record including up to 247 attributes. It is important to note that the number of attributes was reduced by removing identifiable attributes and attributes with a high degree of freedom to minimize data dimensionality.

Ethics

The valid use of the dataset is dedicated to research and was authorized by the Office of Research Assurances of the University of Idaho through an approved Human Research Protections (IRB) application (Authorization 21-235, dated February 1st, 2022). Appendix F provides more information about the IRB.

Preprocessing

In order to perform practical feature engineering in this study, three key preprocessing steps were considered. Firstly, the features were evaluated for their significance in predicting MS DMTs' administration routes through a two-round process. Out of all the features in the dataset, the first round of filtering was conducted from a neurological standpoint after consulting an MS expert. The number of features selected, along with the target variable, was 214. Subsequently, an automated feature

1. Train the model using all features
2. Calculate the feature importance or ranking
3. For each subset *Ti, i = 1, 2, 3,…, n* do
    a. Keep the *Ti* most important features
    b. Train/test model against *Ti* features
    c. Recalculate model performance
    d. Recalculate the importance of ranking of each feature
4. Determine the optimal number of features
5. Use the classification model with the selected optimal set of features

Figure 4.1: RFE algorithm

selection approach was utilized to further reduce the number of features. A subset of the filtered dataset containing records with no missing values was used to accurately perform the second round of feature selection. This was a critical step to ensure the optimal functioning of the automated feature selection tool. The feature selection was conducted using the recursive feature elimination (RFE) algorithm (explained in Figure 4.1) from sklearn.feature_selection module.

Handling incomplete records in the raw dataset was crucial before using ML models, and thus missing value imputation was carried out. A model-based imputation method was employed to impute the missing records after the dimensionality of the dataset was reduced based on the selected features. Model-based imputation techniques that utilize ML algorithms have been shown to outperform traditional statistical imputation techniques (Cevallos Valdiviezo and Van Aelst, 2015; Pan *et al.*, 2015; Silva-Ramírez, Pino-Mejías and López-Coello, 2015; Conroy *et al.*, 2016; Raja and Thangavel, 2020; Bertsimas, Pawlowski and Zhuo, no date). This study utilized k-nearest neighbor (KNN) imputation for the imputation process. The missing values of each instance were imputed using the mean value of the n_neighbors nearest neighbors. Additionally, the categorical independent and target variables were encoded using the OrdinalEncoder and LabelEncoder preprocessors from the scikit-learn library. These

steps were crucial to improving the efficiency and performance of the model training and prediction. The combination of feature selection and missing values imputation is common in recent medical research in the MS domain and has been found to enhance the analysis of the complete dataset (Suganthi and Karunakaran, 2019; Liu *et al.*, 2020; Cui, Hu and Liang, 2021; Sharma and Kaur, 2021; Awawdeh, Faris and Hiary, 2022; Christo *et al.*, 2022; Faisal and Tutz, 2022; Kabir and Farrokhvar, 2022; Nagarajan and Dhinesh Babu, 2022; Seba and Benifa, 2022). Thus, this approach was adopted in this study.

It is noteworthy that prior to feature selection and missing value imputation, records of patients who did not specify their current MS DMT were removed to avoid biases in the imputation of missing values for the target feature. A total of 2082 patients reported their current MS DMTs, and the classification models were trained with these instances (753 oral, 682 infusion, and 648 injectable). This research methodology aims to determine the best ML algorithm that can categorize MS DMTs into injectable, infusion, or oral based on historical data. Therefore, before constructing the predictive models, an extra feature with ternary outcomes was added to each sample *i.* This additional attribute refers to one of DMTs' routes of administration mentioned above. In other words, this newly created label indicates whether an MS patient consumes their prescribed DMT as injectable, infusion, or oral. The outcomes were formally defined as follows:

$$y_i^k \begin{cases} 2, \text{patient with injectable DMT} \\ 1, \text{patient with infusion DMT} \\ 0, \text{patient with oral DMT} \end{cases}$$

Classification

This study considered traditional ML algorithms such as RF, SVM, KNN, LR, DT, naive Bayes (NB), perceptron, and stochastic gradient descent (SGD). Also, the performance of voting, bagging, gradient boosting, AdaBoost, stacking, and XGBoost ensembles was evaluated. The dataset was divided into training

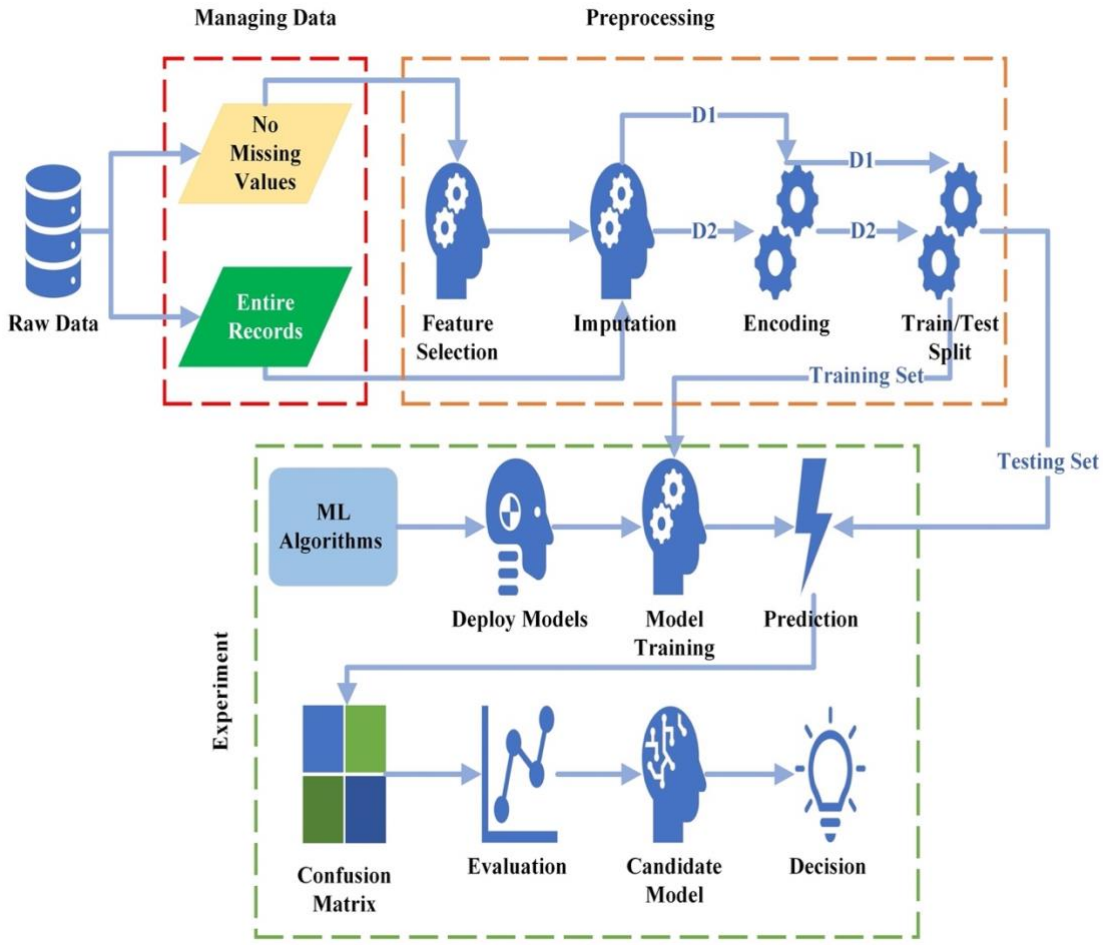


Figure 4.2: Framework design for predicting DMT's route of administration

and testing subsets using five different splitting techniques, and the performance of the models was compared based on accuracy, recall, F1 score, and precision. These evaluation metrics were calculated based on the results of the confusion matrix indicators (True Negative (TN), False Negative (FN), False Positive (FP), and True Positive (TP)). The ratio of correctly classified instances determined the accuracy of the predictions:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Recall measures the proportion of positive instances correctly identified by the classifier out of all positive instances in the data. It is also known as the sensitivity or the TP rate:

$$Recall = \frac{TP}{TP + FN}$$

Besides, precision is the proportion of true positive predictions among all positive predictions made by the model. In other words, It indicates the proportion of positive predictions that are actually correct:

$$Precision = \frac{TP}{TP + FP}$$

F1 score is the harmonic mean of precision and recall. F1 score provides a balance between precision and recall, and a high F1 score indicates a model that has both high precision and high recall, meaning it has a good balance of not falsely classifying negative instances as positive and accurately detecting positive instances:

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

In a multiclass classification problem, where multiple classes exist to be predicted, the confusion matrix is commonly represented as an *N* x *N* matrix, where *N* denotes the total number of classes. The diagonal elements of the matrix correspond to the number of accurate predictions for each class, while the off-diagonal elements represent the misclassifications. Please refer to Figure 3.2 for a detailed demonstration. Accuracy is computed by dividing the total number of correct predictions by the total number of predictions, while other evaluation metrics are computed through macro averaging. Macro averaging is an established method in data analysis and ML to assess a classification model's performance by calculating the model's average performance metric across multiple categories or classes. To carry out macro averaging, the performance metric for each class in the dataset is computed individually. Next, the average of these metrics is determined across all the classes, thereby providing an overall evaluation of the model's performance across different classes.

The following methodology was employed in the present study to identify the most suitable classifier for categorizing patients' DMTs into injectable, infusion, and oral DMT groups based on historical data. Moreover, the analysis was conducted using the scikit-learn package in Python 3.7. RFE model was trained using the complete dataset to select the most important features. The missing values in

the selected features were imputed using KNNImputer. Subsequently, the traditional ML models (RF, SVM, KNN, LR, DT, NB, perceptron, and SGD) were trained using the imputed dataset and evaluated using default parameters. The top classifiers were then selected for further examination. The ensemble learning methods (voting, bagging, AdaBoost, and stacking) were trained using the imputed dataset and their default parameters, with the top traditional models serving as their base estimators. The performance of all models was compared to determine the best classifier. The steps outlined above were carried out using three test:train ratios (10:90, 20:80, and 30:70) and two cross validation (CV) methods (10-folds and 5-folds). Ultimately, the prediction results after training all classifiers on the entire feature set determined from a neurological perspective was observed. In addition, classification results were examined after training all models again but with only the subset of features selected by RFE algorithm. The overall architecture of the prediction framework is illustrated in Figure 4.2.

Results

Dimensionality Reduction

Following the recommendations of a consulted MS specialist, 214 features were initially considered. Moreover, feature selection using RFE was done using RF, DT, and ExtraTreesClassifier as the base estimators embedded in RFE method. From the pool of 214 features, 106 features were found by RFE that used RF as the base estimator (using 30:70 rule). Note that the number of features specified above represents features that enabled the study's classifiers to perform better. The prediction models trained with the chosen sets of features have obtained more convenient results, as demonstrated next. Appendix C lists the most important features and their description, while Appendix D lists features by order of their importance. To address the issue of incomplete data in the newly selected sets of features, KNNImputer was utilized with default parameters to complete the missing values.

Classification Results

After applying various splitting strategies, the results obtained from the tested classifiers are outlined in Table 4.1 and Table 4.2. The results showed that the RF model outperformed all other traditional models in most metrics. In fact, RF model trained using 20:80 splits got the highest accuracy at 48.20%, recall at 48.77%, and F1 score at 47.64%. Furthermore, SGD scored the highest precision at 56 % (10:90 splits). Overall, the results of all evaluation metrics of all traditional classifiers using different splitting techniques ranged between 23% and 56%. However, the results of traditional classifiers showed a noticeable disparity in changing the splitting technique. The only models with comparable and consistent results over the different splits were RF, LR, and SVM. Thus, these three classifiers that achieved the topmost results when trained against 106 features were nominated for further investigation.

Notably, the ensemble learning analysis results vary regardless of the trained model and the splitting approach. Remarkably, gradient boosting classifier, trained with 80% of the data, outperformed all other ensemble methods as it increased the accuracy to 49.88% and recall up to 50.22%. The highest precision rate was 50.90% when AdaBoost classifier (used RF as a base estimator) trained with 20:80 splits. Again, gradient boosting classifier got an F1 score at 49.83% when fitted with 20:80 splits. Overall, ensemble classifiers slightly got improved compared to the traditional models, except in one precisian result. Therefore, the findings suggest that ensemble learning methods performed classification tasks well, but this study also proved that traditional models were as good as ensembles through the recorded results. Figure 4.3 and Figure 4.4 overview the overall comparison between the performance of all models. Figure 4.5 demonstrates the confusion matrices of RF (best traditional model) and gradient boosting (best ensemble classifier). Given the different splitting strategies utilized in each model, the confusion matrices shown in Figure 4.5 have different numbers of observations.

Table 4.1: Route of administration prediction results for traditional methods.

Model	10:90				20:80				30:70			
	Accuracy	*Recall*	*Precision*	*F1 Score*	*Accuracy*	*Recall*	*Precision*	*F1 Score*	*Accuracy*	*Recall*	*Precision*	*F1 Score*
KNN	37.80%	38.68%	38.54%	36.63%	35.25%	37.15%	37.68%	28.38%	40.16%	39.69%	40.22%	36.82%
SVM	34.45%	35.37%	56.00%	28.38%	37.17%	34.80%	39.61%	23.48%	32.48%	33.13%	34.48%	20.47%
LR	34.45%	34.06%	23.65%	nan	42.45%	42.41%	41.90%	40.93%	41.60%	41.38%	41.62%	39.94%
DT	37.32%	37.90%	34.51%	34.42%	48.20%	48.77%	50.28%	47.64%	42.56%	42.94%	44.14%	42.01%
RF	41.63%	42.26%	42.31%	41.52%	39.81%	40.01%	40.06%	39.85%	43.52%	43.57%	43.70%	43.53%
NB	35.89%	36.06%	35.89%	35.92%	44.60%	44.71%	44.64%	44.46%	41.60%	41.73%	42.05%	41.56%
Perceptron	45.93%	46.05%	46.58%	45.89%	41.49%	42.44%	42.34%	39.34%	40.80%	41.48%	41.65%	37.88%
SGD	40.67%	41.47%	41.46%	39.67%	36.93%	37.41%	36.78%	36.19%	35.52%	35.62%	35.27%	35.02%
Stacking	42.72%	43.37%	48.86%	40.31%	43.41%	44.53%	53.46%	40.07%	42.72%	43.37%	48.86%	40.31%
Voting	43.84%	44.36%	45.66%	42.43%	47.00%	47.76%	49.46%	46.04%	43.84%	44.36%	45.66%	42.43%
AdaBoost	45.45%	46.16%	49.70%	43.95%	44.36%	45.18%	50.90%	42.77%	44.32%	44.91%	50.62%	42.50%
Bagging	44.02%	44.20%	44.33%	44.06%	42.93%	43.69%	46.14%	42.33%	43.36%	43.81%	46.34%	42.46%
XGBoost	46.89%	47.14%	47.20%	46.90%	48.92%	49.22%	49.60%	49.02%	44.96%	45.11%	45.34%	44.94%
Gradient	45.93%	46.25%	46.80%	46.04%	49.88%	50.22%	50.80%	49.83%	45.60%	45.80%	46.69%	45.52%

Table 4.2: route of administration prediction results for ensemble methods.

Model	5-Fold CV				10-Fold CV			
	Accuracy	*Recall*	*Precision*	*F1 Score*	*Accuracy*	*Recall*	*Precision*	*F1 Score*
KNN	37.70%	37.57%	38.32%	37.05%	38.90%	38.16%	39.06%	37.66%
SVM	36.17%	36.75%	36.39%	33.61%	35.78%	36.02%	35.90%	35.75%
LR	39.43%	40.35%	39.84%	38.01%	38.81%	39.71%	39.34%	37.35%
DT	42.60%	42.00%	42.95%	41.66%	42.27%	41.76%	42.09%	41.37%
RF	40.01%	39.84%	39.81%	39.82%	38.57%	38.48%	38.54%	38.49%
NB	41.45%	41.30%	41.38%	41.31%	42.22%	42.08%	42.11%	42.09%
Perceptron	41.55%	41.11%	40.85%	40.55%	42.17%	41.75%	41.76%	41.17%
SGD	34.58%	34.92%	34.88%	33.90%	36.70%	37.00%	36.71%	35.98%
Stacking	38.71%	38.09%	38.07%	36.91%	41.64%	41.02%	41.49%	39.89%
Voting	43.08%	42.66%	42.88%	42.10%	43.47%	43.11%	43.19%	42.54%
AdaBoost	43.37%	42.40%	44.12%	41.24%	43.52%	42.36%	43.70%	40.31%
Bagging	43.85%	43.06%	43.70%	42.17%	43.95%	43.15%	43.93%	42.21%
XGBoost	44.38%	44.09%	44.26%	44.10%	45.87%	45.65%	45.71%	45.61%
Gradient	44.91%	44.67%	44.84%	44.67%	44.91%	44.60%	44.79%	44.58%

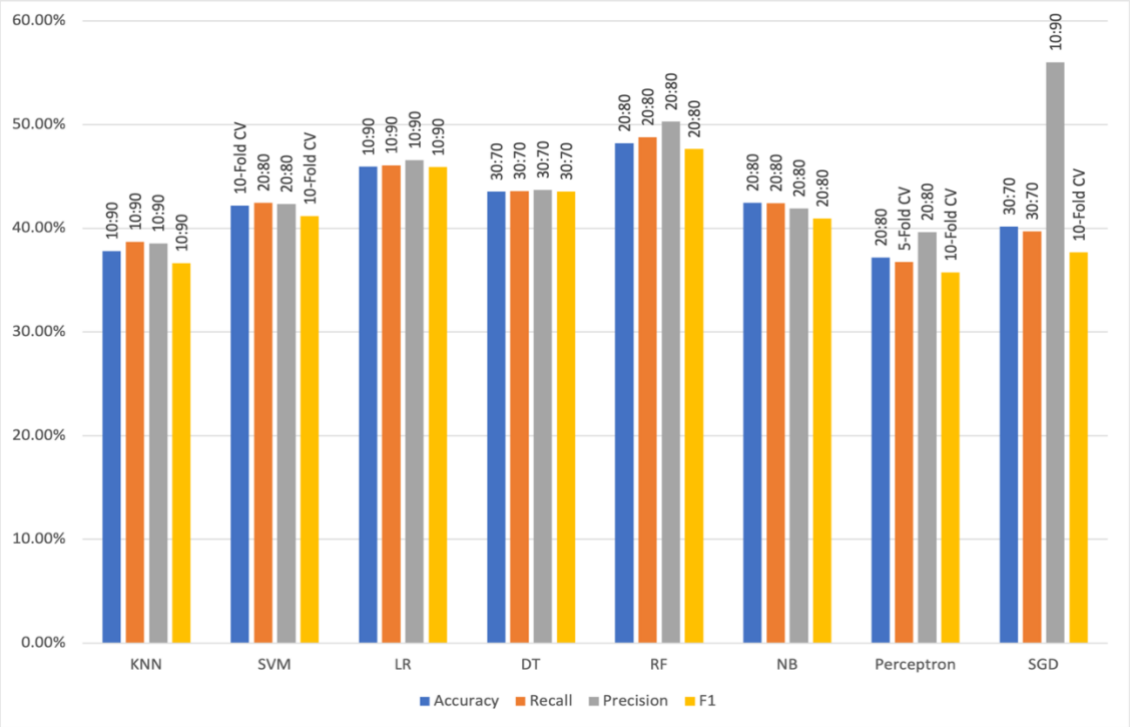


Figure 4.3: Overall highest results of traditional models for predicting route of administration

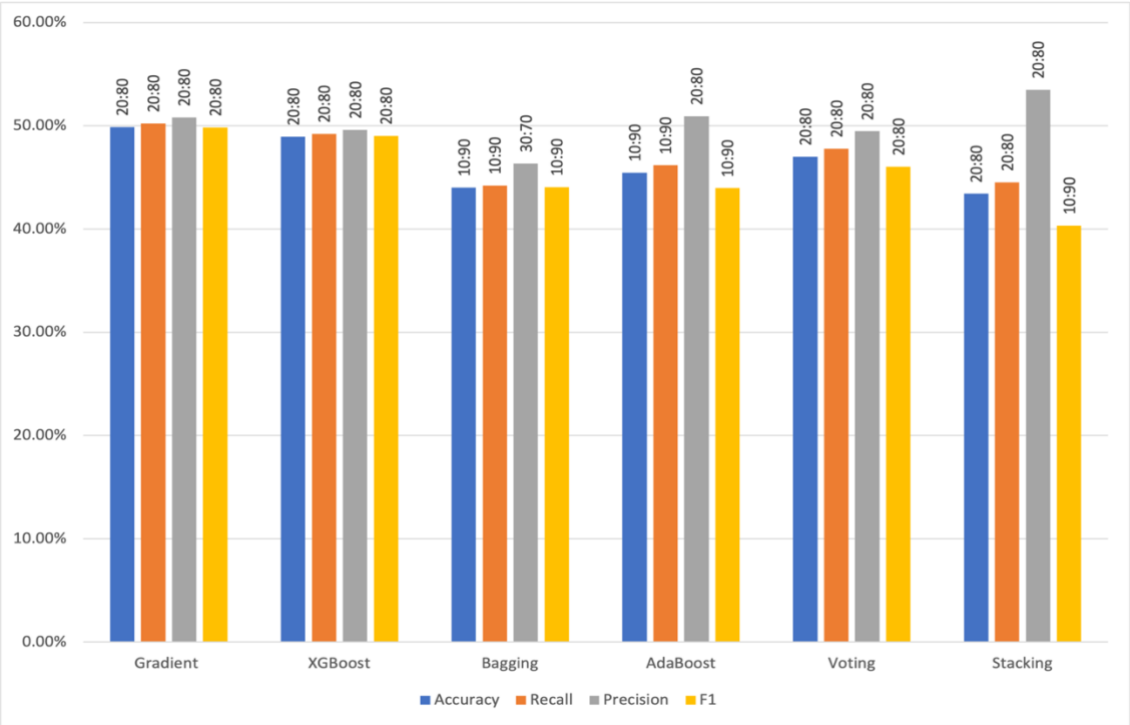


Figure 4.4: Overall highest results of ensemble models for predicting route of administration

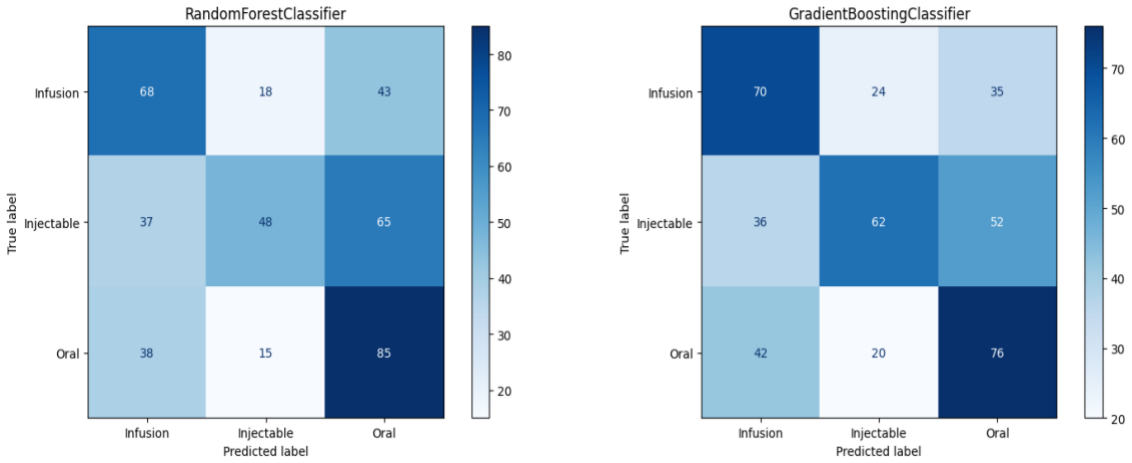


Figure 4.5: Confusion matrices for route of administration best prediction models

Discussion

The initial phase of the study involved evaluating the prediction accuracy, recall, precision, and F1 score of traditional models with default parameters, under various configurations of feature selection in addition to model-based missing value imputation. This was an essential step for understanding the impact of feature selection and missing value imputation on the performance of the traditional models. Subsequently, the prediction results of all classifiers were analyzed after training these classifiers against the best subset of features. Accordingly, all models performed well when the selected subset was considered compared to the results observed after using the entire feature set. Additionally, missing value imputation using model-based methods, such as KNNImputer, was critical as ML models cannot produce accurate predictions with the existence of missing values. KNNImputer was found to preserve the data's amount and variation while maintaining neutrality.

Even though the proceedings presented by (Clark *et al.*, 2019; Talwar *et al.*, 2021) correspond to the objective of this study, they need more application of ML as the first focused mainly on descriptive statistics, and the latter utilized only one ML model. Comparing the performance of several ML methods should be sufficient when choosing ML as the primary solution because each algorithm

performs variously depending on the problem and the dataset being used (Alshamrani *et al.*, 2020). In contrast, the study provided by (Li, Huang and Aparasu, 2022) utilized a variety of ML models, but the objective did not correspond exclusively to the main aim of this analysis. In addition to these studies, a study that attempted to predict which DMTs would be most effective for individuals with another disease was explored. To our knowledge, the only work found was presented by (Khalaf *et al.*, 2017), where the authors made the prediction task efficiently. However, the objective went in a different direction than the main aim presented in this chapter.

The present study offers a novel contribution to the MS domain by making good use of different ML models to facilitate decisions associated with MS DMTs. Explicitly, the outcomes examined in the proposed study significantly promoted ML as the ground solution to automate the shared decision-making process in the presented matter. Namely, the performance of the analyzed models in the presented work surpassed the findings of the proceedings presented in the literature, especially those related to the field of MS. Although the objective of (Li, Huang and Aparasu, 2022) was marginally different as the researchers tried to estimate the probability of DMTs being switched, the results of this analysis exceeded their findings. Besides, the results of (Clark *et al.*, 2019; Talwar *et al.*, 2021) were inconclusive compared to the findings demonstrated earlier as the above mentioned proceeding paper did not provide enough evidence of using the necessary model evaluation metrics that indicate the usefulness of the methodology. In this study, the highest reported accuracy rates were comparable across all models as they were around 37% - 49%, while the highest precision was noticeably disparate, with proportions around 39% - 56%. The highest recall ratio was around 36% - 50%, and the highest F1 scores were 35% - 49%. Ultimately, the prediction performance of all classifiers was assessed after training the models on the entire feature set. As hypothesized, the overall performance of all models was lower when all features were considered. This finding emphasizes the important role of feature selection in preparing ML models for prediction tasks.

Despite the satisfactory results, several limitations must be addressed in future research. Primarily, the dataset's characteristics, including the variables and the number of observations, need to be more medically sufficient due to the removal of records with missing target values. Furthermore, the leading cause of the limitation is the limited availability of more comprehensive MS data. For instance, it could be suggested from the results that additional features correlated with MS DMTs, other than the recorded features in the used dataset, should be considered when forming decisions related to MS DMTs. Comprehensively, it is trusted that the proposed framework and the achieved outcomes are deliberated as an actual first step in boosting ML applications in MS's clinical decisions associated with DMTs. Hence, this is solid evidence that this contribution has superiority over the work presented in the literature.

In light of this, further research is imperative to accelerate the integration of ML into the clinical routines and decision-making processes of MS. A particular focus should be placed on implementing automated tools that simplify decision-making, particularly for newly diagnosed patients. As an illustration, examining decision support techniques that can accurately determine the best DMTs for MS patients based on clinical baseline data without human involvement is suggested. Furthermore, future efforts should be directed toward implementing decision-support technologies to analyze MS progression and identify occurrence patterns. Moreover, further research utilizing model-based techniques is needed to highlight the underlying causes of MS and its impact on affected individuals' quality of life, specifically over the long-term. Nevertheless, it is essential to prioritize the acquisition of data and additional information regarding MS. These considerations should be considered to improve MS's clinical practices.

Chapter 5 Using Knowledge Graph to Improve Informed Multiple Sclerosis Diagnosis and Treatment Decisions

Introduction

Multiple sclerosis (MS) is a chronic and demyelinating inflammatory disorder that affects the central nervous system of the human body caused by various genetic and environmental factors (Trapp *et al.*, 1998; Sospedra and Martin, 2005; Ascherio and Munger, 2016; Reich, Lucchinetti and Calabresi, 2018; deAndrés-Galiana *et al.*, 2019). Several environmental risk factors, including geographic latitude, vitamin D deficiency, exposure to tobacco, obesity, and infection with the Epstein-Barr Virus, are known to enable the progression of MS. However, the precise causes that trigger MS are still unidentified (Trapp *et al.*, 1998). Persons with MS may encounter simultaneous yet independent symptoms during the lifetime of this non-life-threatening disease, such as sensory, visual, motor, cognitive, and cerebellar disorders (Miri Ashtiani *et al.*, 2018). Moreover, these symptoms and their impacts vary between persons with MS as they are unpredictable (Goldenberg, 2012). In general, MS has four main categories that represent its progression level: relapsing remitting MS (RRMS), secondary progressive MS (SPMS), primary progressive MS (PPMS), and progressive relapsing MS (PRMS) (Goldenberg, 2012). All things considered, examining MS's natural history in great detail could reveal a lot about this disease.

Magnetic resonance imaging (MRI), lumbar punctures, and blood tests are the major medical tests that help confirm MS cases, considering the 2017 revised McDonald criteria (Ghasemi, 2017; Thompson *et al.*, 2018; McNicholas *et al.*, 2019). After diagnosis, physicians prescribe disease-modifying therapies (DMTs) to control the potential progression and relapses of MS, in addition to treatments for managing its symptoms since this disorder cannot be prevented or cured (Montalban *et al.*, 2018). MS has several commonalities with other disorders that share several clinical features, making decisions regarding the treatment and the diagnosis difficult, despite the available diagnosis criteria (Hauer, Perneczky and

Sellner, 2021). Thus, decision-making about MS diagnosis and treatment is highly critical and requires intensive experience and knowledge, even though the decisions' quality may remain doubtful due to MS-related uncertainties (Alshamrani *et al.*, 2020). MS is a preference-sensitive condition requiring full participation from MS specialists and patients in decision-making (Colligan, Metzler and Tiryaki, 2017; Alshamrani *et al.*, 2020). Having an accessible platform that can provide intuitive access to information will facilitate and ease the decision-making process for affected persons and physicians mediating the management of this disease (Colligan, Metzler and Tiryaki, 2017). Hence, decision-makers should possess knowledge about the current state of the condition and the possible benefits and risks of all available therapeutic options to attain the best decisions.

A knowledge graph accumulates and conveys domain knowledge through entities of interest and their relationships by representing real-world knowledge (Hogan *et al.*, 2021). In ancient history, the knowledge graph concept was recognized as the core indication of diagrammatic knowledge representation (Gutierrez and Sequeda, 2021). Knowledge graph in the era of big data is considered as the integration of knowledge and data at a large scale with diversified formats (Gutierrez and Sequeda, 2021). The consideration of known general biological mechanisms in patient-specific health data analytics has recently been acknowledged as a necessity (Gustafsson *et al.*, 2014). Acquiring and specifying knowledge in formal, simple, powerful, and incremental ways and then applying applicable reasoning to that knowledge would be a significant outcome in health informatics (Riaño *et al.*, 2012).

Online resources provide extensive medical information and grant access to feasible explanations about diseases and symptoms to people with no medical background or knowledge. Though, the web as a source of information could potentially increase people's anxiety due to the unreliability of the information (White and Horvitz, 2009). Patients seeking information about their conditions need automated tools to support rational medical reasoning procedures (Rotmensch *et al.*, 2017). However,

comprehensive information about MS and the availability of credible sources are globally scarce for persons with MS and anyone eager to know more about this condition (Browne *et al.*, 2014). Persons with MS increasingly need reliable information about this disorder, including its symptoms and ways to control and manage the disease, to name a few (Meca-Lallana *et al.*, 2017).

Knowledge graphs, which naturally close the gap between basic science research and medical practice, can fill this demand as they link data from several biological and medical concept classes (Nicholson and Greene, 2020; Nelson *et al.*, 2022). In fact, such an idea would be phenomenal in the MS domain as it reinforces sharing and reusing MS knowledge among the shared decision-makers for nominating practical decisions. Thus, the knowledge graph is feasible to support shared decision-making in MS diagnosis and treatment. Thus, this chapter aims to develop the MS knowledge graph (MSKG) and make it an accessible, reliable, easy-to-use source of information. The main intention is that MSKG would empower patients in informed decisions associated with MS diagnosis and treatment and strengthen the patient-physician relationship for a more efficient and satisfactory medical decision-making experience.

The remainder of this chapter is organized as follows: Section 2 presents the related work of knowledge graphs in health informatics. Section 3 explains the methodology of MSKG construction. Section 4 describes the results. Section 5 discusses remarkable findings and trends. Finally, Section 6 concludes this study.

Related Work

Note that articles that demonstrate the application of knowledge graphs in MS research as a domain knowledge representation mechanism are included, in addition to articles that examine knowledge graphs as knowledge representation techniques in the medical domain that indirectly consider MS. On

the other hand, articles that investigate using graph-based techniques, graph theories, and knowledge graphs for non-MS purposes are excluded.

By embedding individual patient data into a biomedical knowledge graph, Nelson et al. (Nelson *et al.*, 2022) presented a method for integrating domain knowledge into clinical diagnosis to feasibly enable the detection of MS up to five years prior to their confirmed diagnosis in the clinic. The researchers presented SPOKEsigs, a scheme to embed electronic health record data onto SPOKE knowledge graph (Sanders, Pearce and Baranzini, 2020) to obtain high-dimensional health status profiles. To determine who is at risk of MS, SPOKEsigs were calculated for hundreds of thousands of people, and a random forest (RF) classifier was trained. Fecho et al. (Fecho *et al.*, 2021) implemented an open biomedical knowledge graph-based system known as ROBOKOP (Reasoning Over Biomedical Objects linked in Knowledge Oriented Pathways). This knowledge graph is a biomedical graph-based question-answering system that enables users to find the relationships between the terms found in the provided questions by the users. ROBOKOP represents a wide range of biological entities and predicates. For instance, ROBOKOP addresses the association between carbon monoxide and multiple sclerosis as an example of the impact of workplace exposures on immune-mediated diseases.

Rotmensch et al. (Rotmensch *et al.*, 2017) constructed a health knowledge graph that associates diseases and symptoms directly from electronic medical records (EMR). Their work examined and observed the EMR data to plot a graph to link symptoms to the diseases that caused them. It is worth mentioning that MS frequently appeared in the EMR data of the emergency department. Desarkar et al. (Desarkar *et al.*, 2013) discussed capturing the users' prepositional knowledge, which denotes the representation of facts about diverse medical topics. This proposed framework was known as Med-Tree, intended for decision-making purposes in several healthcare domains. The total number of medical topics was 21, and all were posted and discussed by patients on the PatientsLikeMe social

network. MS ranked sixth among the 21 topics based on the PatientsLikeMe website's intensity of interests and interactions.

Several reasoning and inference agents and prediction techniques can make good use of the domain knowledge conceptualized in knowledge graphs to automate the decision-making process. This would augment medical decisions about MS and conceivably have automated medical justifications that could surpass qualified MS specialists. However, there is no clear intention in the literature to enrich the MS research domain with knowledge graphs. To our knowledge, the potential of the knowledge graph principle has yet to be heavily exploited in the research of MS based on the conducted analysis. Based on this understanding, MS needs proper attention like other chronic diseases, regardless of its uncommonness and data scarcity, especially from the computer and data science perspective. Hence, representing and visualizing MS's fundamental information using a knowledge graph would be a novel endeavor to complement the existing research. This would promote the availability of an MS public knowledge base, improve the quality of MS diagnosis and treatment decisions, and raise public awareness about MS.

Methodology

It was observed that various pieces of knowledge were coded since the emergence of the first ontology engineering tools (Gutierrez and Sequeda, 2021). Ontology, as an aspect, has become gradually significant due to the widespread applications of knowledge graphs, machine learning (ML), natural language processing, and the massive volume of data generated on a daily basis (Kendall and McGuinness, 2019). An ontology is a formal specification of the shared conceptualization of a knowledge domain (Ma *et al.*, 2014). Therefore, ontology engineering would be used to form a suitable knowledge graph for securing the objective of this study.

Figure 5.1: Methodology design for constructing MSKG

In this work, the MSKG is constructed with the aid of ontology engineering by means of Web Ontology Language (OWL2) and expressed in Terse RDF Triple Language (Turtle) using Protégé 5.5. The MSKG addresses the following major questions in the ontology engineering process: (1) what is MS? (2) which part of the human body does it affect? (3) what are the common and possible symptoms? (4) what are the potential causes? (5) what are the licensed DMTs to control it, and what are their side effects? (6) where should patients be treated and by whom? (7) what are the current and standard diagnostic procedures and strategies? and (8) what is the role of MS specialists in such a condition? Figure 5.1 demonstrates the methodology design.

All key concepts and relationships presented in the MSKG were derived from the authors' conceptual understanding based on subjective experience and knowledge. In other words, the inclusion of the terms specified in the MSKG was driven by the interest in representing MS domain knowledge from the perspective of its general overview and basic facts. This comprehension was gained from analyzing various MS knowledge sources such as informal discussions with neurologists and MS specialists, scientific articles, existing healthcare ontologies, and informative online web contents and materials (e.g., the National Multiple Sclerosis Society, MS International Federation, Above MS by Biogen, Multiple Sclerosis Association of America, Multiple Sclerosis Foundation, National Institute of Neurological Disorders and Stroke, and PatientsLikeMe).

Result

The MSKG is structured to explain MS and several common factors related to it for enhancing the MS shared decision-making experience. Furthermore, the sketched MSKG formed a starting point for those who want to gain fundamental knowledge about MS. Thus, newly diagnosed or at high-risk MS patients, their first-blood relatives, and individuals interested in knowing more about MS are the targeted potential audience of the presented knowledge graph. All expressions presented in MSKG are based on their importance to the inclusive understanding of MS and how they are significantly correlated with the conceptual overview of MS.

The "MultipleSclerosis" concept was used as the basis for implementing the class hierarchy in MSKG. Accordingly, all other concepts were organized into preliminary and accumulated information. The preliminary information signifies the prime facts that are rationalized to specify the initial knowledge about MS. This, in turn, would determine additional concepts that enlarge the MSKG framework in order to tolerate a more comprehensive representation of the MS domain. On the other hand, the

accumulated information represents details that expand the MS knowledge base. This expansion has identified the major concepts in MSKG. Substantially, MSKG is a centralized network of terms where MS is the central concept, and all concepts are related to it directly or indirectly.

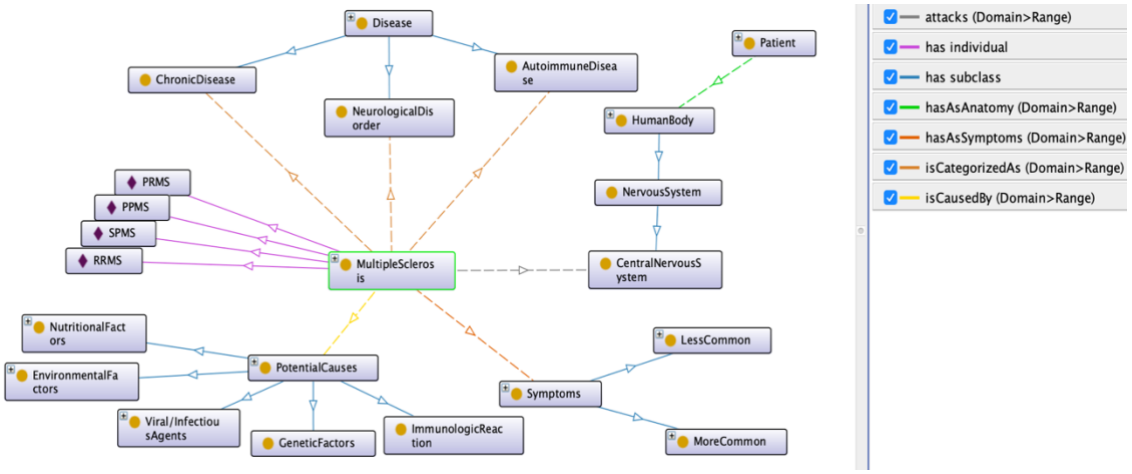


Figure 5.2: Key characteristics of MS

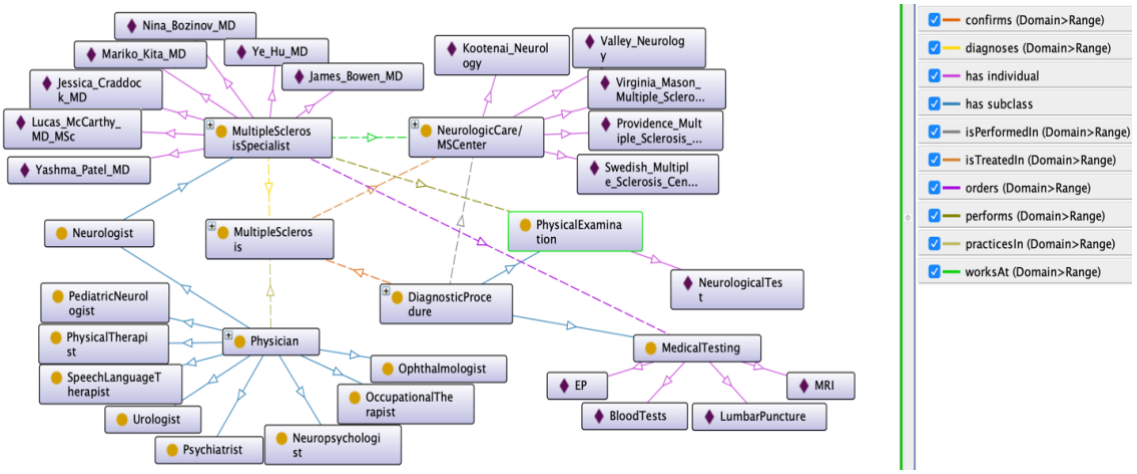


Figure 5.3: MS confirmation procedures

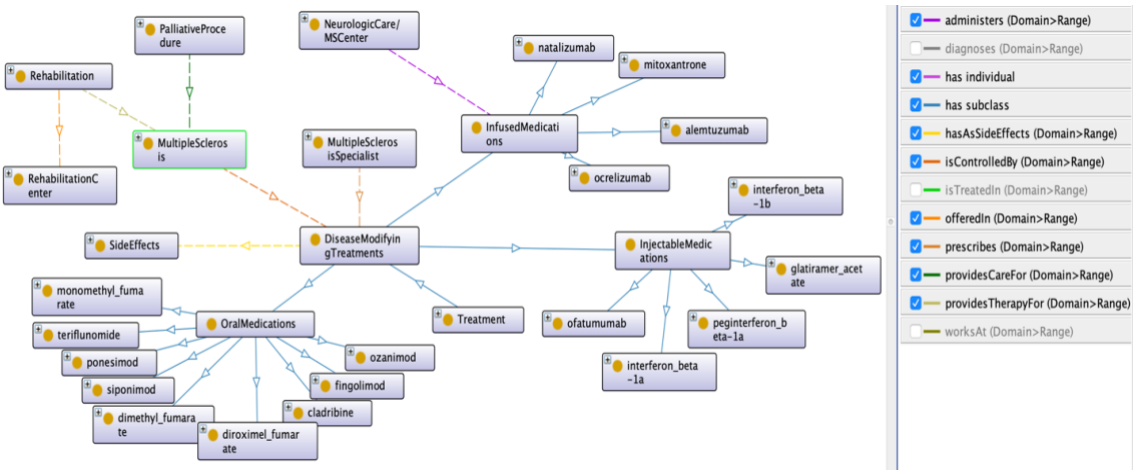


Figure 5.5: MS treatment alternatives and plans

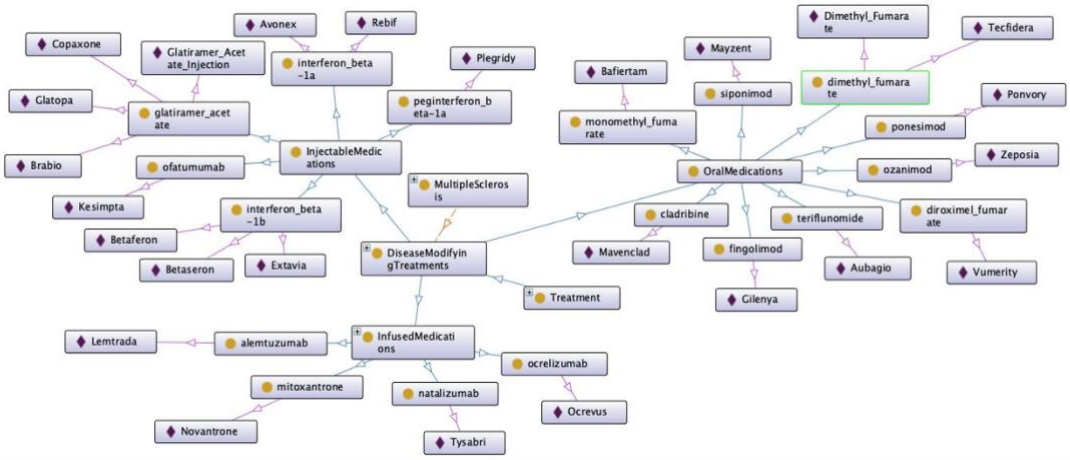


Figure 5.4: Class individuals of MS DMTs

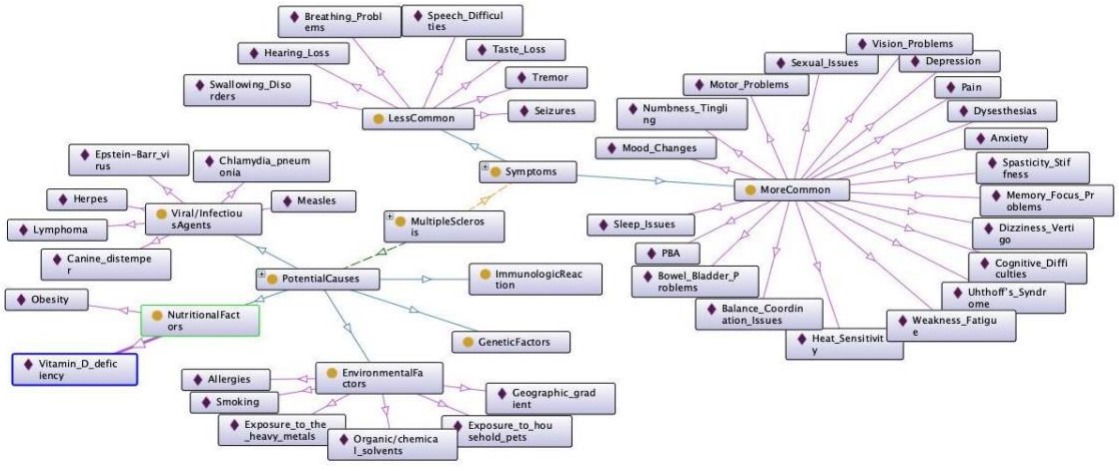


Figure 5.6: MS symptoms and potential causes

The MSKG consists of 63 classes, 20 object properties, and 123 instances. It should be noted that all object properties presented in MSKG are specified to indicate the concepts' extent of linkage to MS. The MSKG classes disclose MS in addition to its common and uncommon symptoms, potential etiology, influence on the central nervous system, and classification under disease taxonomy. Moreover, these classes address the responsibilities of MS specialists and several other healthcare providers in MS practices, the assignment of diagnostic procedures (various medical tests and physical examinations) for confirming MS, the provision of care in neurologic care/MS and rehabilitation centers, and the certified medical procedures that improve MS patients' quality of life. Besides, these classes examine the role of DMTs in controlling MS, their types, and their possible side effects. It is worth mentioning that instances support several of these classes to exemplify the benefits of adopting such a knowledge graph in the MS research domain. Figures 5.2-5.6 conceptualize MSKG using OntoGraf. Note that the MSKG class hierarchy is sketched using OWLViz and can be found in Appendix E.

Discussion

The MS decision-making process demands the involvement of healthcare providers and patients because this condition is a preference-sensitive condition, as noted earlier in this chapter. Confirming MS cases is challenging due to the similarities shared with other diseases of the central nervous system, despite the availability of the updated McDonald criteria for MS diagnosis. Furthermore, pharmaceutical companies offer different MS DMTs with potential benefits and life-threatening risks. Thus, sufficient knowledge and experience are necessary for all participants when forming decisions associated with MS. Therefore, ontology engineering is critical for composing reliable knowledge to provide a shared understanding of the MS domain.

Web users and medical professionals agree that the contents presented online may increase the apprehensions of people with little or no knowledge about several medical fields. Thereby, research has been conducted to enrich people's knowledge in various health domains. On the contrary, MS's key facts are not sufficiently presented in cyberspace. For instance, the research analyses explored by (Desarkar *et al.*, 2013; Rotmensch *et al.*, 2017) consider MS knowledge slightly and are not entirely dedicated to conceptually representing MS domain knowledge. The researchers of (Rotmensch *et al.*, 2017) concentrated on constructing a health knowledge graph to find the associations between major diseases and their common symptoms. On the other hand, the research investigation made by (Desarkar *et al.*, 2013) complies with the objective of this chapter. However, the knowledge graph of that work highlights common information related to numerous medical topics, not just MS. From the ontology engineering perspective, ontology-based knowledge representation of the MS illustrated in (*multiple sclerosis*, no date; *Multiple sclerosis ontology - Summary | NCBO BioPortal*, no date) explains MS from a pure medical viewpoint, which implicitly requires in-depth medical background and is not intended for non-experts. Likewise, the MS Ontology proposed by (Malhotra *et al.*, 2015) is hard to follow due to the presence of non-preliminary information, which may impede a comprehensive understanding of the domain.

Nevertheless, recent studies have emphasized MS domain research using knowledge graphs for decision-making purposes. SPOKEsigs scheme presented in (Nelson *et al.*, 2022) incorporates domain knowledge into clinical diagnosis to potentially detect MS up to five years before a verified diagnosis in the clinic to determine individuals at risk of developing MS, with the aid of the RF algorithm. Likewise, (Fecho *et al.*, 2021) concentrated on finding rational relationships among entities provided in users' questions to the question-answering ROBOKOP knowledge graph. This knowledge graph has successfully found a relationship between MS and the impact of workplace exposures, which would confirm one of the leading causes of MS. However, both studies did not specify the comprehensive domain knowledge associated with MS. Specifically, (Nelson *et al.*, 2022) built a knowledge graph for

prediction purposes, while (Fecho *et al.*, 2021) aimed to make a more generalized knowledge-based schema that focuses on linking entities more than providing inclusive knowledge.

MS knowledge representation is needed to simplify decision-making in the domain using knowledge graph techniques (Alshamrani *et al.*, 2020). Therefore, the novelty of the research presented in this chapter is manifested in the comprehensive justification of the shared conceptualization of the MS domain. The imperative knowledge about MS is emphasized in the MSKG. For this reason, MSKG is superior to the knowledge graphs analyzed above as it significantly covers the most convenient medical concepts related to MS. It should be noted that MSKG includes entities provided in Schema.org, mainly under MedicalEntity which is an entity related to healthcare and medical practices (*Schema.org - Schema.org*, no date). As stated in (*Documentation - schema.org*, no date), Schema.org entities related to healthcare are devoted to defining ontologies that describe medical knowledge, which is often challenging for people to find and navigate online. The scope of sub-entities of MedicalEntity is to serve individuals who target health and medical knowledge, as intended in MSKG. Hence, MSKG would be persuasive for those who rely on Schema.org for extracting health and medical-type concepts. In short, MSKG would be a valuable contribution to MS research as it supports MS knowledge for shared decision-making purposes. From a medical point of view, MSKG allows one to easily find the answers to questions related to MS. By following the guided steps, MSKG will connect individuals to reliable sources in one single platform with confidence and without spending extra time browsing the web.

Due to the rarity of its cases, as indicated in Google Knowledge Panel when searching for it, MS needs to get proper research devotion compared to other chronic disorders, especially from the computer and data science perspective. Therefore, future work should utilize more knowledge base and knowledge representation approaches to understand MS progression, occurrence patterns, etiology, and long-term

effects on the quality of life of MS patients. Hypothetically, exploring the potential correlations between MS and other disorders that share commonalities should also be examined. Addressing these suggestions would augment the MS decision-making process.

Chapter 6 Conclusion, Limitations, and Future Recommendations

MS is a chronic neurological disease that hits the human body's brain and spinal cord in the long-term. It is the most prevalent inflammatory demyelinating autoimmune condition. As its underlying causes are indeed undetermined, MS is regarded as a complicated disease. Therefore, the decision-making process regarding this phenomenon is critical and considered shared decision-making. Automated solutions proposing the best alternatives that could make shared decision-making easier are highly in demand in the MS field. DSS technologies in the MS realm are favorable, especially for enhancing shared decision-making. Recent studies demonstrate the increasing need for DSSs in the MS domain to aid decision-makers in nominating the right decision among several alternatives.

To the best of the contributor’s knowledge, DSS is still not extensively exploited in MS's clinical settings due to the lack of collaboration between computer and data scientists and MS specialists, in addition to the MS specialists' low confidence level in model-based decision support technologies. Therefore, the overall objective of this dissertation is to design a DSS blueprint or prototype that emulates an MS expert to accurately identify the type of MS in affected patients and make appropriate DMT recommendations. This was done to improve MS diagnosis and treatment accuracy and efficiency. The ultimate goal is to provide a tool that can support the work of MS specialists, improve the accuracy of MS diagnosis, enhance the quality of care for patients with MS, and aid MS patients in forming decisions regarding their status. The proposed DSS uses domain knowledge representation and ML techniques to replicate MS experts, including their decision-making, reasoning, and judgment.

Addressing Research Questions 1,2, and 5

MS patients typically experience multiple MS episodes, ranging from mild to more aggressive. Because the course of MS has wide variances and is unpredictable, it is currently almost difficult to assess the severity and progression of MS over time in any affected patients. Given the clinical variability of this

condition, it is imperative to ascertain the MS progression level in affected patients. Predictions of MS development are beneficial in making the appropriate decisions, such as setting up therapies to the patient's needs. However, predicting the course of MS is challenging, and no tested prediction methods are currently available to anticipate how MS will progress. Furthermore, because MS progression and occurrence patterns are vague, it is complicated to forecast the disease's course. To address this concern, investigating several supervised ML models useful in predicting the type of MS in affected people and assessing the models' effectiveness using a real-world MS dataset addressed in this dissertation. To fulfill the research questions 1, 3, and 5 highlighted in Chapter 3, a framework that predicts MS types and contrasts the results of various traditional and ensemble classifiers was presented.

Several traditional models were first trained after performing feature selection and missing value imputation to rationally maintain the dimensionality and retain the data's original volume. Then, the best models were considered for further investigation for formulating several ensemble classifiers. Note that all models were trained using different splitting strategies. The models with the highest results were then augmented using hyperparameter tuning to enhance the top models' accuracy, sensitivity, and specificity. RF and DT showed constantly promising results in the three utilized evaluation metrics. Specifically, RF provided the highest accuracy and specificity results, while DT got the best sensitivity result. As for ensemble methods, bagging classifier achieved the top accuracy rate, AdaBoost classifier scored the best sensitivity rate, and gradient boosting got the highest specificity proportions. Regardless of the evaluation metric and splitting strategy used, the outcomes of all ensemble approaches were remarkably convergent, unlike traditional models. Ultimately, using hyperparameter tuning did not yield any considerable increases in all metrics. Nevertheless, these findings may demonstrate that ensemble classifiers generally outperform traditional classifiers in this quest.

The results are hypothetically accepted medically because the utilized models could be used as specific tests to rule in the presence of a certain MS type, supporting SPecific tests rule IN medical concept.

This part was designed as a proof-of-concept for the idea that ML algorithms may be used to predict MS progression. Patients' needs could be managed using this process as if they were in a genuine hospital setting. Debatably, the proposed methodology and the obtained results might be an encouraging step toward boosting ML applications in MS's clinical diagnosis and prognosis.

Addressing Research Questions 2,4, and 5

MS DMTs have become available for controlling MS, with varying modes of administration, risks, benefits, and side effects. An accurate diagnosis of MS and selecting the best DMT regimen are critical for improving the quality of life of MS patients. The selection of the best DMT, and the route of administration, depends on evaluating the risk of further progression of MS and considering the risk versus efficacy of a specific DMT. Automating the shared decision-making process between physicians and patients is required to tailor treatment plans based on individual patient conditions and symptoms.

Therefore, a framework for a model-based system that expedites the shared decision-making process using ML algorithms is demonstrated in this dissertation. To address research questions 2, 4, and 5, a framework that predicts the best route of administration suitable for MS patients who receive prescribed DMTs is presented in Chapter 4. The outcomes of this framework's underlying traditional and ensemble models were evaluated and compared against real-world MS datasets. The results of the study showed that the proposed methodology has the potential to accurately predict the best route of administration (oral, injectable, or infusion) for delivering DMTs to MS patients. The top-performing models were gradient boosting and RF.

The outcomes provide evidence supporting the potential usefulness of ML algorithms in enhancing the shared decision-making process associated with selecting the best MS DMTs. This part demonstrated the feasibility of utilizing ML algorithms to predict the most appropriate routes of administration for

DMTs prescribed to MS patients. The results obtained from this proof-of-concept investigation are viewed as an optimistic initial step in applying ML techniques to enhance shared clinical decision-making for DMT prescriptions in MS.

Addressing Research Question 6

As mentioned earlier, making critical decisions in this medical domain is vital. Again, MS diagnosis and treatment decisions are considered shared since MS is a preference-sensitive condition. Accordingly, the decision-makers should have the proper and comprehensive knowledge about MS. Improving MS decisions is feasible by enhancing the MS knowledge base and representation within this medical domain. This would be done by representing the necessary fundamental knowledge of MS through a knowledge graph. Representing MS domain knowledge would be extraordinarily beneficial, specifically for promoting rational MS decisions through distributing and reusing MS domain knowledge among the shared decision-makers.

The knowledge graph notion is not thoroughly studied for representing MS domain knowledge. Therefore, an easy-to-use and reliable source of MS information called MSKG is introduced in Chapter 5. MSKG targets newly diagnosed individuals with MS, their first-degree relatives, people at high risk of developing MS, and individuals interested in acquiring knowledge about MS. MSKG was constructed using ontology engineering approach to describe MS and several common factors related to it. Compared with existing knowledge graphs and ontologies associated with MS, MSKG is superior as it comprehensively shares the most convenient medical concepts related to MS.

Limitations

It is important to remember that developing a DSS to mimic an MS expert is a complex and time-consuming process requiring medical knowledge, data science, and computer science expertise.

Working with a team of experts from these fields is recommended in such research. However, most of the work presented in this dissertation was singly orchestrated by the Ph.D. candidate despite all the valuable collaborations with coauthors, colleagues, and experts. Consequently, the presented work has several limitations. Mainly, the need for more sufficient computer resources was a major obstacle. Even though this dissertation's outcomes are fairly promising, the availability of more powerful computing resources would elevate data preprocessing and complex algorithm application, promoting better results acceptable in MS's daily clinical routine and practices.

Finding relevant yet sufficient MS data was the most challenging task. Several healthcare institutions and medical research centers have been contacted to acquire the necessary data. Only one institution agreed to collaborate by providing the data used in the studies described in Chapter 3 and Chapter 4. Although, obtaining the data took considerable time and financial resources. Additionally, DSS relies on high-quality data to generate accurate and reliable decisions. In the case of MS, the quality of data available to DSS may be limited. The quality of the data acquired from the iConquerMS™ initiative was problematic, reflected in the prediction outcomes as indicated by the utilized evaluation metrics. For instance, the initial objective proposed in Chapter 4 was to predict the best DMT that could be recommended to a patient. However, the preliminary prediction results were very trivial and unacceptable because historical data is incomplete as the number of records of patients who reported their current DMTs is insufficient. Most of the investigation was spent on cleansing and preprocessing the data.

The number of existing works related to the objectives demonstrated in the previous chapters is minimal compared to other diseases such as Alzheimer's, various cancers, or diabetes, to name a few. Furthermore, funding and grant opportunities were insubstantial and restricted. All funding opportunities publicly announced by the College of Graduate Studies and the College of Engineering require US citizenship or are not allocated for health informatics and healthcare domain research. The

only grant acquired and approved for this dissertation was the Data Access Grant provided by IMCI. On top of that, much time and effort were devoted to self-education on the MS medical perspective to advance medically compelling and persuasive arguments.

In general, it is difficult and challenging to perform such prediction tasks using ML to develop useful DSS that could fulfill the presented objective because almost no theory can guarantee a solution. Although it is possible to develop some intuition about which strategies work the best for what type of tasks, in the end, it requires lots of trial and error to see which strategies work in practice. That is why it is important to be able to evaluate many strategies. Many different strategies and models have been examined and evaluated in this work using different subsets of the input features, different model-based feature selection and missing value imputation techniques, and different ML models. This was an exhaustive and time-consuming effort. Despite all these limitations, these circumstances allowed presenting a novel and empirical work.

Future Recommendations

Future work should focus on applying DSS technologies to understand the MS's occurrence patterns, emphasize MS etiology, highlight the MS's long-term effects on the affected individuals' quality of life, and find the correlation between MS and other disorders. Implementing automated decision-making systems, particularly for newly diagnosed MS patients, should be given priority in future studies. ML and deep learning technologies have the potential to be helpful in MS research and the MS domain. Further research is necessary to fully realize the potential benefits of ML and deep learning in clinical practices and decision-making for MS. Future work should also contribute toward expanding MSKG's base knowledge to outline disorders that might be correlated with MS. Thus, MSKG would be improved by utilizing knowledge-based reasoning methods to recognize MS patterns in terms of progression, occurrence, etiology, and long- and short-term effects during the disease course.

Considering these recommendations would enrich the research of knowledge graph applications in the MS domain. Lastly and most importantly, scientists conducting research in the MS domain require extensive data, but access to them is restricted. Thus, MS data should be available and easily accessible. In conclusion, DSSs technologies have the potential to be pragmatic in the MS domain and research. Thus, considering these recommendations would enrich the MS field.

References

Alfano, B. *et al.* (2007) 'An Ontology Approach for Classification of Abnormal White Matter in Patients with Multiple Sclerosis', in A. Holzinger (ed.) *HCI and Usability for Medicine and Health Care*. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), pp. 389–402. Available at: https://doi.org/10.1007/978-3-540-76805-0_34.

Almasi, F. *et al.* (no date) 'Designing a Clinical Decision Support System for Diagnosing Multiple Sclerosis (MS)'.

Alshamrani, R. *et al.* (2020) 'Model-Driven Decision Making in Multiple Sclerosis Research: Existing Works and Latest Trends', *Patterns*, 1(8), p. 100121. Available at: https://doi.org/10.1016/j.patter.2020.100121.

Arani, L. *et al.* (2018) 'Intelligent Computer Systems for Multiple Sclerosis Diagnosis: a Systematic Review of Reasoning Techniques and Methods', *Acta Informatica Medica*, 26(4), p. 258. Available at: https://doi.org/10.5455/aim.2018.26.258-264.

Arroyo, R. *et al.* (2017) 'Conjoint analysis to understand preferences of patients with multiple sclerosis for disease-modifying therapy attributes in Spain: a cross-sectional observational study', *BMJ Open*, 7(3), p. e014433. Available at: https://doi.org/10.1136/bmjopen-2016-014433.

Ascherio, A. and Munger, K. (2016) 'Epidemiology of Multiple Sclerosis: From Risk Factors to Prevention—An Update', *Seminars in Neurology*, 36(02), pp. 103–114. Available at: https://doi.org/10.1055/s-0036-1579693.

Awawdeh, S., Faris, H. and Hiary, H. (2022) 'EvoImputer: An evolutionary approach for Missing Data Imputation and feature selection in the context of supervised learning', *Knowledge-Based Systems*, 236, p. 107734. Available at: https://doi.org/10.1016/j.knosys.2021.107734.

BERNERS-LEE, T., HENDLER, J. and LASSILA, O. (2001) 'THE SEMANTIC WEB', *Scientific American*, 284(5), pp. 34–43.

Bertsimas, D., Pawlowski, C. and Zhuo, Y.D. (no date) 'From Predictive Methods to Missing Data Imputation: An Optimization Approach'.

Bowen, J. *et al.* (2020) 'Treatment Patterns Among Patients with Multiple Sclerosis Initiating Second-Line Disease-Modifying Therapy', *Advances in Therapy*, 37(7), pp. 3163–3177. Available at: https://doi.org/10.1007/s12325-020-01367-1.

Browne, P. *et al.* (2014) 'Atlas of Multiple Sclerosis 2013: A growing global problem with widespread inequity', *Neurology*, 83(11), pp. 1022–1024. Available at: https://doi.org/10.1212/WNL.0000000000000768.

Brownlee, W.J. *et al.* (2017) 'Diagnosis of multiple sclerosis: progress and challenges', *The Lancet*, 389(10076), pp. 1336–1346. Available at: https://doi.org/10.1016/S0140-6736(16)30959-X.

Cevallos Valdiviezo, H. and Van Aelst, S. (2015) 'Tree-based prediction on incomplete data using imputation or surrogate decisions', *Information Sciences*, 311, pp. 163–181. Available at: https://doi.org/10.1016/j.ins.2015.03.018.

Chiaravalloti, N.D. and DeLuca, J. (2008) 'Cognitive impairment in multiple sclerosis', *The Lancet Neurology*, 7(12), pp. 1139–1151. Available at: https://doi.org/10.1016/S1474-4422(08)70259-X.

Christo, V.R.E. *et al.* (2022) 'Feature Selection and Instance Selection from Clinical Datasets Using Co-operative Co-evolution and Classification Using Random Forest', *IETE Journal of Research*, 68(4), pp. 2508–2521. Available at: https://doi.org/10.1080/03772063.2020.1713917.

Clark, L.A. *et al.* (2019) 'PSY20 UNDERSTANDING DISEASE-MODIFYING THERAPY ADMINISTRATION ROUTE SUITABILITY IN DIFFERENT MULTIPLE SCLEROSIS PATIENT SEGMENTS IN THE 5EU AND US', *Value in Health*, 22, p. S378. Available at: https://doi.org/10.1016/j.jval.2019.04.1845.

Col, N. *et al.* (2019) 'A Novel Tool to Improve Shared Decision Making and Adherence in Multiple Sclerosis: Development and Preliminary Testing', *MDM Policy & Practice*, 4(2), p. 238146831987913. Available at: https://doi.org/10.1177/2381468319879134.

Colligan, E., Metzler, A. and Tiryaki, E. (2017) 'Shared decision-making in multiple sclerosis', *Multiple Sclerosis Journal*, 23(2), pp. 185–190. Available at: https://doi.org/10.1177/1352458516671204.

Confavreux, C. and Vukusic, S. (2006) 'Natural history of multiple sclerosis: a unifying concept', *Brain*, 129(3), pp. 606–616. Available at: https://doi.org/10.1093/brain/awl007.

Conroy, B. *et al.* (2016) 'A dynamic ensemble approach to robust classification in the presence of missing data', *Machine Learning*, 102(3), pp. 443–463. Available at: https://doi.org/10.1007/s10994-015-5530-z.

Cui, N., Hu, J. and Liang, F. (2021) 'Complementary dimension reduction', *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 14(1), pp. 31–40. Available at: https://doi.org/10.1002/sam.11484.

D'Amico, E. *et al.* (2016) 'A Personalized Approach in Progressive Multiple Sclerosis: The Current Status of Disease Modifying Therapies (DMTs) and Future Perspectives', *International Journal of Molecular Sciences*, 17(10), p. 1725. Available at: https://doi.org/10.3390/ijms17101725.

De Falco, I., Scafuri, U. and Tarantino, E. (2016) 'A Differential Evolution approach for classification of Multiple Sclerosis lesions', in *2016 IEEE Symposium on Computers and Communication (ISCC)*. *2016 IEEE Symposium on Computers and Communication (ISCC)*, Messina, Italy: IEEE, pp. 141–146. Available at: https://doi.org/10.1109/ISCC.2016.7543729.

Desarkar, M.S. *et al.* (2013) 'Med-Tree: A user knowledge graph framework for medical applications', in *13th IEEE International Conference on BioInformatics and BioEngineering*. *2013 IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE)*, Chania, Greece: IEEE, pp. 1–4. Available at: https://doi.org/10.1109/BIBE.2013.6701564.

Dobson, R. and Giovannoni, G. (2019) 'Multiple sclerosis – a review', *European Journal of Neurology*, 26(1), pp. 27–40. Available at: https://doi.org/10.1111/ene.13819.

Documentation - schema.org (no date). Available at: https://schema.org/docs/meddocs.html (Accessed: 5 October 2021).

Dogan, A.F. and Duru, D.G. (2019) 'Comparison of Machine Learning Techniques on MS Lesion Segmentation', in *2019 Medical Technologies Congress (TIPTEKNO)*. *2019 Medical Technologies Congress (TIPTEKNO)*, Izmir, Turkey: IEEE, pp. 1–4. Available at: https://doi.org/10.1109/TIPTEKNO.2019.8895202.

Efendi, H. (2016) 'Clinically Isolated Syndromes: Clinical Characteristics, Differential Diagnosis, and Management', *Noro Psikiyatri Arsivi*, 52(Suppl.1), pp. 1–11. Available at: https://doi.org/10.5152/npa.2015.12608.

Ekşi, Z. *et al.* (2020) 'Differentiation of relapsing-remitting and secondary progressive multiple sclerosis: a magnetic resonance spectroscopy study based on machine learning', *Arquivos de Neuro-Psiquiatria*, 78(12), pp. 789–796. Available at: https://doi.org/10.1590/0004-282x20200094.

Eom, S. and Kim, E. (2006) 'A survey of decision support system applications (1995–2001)', *Journal of the Operational Research Society*, 57(11), pp. 1264–1278. Available at: https://doi.org/10.1057/palgrave.jors.2602140.

Eskyte, I. *et al.* (2019) 'Understanding treatment decisions from the perspective of people with relapsing remitting multiple Sclerosis: A critical interpretive synthesis', *Multiple Sclerosis and Related Disorders*, 27, pp. 370–377. Available at: https://doi.org/10.1016/j.msard.2018.11.016.

Esposito, M., De Falco, I. and De Pietro, G. (2011) 'An evolutionary-fuzzy DSS for assessing health status in multiple sclerosis disease', *International Journal of Medical Informatics*, 80(12), pp. e245–e254. Available at: https://doi.org/10.1016/j.ijmedinf.2011.09.003.

Esposito, M. and De Pietro, G. (2011) 'An ontology-based fuzzy decision support system for multiple sclerosis', *Engineering Applications of Artificial Intelligence*, 24(8), pp. 1340–1354. Available at: https://doi.org/10.1016/j.engappai.2011.02.002.

Faisal, S. and Tutz, G. (2022) 'Nearest neighbor imputation for categorical data by weighting of attributes', *Information Sciences*, 592, pp. 306–319. Available at: https://doi.org/10.1016/j.ins.2022.01.056.

Fecho, K. *et al.* (2021) 'A Biomedical Knowledge Graph System to Propose Mechanistic Hypotheses for Real-World Environmental Health Observations: Cohort Study and Informatics Application', *JMIR Medical Informatics*, 9(7), p. e26714. Available at: https://doi.org/10.2196/26714.

Finkelstein, J., Wood, J. and Shan, Y. (2011) 'Implementing physical telerehabilitation system for patients with multiple sclerosis', in *2011 4th International Conference on Biomedical Engineering and Informatics (BMEI)*. *2011 4th International Conference on Biomedical Engineering and Informatics (BMEI)*, Shanghai, China: IEEE, pp. 1883–1886. Available at: https://doi.org/10.1109/BMEI.2011.6098758.

Fiorini, S. *et al.* (2015) 'A machine learning pipeline for multiple sclerosis course detection from clinical scales and patient reported outcomes', in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan: IEEE, pp. 4443–4446. Available at: https://doi.org/10.1109/EMBC.2015.7319381.

Fraccaro, P. *et al.* (2015) 'Behind the screens: Clinical decision support methodologies – A review', *Health Policy and Technology*, 4(1), pp. 29–38. Available at: https://doi.org/10.1016/j.hlpt.2014.10.001.

deAndrés-Galiana, E.J. *et al.* (2019) 'Analysis of defective pathways and drug repositioning in Multiple Sclerosis via machine learning approaches', *Computers in Biology and Medicine*, 115, p. 103492. Available at: https://doi.org/10.1016/j.compbiomed.2019.103492.

Ghasemi, N. (2017) 'Multiple Sclerosis: Pathogenesis, Symptoms, Diagnoses and Cell-Based Therapy', *CELL JOURNAL*, 19(1), p. 10.

Goldenberg, M.M. (2012) 'Multiple Sclerosis Review', *P & T: A Peer-Reviewed Journal for Formulary Management*, 37(3), pp. 175–184.

Granberg, T. *et al.* (2013) 'Radiologically isolated syndrome – incidental magnetic resonance imaging findings suggestive of multiple sclerosis, a systematic review', *Multiple Sclerosis Journal*, 19(3), pp. 271–280. Available at: https://doi.org/10.1177/1352458512451943.

Gustafsson, M. *et al.* (2014) 'Modules, networks and systems medicine for understanding disease and aiding diagnosis', *Genome Medicine*, 6(10), p. 82. Available at: https://doi.org/10.1186/s13073-014-0082-6.

Gutierrez, C. and Sequeda, J.F. (2021) 'Knowledge graphs', *Communications of the ACM*, 64(3), pp. 96–104. Available at: https://doi.org/10.1145/3418294.

Hadzic, M., Chen, M. and Dillon, T.S. (2008) 'Towards the Mental Health Ontology', in *2008 IEEE International Conference on Bioinformatics and Biomedicine*. *2008 IEEE International Conference on Bioinformatics and Biomedicine*, Philadelphia, PA, USA: IEEE, pp. 284–288. Available at: https://doi.org/10.1109/BIBM.2008.59.

Hauer, L., Perneczky, J. and Sellner, J. (2021) 'A global view of comorbidity in multiple sclerosis: a systematic review with a focus on regional differences, methodology, and clinical implications', *Journal of Neurology*, 268(11), pp. 4066–4077. Available at: https://doi.org/10.1007/s00415-020-10107-y.

Hillert, J. and Stawiarz, L. (2015) 'The Swedish MS registry - clinical support tool and scientific resource', *Acta Neurologica Scandinavica*, 132, pp. 11–19. Available at: https://doi.org/10.1111/ane.12425.

Ho, J.C., Ghosh, J. and Unnikrishnan, K.P. (2013) 'Risk Prediction of a Multiple Sclerosis Diagnosis', in *2013 IEEE International Conference on Healthcare Informatics*. *2013 IEEE International Conference on Healthcare Informatics (ICHI)*, Philadelphia, PA, USA: IEEE, pp. 175–183. Available at: https://doi.org/10.1109/ICHI.2013.31.

Hogan, A. *et al.* (2021) 'Knowledge Graphs', *arXiv:2003.02320 [cs]* [Preprint]. Available at: http://arxiv.org/abs/2003.02320 (Accessed: 25 April 2021).

Ion-Mărgineanu, A. *et al.* (2017) 'Machine Learning Approach for Classifying Multiple Sclerosis Courses by Combining Clinical Data with Lesion Loads and Magnetic Resonance Metabolic Features', *Frontiers in Neuroscience*, 11, p. 398. Available at: https://doi.org/10.3389/fnins.2017.00398.

Jensen, M. *et al.* (2013) 'The neurological disease ontology', *Journal of Biomedical Semantics*, 4(1), p. 42. Available at: https://doi.org/10.1186/2041-1480-4-42.

Kabir, S. and Farrokhvar, L. (2022) 'Non-linear missing data imputation for healthcare data via index-aware autoencoders', *Health Care Management Science*, 25(3), pp. 484–497. Available at: https://doi.org/10.1007/s10729-022-09597-1.

Karaca, Y. *et al.* (2017) 'The Differential Diagnosis of Multiple Sclerosis Using Convex Combination of Infinite Kernels', *CNS & Neurological Disorders - Drug Targets*, 16(1), pp. 36–43. Available at: https://doi.org/10.2174/1871527315666161024142439.

Karaca, Y. and Hayta, Ş. (2015) 'The Significance of Artificial Neural Networks Algorithms Classification in the Multiple Sclerosis and Its Subgroups', *IARJSET*, 2(12), pp. 1–7. Available at: https://doi.org/10.17148/IARJSET.2015.21201.

Kasper, J. *et al.* (2012) 'Applying the theory of planned behaviour to multiple sclerosis patients' decisions on disease modifying therapy – questionnaire concept and validation', *BMC Medical Informatics and Decision Making*, 12(1), p. 60. Available at: https://doi.org/10.1186/1472-6947-12-60.

Kawamoto, K. *et al.* (2005) 'Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success', *BMJ*, 330(7494), p. 765. Available at: https://doi.org/10.1136/bmj.38398.500764.8F.

Kendall, E.F. and McGuinness, D.L. (2019) *Ontology Engineering*.

Khalaf, M. *et al.* (2017) 'Machine learning approaches to the application of disease modifying therapy for sickle cell using classification models', *Neurocomputing*, 228, pp. 154–164. Available at: https://doi.org/10.1016/j.neucom.2016.10.043.

Kobelt, G. *et al.* (2017) 'New insights into the burden and costs of multiple sclerosis in Europe', *Multiple Sclerosis Journal*, 23(8), pp. 1123–1136. Available at: https://doi.org/10.1177/1352458517694432.

Kragt, J. *et al.* (2009) 'Higher levels of 25-hydroxyvitamin D are associated with a lower incidence of multiple sclerosis only in women', *Multiple Sclerosis Journal*, 15(1), pp. 9–15. Available at: https://doi.org/10.1177/1352458508095920.

Kurtzke, J.F. (1983) 'Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS)', *Neurology*, 33(11), pp. 1444–1444. Available at: https://doi.org/10.1212/WNL.33.11.1444.

Langhorne, A.L., Thomas, P. and Kolaczkowski, L. (2013) 'Untangling the Web of e-Health: Multiple Sclerosis Patients' Perceptions of Online Health Information, Information Literacy, and the Impact on Treatment Decision Making', in A.A. Ozok and P. Zaphiris (eds) *Online Communities and Social Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), pp. 304–312. Available at: https://doi.org/10.1007/978-3-642-39371-6_35.

Lassmann, H., Brück, W. and Lucchinetti, C.F. (2007) 'The Immunopathology of Multiple Sclerosis: An Overview', *Brain Pathology*, 17(2), pp. 210–218. Available at: https://doi.org/10.1111/j.1750-3639.2007.00064.x.

Lee Mortensen, G. and Rasmussen, P.V. (2017) 'The impact of quality of life on treatment preferences in multiple sclerosis patients', *Patient Preference and Adherence*, Volume 11, pp. 1789–1796. Available at: https://doi.org/10.2147/PPA.S142373.

Leray, E. *et al.* (2016) 'Epidemiology of multiple sclerosis', *Revue Neurologique*, 172(1), pp. 3–13. Available at: https://doi.org/10.1016/j.neurol.2015.10.006.

Li, J., Huang, Y. and Aparasu, Rr. (2022) 'P23 Machine Learning Approaches to Evaluate Treatment Switching in Patients with Multiple Sclerosis: Analyses of Electronic Medical Records', *Value in Health*, 25(7), p. S291. Available at: https://doi.org/10.1016/j.jval.2022.04.031.

Lim, S., Tucker, C.S. and Kumara, S. (2017) 'An unsupervised machine learning model for discovering latent infectious diseases using social media data', *Journal of Biomedical Informatics*, 66, pp. 82–94. Available at: https://doi.org/10.1016/j.jbi.2016.12.007.

Linder, R. *et al.* (2009) 'Computer-Aided Diagnosis of Multiple Sclerosis', *Computational and Mathematical Methods in Medicine*, 10(1), pp. 39–47. Available at: https://doi.org/10.1080/17486700802070724.

Liu, C.-H. *et al.* (2020) 'The Feature Selection Effect on Missing Value Imputation of Medical Datasets', *Applied Sciences*, 10(7), p. 2344. Available at: https://doi.org/10.3390/app10072344.

Loizou, C.P. *et al.* (2011) 'Multiscale Amplitude-Modulation Frequency-Modulation (AM–FM) Texture Analysis of Multiple Sclerosis in Brain MRI Images', *IEEE Transactions on Information Technology in Biomedicine*, 15(1), pp. 119–129. Available at: https://doi.org/10.1109/TITB.2010.2091279.

Lopez, C. *et al.* (2018) 'An unsupervised machine learning method for discovering patient clusters based on genetic signatures', *Journal of Biomedical Informatics*, 85, pp. 30–39. Available at: https://doi.org/10.1016/j.jbi.2018.07.004.

Lublin, F.D. *et al.* (2014) 'Defining the clinical course of multiple sclerosis: The 2013 revisions', *Neurology*, 83(3), pp. 278–286. Available at: https://doi.org/10.1212/WNL.0000000000000560.

Lynch, C.J. and Liston, C. (2018) 'New machine-learning technologies for computer-aided diagnosis', *Nature Medicine*, 24(9), pp. 1304–1305. Available at: https://doi.org/10.1038/s41591-018-0178-4.

Ma, K.C. (2018) *Imaging Informatics-Based Electronic Patient Record and Analysis System for Multiple Sclerosis Research, Treatment, and Disease Tracking*. PhD Thesis. University of Southern California.

Ma, X. *et al.* (2014) 'Ontology dynamics in a data life cycle: Challenges and recommendations from a Geoscience Perspective', *Journal of Earth Science*, 25(2), pp. 407–412. Available at: https://doi.org/10.1007/s12583-014-0408-8.

Mahad, D.H., Trapp, B.D. and Lassmann, H. (2015) 'Pathological mechanisms in progressive multiple sclerosis', *The Lancet Neurology*, 14(2), pp. 183–193. Available at: https://doi.org/10.1016/S1474-4422(14)70256-X.

Malhotra, A. *et al.* (2015) 'Knowledge Retrieval from PubMed Abstracts and Electronic Medical Records with the Multiple Sclerosis Ontology', *PLOS ONE*. Edited by V. Grolmusz, 10(2), p. e0116718. Available at: https://doi.org/10.1371/journal.pone.0116718.

Mateos-Pérez, J.M. *et al.* (2018) 'Structural neuroimaging as clinical predictor: A review of machine learning applications', *NeuroImage: Clinical*, 20, pp. 506–522. Available at: https://doi.org/10.1016/j.nicl.2018.08.019.

McDonald, W.I. *et al.* (2001) 'Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis', *Annals of Neurology*, 50(1), pp. 121–127. Available at: https://doi.org/10.1002/ana.1032.

McGinley, M.P., Goldschmidt, C.H. and Rae-Grant, A.D. (2021) 'Diagnosis and Treatment of Multiple Sclerosis: A Review', *JAMA*, 325(8), p. 765. Available at: https://doi.org/10.1001/jama.2020.26858.

McGinnis, R.S. *et al.* (2017) 'A machine learning approach for gait speed estimation using skin-mounted wearable sensors: From healthy controls to individuals with multiple sclerosis', *PLOS ONE*. Edited by J.-M. Yang, 12(6), p. e0178366. Available at: https://doi.org/10.1371/journal.pone.0178366.

McNicholas, N. *et al.* (2019) 'New versus old: Implications of evolving diagnostic criteria for relapsing–remitting multiple sclerosis', *Multiple Sclerosis Journal*, 25(6), pp. 867–870. Available at: https://doi.org/10.1177/1352458518770088.

Meca-Lallana, V. *et al.* (2017) 'Patients' Information Sources and Needs in Multiple Sclerosis: the Infoseek-MS Questionnaire', *Value in Health*, 20(9), p. A726. Available at: https://doi.org/10.1016/j.jval.2017.08.1964.

Milo, R. and Kahana, E. (2010) 'Multiple sclerosis: Geoepidemiology, genetics and the environment', *Autoimmunity Reviews*, 9(5), pp. A387–A394. Available at: https://doi.org/10.1016/j.autrev.2009.11.010.

Miri Ashtiani, S.N. *et al.* (2018) 'Altered topological properties of brain networks in the early MS patients revealed by cognitive task-related fMRI and graph theory', *Biomedical Signal Processing and Control*, 40, pp. 385–395. Available at: https://doi.org/10.1016/j.bspc.2017.10.006.

Montalban, X. *et al.* (2018) 'ECTRIMS/EAN Guideline on the pharmacological treatment of people with multiple sclerosis', *Multiple Sclerosis Journal*, 24(2), pp. 96–120. Available at: https://doi.org/10.1177/1352458517751049.

Mowry, E.M. *et al.* (2018) 'Incorporating machine learning approaches to assess putative environmental risk factors for multiple sclerosis', *Multiple Sclerosis and Related Disorders*, 24, pp. 135–141. Available at: https://doi.org/10.1016/j.msard.2018.06.009.

multiple sclerosis (no date). Available at: https://www.ebi.ac.uk/ols/ontologies/efo/terms?short_form=EFO_0003885 (Accessed: 4 October 2021).

Multiple sclerosis ontology - Summary | NCBO BioPortal (no date). Available at: https://bioportal.bioontology.org/ontologies/MSO/?p=summary (Accessed: 4 October 2021).

Nagarajan, G. and Dhinesh Babu, L.D. (2022) ‘Missing data imputation on biomedical data using deeply learned clustering and L2 regularized regression based on symmetric uncertainty’, *Artificial Intelligence in Medicine*, 123, p. 102214. Available at: https://doi.org/10.1016/j.artmed.2021.102214.

Nelson, C.A. *et al.* (2022) ‘Embedding electronic health records onto a knowledge network recognizes prodromal features of multiple sclerosis and predicts diagnosis’, *Journal of the American Medical Informatics Association*, 29(3), pp. 424–434. Available at: https://doi.org/10.1093/jamia/ocab270.

Nicholson, D.N. and Greene, C.S. (2020) ‘Constructing knowledge graphs and their biomedical applications’, *Computational and Structural Biotechnology Journal*, 18, pp. 1414–1428. Available at: https://doi.org/10.1016/j.csbj.2020.05.017.

Obermeyer, Z. and Lee, T.H. (2017) ‘Lost in Thought — The Limits of the Human Mind and the Future of Medicine’, *New England Journal of Medicine*, 377(13), pp. 1209–1211. Available at: https://doi.org/10.1056/NEJMp1705348.

Okuda, D.T. *et al.* (2009) ‘Incidental MRI anomalies suggestive of multiple sclerosis: The radiologically isolated syndrome’, *Neurology*, 72(9), pp. 800–805. Available at: https://doi.org/10.1212/01.wnl.0000335764.14513.1a.

Pan, R. *et al.* (2015) ‘Missing data imputation by K nearest neighbours based on grey relational structure and mutual information’, *Applied Intelligence*, 43(3), pp. 614–632. Available at: https://doi.org/10.1007/s10489-015-0666-x.

Peeters, L.M. (2018) ‘Fair data for next-generation management of multiple sclerosis’, *Multiple Sclerosis Journal*, 24(9), pp. 1151–1156. Available at: https://doi.org/10.1177/1352458517748475.

Pérez del Palomar, A. *et al.* (2019) ‘Swept source optical coherence tomography to early detect multiple sclerosis disease. The use of machine learning techniques’, *PLOS ONE*. Edited by S. Bhattacharya, 14(5), p. e0216410. Available at: https://doi.org/10.1371/journal.pone.0216410.

Pinto, M.F. da S. (2018) *Evaluation and prediction of Multiple Sclerosis Disease Progression*. M.S. Thesis. Universidade de Coimbra.

Pota, M., Esposito, M. and Pietro, G.D. (2012) ‘Hybridization of possibility theory and supervised clustering to build DSSs for classification in medicine’, in *2012 12th International Conference on Hybrid Intelligent Systems (HIS)*. *2012 12th International Conference on Hybrid Intelligent Systems (HIS)*, Pune, India: IEEE, pp. 490–495. Available at: https://doi.org/10.1109/HIS.2012.6421383.

Pourakbari, H. *et al.* (2014) ‘Proper Features Extraction from the Multiple Sclerosis Disease Postural Disorders for Decision Support System Definition’, *Applied Mechanics and Materials*, 666, pp. 230–234. Available at: https://doi.org/10.4028/www.scientific.net/AMM.666.230.

Power, D.J. (2002) *Decision support systems: concepts and resources for managers*. Westport, Conn: Quorum Books.

Power, D.J. and Sharda, R. (2007) ‘Model-driven decision support systems: Concepts and research directions’, *Decision Support Systems*, 43(3), pp. 1044–1061. Available at: https://doi.org/10.1016/j.dss.2005.05.030.

Power, M., Fell, G. and Wright, M. (2013) 'Principles for high-quality, high-value testing', *Evidence Based Medicine*, 18(1), pp. 5–10. Available at: https://doi.org/10.1136/eb-2012-100645.

Rahn, A.C. *et al.* (2015) 'Evaluator-blinded trial evaluating nurse-led immunotherapy DEcision Coaching In persons with relapsing-remitting Multiple Sclerosis (DECIMS) and accompanying process evaluation: study protocol for a cluster randomised controlled trial', *Trials*, 16(1), p. 106. Available at: https://doi.org/10.1186/s13063-015-0611-7.

Rahn, A.C. *et al.* (2018) 'Nurse-led immunotreatment DEcision Coaching In people with Multiple Sclerosis (DECIMS) – Feasibility testing, pilot randomised controlled trial and mixed methods process evaluation', *International Journal of Nursing Studies*, 78, pp. 26–36. Available at: https://doi.org/10.1016/j.ijnurstu.2017.08.011.

Raja, P.S. and Thangavel, K. (2020) 'Missing value imputation using unsupervised machine learning techniques', *Soft Computing*, 24(6), pp. 4361–4392. Available at: https://doi.org/10.1007/s00500-019-04199-6.

Reddel, S.W. *et al.* (2019) 'Successful implementation of an automated electronic support system for patient safety monitoring: The alemtuzumab in multiple sclerosis safety systems (AMS3) study', *Multiple Sclerosis Journal*, 25(8), pp. 1124–1131. Available at: https://doi.org/10.1177/1352458518783673.

Reich, D.S., Lucchinetti, C.F. and Calabresi, P.A. (2018) 'Multiple Sclerosis', *New England Journal of Medicine*. Edited by D.L. Longo, 378(2), pp. 169–180. Available at: https://doi.org/10.1056/NEJMra1401483.

Riaño, D. *et al.* (2012) 'An ontology-based personalization of health-care knowledge to support clinical decisions for chronically ill patients', *Journal of Biomedical Informatics*, 45(3), pp. 429–446. Available at: https://doi.org/10.1016/j.jbi.2011.12.008.

Rieckmann, P. *et al.* (2013) 'Future MS care: a consensus statement of the MS in the 21st Century Steering Group', *Journal of Neurology*, 260(2), pp. 462–469. Available at: https://doi.org/10.1007/s00415-012-6656-6.

Rothman, B., Leonard, J.C. and Vigoda, M.M. (2012) 'Future of Electronic Health Records: Implications for Decision Support', *Mount Sinai Journal of Medicine: A Journal of Translational and Personalized Medicine*, 79(6), pp. 757–768. Available at: https://doi.org/10.1002/msj.21351.

Rotmensch, M. *et al.* (2017) 'Learning a Health Knowledge Graph from Electronic Medical Records', *Scientific Reports*, 7(1), p. 5994. Available at: https://doi.org/10.1038/s41598-017-05778-z.

Saccà, V. *et al.* (2019) 'Evaluation of machine learning algorithms performance for the prediction of early multiple sclerosis from resting-state FMRI connectivity data', *Brain Imaging and Behavior*, 13(4), pp. 1103–1114. Available at: https://doi.org/10.1007/s11682-018-9926-9.

Salem, M. *et al.* (2018) 'A supervised framework with intensity subtraction and deformation field features for the detection of new T2-w lesions in multiple sclerosis', *NeuroImage: Clinical*, 17, pp. 607–615. Available at: https://doi.org/10.1016/j.nicl.2017.11.015.

Sánchez, D. *et al.* (2015) ‘Semantic variance: An intuitive measure for ontology accuracy evaluation’, *Engineering Applications of Artificial Intelligence*, 39, pp. 89–99. Available at: https://doi.org/10.1016/j.engappai.2014.11.012.

Sanders, G., Pearce, R. and Baranzini, S. (2020) *Topological Analysis of The SPOKE Graph*. LLNL-SR--815089, 1669224, 1024157, p. LLNL-SR--815089, 1669224, 1024157. Available at: https://doi.org/10.2172/1669224.

Schema.org - Schema.org (no date). Available at: https://schema.org/docs/full.html (Accessed: 4 October 2021).

Seba, P.A. and Benifa, J.V.B. (2022) ‘Relevancy contemplation in medical data analytics and ranking of feature selection algorithms’, *ETRI Journal*, p. etrij.2022-0018. Available at: https://doi.org/10.4218/etrij.2022-0018.

Seccia, R. *et al.* (2020) ‘Considering patient clinical history impacts performance of machine learning models in predicting course of multiple sclerosis’, *PLOS ONE*. Edited by F. Lolli, 15(3), p. e0230219. Available at: https://doi.org/10.1371/journal.pone.0230219.

Shanmuganathan, M. *et al.* (2020) ‘Review of advanced computational approaches on multiple sclerosis segmentation and classification’, *IET Signal Processing*, 14(6), pp. 333–341. Available at: https://doi.org/10.1049/iet-spr.2019.0543.

Sharma, M. and Kaur, P. (2021) ‘A Comprehensive Analysis of Nature-Inspired Meta-Heuristic Techniques for Feature Selection Problem’, *Archives of Computational Methods in Engineering*, 28(3), pp. 1103–1127. Available at: https://doi.org/10.1007/s11831-020-09412-6.

Siddiqui, M.F., Reza, A.W. and Kanesan, J. (2015) ‘An Automated and Intelligent Medical Decision Support System for Brain MRI Scans Classification’, *PLOS ONE*. Edited by G.P.S. Raghava, 10(8), p. e0135875. Available at: https://doi.org/10.1371/journal.pone.0135875.

Silva-Ramírez, E.-L., Pino-Mejías, R. and López-Coello, M. (2015) ‘Single imputation with multilayer perceptron and multiple imputation combining multilayer perceptron and k-nearest neighbours for monotone patterns’, *Applied Soft Computing*, 29, pp. 65–74. Available at: https://doi.org/10.1016/j.asoc.2014.09.052.

Sormani, M.P. *et al.* (2021) ‘Disease-Modifying Therapies and Coronavirus Disease 2019 Severity in Multiple Sclerosis’, *Annals of Neurology*, 89(4), pp. 780–789. Available at: https://doi.org/10.1002/ana.26028.

Sospedra, M. and Martin, R. (2005) ‘IMMUNOLOGY OF MULTIPLE SCLEROSIS’, *Annual Review of Immunology*, 23(1), pp. 683–747. Available at: https://doi.org/10.1146/annurev.immunol.23.021704.115707.

Suganthi, M. and Karunakaran, V. (2019) ‘Instance selection and feature extraction using cuttlefish optimization algorithm and principal component analysis using decision tree’, *Cluster Computing*, 22(S1), pp. 89–101. Available at: https://doi.org/10.1007/s10586-018-1821-z.

Sweeney, E.M. *et al.* (2014) ‘A Comparison of Supervised Machine Learning Algorithms and Feature Vectors for MS Lesion Segmentation Using Multimodal Structural MRI’, *PLoS ONE*. Edited by B. Draganski, 9(4), p. e95753. Available at: https://doi.org/10.1371/journal.pone.0095753.

Talwar, A. *et al.* (2021) 'PND67 Factors Associated with Prescribing of Disease-Modifying Agents Among Older Adults with Multiple Sclerosis: A Real-World Analysis of Electronic Medical Records', *Value in Health*, 24, pp. S171–S172. Available at: https://doi.org/10.1016/j.jval.2021.04.851.

Taschler, B. *et al.* (no date) 'Classification of Multiple Sclerosis Patients from the Geometry and Texture of White Matter Lesions'.

the Sylvia Lawry Centre for Multiple Sclerosis Research *et al.* (2007) 'Prognosis of the individual course of disease - steps in developing a decision support tool for Multiple Sclerosis', *BMC Medical Informatics and Decision Making*, 7(1), p. 11. Available at: https://doi.org/10.1186/1472-6947-7-11.

Thompson, A.J. *et al.* (2018) 'Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria', *The Lancet Neurology*, 17(2), pp. 162–173. Available at: https://doi.org/10.1016/S1474-4422(17)30470-2.

Tintore, M. *et al.* (2015) 'Defining high, medium and low impact prognostic factors for developing multiple sclerosis', *Brain*, 138(7), pp. 1863–1874. Available at: https://doi.org/10.1093/brain/awv105.

Trapp, B.D. *et al.* (1998) 'Axonal Transection in the Lesions of Multiple Sclerosis', *New England Journal of Medicine*, 338(5), pp. 278–285. Available at: https://doi.org/10.1056/NEJM199801293380502.

Tuggener, L. *et al.* (2019) 'Automated Machine Learning in Practice: State of the Art and Recent Results', in *2019 6th Swiss Conference on Data Science (SDS)*. *2019 6th Swiss Conference on Data Science (SDS)*, Bern, Switzerland: IEEE, pp. 31–36. Available at: https://doi.org/10.1109/SDS.2019.00-11.

Veloso, M. (2013) 'An agent-based simulation model for informed shared decision making in multiple sclerosis', *Multiple Sclerosis and Related Disorders*, 2(4), pp. 377–384. Available at: https://doi.org/10.1016/j.msard.2013.04.001.

Veloso, M. (2014) 'A web-based decision support tool for prognosis simulation in multiple sclerosis', *Multiple Sclerosis and Related Disorders*, 3(5), pp. 575–583. Available at: https://doi.org/10.1016/j.msard.2014.04.005.

Visser, L.A. *et al.* (2020) 'Patient needs and preferences in relapsing-remitting multiple sclerosis: A systematic review', *Multiple Sclerosis and Related Disorders*, 39, p. 101929. Available at: https://doi.org/10.1016/j.msard.2020.101929.

Wagner, E.H., Austin, B.T. and Korff, M.V. (1996) 'Organizing Care for Patients with Chronic Illness', *The Milbank Quarterly*, 74(4), p. 511. Available at: https://doi.org/10.2307/3350391.

Wang, Y. *et al.* (2018) 'Classification of multiple sclerosis and non-specific white matter lesions using spherical harmonics descriptors', in *Proceedings of the 3rd International Workshop on Interactive and Spatial Computing - IWISC '18*. *the 3rd International Workshop*, Richardson, Texas: ACM Press, pp. 97–102. Available at: https://doi.org/10.1145/3191801.3191806.

White, R.W. and Horvitz, E. (2009) 'Cyberchondria: Studies of the escalation of medical concerns in Web search', *ACM Transactions on Information Systems*, 27(4), pp. 1–37. Available at: https://doi.org/10.1145/1629096.1629101.

Wiendl, H. *et al.* (2021) ‘Multiple Sclerosis Therapy Consensus Group (MSTCG): position statement on disease-modifying therapies for multiple sclerosis (white paper)’, *Therapeutic Advances in Neurological Disorders*, 14, p. 175628642110396. Available at: https://doi.org/10.1177/17562864211039648.

Wiens, J. *et al.* (2019) ‘Do no harm: a roadmap for responsible machine learning for health care’, *Nature Medicine*, 25(9), pp. 1337–1340. Available at: https://doi.org/10.1038/s41591-019-0548-6.

Winkelmann, A. *et al.* (2016) ‘Disease-modifying therapies and infectious risks in multiple sclerosis’, *Nature Reviews Neurology*, 12(4), pp. 217–233. Available at: https://doi.org/10.1038/nrneurol.2016.21.

World Health Organization and Multiple Sclerosis International Federation (2008) ‘Atlas : multiple sclerosis resources in the world 2008’, p. 51.

Wottschel, V. *et al.* (2015) ‘Predicting outcome in clinically isolated syndrome using machine learning’, *NeuroImage: Clinical*, 7, pp. 281–287. Available at: https://doi.org/10.1016/j.nicl.2014.11.021.

Yperman, J. *et al.* (2020) ‘Machine learning analysis of motor evoked potential time series to predict disability progression in multiple sclerosis’, *BMC Neurology*, 20(1), p. 105. Available at: https://doi.org/10.1186/s12883-020-01672-w.

Zhang, Y. *et al.* (2016) ‘Comparison of machine learning methods for stationary wavelet entropy-based multiple sclerosis detection: decision tree, *k* -nearest neighbors, and support vector machine’, *SIMULATION*, 92(9), pp. 861–871. Available at: https://doi.org/10.1177/0037549716666962.

Zhao, Y. *et al.* (2017) ‘Exploration of machine learning techniques in predicting multiple sclerosis disease course’, *PLOS ONE*. Edited by S.V. Ramagopalan, 12(4), p. e0174866. Available at: https://doi.org/10.1371/journal.pone.0174866.

Zurita, M. *et al.* (2018) ‘Characterization of relapsing-remitting multiple sclerosis patients using support vector machine classifications of functional and diffusion MRI data’, *NeuroImage: Clinical*, 20, pp. 724–730. Available at: https://doi.org/10.1016/j.nicl.2018.09.002.

List of Publications, Awards, and Grants

Publications

Alshamrani, R., Althbiti, A., Alshamrani, Y., Alkomah, F., Ma, X. 2020. Model-Driven Decision Making in Multiple Sclerosis Research: Existing Works and Latest Trends. In Patterns, vol. 1, no. 8, p. 100121, Nov. 2020, doi: 10.1016/j.patter.2020.100121.

Alshamrani, R., Said-Said, S., Ma, X. 2021. Using Knowledge Graph to Improve Informed Multiple Sclerosis Diagnosis and Treatment Decisions (under review)

Alshamrani, R., Ma, X. 2019. Deep Learning. In: Schintler, L.A., McNeely, C.L. (eds.) Encyclopedia of Big Data. Springer, Cham, Switzerland. 5pp. In Press. https://doi.org/10.1007/978-3-319-32001-4_533-1

Althbiti, A., Alshamrani, R., Alghamdi, T., Lee, S., Ma, X., 2021. Addressing Data Sparsity in Collaborative Filtering Based Recommender Systems Using Clustering and Artificial Neural Network. In: Proceedings of CCWC 2021: The 11th Annual Computing and Communication Workshop and Conference (Virtual), Las Vegas, NV, USA, pp.217-226.

Althbiti, A., Alshamrani, R., Ma, X., 2020. A Literature Review of Data Mining Techniques Used in Collaborative Filtering Recommender Systems. In: Proceedings of The 2020 International Conference on Computational Science and Computational Intelligence, Las Vegas, NV, USA. pp.424-430

Presentations

Alshamrani, R. Ma, X. 2019. Predicting Students' Final Outcomes by Analyzing Early Indicators. Poster presented at 2nd U.S. Semantic Technologies Symposium 2019, Durham, NC

Alshamrani, R. Althbiti, A. Aljuhani, A. Ma, X. 2019. Leveraging Early Indicators for Evaluating Students' Final Grades. Poster presented at UI Research Computing and Data Science Symposium 2019, Moscow, ID

Althbiti, A. Alshamrani, R. Aljuhani, A. Ma, X. 2019. Personalized Courses Recommender System based on Content Approach. Poster presented at UI Research Computing and Data Science Symposium 2019, Moscow, ID

Aljuhani, A. Althbiti, A. Alshamrani, R. Ma, X. 2019. Payment Methods Preference Prediction. Poster presented at UI Research Computing and Data Science Symposium 2019, Moscow, ID

Althbiti, A. Alshamrani, R. Aljuhani, A. Ma, X. 2019. Semantic prediction of the classes of Arabic idioms from rating records. Poster presented at UI Research Computing and Data Science Symposium 2019, Moscow, ID

Awards

Awarded Best Poster at the 2019 University of Idaho Computer Science Industrial Advisory Board Meeting.	2019

Grants

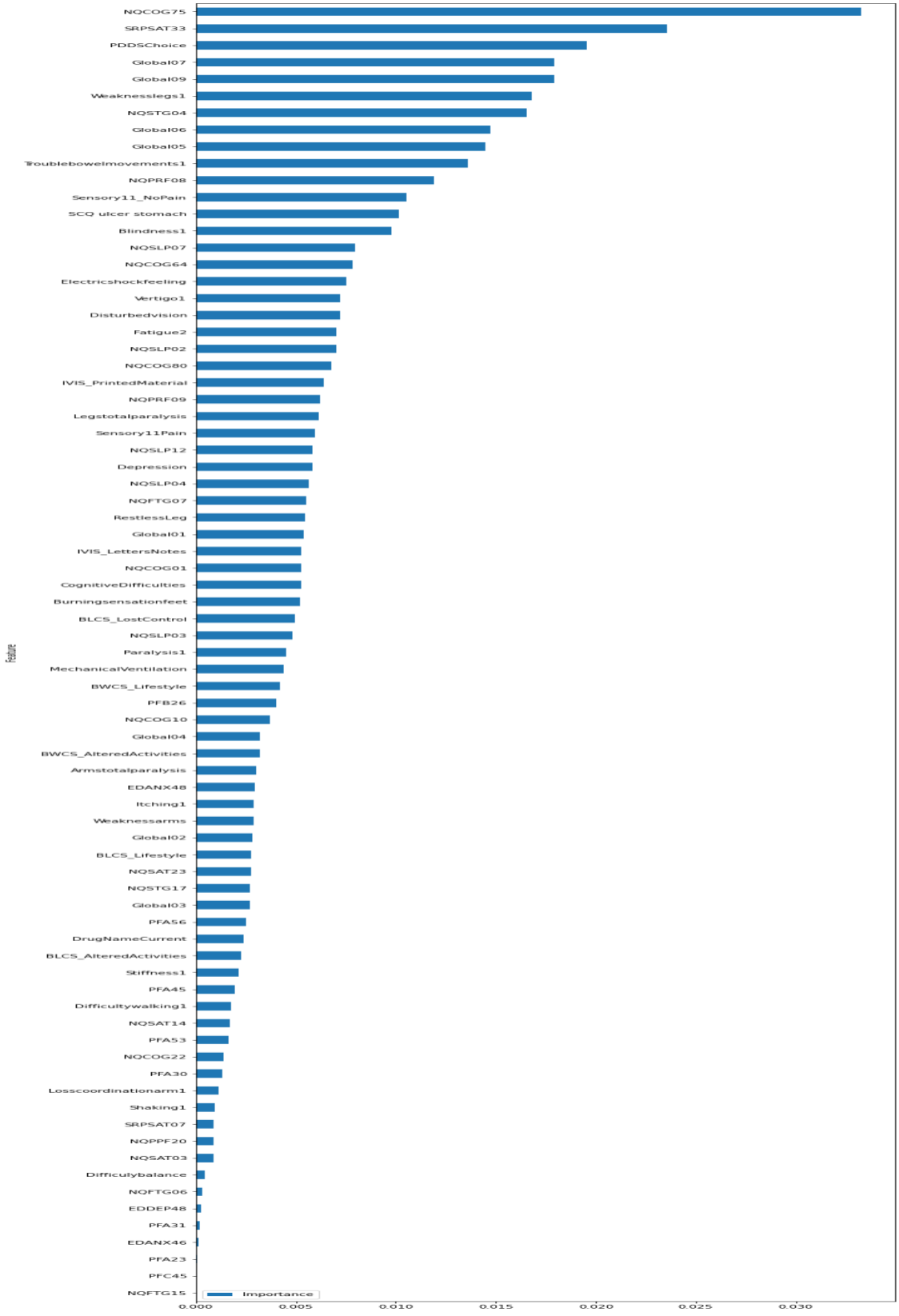
Received $4,000 Data Access Grant from the Institute for Modeling Collaboration and Innovation at the University of Idaho.	2022
Received a waiver of open access fee ($5,200) from the Patterns Journal (Cell Press) to publish a literature review.	2020
Awarded US2TS2019 travel grant from Sloan.	2019

Appendix A: List of Important Features and their Descriptions (Chapter 3)

Feature	Description
NQCOG75	My thinking was slow.
SRPSAT33	I am satisfied with my ability to do things for fun outside my home.
PDDSChoice	Patient-Determined Disease Steps choice
Global07	How would you rate your pain on average?
Global09	In general, please rate how well you carry out your usual social activities and roles.
Weaknesslegs1	Weakness in legs/feet
NQSTG04	Because of my illness, I felt left out of things.
Global06	To what extent are you able to carry out your everyday physical activities such as walking, climbing stairs, carrying groceries, or moving a chair?
Global05	In general, how would you rate your satisfaction with your social activities and relationships
Troublebowelmovements1	Trouble with bowel movements
NQPRF08	I am able to socialize with my friends.
Sensory11_NoPain	Sensory symptoms (excluding pain)
SCQ ulcer stomach	Do you have ulcer or stomach disease?
Blindness1	Blindness or blurry vision in one eye or both
NQSLP07	I had trouble falling asleep.
NQCOG64	I had to read something several times to understand it.
Electricshockfeeling	Electric shock-like feeling when bending neck
Vertigo1	Vertigo
Disturbedvision	Disturbed vision e.g. double vision, objects moving, etc.
Fatigue2	Fatigue
NQSLP02	I had to force myself to get up in the morning.
NQCOG80	I had trouble concentrating.
IVIS_PrintedMaterial	Read or access printed materials, such as books, magazines, newspapers, etc.?
NQPRF09	I am able to do all of my regular activities with friends.
Legstotalparalysis	Total paralysis of legs
Sensory11Pain	Painful sensations
NQSLP12	Pain woke me up.
Depression	Changes in mood or depression considered out of the ordinary
NQSLP04	I was sleepy during the daytime.
NQFTG07	I was too tired to leave the house.
RestlessLeg	Restless leg syndrome
Global01	In general, would you say your health is…
IVIS_LettersNotes	Read or access personal letters or notes?
NQCOG01	Writing notes to yourself, such as appointments or 'to do' lists?
CognitiveDifficulties	Cognitive difficulties, e.g. memory problems
Burningsensationfeet	Burning sensation in feet
BLCS_LostControl	Lost control of your bladder or had an accident?
NQSLP03	I had trouble stopping my thoughts at bedtime.
Paralysis1	Paralysis of half or whole face (i.e. facial drooping with altered smile, difficulty closing an eye tightly or wrinkling forehead)

Feature	Description
MechanicalVentilation	Need for mechanical ventilation
BWCS_Lifestyle	During the past 4 weeks, how much have bowel problems restricted your overall lifestyle?
PFB26	Are you able to shampoo your hair?
NQCOG10	Organizing what you want to say?
Global04	In general, how would you rate your mental health, including your mood and ability to think?
BWCS_AlteredActivities	Altered your activities because of bowel control problems?
Armstotalparalysis	Total paralysis of arms
EDANX48	Many situations made me worry.
Itching1	Itching, not due to other causes e.g. psoriasis, insect bites, etc.
Weaknessarms	Weakness in arms/hands
Global02	In general, would you say your quality of life is…
BLCS_Lifestyle	During the past 4 weeks, how much have bladder problems restricted your overall lifestyle?
NQSAT23	I am disappointed in my ability to socialize with my family.
NQSTG17	I felt embarrassed because of my physical limitations.
Global03	In general, how would you rate your physical health?
PFA56	Are you able to get in and out of a car?
DrugNameCurrent	Name of the current DMT
BLCS_AlteredActivities	Altered your activities because of bladder problems?
Stiffness1	Stiffness / spasms
PFA45	Are you able to get out of bed into a chair?
Difficultywalking1	Difficulty walking / dragging a foot
NQSAT14	I am bothered by limitations in my regular activities with friends.
PFA53	Are you able to run errands and shop?
NQCOG22	Reading and following complex instructions (e.g., directions for a new medication)?
PFA30	Are you able to step up and down curbs?
Losscoordinationarm1	Loss of coordination in arms / hands
Shaking1	Shaking or tremors
SRPSAT07	I am satisfied with how much of my work I can do (include work at home).
NQPPF20	My life had purpose.
NQSAT03	I am bothered by my limitations in regular family activities.
Difficultybalance	Difficulty with balance
NQFTG06	I was too tired to do my household chores.
EDDEP48	I felt that my life was empty.
PFA31	Are you able to get up off the floor lying on your back without help?
EDANX46	I felt nervous.
PFA23	Are you able to go for a walk of at least 15 minutes?
PFC45	Are you able to get on and off the toilet?
NQFTG15	I felt fatigued.

Appendix B: Order of Important Features (Chapter 3)

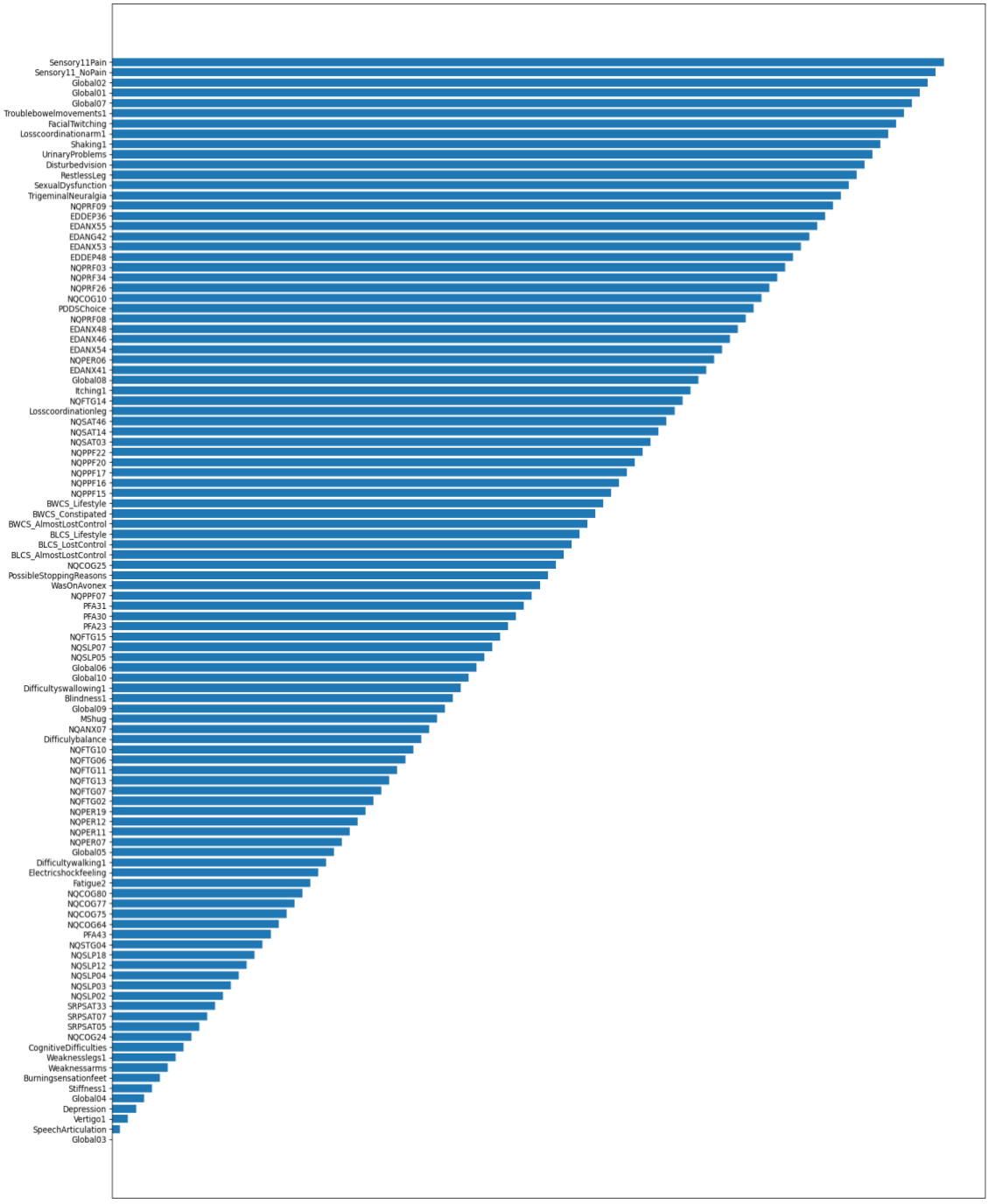


Appendix C: List of Important Features and their Descriptions (Chapter 4)

Feature	Description
Sensory11Pain	Painful sensations
Sensory11_NoPain	Sensory symptoms (excluding pain)
Global02	In general, would you say your quality of life is…
Global01	In general, would you say your health is…
Global07	How would you rate your pain on average?
Troublebowelmovements1	Trouble with bowel movements
FacialTwitching	Facial twitching
Losscoordinationarm1	Loss of coordination in arms / hands
Shaking1	Shaking or tremors
UrinaryProblems	Urinary problems, e.g. unusual urgency or hesitancy
Disturbedvision	Disturbed vision e.g. double vision, objects moving, etc.
RestlessLeg	Restless leg syndrome
SexualDysfunction	Sexual dysfunction, not caused by medication
TrigeminalNeuralgia	Sharp, painful feeling in face not due to trauma or injury (trigeminal neuralgia)
NQPRF09	I am able to do all of my regular activities with friends.
EDDEP36	I felt unhappy.
EDANX55	I had difficulty calming down.
EDANG42	I had trouble controlling my temper.
EDANX53	I felt uneasy.
EDDEP48	I felt that my life was empty.
NQPRF03	I am able to do all of my regular family activities.
NQPRF34	I can keep up with my work responsibilities (include work at home).
NQPRF26	I am able to participate in leisure activities.
NQCOG10	Organizing what you want to say?
PDDSChoice	Patient-Determined Disease Steps choice
NQPRF08	I am able to socialize with my friends.
EDANX48	Many situations made me worry.
EDANX46	I felt nervous.
EDANX54	I felt tense.
NQPER06	I said or did things without thinking.
EDANX41	My worries overwhelmed me.
Global08	How would you rate your fatigue on average?....
Itching1	Itching, not due to other causes e.g. psoriasis, insect bites, etc.
NQFTG14	I felt tired.
Losscoordinationleg	Loss of coordination in legs / feet
NQSAT46	I am satisified with my ability to do household chores or tasks.
NQSAT14	I am bothered by limitations in my regular activities with friends.
NQSAT03	I am bothered by my limitations in regular family activities.
NQPPF22	I felt cheerful.
NQPPF20	My life had purpose.
NQPPF17	My life had meaning.
NQPPF16	I had a sense of balance in my life.
NQPPF15	My life was satisfying.
BWCS_Lifestyle	During the past 4 weeks, how much have bowel problems restricted your overall lifestyle?
BWCS_Constipated	Been constipated?
BWCS_AlmostLostControl	Almost lost control of your bowels or had an accident?
BLCS_Lifestyle	During the past 4 weeks, how much have bladder problems restricted your overall lifestyle?
BLCS_LostControl	Lost control of your bladder or had an accident?
BLCS_AlmostLostControl	Almost lost control of your bladder or had an accident?
NQCOG25	Managing your time to do most of your daily activities?
PossibleStoppingReasons	Reasons to stop a DMT
WasOnAvonex	Was patient using Avonex?

Feature	Description
PFA31	Are you able to get up off the floor lying on your back without help?
PFA30	Are you able to step up and down curbs?
PFA23	Are you able to go for a walk of at least 15 minutes?
NQFTG15	I felt fatigued.
NQSLP07	I had trouble falling asleep.
NQSLP05	I had trouble sleeping because of bad dreams.
Global06	To what extent are you able to carry out your everyday physical activities such as walking, climbing stairs, carrying groceries, or moving a chair?
Global10	How often have you been bothered by emotional problems such as feeling anxious, depressed or irritable?....
Difficultyswallowing1	Difficulty with swallowing
Blindness1	Blindness or blurry vision in one eye or both
Global09	In general, please rate how well you carry out your usual social activities and roles.
MShug	"MS hug" (feeling of tightness in the torso)
NQANX07	I felt nervous when my normal routine was disturbed.
Difficultybalance	Difficulty with balance
NQFTG10	I was frustrated by being too tired to do the things I wanted to do.
NQFTG06	I was too tired to do my household chores.
NQFTG11	I felt that I had no energy.
NQFTG13	I felt exhausted.
NQFTG07	I was too tired to leave the house.
NQFTG02	I had to limit my social activity because I was tired.
NQPER19	I was in conflict with others.
NQPER12	I was bothered by little things.
NQPER11	I was irritable around other people.
NQPER07	I got impatient with other people.
Global05	In general, how would you rate your satisfaction with your social activities and relationships
Difficultywalking1	Difficulty walking / dragging a foot
Electricshockfeeling	Electric shock-like feeling when bending neck
Fatigue2	Fatigue
NQCOG80	I had trouble concentrating.
NQCOG77	I had to work really hard to pay attention or I would make a mistake.
NQCOG75	My thinking was slow.
NQCOG64	I had to read something several times to understand it.
PFA43	Are you able to write with a pen or pencil?
NQSTG04	Because of my illness, I felt left out of things.
NQSLP18	I felt physically tense during the middle of the night or early morning hours.
NQSLP12	Pain woke me up.
NQSLP04	I was sleepy during the daytime.
NQSLP03	I had trouble stopping my thoughts at bedtime.
NQSLP02	I had to force myself to get up in the morning.
SRPSAT33	I am satisfied with my ability to do things for fun outside my home.
SRPSAT07	I am satisfied with how much of my work I can do (include work at home).
SRPSAT05	I am satisfied with the amount of time I spend doing leisure activities.
NQCOG24	Planning for and keeping appointments that are not part of your weekly routine (e.g. a therapy or doctor's appointment, or a social gathering with friends and family)?
CognitiveDifficulties	Cognitive difficulties, e.g. memory problems
Weaknesslegs1	Weakness in legs/feet
Weaknessarms	Weakness in arms/hands
Burningsensationfeet	Burning sensation in feet
Stiffness1	Stiffness / spasms
Global04	In general, how would you rate your mental health, including your mood and ability to think?
Depression	Changes in mood or depression considered out of the ordinary
Vertigo1	Vertigo
SpeechArticulation	Speech articulation (speech sounds slurred or slowed or loses normal rhythm)
Global03	In general, how would you rate your physical health?
NQPPF07	Many areas of my life were interesting to me.

Appendix D: Order of Important Features (Chapter 4)



Appendix E: MSKG Class Hierarchy Using OWLViz (Chapter 5)

Appendix F: IRB Approval Letter

University of Idaho
Office of Research Assurances
Institutional Review Board
875 Perimeter Drive, MS 3010
Moscow, ID 83844-3010
Phone: 208-885-6162
Fax: 208-885-6014
Email: irb@uidaho.edu

February 25, 2022

To: Xiaogang Ma

Cc: Rayan Alshamrani

From: University of Idaho Institutional Review Board

Approval Date: February 25, 2022

Title: Enhancing The Shared Decision-Making Experience of Multiple Sclerosis Via Model-Driven Decision Support

Protocol: 21-235, Reference: 016237

Exempt under Category 2 at 45 CFR 46.104(d)(2).

On behalf of the Institutional Review Board at the University of Idaho, I am pleased to inform you that the protocol for this research project has been certified as exempt under the category listed above.

This certification is valid only for the study protocol as it was submitted. Studies certified as Exempt are not subject to continuing review and this certification does not expire. However, if changes are made to the study protocol, you must submit the changes through VERAS for review before implementing the changes. Amendments may include but are not limited to, changes in study population, study personnel, study instruments, consent documents, recruitment materials, sites of research, etc.

As Principal Investigator, you are responsible for ensuring compliance with all applicable FERPA regulations, University of Idaho policies, state and federal regulations. Every effort should be made to ensure that the project is conducted in a manner consistent with the three fundamental principles identified in the Belmont Report: respect for persons; beneficence; and justice. The Principal Investigator is responsible for ensuring that all study personnel have completed the online human subjects training requirement. Please complete the *Continuing Review and Closure Form* in VERAS when the project is completed.

You are required to notify the IRB in a timely manner if any unanticipated or adverse events occur during the study, if you experience an increased risk to the participants, or if you have participants withdraw or register complaints about the study.

IRB Exempt Category (Categories) for this submission:
Category 4: Secondary research for which consent is not required: Secondary research uses of identifiable private information or identifiable biospecimens, if at least one of the following criteria is met: i. The identifiable private information or identifiable biospecimens are publicly

University of Idaho
Office of Research Assurances
Institutional Review Board
875 Perimeter Drive, MS 3010
Moscow, ID 83844-3010
Phone: 208-885-6162
Fax: 208-885-6014
Email: irb@uidaho.edu

available; ii. Information, which may include information about biospecimens, is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained directly or through identifiers linked to the subjects, the investigator does not contact the subjects, and the investigator will not re-identify subjects; iii. The research involves only information collection and analysis involving the investigator's use of identifiable health information when that use is regulated under 45 CFR parts 160 and 164, subparts A and E [HIPAA], for the purposes of "health care operations" or "research" as those terms are defined at 45 CFR 164.501 or for "public health activities and purposes" as described under 45 CFR 164.512(b); or iv. The research is conducted by, or on behalf of, a Federal department or agency using government-generated or government-collected information obtained for nonresearch activities, if the research generates identifiable private information that is or will be maintained on information technology that is subject to and in compliance with section 208(b) of the E-Government Act of 2002, 44 U.S.C. 3501 note, if all of the identifiable private information collected, used, or generated as part of the activity will be maintained in systems of records subject to the Privacy Act of 1974, 5 U.S.C. 552a, and, if applicable, the information used in the research was collected subject to the Paperwork Reduction Act of 1995, 44 U.S.C. 3501 et seq.