Interval Type-2 Fuzzy Logic Learning Rules-based Approach for Health and Wellbeing

by
Sachin Kapase

In partial fulfillment of the requirements for the degree of

MSc in

Advanced Computer Science with Artificial Intelligence

University of Strathclyde

Department of

Computer and Information Sciences

August 2023

Abstract:

Health is an important factor that greatly influences the overall well-being and academic performance of young students. Optimal health and mental state significantly enhance students' abilities to manage stress, emotions, and the various challenges they face during their educational journey. This not only fosters better focus, concentration, and motivation but also enhances their learning outcomes. This is mainly due to the various factors contributing to health conditions and the inherent uncertainty in health predictions. Traditional approaches often lack the required flexibility and adaptability in handling such complex issues. The aim of this project titled, "Interval Type-2 Fuzzy Logic Learning Rules-based Approach for Health and Wellbeing" is to leverage Artificial Intelligence technologies, specifically Fuzzy Logic System (FLS), to better understand and address health issues among students.

The focus is on using Type-1 Fuzzy Logic rule-based learning from the data and Interval Type-2 Fuzzy Logic learning from the data algorithms to evaluate and predict health conditions. A fuzzy Logic System (FLS) is an AI algorithm designed to deal with uncertainty and vagueness, making it ideal for health predictions. The method involves the use of five distinct health prediction models - Type-1 Fuzzy Logic learning from the data, Type-1 Fuzzy Logic with auto rules generation which is the traditional fuzzy rules-based system, Interval Type-2 Fuzzy Logic learning from the data, Interval Type-2 Fuzzy Logic with auto rule generation through experimentation, and a Machine Learning model. These models are applied, and their results are compared to evaluate their effectiveness. The interval type-2 fuzzy logic rule-based learning algorithm, known for its superior handling of uncertainty, is predicted to yield the most effective results. Findings from the study reveal that the Interval Type-2 Fuzzy Logic based on learning rules from the data approach achieved the highest accuracy in health prediction among students. This is followed by Type-1 Fuzzy Logic and Interval Type-2 Fuzzy Logic with auto rule generation. Surprisingly, the Machine Learning model delivered the lowest accuracy. The promising results from this study indicate the potential of using a Fuzzy Logic System, specifically Interval Type-2, in health prediction systems. Future work will aim at refining the use of this AI technology and potentially exploring its combination with other AI techniques to enhance its predictive accuracy and usability. The long-term objective is to provide critical insights into health trends among students, inform strategies to improve their overall well-being, and thereby enhance their academic performance.

Declaration:

I hereby declare that this dissertation, titled "Interval Type-2 Fuzzy Logic Learning Rules-based Approach for Health and Wellbeing," submitted in partial fulfillment of the requirements for the degree of MSc in Advanced Computer Science with Artificial Intelligence at the University of Strathclyde, is entirely my own original work. I confirm that I have composed this dissertation independently and have appropriately acknowledged all sources and references used. I further affirm that I have obtained the necessary ethics approval from the relevant authority for conducting my research. I hereby grant permission to the University of Strathclyde, Department of Computer, and Information Sciences, to distribute copies of this dissertation at a cost, exclusively for private study or research purposes. Additionally, I authorize the University of Strathclyde to archive a copy of this dissertation in a publicly accessible repository.

I, therefore, certify that this dissertation has an overall word count of 15,902 (excluding the title page, declaration, abstract, table of contents, list of illustrations, citations, and appendix). Please analyze this as a Type 1 2 3 4 **5** Dissertation.

Signature:

Date: 14/08/2023

Acknowledgments:

The successful completion of this project has required a significant investment of time and effort, and I am grateful for the unwavering support and motivation I have received from various sources. First and foremost, I extend my heartfelt appreciation to my family for their constant support and provision of the necessary resources that enabled me to produce my best work.

I am indebted to Dr. Nur Muhamad Naim for their invaluable support and guidance throughout this project. Without their expertise, mentorship, and encouragement, I would not have been able to develop the models and achieve the results presented in this dissertation. I extend my sincere thanks to Dr. Nur Muhamad Naim for their valuable teachings and unwavering support throughout the course of this project.

Table of Contents:

1. Introduction:

The importance of health for young people and students cannot be overstated. Mental well-being plays a crucial role in their overall development, academic performance, and quality of life. A positive mental state allows young individuals to effectively cope with the challenges, pressures, and stresses they encounter during their formative years. First and foremost, prioritizing health promotes optimal well-being. It equips young people with the tools to navigate and manage their emotions, relationships, and personal growth. When students have good health, they are better able to concentrate, focus, and retain information, leading to improved academic performance and learning outcomes. A healthy mind fosters creativity, critical thinking, and problem-solving skills, enabling students to excel in their studies and future endeavors. Moreover, health significantly impacts social interactions and relationships. Positive mental well-being enhances self-esteem, confidence, and resilience, allowing young people to build strong and healthy connections with their peers, teachers, and family members. It helps them develop effective communication skills, empathy, and a sense of belonging, which are crucial for personal and social development (E. S. Singh, 2015).

Established machine learning methods, such as neural networks and support vector machines (SVM), have extensively been employed in healthcare, primarily for extracting meaningful features from organized healthcare datasets. These techniques, rooted in the principles of statistical learning, have demonstrated their capability to manage data with high dimensions and to forecast results based on complicated variable relationships. For instance, the use of neural networks, due to their ability to discern complex patterns and relationships, has become widespread in predicting diseases, analyzing medical imaging, and stratifying patient risk, among other applications (Jiang, 2017). However, these conventional methods are not without shortcomings. The resource-intensive nature of neural networks and SVMs can be challenging, and these methods often lack transparency and interpretability. While natural language processing (NLP) is potent, it still grapples with comprehending context, subtle language variations, and ambiguities inherent in human language. To overcome these challenges, emerging AI technologies, such as Fuzzy Logic Systems, are being explored. These technologies are particularly valued for their competence in managing the inherent ambiguity and uncertainty in health predictions (J. Betty Jane, n.d.). In recent times, fuzzy logic has emerged as a notable tool for decision-making in intricate systems. At its core, fuzzy logic is a mathematical approach designed to represent unclear or vague information, broadening its utility across numerous applications. Unlike traditional sets that categorize elements as either inside or outside a set, fuzzy logic operates on the premise of fuzzy sets. These sets permit partial membership, allowing an element to be part of a set to varying degrees. This nuanced approach provides a means to address real-world challenges that traditional logic struggles with. Fuzzy logic's versatility finds applications in diverse areas, from engineering to economics. Engineers utilize it for tasks like optimization, control systems, and decision processes. Computer scientists tap into its potential for areas such as data mining, natural language processing, and pattern identification. Meanwhile, in economics, it aids in forecasting, evaluating risks, and decision-making. A standout feature of fuzzy logic is its adeptness at managing ambiguous or uncertain data. This proves invaluable in scenarios where conventional logic falls short, like interpreting human language or making decisions amidst uncertainty. Moreover, its capability to navigate non-linear relationships between variables offers a robust framework for understanding sophisticated systems (Moraga, 2005).

Fuzzy logic is a mathematical logic approach that allows for degrees of truth instead of strict true/false values. Traditional binary logic relies on the notion that a statement is either true or false without any middle ground. However, many real-world situations involve uncertainty or imprecision, making it challenging to assign a clear true/false value. This is where fuzzy logic comes into play. In fuzzy logic, truth values are expressed as degrees of membership in a fuzzy set. A fuzzy set consists of objects that possess varying degrees of membership within the set. For instance, a fuzzy set of "tall people" might include individuals who are very tall with a high degree of membership, as well as those who are moderately tall with a lower degree of membership. This nuanced approach allows for a more comprehensive representation of reality compared to traditional binary logic (A. Torres-Iglesias and J. J. Nieta, 2006).

In the context of healthcare, fuzzy logic can be employed to model and analyze complex medical data that may not neatly fit into conventional binary categories. For instance, fuzzy logic can analyze patient symptoms and determine the likelihood of a specific diagnosis by assessing the similarity between the patient's symptoms and those associated with various diseases. Fuzzy logic is also valuable for modeling intricate biological systems like gene regulatory networks and analyzing extensive datasets in bioinformatics research. The data consists of the young adult behavior that helps to predict health problems. Impact of mental illness on family relationships, as well as the impact of family relationships on health. health problems and illnesses can have a significant effect on families and relationships, with issues such as a lack of diagnosis, treatment use and efficacy, chronicity, and propensity all impacting a family's experience of mental illness. Likewise, relationship and family problems can have a significant impact on health. Both living with and/or caring for someone under these circumstances not only impacts the well-being of the person with the illness but those around them.

Fuzzy logic is a computational approach that differs from traditional Boolean logic by introducing a "degree of truth" rather than strict binary values of 0 and 1. It was developed by Dr. Lotfi Zadeh at the University of California during the 1960s. Dr. Zadeh recognized the significance of having more precise values that lie between the extremes of true and false. This led to the creation of Fuzzy Logic, which provides a reasoning method specifically designed to handle uncertain situations. By incorporating degrees of truth, Fuzzy Logic enables a more nuanced and flexible approach to computing and decision-making (Zadeh, 1965).

Fuzzy logic encompasses two main types: Type-1 fuzzy logic and Type-2 fuzzy logic, which further includes Interval Type-2. Each type assigns quantified values based on its specific characteristics. For this case study, Type-1, and Interval Type-2 fuzzy logic are used. The process for both types is similar, with the difference being that Interval Type-2 includes a type-reducer that acts as a converter to transform the output into Type-1 fuzzy logic. This ensures the fuzzy logic produces membership function grades for the given situation (N. Gupta, 2014). Type-1 fuzzy logic is a classical set that has limitations in analyzing real-world models due to the limited linguistic terms for measuring uncertainty. In the prediction of young adults' health using Type-1 fuzzy logic, linguistic variables such as "low, medium, and high" are employed to describe the health index. The resulting membership function value ranges from 0 to 1, with 0 indicating poor health and 1 indicating good health. Interval Type-2 fuzzy logic, on the other hand, provides a more convenient way to evaluate linguistic variables by utilizing the footprint of uncertainty to measure uncertainty. The linguistic variables used for Interval Type-2, like Type-1, are "low, medium, and high." However, Interval Type-2 introduces additional linguistic variables such as "very_low, low, average, high, very_high" This allows Interval Type-2 fuzzy logic to generate more precise results compared to Type-1 fuzzy logic (N. Gupta, 2014).

Researchers are increasingly focusing on applying Type-2 fuzzy logic systems to measure health conditions due to the superiority of Interval Type-2 fuzzy logic over Type-1 fuzzy logic. Interval Type-2 excels in handling uncertainty through modeling linguistic variables and utilizing the footprint of uncertainty (FOU) to determine fuzzy set membership functions. Additionally, the type-reducer employed in Interval Type-2 fuzzy logic provides better approximation results compared to Type-1 fuzzy logic in numerical experiments. The type-reducer converts Interval Type-2 results into crisp Type-1 fuzzy logic values, ensuring the evaluation of the young adult's health index based on behavioral factors that may contribute to health problems (Carreon-Ortiz, 2023).

The membership function plays a crucial role in both Type-1 and Interval Type-2 fuzzy logic systems, as it is utilized throughout the entire fuzzy logic process. It serves to evaluate the antecedent values used in the rule-based process and the consequent values employed in the inference process. Ultimately, the membership function generates the output for Type-1 and Interval Type-2 fuzzy logic (E. S. Singh, 2015). Unlike a binary approach that simply determines whether a young adult has health issues or not, the fuzzy logic system provides a degree or value representing the health index of the individual. In this case study, a lower degree of health index indicates a higher likelihood of young adults being affected by health problems. The results generated by Type-1 and Interval Type-2 fuzzy logic are subjected to evaluation.

In this study, a dataset of 1045 entries were utilized, out of which 350 data points were selected for analysis of the produced results. These 350 test cases were carefully chosen to encompass diverse patterns and traits. The outcomes of the tested data were categorized into three groups: bad, average, and good health Index. This evaluation aims to assess the correlation between the obtained results and expert opinions. This unique research aims to explore the potential of fuzzy logic in assessing and understanding health-related data from a holistic perspective. By harnessing the capabilities of fuzzy logic, we can develop comprehensive health and well-being systems that consider the multifaceted nature of human health. This study compared a total of four models: Type-1 fuzzy logic rule based learning system, Type-1 fuzzy logic with auto rule generation, Type-2 fuzzy logic rule based learning system, and Type-2 fuzzy logic with auto rule generation. A comprehensive analysis of the results can be found in the analysis and results section. The student performance dataset has been utilized for this case study. The dataset was compiled by Paulo Cortez of the University of Minho. Using two Portuguese institutions' student accomplishment data. The attributes include test scores, demographic, interpersonal, and academic variables. The data was collected through school reports and surveys. Two datasets are provided (por) about the performance in Portuguese language proficiency and mathematics (mat). There are 1,044 instances (students) in this data, including 395 students in math classes and 649 students in Portuguese language classes.

The student behaviors that are influenced by academic performance are included in the dataset. To forecast the student outcomes for their case study, the dataset uses roughly 33 features as an attribute. Therefore, as secondary school students are also considered to be young adults, the case study could analyze and construct the young adult health index using the standard dataset. The dataset comprises 33 attributes, consisting of 32 feature columns and a label column called G3. The attributes provide valuable information about the students. The dataset includes the following information: Id, which represents the student's identification number; School, indicating the school the student attends; Class, denoting the student's class; Sex, representing the gender of the student; Age, indicating the age of the student; Address, providing the student's home address; Famsize, specifying the total number of family members in the student's family; Pstatus, indicating the status of parent's cohabitation; Medu, representing the mother's level of education; Fedu, denoting the father's level of education; Mjob, specifying the profession of the mother; Fjob, indicating the profession of the father; Reason, representing the reason for selecting the school for education; Guardian, denoting the guardian of the student; Traveltime, indicating the travel time from home to school; Studytime, representing the weekly study time assigned to the student; Failures, denoting the total count of past class failures. The dataset also includes attributes such as Schoolsup, Famsup, Paid, Activities, Nursery, Higher, Internet, Romantic, Famrel, Freetime, Goout, Dalc, Walc, health, Absences, and G3, each providing specific information related to the student’s school life, behavior, and academic performance(archive.ics.uci.edu., n.d.). The outcomes show that the Interval Type 2 Fuzzy Logic rule-based learning system stood out, outperforming all other methods, and exemplifying its robust ability to navigate uncertainty and manage intricate scenarios. Following closely, the Type 1 Fuzzy Logic rule-based learning system emerged as the next most effective in terms of predictive capacity. On the other hand, the Type 2 Fuzzy Logic System with auto rule generation and the Type 1 Fuzzy Logic System with auto rule generation lagged in their performance. Notably, Machine Learning algorithms were overshadowed, falling short in comparison to the prowess of the Fuzzy Logic Systems.

1.2 Objective:

Health issues among young adults have become a growing concern today. The ability to accurately predict and assess the health index of this population is crucial for early intervention and effective support. However, traditional diagnostic methods often rely on subjective assessments, leading to potential inaccuracies and delays in providing appropriate care. To address this challenge, this study aims to utilize Fuzzy Logic Systems (FLS) to develop a reliable and objective approach for predicting the health index of young adults. The primary problem revolves around identifying the factors that significantly impact the health index. By exploring various socio-demographic, behavioral, and environmental attributes, this research seeks to establish a comprehensive understanding of the underlying determinants. Furthermore, the study will employ two types of Fuzzy Logic Systems, namely Type-1 and Interval Type-2, to evaluate their effectiveness in predicting the health index. By comparing the performance of these two approaches, the research aims to determine which type of FLS yields more accurate and reliable results.
Additionally, the study will utilize a standardized dataset specifically curated for this research, encompassing a wide range of young adult behaviors and indicators. Through rigorous analysis and experimentation, the research intends to establish a robust model that can effectively predict the health index based on these behavioral attributes. To determine the membership function range for the dataset of the health index, fuzzy logic will be used. These Type-1 and Interval Type-2 algorithms would produce analysis since the Fuzzy Logic system is made up of both types of algorithms. The findings from this study will contribute to advancing the field of health assessment and intervention by providing a data-driven and objective methodology for evaluating the health index of young adults. Timely, this research aims to enhance early detection, intervention, and support for individuals at risk of health challenges in their young adult years.

This dissertation aims to address the following research questions to investigate this area:

1. What is the fuzzy logic system and how does the basic framework of fuzzy logic work?

   Fuzzy logic is a method that manages indefinite or unclear data to form conclusions. Contrasting standard logic systems which operate on absolute binary values (true or false), fuzzy logic embraces more uncertainty, mirroring the intricacies of human reasoning. Its fundamental structure includes input values, which are processed following certain guidelines, resulting in an output.

2. How are the factors that influence the health index of young adults explored?

   The aim is to delve into the attributes from the dataset that might significantly affect the health index of young adults. These attributes encompass time spent socializing, study duration, free time, parents' marital status, and the job profiles of both parents. By scrutinizing these variables, we can understand their impact on the health of young adults. For example, determining if longer study periods or limited free time are associated with a decreased health index. Likewise, evaluating how the parental situation or their employment roles could influence a young adult's health could yield essential insights to enhance health prognoses.

3. How can the prediction of the health index of young adults be implemented and compared using Type-1 Fuzzy Logic versus the proposed Interval Type-2 Fuzzy Logic System with rules-based learning from data?

   The aim is to focus on applying and contrasting two distinct fuzzy logic models - Type-1 and Interval/Type-2 - for forecasting the health index of young adults. Type-1 Fuzzy Logic, characterized by well-defined set membership values, offers a more straightforward but possibly less detailed solution. Conversely, the suggested Interval/Type-2 Fuzzy Logic system, with its capacity to manage more uncertainty, could yield more precise predictions. These systems will be based on rules learned from the collected data. By juxtaposing these two methodologies, the project seeks to discern which system most accurately and efficiently predicts the health indices for young adults.

2. Literature Review:

The significance of promoting health and well-being has long been a subject of investigation and interest. With the advancement of technology, fuzzy logic has emerged as a promising technique for developing health and well-being systems. Fuzzy logic, a mathematical approach, effectively addresses the challenges posed by imprecise and unpredictable data. Its application extends across diverse industries, including engineering, finance, and healthcare. By leveraging fuzzy logic, it becomes possible to analyze and comprehend data associated with various health metrics, encompassing health, blood pressure, heart rate, glucose levels, and other factors that contribute to overall wellness and health.

Research conducted by (Ekong, 2013) Victor E. Ekong focused on utilizing fuzzy inference systems (FIS) to forecast levels of depression risk. Fuzzy inference systems employ fuzzy logic, an artificial intelligence technique that aids decision-making with imprecise or uncertain data. This study employed the FLS model to predict the severity of depression risk by incorporating expert knowledge in the form of fuzzy rules and gathering physiological and psychological parameters from patients. The findings demonstrated the accuracy of the FLS model in forecasting depression risk levels, indicating its potential as a valuable tool for medical diagnosis and disease management. Furthermore, the report emphasizes the possibility of extending the fuzzy system approach to other medical domains such as cardiology, neurology, and epidemiology. Consequently, the FLS model could be modified to predict the severity levels of patient encounters within these medical fields. In conclusion, the report provides valuable insights into the application of fuzzy inference systems in medical diagnosis, while also highlighting their potential in various other healthcare domains.

The review article penned by Santosh Shinde and P. Raja Rajeswari provides an overview of contemporary research on smart health risk prediction systems leveraging machine learning. The authors delve into the range of machine learning algorithms utilized in health risk prediction, including decision trees, support vector machines, and neural networks. The discussion extends to identifying the shortcomings of current methodologies and suggesting potential avenues for future exploration. Shinde and Rajeswari underscore the critical role of precise health risk prediction systems; these systems can assist healthcare professionals in pinpointing patients with a higher likelihood of disease onset, enabling timely interventions for disease prevention or management.

The integration of such systems with wearable devices for real-time health monitoring also forms a significant part (Shinde, 2018).

The author Aggarwal, A., Singhal, R. S. conducted a study focused on detecting depression through facial expressions. They introduced a new method that combines fuzzy logic and deep learning techniques to enhance the precision and effectiveness of depression detection. Through experiments, the authors evaluated the performance of their proposed approach and compared it with existing methods. They utilized a dataset consisting of facial expressions from individuals with and without depression to train and test their model. The authors offered a comprehensive explanation of the fuzzy logic and deep learning techniques utilized in their approach. Fuzzy logic was employed to capture the inherent uncertainty and ambiguity present in facial expressions, while deep learning techniques were used to extract features from the expressions and classify them as indicative of depression or not. Furthermore, the authors discussed potential applications of their approach in the field of health, along with the limitations of their research and potential avenues for future exploration. The study presents a comprehensive and innovative approach to detecting depression through facial expressions. The proposed method has the potential to significantly enhance the accuracy and efficiency of depression detection, thereby holding important implications for the diagnosis and treatment of depression (Bedi, 2023).

In the case study (Facchinetti, 2012), the author delves into the utilization of fuzzy logic for measuring well-being, considering its intricate and multifaceted characteristics. The challenges associated with measuring well-being are discussed, including the difficulty of quantifying subjective experiences and the necessity of considering various dimensions of well-being. The author argues that a fuzzy approach can effectively tackle these challenges by enabling a more nuanced and adaptable measurement of well-being. Additionally, the authors present a case study that exemplifies the application of fuzzy logic in measuring well-being in Italy. They construct a fuzzy inference system by collaborating with experts and utilizing data from the Italian Statistical National Institute survey on health conditions from 2004 to 2008. The system incorporates variables associated with health-related quality of life, obesity, and other health indicators. Through this case study, the authors illustrate how the fuzzy logic approach can be employed to identify the factors contributing to well-being and their interconnectedness. This paper makes a valuable contribution to well-being research by emphasizing the significance of adopting a multidimensional and flexible approach to measurement. The utilization of fuzzy logic aids researchers in comprehending the complex nature of well-being and in developing more effective interventions to promote it.

The study proposed in this paper by (Chung, 2022) review of machine learning techniques that are being used to predict, diagnose, and identify mental health problems. The authors conducted a comprehensive search of multiple databases and identified a total of 142 research articles and papers related to this field. After screening for duplicates and eligibility, a total of 30 research studies were included in the review. The authors provide a taxonomy of machine learning techniques that are commonly used in mental health prediction, including supervised learning, unsupervised learning, and deep learning. They also discuss the applications of these techniques in various mental health domains, such as depression, anxiety, PTSD, and schizophrenia. The authors highlight the potential benefits of using machine learning in mental health, such as improving the accuracy and efficiency of diagnoses, identifying at-risk individuals, and personalizing treatments. However, the authors also discuss the challenges and limitations of applying machine learning techniques in this area. These include issues related to data quality, privacy, and bias, as well as the need for interpretability and transparency in the decision-making process. The authors also note that the current state of research in this area is limited by small sample sizes, lack of standardization, and limited generalizability. the authors conclude that machine learning has the potential to revolutionize the field of mental health by providing new insights and tools for prediction, diagnosis, and treatment. However, they emphasize the need for careful consideration of ethical and practical issues, as well as the need for further research to address the current limitations and gaps in the field. The study provides a valuable contribution to the state of the art in this area and offers potential research directions that could assist researchers in gaining knowledge about the methods and applications of big data in the mental health field (Chung, 2022).

In the study author has developed a cognitive diagnostic model, utilizing Fuzzy Logic, to evaluate students' learning in educational settings. This model aims to capture the expert knowledge of experienced professors and employs three input linguistic variables (progression of marks, level of test approval, and final grade relative to the average) to estimate the level of knowledge as an output linguistic variable. The model is constructed based on fuzzy rules, which allow for a representation of interpretable knowledge that is both well-defined and intuitive. This approach offers a high degree of flexibility, enabling the inclusion of new linguistic variables and the enhancement of qualitative and quantitative criteria in student assessment. For instance, the author found that, according to the opinion of experienced teachers, four out of the 27 rules in the developed model could be discarded.

To test the model, it was applied in two courses, Numerical Analysis, and Computation, with approximately 150 and 350 students per year, respectively. The diagnostic system was used to identify a group of students from each course for whom final grading was challenging. In 2010, the system's assessment coincided with the teachers' assessment in 87% of the cases. After refining some rules, in 2011, the two assessments coincided in 89% of the cases. As a future research direction, the study author plans to incorporate new evidence into the diagnostic model, such as the test dates and the difficulty level of topics. Additionally, new linguistic variables, such as the level of interest and motivation, will be integrated. The study author's work presents a promising approach to assessing students' learning in educational contexts. By capturing the expertise of experienced professors and employing fuzzy rules, the model offers a flexible and interpretable method for evaluating students' knowledge and improving the criteria used in student assessment (Constanza Huapaya1, n.d.).

The utilization of fuzzy logic in predicting the health index of young adults has proven beneficial in preventing health issues. The diagnosis of health conditions employs a Fuzzy Logic System (FLS) algorithm, which incorporates both Type-1 fuzzy logic and Interval Type-2 fuzzy logic. To ensure the smooth functioning of the Fuzzy Logic system, a set of rules is developed based on the dataset. Fuzzy logic has significantly contributed to addressing problems characterized by inaccuracy and uncertainty (E. S. Singh, 2015). The set of rules in fuzzy logic operates similarly to Boolean logic, where the IF part serves as the cause for the THEN part (Mendel, 2000) . The distinguishing factor between Boolean and fuzzy logic lies in the ability of fuzzy logic to quantify uncertainty and assign values to the outcomes, unlike Boolean logic which only provides true or false values. This capability enables fuzzy logic to handle data that goes beyond binary values, incorporating linguistic responses to measure uncertainty more accurately (Ahna Ballonoff Suleiman, 2016).

Uncertain rule-based fuzzy logic systems introduction and new directions" (M., 2003) by Dr. Jerry Mendel is a comprehensive book introducing the concept of type-2 fuzzy sets and rule-based fuzzy systems. This novel approach aims to address uncertainty more effectively than traditional fuzzy logic systems. The book is divided into four parts. The first part offers background information on uncertainty, and membership functions, and presents two case studies on time series forecasting and knowledge discovery using surveys. Part 2 delves into the fundamentals of type-2 fuzzy sets, covering operations, relations, composition, and rule semantics. Part 3 focuses on the design of type-2 fuzzy systems, including the application of least-squares and steepest-descent methods. Finally, Part 4 explores various applications of type-2 fuzzy systems, such as control, decision-making, and pattern recognition. By employing type-2 fuzzy sets and rule-based fuzzy systems, the study proposes that these methods can find applications in diverse fields, including robotics, control systems, and decision-making systems. Through numerous examples and case studies, the book demonstrates the effectiveness and potential of this approach, hinting at the possibility of transforming the field of fuzzy logic and opening new possibilities in different domains.

In the research study, the (Monard, 2009) conducted a comparative analysis of four traditional machine learning algorithms (MLP, NaiveBayes, OneRule, and ZeroRule), along with two fuzzy logic methods (DoC-Based and Wang & Mendel) in the context of classification problems. The experimentation was carried out using four distinct datasets and employed a 5-fold cross-validation strategy. According to the results, the DoC-Based approach demonstrated excellent performance in both accuracy and comprehensibility. On the other hand, while the method was also found to be highly interpretable, it did so at a considerable cost to its accuracy. The MLP and NaiveBayes techniques achieved good levels of accuracy, although they lacked interpretability. Additionally, the Wang & Mendel method displayed commendable accuracy but was less efficient due to the generation of a large volume of rules compared to the other methods. The OneRule and ZeroRules techniques showed a lack of accuracy.

Author (Umoh, 2021) in his work on the Interval Type-2 Fuzzy Framework for Healthcare Monitoring and Prediction, delves into the use of advanced fuzzy logic in the healthcare domain. They start by stressing the significance of health monitoring and forecasting, especially for chronic diseases and personalized medical care. The paper then explains the concept of fuzzy logic, demonstrating its capability to encapsulate uncertainty in health data. The Interval Type-2 fuzzy framework is introduced as an enhanced fuzzy logic form that can manage heightened uncertainty levels. The paper presents the distinct components of the Interval Type-2 fuzzy framework, such as membership functions, the rule base, and the defuzzification process. They also provide examples of how this framework can be effectively used in health monitoring and prediction, citing diseases like diabetes and heart disease as case studies. Throughout the document, the authors emphasize the superiority of the Interval Type-2 fuzzy framework over other healthcare monitoring and prediction techniques, including conventional fuzzy logic and machine learning. They also bring to light certain challenges tied to the framework's implementation in health settings, like the requirement for abundant data and the complex interpretation of results. The paper plays a crucial role in the healthcare monitoring and prediction arena by bringing forward a sophisticated form of fuzzy logic that is equipped to handle greater uncertainty levels. The authors demonstrate the promise of the Interval Type-2 fuzzy framework in enhancing the precision and customization of health monitoring and prediction. This could pave the way for improved patient health outcomes in the long run.

The authors have put forth a proposal for a fuzzy logic system designed to analyze a student's lifestyle. The objective of this system is to assist students in effectively managing their time across different domains, including education, health, social activities, and daily chores. The authors recognize that a well-balanced lifestyle is crucial for students to excel academically, physically, and socially, but only a limited number of students actively engage in such practices. To address this, the authors propose a novel approach that utilizes fuzzy logic to analyze a student's daily time allocation in various categories. The system employs a GPS-enabled Android application to track the student's location and activities. The collected data is then processed using fuzzy logic techniques to generate a comprehensive report on the student's daily time distribution. This report offers insights into the time spent by the student in different domains, enabling the identification of areas where improvements can be made. Based on the analysis of the obtained results, the system provides regular recommendations and suggestions to help students monitor their own habits and maintain a balanced lifestyle. The authors present a distinct approach to address student lifestyle concerns through the utilization of fuzzy logic. The proposed system has the potential to support students in effectively managing their time and achieving a balanced lifestyle. Ultimately, this can assist them in making informed career choices aligned with their interests (Ghosh, 2017).

In the study (Nguyen, 2015), the researchers suggest a novel method for the classification of medical data by integrating wavelet transformation with an interval type-2 fuzzy logic system. The goal of this proposed method is to enhance the precision of medical data classification, a critical factor in medical diagnosis and treatment. Initially, the authors utilize wavelet transformation on the medical data to isolate features pertinent to classification. Following this, an interval type-2 fuzzy logic system is employed to categorize the data using these identified features. The use of an interval type-2 fuzzy logic system is advantageous due to its ability to deal with data uncertainty and inaccuracy, characteristics often found in medical data. The effectiveness of this proposed methodology is assessed using multiple sets of medical data, including electroencephalogram (EEG) and heart rate variability (HRV) signals. The findings illustrate that the proposed method surpasses other classification techniques such as the support vector machine (SVM) and k-nearest neighbor (k-NN) in terms of accuracy. This new method carries the potential to enhance the precision of medical data classification, potentially leading to improved medical diagnosis and treatment.

The author (Kazemzadeh, 2014) explores the conceptual understanding of emotional words through a computational model. Using an interval approach to fuzzistics, they employ interval type-2 fuzzy sets to estimate membership functions of fuzzy sets for emotion vocabularies.

Understanding the Conceptual Meaning of Emotional Words:

The authors compute similarity and subset hood measures among different emotion words and establish correspondences between various emotional vocabularies. Data for their study was collected through a survey on Amazon Mechanical Turk (AMT), involving 137 purported English-speaking participants from the United States. The survey used sets of 30 stimuli, each receiving an average of 38.5 ratings. They propose two models to represent the truth values of answers about emotions. Model 1 is based on Valence, Activation, and Dominance, while the second model uses interval type-2 fuzzy sets (IT2 FSs). The authors found that the second model was essential to capture more nuanced meanings, particularly when handling a large vocabulary of emotional words. The author's work is an application of Computing with Words (CWW), a paradigm considering words as input and output objects of computation(Kazemzadeh, 2014).

Comparison of Type-1 and Interval Type-2 Fuzzy Logic Systems in Optimization:

The second part of the research involves a comparative analysis of the performance of Type-1 and Interval Type-2 Fuzzy Logic Systems within the Fuzzy Discrete Mycorrhiza Optimization Algorithm. This algorithm, used for solving complex problems, simulates the behavior of mycorrhiza, a symbiotic relationship between fungi and plants. The authors assess the performance of Type-1 and Interval Type-2 Fuzzy Logic Systems in optimizing this algorithm. They evaluate various metrics such as convergence rate, accuracy, and robustness. The findings show that Interval

Type-2 Fuzzy Logic Systems display superior performance in terms of convergence rate and robustness compared to Type-1 Fuzzy Logic Systems. Consequently, the study concludes that integrating Interval Type-2 Fuzzy Logic Systems into optimization algorithms can enhance performance and accuracy. They suggest this research holds potential for real-world applications like optimizing complex systems in engineering, finance, and other sectors. This study, therefore, provides valuable insights into the use of fuzzy logic systems within optimization algorithms and highlights the benefits of employing Interval Type-2 Fuzzy Logic Systems. (Kazemzadeh, 2014).

3. Methodology:

This section will discuss the methodology used to conduct this project. Details about the dataset sourced for this study are provided, followed by an explanation regarding the choice of specific data points, commonly referred to as "feature selection". The importance and role of health metrics in this analysis are highlighted. An overview of the type 1 fuzzy logic rule-based learning system and the Interval type 2 fuzzy logic rule-based learning system is given. Additionally, a method that automatically creates rules, known as the Fuzzy Logic System with Auto Rule Generations, is introduced. Lastly, the use of various machine learning algorithms in the study is discussed.

3.1 Data:

The dataset for this case study has been collected from the UCI Machine Learning Repository, which is the official website of the University of California, Irvine (UCI), and serves as a highly regarded and extensively utilized online platform. It offers a vast collection of datasets catered to researchers and practitioners working within the domains of machine learning and data mining. The repository is renowned for its comprehensive range of datasets, covering diverse fields and facilitating research and practical applications in the machine learning field. It provides free access to datasets accompanied by detailed descriptions and attribute information and often includes benchmark datasets for evaluating machine learning algorithms. As a prominent resource, the UCI Machine Learning Repository plays a pivotal role in advancing research and educational endeavors in the machine learning community (archive.ics.uci.edu., n.d.). This case study utilizes the student performance dataset, which was sourced from Paulo Cortez at the University of Minho in Guimarães, Portugal. This dataset examines the academic performance of secondary school students from two Portuguese schools. The data was collected through school reports and surveys, encompassing various attributes such as student grades, demographic information, social factors, and educational factors. The dataset focuses on student performance in two distinct subjects: Portuguese language and mathematics (referred to as "por" and "mat" datasets, respectively). The dataset has total of 33 attributes as follows: Id, school, class, sex, age, address, famsize, Pstatus, Medu, Fedu, Mjob, Fjob, reason, guardian, traveltime, studytime, failures, schoolsup, famsup, paid, activities, nursery, higher, internet, romantic, famrel, freetime, goout, Dalc, Walc, health, absences, G1, G2, G3.

```
df.head()
```

	school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	...	famrel	freetime	goout	Dalc	Walc	health	absences	G1	G2	G3
0	GP	F	18	U	GT3	A	4	4	at_home	teacher	...	4	3	4	1	1	3	4	0	11	11
1	GP	F	17	U	GT3	T	1	1	at_home	other	...	5	3	3	1	1	3	2	9	11	11
2	GP	F	15	U	LE3	T	1	1	at_home	other	...	4	3	2	2	3	3	6	12	13	12
3	GP	F	15	U	GT3	T	4	2	health	services	...	3	2	2	1	1	5	0	14	14	14
4	GP	F	16	U	GT3	T	3	3	other	other	...	4	3	2	1	2	5	0	11	13	13

Fig 3.1 Dataset

3.2 Fuzzy Logic System:

A fuzzy logic system is a computational framework that utilizes fuzzy set theory to handle imprecise and uncertain information. It differs from traditional binary logic by allowing for the representation and manipulation of degrees of truth, enabling more nuanced and adaptable reasoning. In a fuzzy logic system, variables are described using linguistic terms, which account for the imprecision found in natural language. These terms are associated with membership functions that indicate the degree of membership of an element to a particular label. The shape of these functions can vary and provide a mapping between the input variable and its membership to each linguistic term.

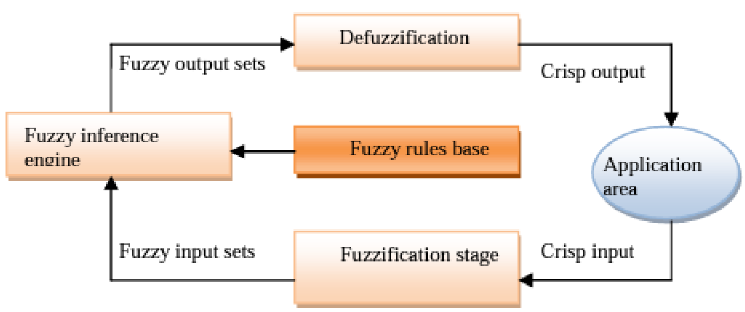

Fig 3.2 Workflow of Fuzzy logic from (Huapaya, n.d.)

Rules in a fuzzy logic system express the relationships between the linguistic terms of input variables and the output variable. These rules use "IF-THEN" conditional statements. The "IF" part specifies the conditions based on the linguistic terms of the input variables, while the "THEN" part represents the conclusion in terms of the output variable (Huapaya, n.d.). Fuzzy logic systems employ fuzzy inference mechanisms to determine the degree of truth for the output variable based on the input variables and defined rules. Various inference methods, such as Mamdani or Sugeno, can be used to compute the fuzzy output. Finally, the output is defuzzified to obtain a crisp value or decision.

Fuzzy logic systems have diverse applications in control systems, pattern recognition, decision-making, and artificial intelligence. They excel in handling uncertainty and imprecision, allowing for more human-like reasoning in complex domains. By capturing and processing fuzzy information, fuzzy logic systems enable flexible and adaptive problem-solving approaches. The chosen features play a vital role in understanding and predicting the health index of young adults. These criteria form the basis for constructing a robust set of rules that will be utilized in the fuzzy logic system. By leveraging information, articles, and traits related to the selected criteria, the fuzzy logic system can provide a precise health index while accounting for uncertainties during the prediction process. Reference (E. S. Singh, 2015) explains how the mapping of fuzzy logic with health symptoms ultimately yields the outcome of identifying health problems. However, this case study takes a different approach by focusing on predicting the health index based on criteria related to young adult behavior, rather than mapping to specific health symptoms or problems. The attributes or criteria are carefully selected from an online standard dataset, ensuring their relevance in predicting the health index of young adults. The chosen criteria encompass various aspects of young adult behavior that can contribute to the occurrence of health problems and subsequently impact the health index. These criteria include family relations, social interactions with friends, health status, academic performance, school attendance, and examination grades (Chung, 2022).

Furthermore, attaining high scores for these features can help young adults prevent health problems. By achieving high scores, young adults can maintain good health conditions and reduce the likelihood of experiencing stress, depression, or anxiety. The case study results also enable the evaluation of the membership function degree within the fuzzy logic system by generating the health index of young adults (E. S. Singh, 2015). The algorithms employed in the fuzzy logic system contribute to the advancement of the medical field, enabling smarter decision-making processes comparable to human cognition.

3.3 FLS in Health Prediction:

The application of fuzzy logic, encompassing Type-1 fuzzy logic fuzzy logic rule-based learning and Interval Type-2 fuzzy logic rule-based learning, has played a crucial role in predicting the health index of young adults, aiding in the prevention of health problems. By establishing a set of rules based on the dataset, the Fuzzy Logic System operates effectively, providing a more accurate approach to handle uncertainty and enhance the outcomes. In this case study, the focus is on predicting the health index of young adults using the student performance dataset. The dataset chosen for this study is sourced from an online standard dataset, which facilitates the experimentation related to health index prediction. Out of the 33 attributes available in the dataset, five specific criteria are selected and utilized throughout this case study. These selected features primarily revolve around the behavior of young adults, including aspects such as family relations, social interactions, health status, academic performance, attendance, and examination grades. It is expected that the evaluation of these behavioral criteria will enable the assessment of the health index of young adults and yield meaningful output results at the conclusion of the case study.

```
features = ['goout', 'freetime','studytime','Pstatus', 'famrel', 'health']
data = data[features]
```

```
data.head()
```

	goout	freetime	studytime	Pstatus	famrel	health
0	4	3	2	1	4	3
1	3	3	2	0	5	4
2	2	3	2	0	4	3
3	2	2	3	0	3	4
4	2	3	2	0	4	4

Fig 3.3.1 Feature Selection

The fuzzy logic system employed in this study assesses the health index of young adults based on their physical and emotional behavior (E. S. Singh, 2015). The prediction of the health index relies on detailed analysis encompassing information, articles, and traits. The fuzzy logic system calculates and evaluates various criteria that may contribute to the likelihood of health problems. This evaluation is conducted by generating membership function degrees for the health index of young adults, enabling predictions of whether their health index is bad, average, or good. Consequently, mapping the criteria plays a crucial role in determining the health index of young adults. Lower mappings with the criteria indicate a lower health index, which may indicate potential health problems. The evaluation of the health index utilizes a fuzzy logic system, generating a membership degree that quantifies the health index of young adults. The methodology of this study comprises three components, namely the prediction flow chart for the health index, the knowledge pertaining to health problems, and the analysis algorithms of the fuzzy logic system. The fuzzy logic system applied in this study enables the assessment and prediction of the health index of young adults. By analyzing various criteria and employing fuzzy logic techniques, this methodology provides valuable insights into the mental well-being of young adults, aiding in the identification and understanding of potential health concerns.

Feature Selection:

This section has been specifically crafted to delve into the process of feature selection from the dataset for the case study at hand. The focus primarily lies on those variables which have demonstrated a high degree of correlation with the 'health' parameter, as clearly represented in Figure 3.2.

```
plt.figure(figsize=(6, 6))
heatmap = sns.heatmap(data.corr()[['health']].sort_values(by='health', ascending=False), vmin=-1, vmax=1, annot=True, cmap='BrBG'
heatmap.set_title('Features Correlating with the health of student', fontdict={'fontsize':18}, pad=16);
```

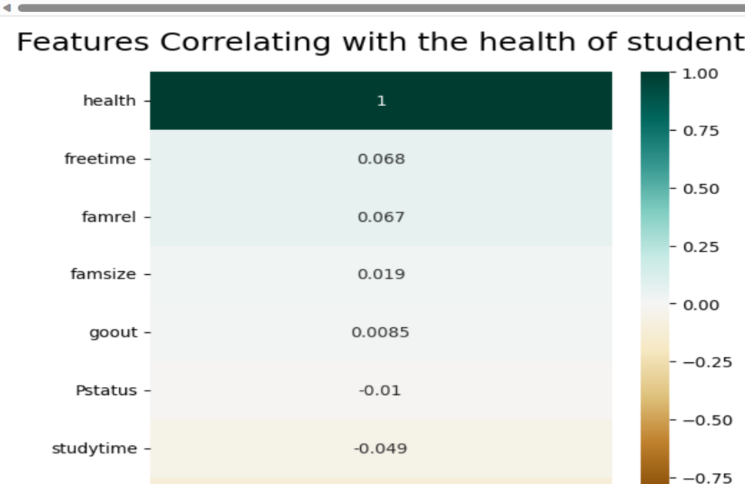

Fig 3.3.2 Correlation with health parameter

Among these variables, a few key ones are free time, family relations, the time spent in social outings (going out time), the amount of time dedicated to studying (study time), and the status of parents (parental status). Each of these features has shown a significant correlation with 'health', thereby making them critical considerations in our study. The rationale behind limiting our focus to these specific features is rooted deeply in the findings of the case study. It has been observed that when too many features are introduced into the system, it tends to create a situation of complexity. This often results in the overlapping of rules within the fuzzy logic system. This overlap can diminish the system's effectiveness, as it could lead to confusion, and possibly, conflicting results. Hence, by concentrating on a carefully selected subset of features that are highly correlational with 'health', we maintain the system's clarity and ensure that each rule within the fuzzy logic system can perform its role effectively without any encumbrances.

3.4 Prediction workflow:

To ensure the effectiveness of the case study and the resulting membership values, careful selection, and inclusion of criteria from a standard online dataset for young adults are essential. This approach aims to minimize the potential risk of health problems by implementing preventive measures based on the health index generated through the fuzzy logic system. By utilizing a reputable online dataset, the study can accurately predict health issues in young adults and provide valuable insights for health prevention strategies.

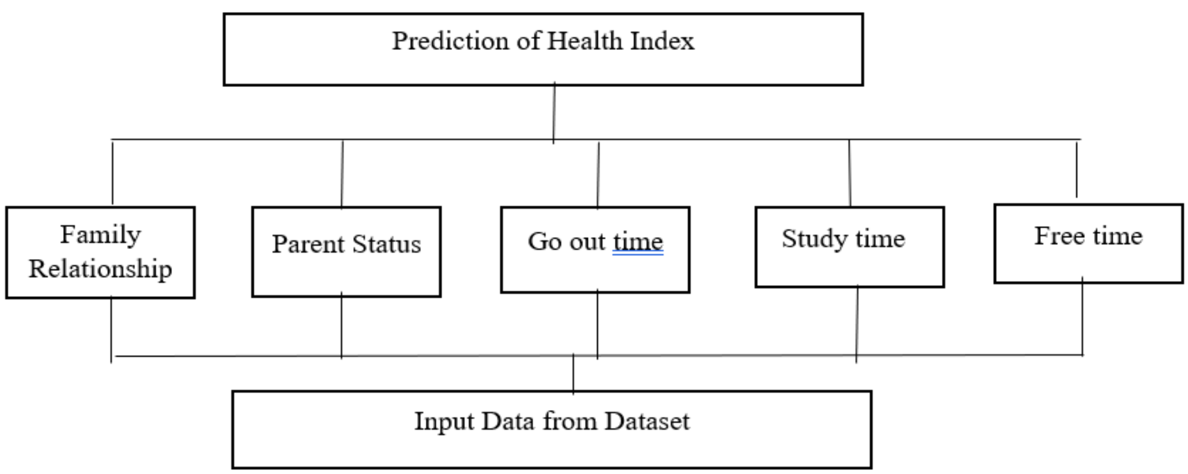

Fig 3.4 Structure of Prediction Workflow

According to reference (Elly Robinson, 2008), the significance of family relationships in the lives of young adults is emphasized, as they provide the support and encouragement necessary to navigate life's challenges. This implies that having a strong bond with family members empowers young adults to face difficulties with resilience. Conversely, families burdened with problems and challenges may lead to emotional distance between young adults and their families. Another criterion, described in reference (Julia Velten, 2018), pertains to spending time with family and friends. Engaging in social activities and going out with loved ones can potentially reduce the risk of health problems by allowing individuals to focus on enjoying life rather than being overwhelmed by stress. As stated in reference (Elly Robinson, 2008), the quality of family relationships holds great importance for young adults, providing them with support and courage. Similarly, reference (Julia Velten, 2018) highlights the benefits of spending time with family and friends, which can contribute to improved mental well-being by promoting a focus on enjoying life instead of succumbing to stress.

The study (Julia Velten, 2018) indicates that establishing a regular social rhythm, which involves engaging in social activities like going out, is indicative of better health among Chinese and German students. This suggests that maintaining a consistent schedule of social engagements, including going out, plays a crucial role in promoting positive mental well-being. Conversely, the presence of social rhythm irregularity, often observed among college students with varying schedules and part-time jobs, is strongly associated with increased health issues. This implies that an inconsistent social rhythm may have negative implications for health. In essence, the study emphasizes the significance of lifestyle choices in fostering good health. While certain lifestyle factors have been identified as predictors of future health problems, research on predicting future positive health outcomes remains relatively scarce. Therefore, further investigations are necessary to identify the lifestyle choices that are most beneficial for health and to explore the relative importance of various behaviors in achieving positive health outcomes.

The (Hannighofer, 2017) study presents that explores the influence of family type, comprising family status combined with relationship quality, on the health of both mothers and their children. The research reveals that family type is a more significant determinant than merely relationship status (whether a relationship is present or absent) when considering the mental well-being of mothers and their children. The quality of the relationship emerges as a key factor, exerting a long-term impact on the health of both mothers and children. Notably, certain factors associated with improved health in mothers do not yield the same outcomes for their children. The study highlights that children from unstable relationships and those of single mothers tend to exhibit a higher prevalence of health problems compared to children from stable two-parent families. This emphasizes the importance of family structure and stability in relation to children's health. Additionally, the study underscores the close association between the health of mothers and their children, indicating the need to consider both when examining the impact of family status on health. The study underscores the influence of family type, particularly relationship quality, on the health of mothers and their children. It reveals the significance of family structure and stability in shaping children's mental well-being and emphasizes the interdependence between the health of mothers and their children.

3.5 Fuzzy Logic Auto Rule Generation:

Auto rule generation in fuzzy logic systems offers numerous advantages that make it a valuable approach in various applications. One of its main benefits is the automation of the rule creation process based on input-output data, which significantly saves time and effort compared to manual rule crafting. This efficiency is particularly advantageous when dealing with complex systems or large datasets. Moreover, auto-rule generation methods demonstrate remarkable flexibility and can effectively handle diverse data types and complexities (M. C. P. de Souto, 2014), making them applicable to a wide range of domains and applications. By relying on data-driven methods, auto rule generation helps to mitigate human bias and subjectivity that may arise during manual rule creation. However, despite these advantages, there are also certain drawbacks associated with auto-rule generation in fuzzy logic systems. For instance, a notable disadvantage is the lack of interpretability in the automatically generated rules. The resulting fuzzy rule bases can be intricate and challenging for humans to comprehend. This lack of transparency may pose challenges in applications where a clear understanding of the decision-making process is crucial.

Fig 3.5 shows the workflow for Fuzzy Logic System (FLS) with Auto Rules Generation. The figure illustrates the process of generating fuzzy rules automatically based on input-output data. The input data is obtained from a dataset that contains information about family, relationships, parent status, go-out time, study time, and free time. This input data is then used to predict the health index of the student. The FLS with Auto Rules Generation is a data-driven method that automates the rule creation process based on input-output data. This approach saves time and effort compared to manual rule crafting, especially when dealing with complex systems or large datasets. However, one of the drawbacks of this approach is the lack of interpretability in the automatically generated rules. The resulting fuzzy rule bases can be intricate and challenging for humans to comprehend. In the study (Surmann, 2002) author points out that an issue with automatically generating fuzzy logic rule bases using genetic algorithms is the possibility of fuzzy sets and rules overlapping in a messy manner, which may not align with common sense.

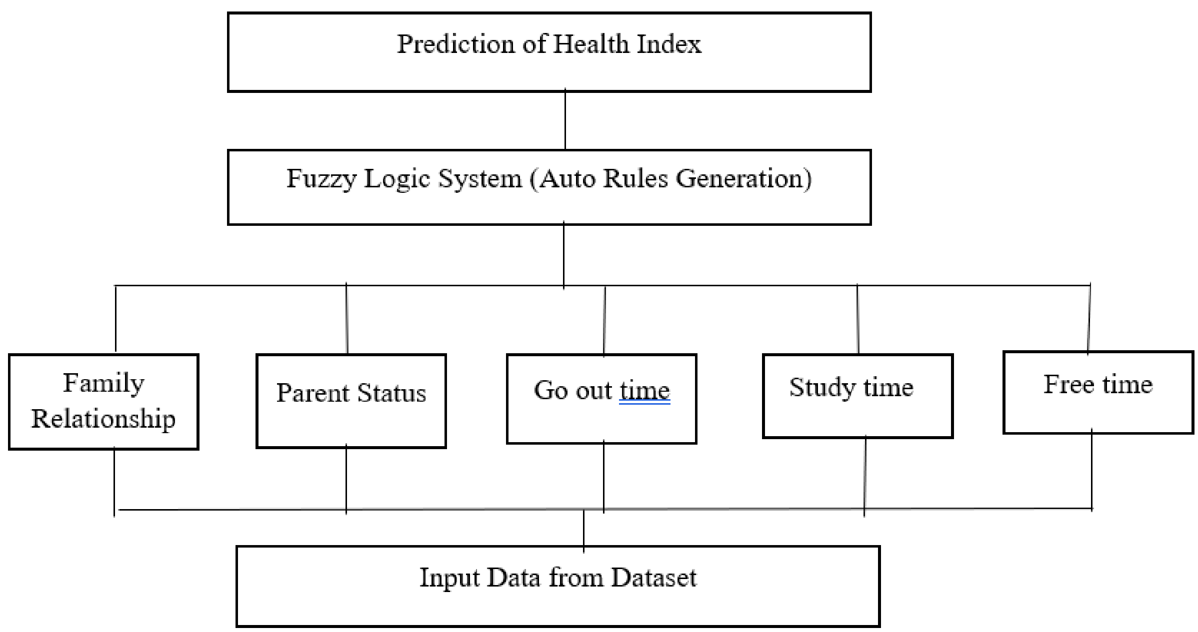

Fig 3.5 Workflow for FLS with Auto Rules Generation

Another drawback is the potential risk of overfitting the data. When the training dataset contains noise or outliers, the generated rules may become excessively tailored to the training data, resulting in subpar performance when applied to new, unseen data. For example, consider a technique for forecasting the stock market. An automatically created rule would read, "IF the date is within this month AND stock is X, THEN stock will rise," if the system was solely educated on data from the previous month in which a certain stock constantly increased. Because it is too precisely customized to the pattern of that month, this rule may succeed in the training data but fall short in the real world. Let’s look at another example. Imagine a medical diagnostic system that would help physicians identify a rare disease using a range of patient symptoms and test outcomes. Either an Interval Type-2 rule-based learning approach or a fuzzy logic auto-rule generation system are used by the system. Automatically creating rules can result in an extremely complex system due to the wide range of symptoms and test results for different diseases. It can be difficult for a doctor to understand why a particular diagnosis was made, especially when the system generates a lot of criteria. The auto-generated rules run the danger of overfitting the data if the system is trained on a small sample size (because the disease is uncommon). This implies that while the system may work well on well-known cases, it may incorrectly diagnose brand-new patients with slightly varied symptom presentations. It's possible that automatically created rules don't always correspond to expert medical knowledge. The system might, for instance, fail to recognize the relevance of a rare symptom that a skilled physician understands is essential to a correct diagnosis.

3.6 Type 1 Fuzzy Logic System (Rule-Based Learning):

In a Fuzzy Logic System, the notion of rule-based learning involves the creation or refinement of a collection of fuzzy if-then statements derived from input data. These statements act as the decision-making foundation within the system, influencing how it reacts to varied inputs. The core elements of a fuzzy logic system are these fuzzy if-then rules. They depict the fuzzy associations existing between the variables of input and output. The primary aim of implementing rule-based learning within a fuzzy logic system is to establish a rule base capable of accurately correlating inputs to outputs for a specific problem scenario. This procedure enables the fuzzy logic system to "learn" from data, thereby enhancing its performance incrementally. Consequently, it becomes more capable of adapting and efficiently handling complex problems characterized by imprecision and uncertainty. The system generates fuzzy if-then rules that capture the relationships between the input and output variables. These rules are based on the patterns and correlations found in the training data. The primary aim of implementing rule-based learning within a fuzzy logic system is to establish a rule base capable of accurately correlating inputs to outputs for a specific problem scenario. This procedure enables the fuzzy logic system to "learn" from data, thereby enhancing its performance incrementally. Consequently, it becomes more capable of adapting and efficiently handling complex problems characterized by imprecision and uncertainty.

As shown in Figure 3.6.1 Type-1 fuzzy logic is a type of controller that effectively deals with uncertainties and nonlinearity in complex systems. It is employed when traditional controllers struggle to deliver satisfactory performance in such scenarios. In a type-1 fuzzy logic controller, input variables are associated with fuzzy sets using membership functions. These functions determine the degree to which an input variable belongs to a specific fuzzy set. Similarly, the controller's output is also linked to a fuzzy set through a membership function. Fuzzy if-then rules define the relationship between the input variables and the output variable in the controller. These rules are constructed using linguistic variables, which are words or phrases describing the input and output variables. Fuzzy sets define these linguistic variables through membership functions, allowing for the handling of uncertainties and nonlinearity in the system. The rules are assessed using fuzzy logic, which enables flexible reasoning.

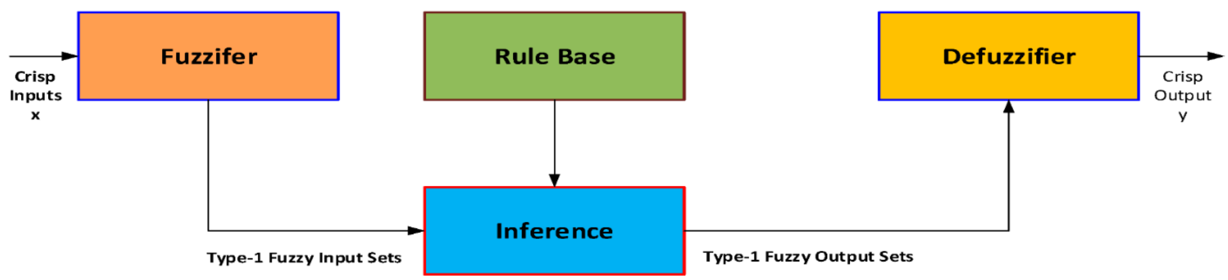

Fig 3.6.1 Type 1 Fuzzy Logic System (Carreon-Ortiz, 2023)

The controller's output is obtained by combining the outputs of all the rules through a fuzzy inference method. The Mamdani method, commonly used in fuzzy logic, utilizes the minimum operator to merge the rule outputs. To obtain a crisp value, the fuzzy output is defuzzified using a defuzzification method. The centroid method, frequently employed, calculates the center of gravity of the fuzzy output. A type-1 fuzzy logic controller relies on fuzzy sets, membership functions, fuzzy if-then rules, and fuzzy logic to handle uncertainties and nonlinearity in a system. It involves mapping input variables to fuzzy sets, evaluating rules using fuzzy logic, aggregating rule outputs, and defuzzifying the fuzzy output to obtain a crisp value (Carreon-Ortiz, 2023).

$$A = \{(x, \mu_A(x)) \mid x \in X\}$$

Eq 3.6 Type 1 Fuzzy Logic System from (Carreon-Ortiz, 2023)

Where: -

$\mu_A(x)$ is the membership function of the fuzzy set

A - x is the input variable

c is the center of the fuzzy set

A - a is the width of the fuzzy set A.

3.7 Interval Type 2 Fuzzy Logic System (Rule-Based Learning):

Interval Type-2 Fuzzy Logic rule-based learning is a variation of fuzzy logic used in healthcare and controlling autonomous mobile robots in dynamic and unfamiliar environments. It extends the concepts of Type-1 Fuzzy Logic, a mathematical framework for handling uncertainty and imprecision in decision-making. Interval Type-2 Fuzzy Logic rule-based learning maintains the same structure and operations as Type-1, including membership functions, rules, t-norms, fuzzification, inference, and defuzzification. However, the key distinction lies in the third block of the fuzzy logic system, where a type reducer and defuzzification components are employed to form the output processing block. This is necessary to accommodate the characteristics of the membership functions, which include additional degrees in the type of fuzzy sets(Kazemzadeh, 2014).

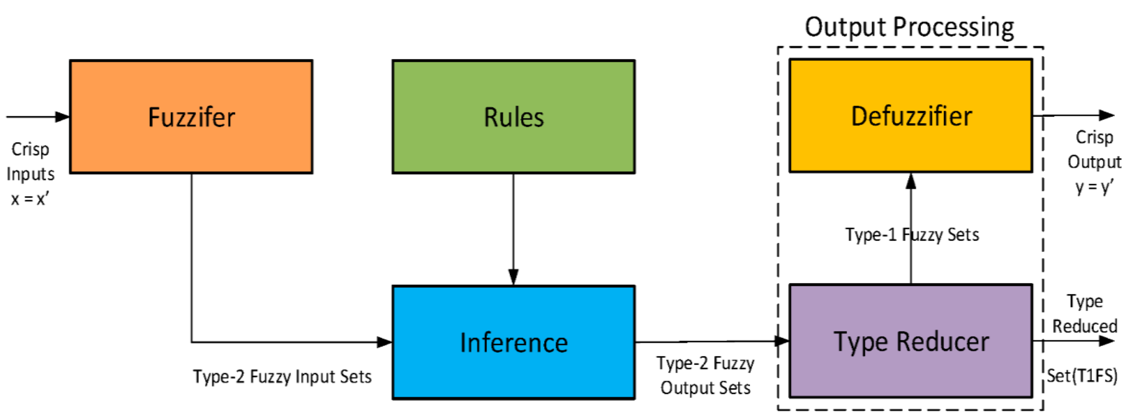

Fig 3.7.1 Interval Type 2 Fuzzy Logic System (Carreon-Ortiz, 2023)

The primary distinction between Type-1 and Interval Type-2 Fuzzy Logic systems lies in their treatment of uncertainty and imprecision. Type-1 Fuzzy Logic employs crisp membership functions that disregard any uncertainty in the input variables. In contrast, to Type-1 fuzzy logic, the rule-based approach in Interval Type-2 fuzzy logic introduces a different membership function structure. Instead of a single membership function, the rule-based approach incorporates upper and lower limits, known as the footprint of uncertainty (FOU). The FOU plays a crucial role in handling uncertainty by calculating the range of possible values for the intermediate points of the membership function. By utilizing the FOU, Interval Type-2 fuzzy logic addresses uncertainty and provides a more comprehensive representation of membership values.

$$\mathrm{FOU}(\tilde{A}) = \sum_{j=1}^{n_A} \bar{A}_e^j = \left[\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)\right] \qquad \forall\, x \in X$$

Eq 3.7.1 The equation for a footprint of uncertainty (Carreon-Ortiz, 2023)

Additionally, there is a disparity in the output processing block of the fuzzy logic system. Type-1 Fuzzy Logic defuzzifies the output to derive a crisp value. In Interval Type-2 Fuzzy Logic, a type reducer is employed to condense the two degrees of uncertainty into a single degree. This reduced degree is then defuzzified to obtain a crisp value.

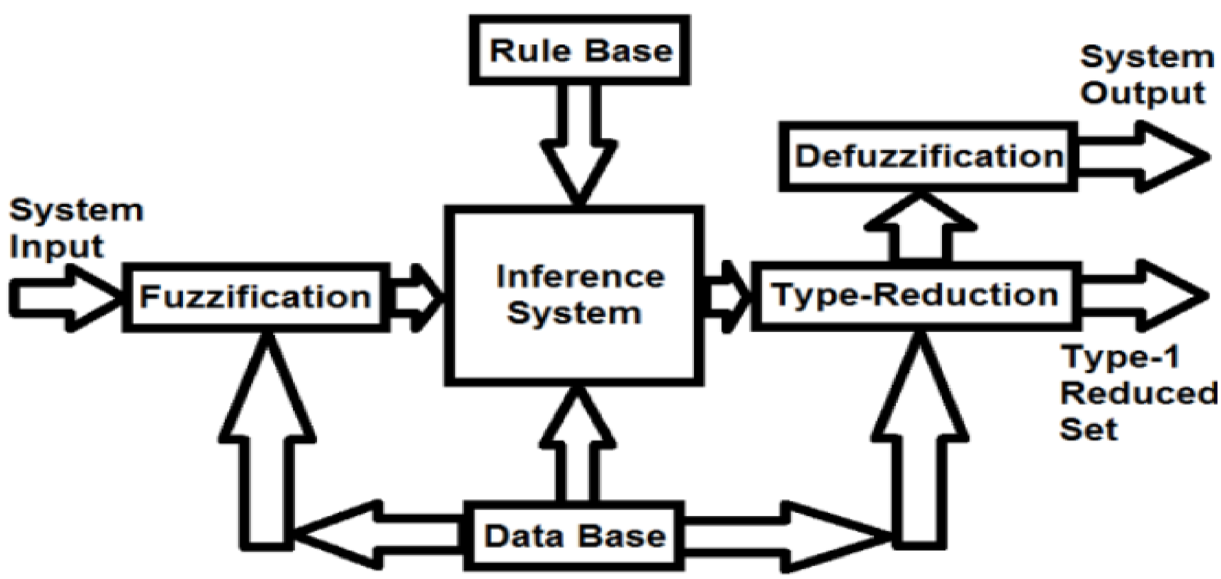

Fig 3.7.2 Workflow for Interval Type 2 FLS (Allawi, 2014)

Figure 3.7.2 illustrates the operation workflow of an Interval Type-2 Fuzzy Logic Rule-Based Learning System (IT2 FLS). The IT2 FLS is a variant of a fuzzy logic controller that leverages interval type-2 fuzzy sets to encapsulate uncertainty associated with input and output variables. Four principal phases constitute the IT2 FLS: fuzzification, rule evaluation, type-reduction, and defuzzification. The fuzzification phase is responsible for transforming precise input variables into fuzzy sets, utilizing membership functions (MFs). These functions establish a mapping of input variables onto fuzzy sets, indicating the degree to which an input variable belongs to each fuzzy set. These fuzzy sets are then passed onto the rule evaluation stage, wherein fuzzy rules are enacted to derive the output fuzzy sets. The output from the inference system of the IT2 FLS is Type-2 fuzzy sets, which necessitate type-reduction before the defuzzifier can convert them into a crisp output. This represents the fundamental structural distinction between Type-1 and Type-2 FLSs. The most frequently employed type-reduction method is the Center of Sets (COS) type-reducer. It calculates the centroid of the interval Type-2 consequent set to derive an interval Type-1 set. Lastly, the defuzzification stage transforms the interval Type-1 fuzzy sets into a clear-cut output variable, employing a defuzzification method.

$$\tilde{A} = \{((x,u),1) | \forall x \in X, \forall u \in J_x \subseteq [0,1]\}$$

Eq 3.7.2 Type 2 Fuzzy Logic System from (Carreon-Ortiz, 2023)

Where: -

μ_Ã(x,u) is the membership function of the type-2 fuzzy set

Ã - x is the input variable.

u is the secondary membership function that defines the possibilities of the primary membership

c(u) is the center of the fuzzy set Ã for a given secondary membership u

a(u) is the width of the fuzzy set Ã for a given secondary membership u.. +3

Type reducer in Interval type 2:

Interval Type-2 Fuzzy Logic Rule-Based Learning Systems (IT2FLS) utilize a type reducer, a function that transforms a type-2 fuzzy set into a type-1 fuzzy set. The primary purpose of the type reducer is to simplify the complexity of type-2 fuzzy sets, making them compatible with traditional fuzzy logic operations. Since type-2 fuzzy sets are more intricate than type-1 fuzzy sets, direct application of conventional fuzzy logic operations is not possible. Therefore, the type reducer is employed to map the type-2 fuzzy set to a type-1 fuzzy set, taking into account the uncertainty inherent in the secondary membership function.

$$y(\mathbf{x}) \equiv \frac{1}{2}\left[\underline{\mu}_{\tilde{B}}(y) + \bar{\mu}_{\tilde{B}}(y)\right]$$

Eq 3.7.3 The equation for Type Reducer (Carreon-Ortiz, 2023)

Several methods exist for type reduction, including the Karnik-Mendel algorithm, centroid method, and weighted average method. These methods differ in their approach to handling uncertainty in the secondary membership function and in their mapping of the type-2 fuzzy set to a type-1 fuzzy set. The type reducer is a crucial element within IT2FLS as it enables the utilization of traditional fuzzy logic operations on type-2 fuzzy sets. This capability is invaluable for modeling complex systems characterized by uncertainty and imprecision.

The system is based on historical data and clinical symptoms of patients and is a combination of forward and backward chaining. The rules are generated using a combination of expert knowledge and data-driven learning. The system uses interval type-2 fuzzy logic to handle the uncertainty of diagnosing spinal cord disorders. The rules are represented in the form of "IF-THEN" statements, where the antecedent is a set of input variables, and the consequent is a set of output variables. The system uses a defuzzification process to convert the fuzzy output into a crisp value (Rahimi Damirchi-Darasi, 2018). When using an Interval Type-2 Fuzzy Logic Rule-Based Learning System to estimate students' stress based on data like sleep hours and assignment loads, patterns emerge. For instance, a rule might state "Reduced sleep with increased assignments correlates with heightened stress." What sets Interval Type-2 apart is its "band" of membership, which captures broader uncertainties. This band's shape and breadth are influenced by the data, mirroring the diverse ways students react. As fresh data pours in, the system updates its rules and the membership bands, ensuring it remains responsive to the nuanced real-world situations it's modeling.

3.8 Membership Function:

In the realm of fuzzy logic, a membership function is an essential element that quantifies the extent to which an element belongs to a fuzzy set. It serves as a mathematical representation of the uncertainty tied to the membership of an element in a fuzzy set. The significance of membership functions lies in their role within modeling and reasoning processes, enabling fuzzy logic to effectively handle imprecise, vague, or subjective concepts. Membership functions establish a mapping between elements from the universe of discourse, which encompasses the range of possible values for a variable, and membership values that span the spectrum from 0 to 1. These membership values signify the degree of membership of the elements within the fuzzy set. The choice of membership function form can vary and include triangular, trapezoidal, Gaussian, or sigmoidal curves. The specific selection is based on the characteristics of the fuzzy set being represented and the linguistic interpretation applied to the variables in question. By employing membership functions, fuzzy logic possesses the capability to accurately model and reason with fuzzy information, thus effectively managing uncertainties. This capacity aids in making informed decisions when confronted with imprecision or subjectivity (Setiawan, 2020).

Triangular Membership Function: This function adopts a triangular shape, defined by three parameters. It assigns a membership value of 1 at the peak of the triangle, gradually decreasing to 0 towards the edges.

Trapezoidal Membership Function: Represented by a trapezoid shape defined by four parameters, this function assigns a membership value of 1 within a specified range and linearly decreases towards 0 at the edges.

Gaussian Membership Function: Utilizing a bell-shaped curve defined by the mean and standard deviation, this function assigns higher membership values to elements closer to the mean, gradually decreasing as the distance increases.

Sigmoidal (S-Shaped) Membership Function: Exhibiting an S-shaped curve defined by two parameters, this function assigns lower membership values at the extremes and higher values around the midpoint.

Generalized Bell Membership Function: This flexible function allows for various shapes based on specified parameters accommodating diverse application.

These membership functions enable the modeling and handling of uncertainties in fuzzy logic, facilitating decision-making processes and providing a framework to represent and reason with imprecise information.

The case study used the Gaussian membership function. The Gaussian membership function is a popular choice in fuzzy logic systems. The Gaussian function has a smooth, bell-shaped curve, allowing for a gradual and continuous transition between membership values. This provides a natural representation of gradual changes in membership. The Gaussian function's parameters, mean and standard deviation, can be adjusted to customize the position and width of the curve, making it adaptable to various applications and interpretations. The Gaussian distribution is well-known and widely used, making it familiar and easy to work with in fuzzy logic systems. The bell-shaped nature of the Gaussian function effectively captures gradual transitions in membership, providing a smooth and intuitive representation of uncertainty. Effective for Uncertainty Modeling: The Gaussian membership function assigns higher membership values to elements close to the mean, gradually decreasing them as elements move away. This makes it well-suited for modeling uncertainty in fuzzy logic systems. The Gaussian membership function is favored in fuzzy logic systems due to its smooth shape, symmetry, adaptability, familiarity, ability to capture gradual transitions, and effectiveness in modeling uncertainty.

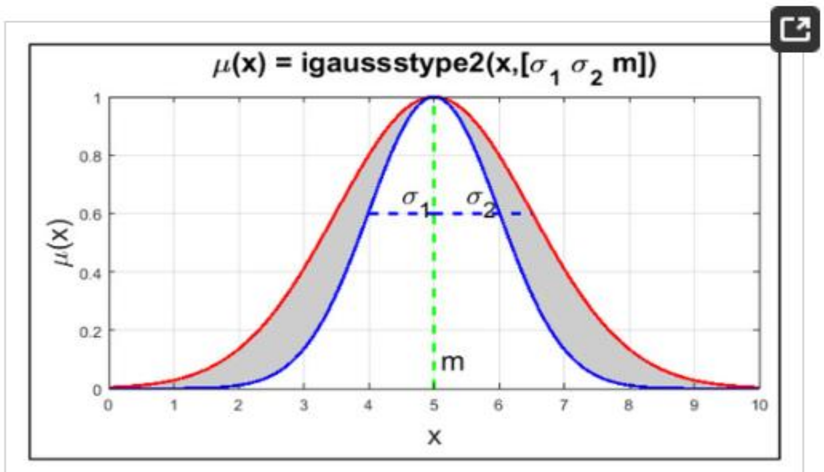

Fig 3.8.1 Gaussian membership function (Carreon-Ortiz, 2023)

Figure 3.8.1 illustrates a Gaussian Membership Function, which incorporates uncertainty in the standard deviation. This function is utilized in the architecture design of an Interval Type-2 Fuzzy Logic System. Within this framework, three Gaussian membership functions, namely Low, Medium, and High, are depicted. The red line represents the Upper Membership Function (UMF), while the blue line represents the Lower Membership Function (LMF). The shaded gray area within the membership function is known as the Footprint of Uncertainty (FOU) (Carreon-Ortiz, 2023).

3.9 Machine Learning:

Machine learning is a subset of artificial intelligence that focuses on developing algorithms and models that enable computers to learn and improve from experience without being explicitly programmed. It involves training systems on data to make predictions, recognize patterns, and solve complex tasks across various domains. In the case study, the performance of fuzzy logic was compared with machine learning. Random Forest and Decision Tree algorithms were used in the comparison.

Decision Tree:

The Decision Tree represents a machine learning algorithm used for classification and regression tasks. Its tree-like structure involves internal nodes as attribute tests, branches as test outcomes, and leaf nodes as class labels or numerical values. Decision Trees' interpretability makes them a favored option across various applications. The Decision Tree algorithm operates by recursively dividing data into subsets based on attribute values. The objective is to build a tree that predicts class labels or numerical values for new instances, relying on their attribute values. Figure 3.10.1 shows the structure of the decision tree. The algorithm employs measures of impurity like entropy or Gini index to select the best attribute for data splitting at each node. The attribute with the highest information gain or lowest impurity is chosen as the splitting attribute. Decision Trees offer numerous advantages, including easy interpretability, handling of categorical and continuous data, and addressing missing data. They are effective for high-dimensional data and support feature selection. Nonetheless, overfitting can be a concern, where the model becomes overly complex, fitting the training data too closely and leading to subpar generalization performance on new data (Song YY, 2015).

Pruning is a strategy employed to minimize the size of a decision tree by eliminating branches that don't significantly influence its precision. It can be executed in two methods: pre-pruning, which halts the construction of the tree prematurely, and post-pruning, which eradicates branches from a fully developed tree. Post-pruning, generally favored due to its ability to generate smaller, more accurate trees, can be classified into two categories: pessimistic and optimistic pruning. The former, including techniques like reduced error pruning (REP) and error-based pruning (EBP), assumes the actual error rate of a pruned tree is worse than the anticipated rate. In contrast, optimistic pruning, encompassing cost-complexity pruning (CCP) and minimum description length (MDL) pruning, operates under the assumption that the true error rate surpasses the estimated one. The effectiveness of pruning can be assessed using a training set, validation set, or test set, with the validation set determining the best-pruned tree and the test set estimating its accuracy. However, reserving some data specifically for the pruning set isn't typically an effective strategy, and certain methods may unexpectedly lead to excessive or insufficient pruning. The impact of pruning is contingent on the volume of training data in comparison to the complexity of the true structure that guides data generation. In situations where data is sparse, some pruning methods could potentially reduce the predictive accuracy of the generated trees (Nieta, 2006).

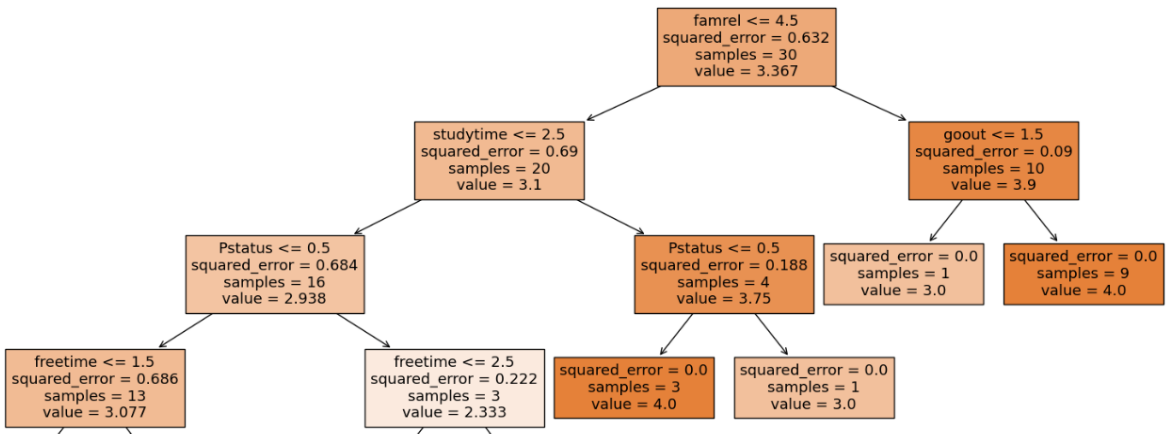

Fig 3.10.1 Decision Tree Regressor

Random Forest:

Random Forest is a versatile machine-learning algorithm for classification and regression tasks. It employs an ensemble learning approach, combining multiple decision trees to enhance prediction accuracy. Each decision tree is trained on a random subset of the data and features, reducing overfitting risks, and improving model performance. The final prediction is derived from aggregating the outputs of all the decision trees in the forest. Random Forests have gained popularity owing to their robustness, high accuracy, and capability to handle large and high-dimensional datasets. A significant advantage lies in their ability to handle missing and noisy data without overfitting, as well as accommodating both categorical and continuous data. Consequently, they find applications in various domains, including finance, healthcare, and image recognition. In finance, they aid in credit scoring, fraud detection, and stock price prediction, while in healthcare, they contribute to disease diagnosis, drug discovery, and patient monitoring. For image recognition tasks, they excel in object detection, face recognition, and scene classification. Overall, Random Forests stands as a powerful and user-friendly machine learning algorithm offering precise predictions for complex data scenarios (Ali, 2012).

Overfitting and Hyper Parameter Tuning:

Overfitting and hyperparameter tuning are fundamental concepts in machine learning with significant impacts on model performance and reliability. Overfitting occurs when a model becomes overly complex, excelling on training data but struggling to generalize to new, unseen data. This phenomenon arises as the model memorizes training data rather than understanding underlying patterns, resulting in subpar real-world performance. Detecting overfitting involves assessing the model's performance on a separate validation or testing dataset to ensure effective generalization. To counter overfitting, diverse techniques like cross-validation, regularization, data augmentation, and feature selection can be applied (Xue Ying, 2019). Figure 3.10.2 represents the workflow for random forest with hyper parameter tuning using grid search cv. In the hyper parameter tuning, models tune the parameter of algorithms such as “max depth”, “iteration”, “n_estimators.

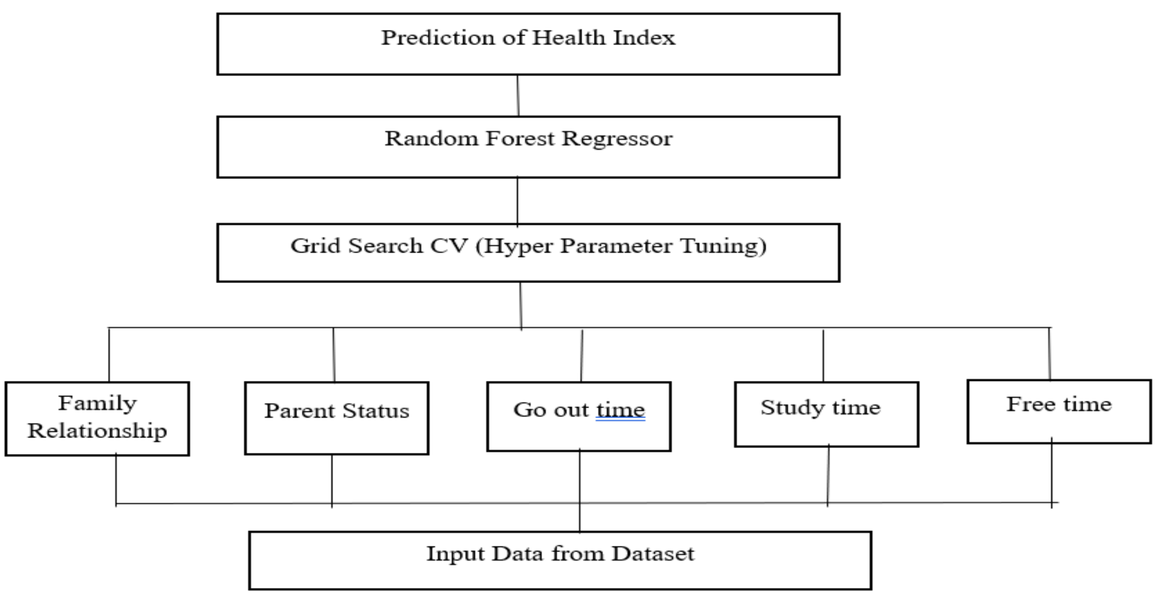

Fig 3.10.2 Workflow for Random Forest with Grid Search CV (Hyper Parameter Tuning)

Hyperparameter tuning involves optimizing the model's hyperparameters, which are pre-set parameters influencing model performance. Examples include learning rate, hidden layer count, and regularization strength. The significance of hyperparameter tuning lies in its capacity to enhance model performance and prevent overfitting. Hyperparameter tuning can be accomplished through approaches like grid search, evaluating different hyperparameter combinations, random search that selects values randomly within ranges, Bayesian optimization using probabilistic models, and genetic algorithms inspired by natural selection principles. Careful management of overfitting and diligent hyperparameter tuning enable machine learning models to achieve enhanced generalization and provide dependable predictions for new, unseen data. These practices play a pivotal role in developing accurate and robust models across diverse applications in the realm of machine learning.

3.10 Result Calculation:

```
count = 0
for i in range(0,350):
    it2out, tr = myIT2FLS.evaluate({"goout": data['goout'][i], "free_time": data['freetime'][i],
                                    "study_time": data['studytime'][i], "parents_status": data['Pstatus'][i],
                                    "family_relationship": data['famrel'][i]
                                   }, min_t_norm, max_s_norm, domain, method="Centroid", algorithm="KM")

    # Obtain the crisp output value
    crisp_output = crisp(tr["health"])
    print('Upper and Lower Limit :', tr)
    print('mental_health_of_the_student:', crisp_output)
    print('Actual_value:', data['health'][i])
    if round(crisp_output) == data['health'][i]:
        count = count + 1
    print()
print(count)
Accuracy = (count/350)*100
print('Accuracy :', round(Accuracy),'%')
```

Fig 3.10.1 Accuracy Calculation in Fuzzy Logic System

Figure 3.10.1 is a code snippet provided; the primary focus is on determining the accuracy of predictions made by the Fuzzy Logic System. Count variable is set to zero, aiming to record the number of times the system's predictions align with the actual data. By iterating through 350 data points, the system assesses various student attributes like their free time, study hours, parental status, and family relationships. For each student, the system outputs a predicted health value. This prediction is then juxtaposed against the actual health data from the dataset. Whenever a match is observed between the rounded off system's prediction and the actual data, the count increases by one. After processing all 350 students, the accuracy is derived by computing the ratio of correctly matched predictions to the total number of students, subsequently converting it into a percentage. This percentage encapsulates the effectiveness and reliability of the Fuzzy Logic System in predicting student health based on the specified attributes.

4. Implementation and Analysis:

This section presents the coding implementation for the case study, where four different fuzzy logic systems were implemented. These systems include Type 1 fuzzy logic rule-based learning, Type 1 fuzzy logic with automatic rules generation, Interval Type 2 fuzzy logic rule-based learning, and Interval Type 2 fuzzy logic with automatic rules generation. The results obtained from these implementations were thoroughly analyzed and compared to draw meaningful conclusions. Throughout the research, the coding implementation demonstrated the effectiveness of each fuzzy logic system. By comparing the results of these systems, valuable insights were gained. The performance of each system was evaluated based on various metrics and criteria. The implementation of Type 1 fuzzy logic rule-based learning showcased its capability in handling uncertainty and imprecision, while the automatic rules generation approach provided additional convenience and efficiency. On the other hand, Interval Type 2 fuzzy logic rule-based learning offered enhanced flexibility and improved representation of uncertainty, which further expanded the capabilities of the system. The inclusion of automatic rules generation in the Interval Type 2 fuzzy logic system brought added advantages in terms of ease of use and adaptability. By examining and comparing the outcomes of these fuzzy logic systems, a comprehensive understanding of their strengths and limitations was obtained. The research aimed to identify the most suitable approach for the specific case study, considering factors such as accuracy, robustness, and computational efficiency. The code for all four fuzzy logic systems has been mentioned in the appendix section.

4.1 Type 1 Fuzzy Logic System:

The code provided employs the scikit-fuzzy library to create a fuzzy logic system. It starts by importing the necessary libraries and modules. The input and output variables are defined using the ctrl.Antecedent and ctrl.Consequent classes, respectively, assigning names and value ranges to each variable. Membership functions, representing the degree of membership in fuzzy sets, are then associated with the variables. Gaussian membership functions are used, specified by their names and parameters that determine their shape and location as shown in the figure 4.1.1.

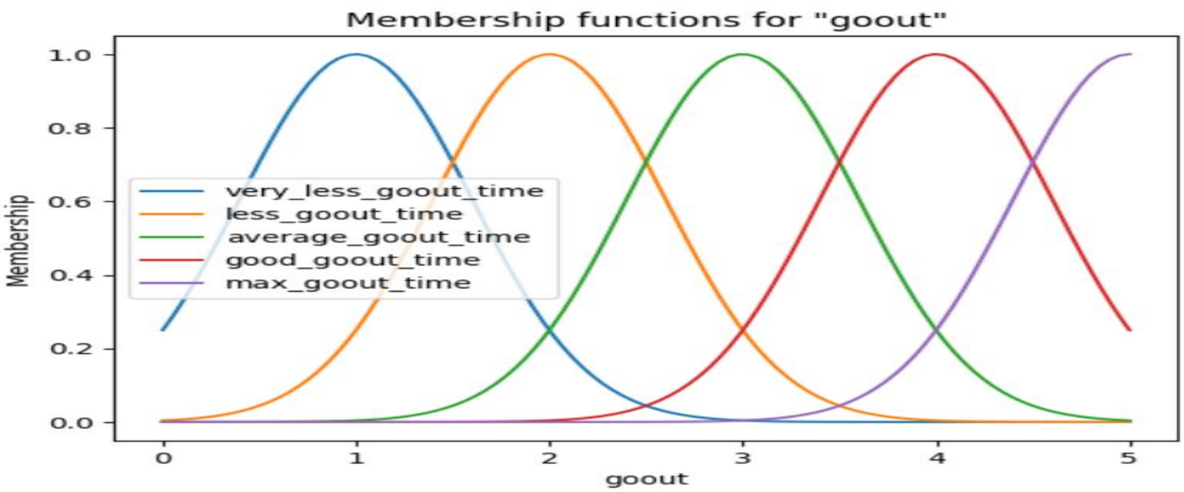

Fig 4.1.1 Membership Function for Go out Time.

Next, fuzzy logic rules are defined using the ctrl.Rule class, which combines input fuzzy sets using logical operators and associates them with output fuzzy sets. These rules represent the relationships between the input and output variables. A control system is created and initialized with the defined rules using the ctrl.ControlSystem class.

```
rule8 = ctrl.Rule(goout['average_goout_time'] &
                  freetime['average_freetime'] &
                  studytime['very_less_studytime'] &
                  Pstatus['together'] &
                  famrel['average_relationship'],
                  health['good'])


rule9 = ctrl.Rule(goout['good_goout_time'] &
                  freetime['less_freetime'] &
                  studytime['very_less_studytime'] &
                  Pstatus['together'] &
                  famrel['good_relationship'],
                  health['good'])

health_ctrl = ctrl.ControlSystem([rule1, rule2, rule3, rule4, rule5, rule6, rule7, rule8, rule9])
```

Fig 4.1.2 Rules for Type 1 Fuzzy Logic

This system forms the basis for the subsequent computations. A control system simulation object is instantiated using the ctrl.ControlSystemSimulation class, allowing for input values to be provided and output values to be computed based on the fuzzy logic rules.

```
Input: goout=4, studytime=2,  freetime=0, Pstatus=4,famrel=4,
Actual_value : 3
Predicted value Output: health=2.85089435545545
```

Fig 4.1.3 Output for Type 1 Fuzzy Logic System

The code then iterates over a dataset, setting the input values for the variables using the simulation object. The output value for the 'health' variable is computed, and the actual and predicted values are printed for each iteration. Finally, the code calculates the accuracy of the fuzzy logic system by comparing the predicted 'health' values with the actual values. The accuracy count is determined by correctly predicted values, and the overall accuracy is calculated as a percentage. In summary, this code demonstrates the implementation of a fuzzy logic system for predicting 'health' values based on the given input variables. It defines the variables, membership functions, and rules, and utilizes a control system simulation to compute the output values and evaluate the accuracy of the system.

The code for the Type 1 fuzzy logic rule-based learning system has been mentioned in the Appendix (Type 1 Fuzzy logic System). For the implementation of the Type 1 fuzzy logic system the library used is skfuzzy. scikit-fuzzy, also known as skfuzzy, is a versatile Python library belonging to the scikit-learn suite. It's specifically designed to work with fuzzy logic and is equipped with a variety of tools to facilitate fuzzy logic-based systems and controllers. skfuzzy enables the construction of comprehensive fuzzy systems, including the definition of fuzzy sets, rules, and methodologies for reasoning. It provides support for creating fuzzy control systems, which leverage fuzzy rules to govern processes or make informed decisions based on fuzzy inputs. One of the significant features of skfuzzy is fuzzy arithmetic, supporting addition, subtraction, multiplication, and division with fuzzy numbers. It also includes a function for the fuzzy c-means clustering algorithm, a distinctive clustering method where each data point can be a member of multiple clusters with varying degrees of membership. Scikit-fuzzy offers a variety of functions to generate membership functions, including triangular, trapezoidal, and Gaussian, among others. Furthermore, it includes functionality for fuzzy image processing and morphological operations. Defuzzification, the process of converting a fuzzy set into a single crisp number, is well-supported by scikit-fuzzy. The library offers several methods for this, including centroid, bisector, mean of maximum, smallest of maximum, and largest of maximum.

4.2 Type 2 fuzzy logic system:

Interval Type-2 Fuzzy Logic rule-based learning system (IT2FLS) and evaluates its performance in predicting health values based on input variables. The code defines the domain of the input variables using the numpy linspace function. The domain represents a range of values over which the fuzzy sets will be defined. Following that, the code defines various fuzzy sets for each input and output variable using the IT2FS_Gaussian_UncertStd class from the pyit2fls library. These fuzzy sets are based on Gaussian membership functions with uncertain standard deviation parameters. The IT2FLS object is then created to represent the overall fuzzy logic system. The code proceeds to add input and output variables to the IT2FLS system using the add_input_variable() and add_output_variable() methods, respectively. The input variables include "goout," "family_relationship," "free_time," "study_time," and "parents_status," while the output variable is "health."

```
domain = linspace(0.0, 5.0, 100)

very_bad_relationship = IT2FS_Gaussian_UncertStd(domain, [1, 0.6, 0.4, 0.8])
bad_relationship = IT2FS_Gaussian_UncertStd(domain, [2, 0.6, 0.4, 0.8])
average_relationship = IT2FS_Gaussian_UncertStd(domain, [3, 0.6, 0.4, 0.8])
good_relationship = IT2FS_Gaussian_UncertStd(domain, [4, 0.6, 0.4, 0.8])
very_good_relationship = IT2FS_Gaussian_UncertStd(domain, [5, 0.6, 0.4, 0.8])

very_less_free_time = IT2FS_Gaussian_UncertStd(domain, [1, 0.6, 0.4, 0.8])
less_free_time = IT2FS_Gaussian_UncertStd(domain, [2, 0.6, 0.4, 0.8])
average_free_time = IT2FS_Gaussian_UncertStd(domain, [3, 0.6, 0.4, 0.8])
good_free_time = IT2FS_Gaussian_UncertStd(domain, [4, 0.6, 0.4, 0.8])
max_free_time = IT2FS_Gaussian_UncertStd(domain, [5, 0.6, 0.4, 0.8])


very_less_study_time = IT2FS_Gaussian_UncertStd(domain, [1, 0.6, 0.4, 0.8])
less_study_time = IT2FS_Gaussian_UncertStd(domain, [2, 0.6, 0.4, 0.8])
average_study_time = IT2FS_Gaussian_UncertStd(domain, [3, 0.6, 0.4, 0.8])
good_study_time = IT2FS_Gaussian_UncertStd(domain, [4, 0.6, 0.4, 0.8])
max_study_time = IT2FS_Gaussian_UncertStd(domain, [5, 0.6, 0.4, 0.8])
```

Fig. 4.2.1 Membership Function

The next section of the code adds rules to the IT2FLS system using the add_rule() method. Each rule consists of antecedents (input variables and their associated fuzzy sets) and consequents (output variable and its fuzzy set). The rules are defined based on specific combinations of linguistic terms for each variable. After adding the rules, the code enters a loop that iterates over a dataset. In each iteration, the IT2FLS system evaluates the input values for the variables using the evaluate () method. The input values are obtained from the dataset, and the evaluation is performed using the minimum t-norm and maximum s-norm operators. The method "Centroid" and algorithm "KM" are specified for defuzzification.

```
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", average_free_time),
                   ("study_time", less_study_time),("parents_status", divorced),
                   ("family_relationship", good_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", good_free_time),
                   ("study_time", less_study_time),("parents_status", together),
                   ("family_relationship", bad_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", good_free_time),
                   ("study_time", average_study_time),("parents_status", together),
                   ("family_relationship", average_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", good_free_time),
                   ("study_time", less_study_time),("parents_status", together),
                   ("family_relationship", good_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", average_free_time),
                   ("study_time", less_study_time),("parents_status", together),
                   ("family_relationship", good_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", average_free_time),
                   ("study_time", good_study_time),("parents_status", together),
                   ("family_relationship", good_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", average_goout_time), ("free_time", average_free_time),
                   ("study_time", very_less_study_time),("parents_status", together),
                   ("family_relationship", average_relationship)], [("health", good)])
```

Fig 4.2.2 Rules for Interval Type-2 FLS

The result of the evaluation is an IT2FS object and a transformation result. The crisp output value is obtained by rounding the crisp(tr["health"]) value. The code then compares the crisp output value with the actual value from the dataset. If they match, the count variable is incremented. The health value and actual value are printed for each iteration.

```
it2out, tr = myIT2FLS.evaluate({"goout": data['goout'][i], "free_time": data['freetime'][i],
                                "study_time": data['studytime'][i], "parents_status": data['Pstatus'][i],
                                "family_relationship": data['famrel'][i]
                               }, min_t_norm, max_s_norm, domain, method="Centroid", algorithm="KM")
```

Fig 4.2.3 Evaluation Code

Finally, the accuracy of the IT2FLS system is calculated by dividing the count by the total number of iterations and multiplying by 100. The accuracy is rounded and displayed as a percentage. This code implements an IT2FLS system for predicting health values based on input variables. It defines the fuzzy sets, adds rules to the system, evaluates the system on a dataset, and calculates the accuracy of the predictions. This implementation demonstrates the use of fuzzy logic in modeling and predicting health outcomes.

```
Upper and Lower Limit : {'health': (3.5977429131114396, 4.198044750447484)}
mental_health_of_the_student: 3.897893831779462
Actual_value: 4
```

Fig 4.2.4 Interval Fuzzy Logic Type-2 Output

In the Interval Type-2 Fuzzy Logic System (IT2FLS), the output is obtained as an interval with a lower limit and an upper limit, as depicted in the figure. In type-2 fuzzy logic rule-based learning when the lower limit is 3.6 and the upper limit is 4.2, taking the average of these values (3.9) provides a predicted value. This interval-based representation is valuable because it allows for a more expressive modeling of uncertainty. It enables the capture of different degrees of uncertainty and imprecision in fuzzy membership degrees, making the approach more robust and adaptable to complex and uncertain systems. By considering a range of values within an interval, Type-2 fuzzy logic accommodates situations where the precise membership value of an element is uncertain or subject to variation. This interval-based approach provides a means to represent and reason uncertainty in a more accurate and reliable manner, facilitating better decision-making in real-world applications. The practical Implementation of the interval type 2 fuzzy logic rule-based learning system has been mentioned in the Appendix (Interval Type 2 Fuzzy logic System). The PYIT2FLS library is a Python-based tool specifically designed for working with interval type 2 fuzzy logic rule-based learning systems (IT2-FLSs). Unlike traditional Type-1 Fuzzy Logic Systems, IT2-FLSs are equipped to handle a higher degree of uncertainty, which makes this library especially useful for managing imprecise, incomplete, or noisy data. The library provides various types of membership functions, including triangular and trapezoidal, for defining fuzzy sets. Users can create fuzzy rules, which are logical statements in an IF-THEN format, used to guide the decision-making process in a fuzzy logic system. pyIT2FLS facilitates the creation of a complete fuzzy inference system that uses these rules to make decisions based on input data. Following the fuzzy inference process, the results obtained are typically in the form of fuzzy outputs. pyIT2FLS provides a mechanism to convert these fuzzy results back into crisp outputs, a process known as defuzzification. While pyIT2FLS may not be as widely used or have as extensive community support as some other fuzzy logic libraries, it offers a unique focus on Interval Type-2 fuzzy logic rule-based learning systems, making it a useful tool for applications requiring this more nuanced handling of uncertainty (Haghrah, 2023).

4.3 Type 1 and Interval Type 2 fuzzy logic system with Auto rule generation:

In this case, a fuzzy logic system has been implemented but automatically generates fuzzy logic rules by randomly selecting linguistic terms for each variable. It defines lists of linguistic terms for each variable, such as "famrel_terms" and "health_terms". The "generate_rule()" function randomly selects terms from these lists and combines them to form a rule in the form of "antecedents" and a "consequent". The generated rules are stored in the "rules" list. Finally, the code prints the generated rules using a loop. This automated rule generation can be useful for exploring different rule combinations and their potential impact on the fuzzy logic system's behavior.

```
def generate_rule():
    goout = random.choice(goout_terms)
    famrel = random.choice(famrel_terms)
    freetime = random.choice(freetime_terms)
    studytime = random.choice(studytime_terms)
    Pstatus = random.choice(Pstatus_terms)
    health = random.choice(health_terms)

    return "(goout['{}'] & freetime['{}'] & studytime['{}'] & Pstatus['{}'] & famrel['{}

rules = [generate_rule() for _ in range(9)]
```

Fig 4.3.1 Rule Generation

4.4 Machine Learning Implementation:

The case study involved the implementation of both Decision Tree Regressor and Random Forest Regressor to make health index predictions. The primary objective was to compare and assess the performance of these machine learning algorithms alongside fuzzy logic, all using the same dataset. During the analysis, it was observed that the Decision Tree model exhibited overfitting tendencies, as evidenced by its impressive 95% training accuracy but a notably poor -0.2% testing accuracy. This occurrence is common in machine learning, where algorithms excel in learning relationships within the training dataset but may falter when tested on unseen data, leading to overfitting conditions.

```
print('Traning Accuracy :', dt.score(X_train_scaled, y_train)*100)
print('Testing Accuracy :',dt.score(X_test_scaled, y_test )*100)

Traning Accuracy : 95.08528948212617
Testing Accuracy : -8.183925315554298
```

Fig 4.4.1 Decision Tree Regressor

Overfitting is a common problem in machine learning where a model becomes too complex or extensively trained on a specific dataset, resulting in its inability to generalize well to new, unseen data. Although the model may achieve exceptional performance on the training data, it fails to perform effectively on testing or validation data. The model essentially memorizes the training data instead of learning the underlying patterns, leading to inaccurate predictions when applied to real-world situations. This phenomenon can significantly impact the overall performance and reliability of the model.

In the case study, the Random Forest Regressor was employed on the same dataset, yielding a training accuracy of 87% and a testing accuracy of 20%. It was evident that the Random Forest Regressor outperformed the Decision Tree Regressor. However, a concerning issue arose due to the presence of overfitting, wherein the model became excessively complex and memorized the training data, resulting in poor generalization to new data. The significance of testing accuracy cannot be stressed enough, especially when dealing with health parameters, as accurate predictions directly impact individuals' well-being. Unfortunately, the observed overfitting in the model raised concerns regarding its reliability and capability to provide accurate predictions in real-world scenarios.

```
: print('traning accuracy :',rf.score(X_train_scaled,y_train)*100,'%')
  print('traning accuracy :',rf.score(X_test_scaled, y_test )*100,'%')

  traning accuracy : 87.2417789748348 %
  traning accuracy : 19.70597167402873 %
```

Fig 4.4.2 Random Forest Regressor

To address overfitting, hyperparameter tuning was conducted, aiming to fine-tune the model's parameters for better generalization. Therefore, the training accuracy decreased to 53%, and the testing accuracy decreased to 12%. While overfitting was reduced to some extent, the testing accuracy remained disappointingly low, indicating that the model's performance was still far from an optimal solution.

```
print('traning accuracy :', rf_with_param_tune.score(X_train_scaled, y_train)*100)
print('testing accuracy :', rf_with_param_tune.score(X_test_scaled, y_test)*100)

traning accuracy : 53.96128642275057
testing accuracy : 11.663437882105153
```

Fig 4.4.3 Random Forest Regressor with Hyper Parameter Tunning

Upon comparing the outcomes, it became evident that Fuzzy Logic clearly outperformed the machine learning algorithms in this case. Fuzzy Logic, renowned for its ability to handle uncertainty and imprecision, demonstrated superior performance, likely due to its inherent interpretability and capacity to capture complex data relationships without being susceptible to overfitting. In contrast, despite the appeal of machine learning algorithms like Random Forest

Regressor, the presence of overfitting hindered their effectiveness, making them less suitable for this specific health parameter prediction task. As such, Fuzzy Logic emerged as the more robust and reliable choice for this case study.

The case study includes a practical coding implementation of both machine learning algorithms, Decision Tree Regressor and Random Forest Regressor, which can be found in the Appendix section. The provided code demonstrates how these algorithms are applied to predict the health index using the given dataset. It covers various essential steps such as data preprocessing, dataset splitting for training and testing, model initialization, training, and performance evaluation. The implementation of the Decision Tree Regressor showcases the creation and training of a Decision Tree model using Python libraries like scikit-learn. Similarly, the Random Forest Regressor code illustrates the application of the Random Forest algorithm for regression tasks. Additionally, the code exemplifies the hyperparameter tuning process, where different combinations of hyperparameters are tested to achieve improved model performance and mitigate overfitting issues.

4.5 Comparison:

The following table presents a comparison of the four Fuzzy Logic Systems (FLS) and Machine learning algorithms in terms of their accuracy. In the case of FLS for testing purposes, 350 input values were selected from the dataset, and for machine learning 200 inputs were selected from the dataset. The predicted health index values were then compared to the actual values provided in the dataset. The predicted health index values were categorized into three levels: 2 for bad health, 3 for average health, and 4 for good health. In the evaluation of the fuzzy logic system, a set of 350 records was utilized for testing purposes. However, for the assessment of the machine learning algorithms, only 200 records were used. The decision to limit the machine learning evaluation to 200 records was made after observing that the inclusion of more than 200 records led to poorer performance outcomes. observing table 4.4, it can clearly be seen as FLS easily outperforms machine learning in this case study. The table presents a comparison of different techniques and their corresponding accuracy performances. The Decision Tree Regressor demonstrates a negative accuracy of -8%, suggesting potential overfitting.

Techniques	No. of records	Accuracy (%)
Decision Tree Regressor	200	-8
Random Forest Regressor with Hyperparameter Tuning	200	11
Random Forest Regressor	200	20
Type-1 (auto rules generation)	350	27
Type-2 (auto rules generation)	350	29
Type- 1 (Rule-Based Learning from data)	350	46
Type- 2 (Rule-Based Learning from data)	350	65

Table 4.4 Comparison of Result

In contrast, the Random Forest Regressor with Hyperparameter Tuning achieves an accuracy of 20%, while the standard Random Forest Regressor achieves 11%. Both auto rules generation techniques (Type-1 and Type-2) show improved accuracies of 27% and 29%, respectively. However, the rule-based learning from data (Type-1) surpasses them with even higher accuracies of 46% and 65% in separate instances, indicating its superior effectiveness among the considered techniques. This comparison allows us to evaluate the performance of each FLS in accurately predicting health levels based on the input variables. By examining the accuracy results, we can assess the effectiveness of the different FLS approaches in capturing and representing the complex relationships between the input variables and health outcomes.

5. Conclusion and Future Work:

5.1 Conclusion:

The project "Interval Type-2 Fuzzy Logic Learning Rules-based Approach for Health and Wellbeing" aimed to explore and compare different types of fuzzy logic systems for predicting students' health Index. Fuzzy logic offers a unique way to manage ambiguous data, making decisions that resemble human thought processes. This research delved into elements affecting the health index of young adults, considering aspects like time spent socializing, study habits, and family background. We evaluated two fuzzy logic models: Type-1 fuzzy logic rule-based learning and Interval type-2 rule-based learning, to predict these health metrics. Initial observations indicate that while Type-1 provides a direct approach, the Interval Type-2 model potentially offers better accuracy, particularly with complex data sets. These insights set the foundation for developing more precise and tailored health prediction strategies for young individuals. It utilized data from two schools in Portugal, compiled through academic records and questionnaires covering a range of aspects, such as students' marks, demographic details, societal influences, and educational factors. Traditional health prediction models often struggle to manage the uncertainty and complexity inherent in human health data. This challenge becomes especially critical when predicting student health, which can be influenced by numerous interrelated factors. Hence, a more sophisticated model was needed.

The objective was to develop and compare different types of fuzzy logic systems, including Type-1 fuzzy logic rule-based learning system, Type-1 fuzzy logic system with auto rule generation, Interval type-2 fuzzy logic rule-based learning system, and Interval type-2 fuzzy logic system with auto rule generation. Furthermore, these systems were compared to three Machine Learning algorithms: Decision Tree Regressor, Random Forest Regressor, and Random Forest Regressor with Parameter Tuning. The goal was to identify the system that best predicted the health indices of students. Four fuzzy logic systems and three machine learning models were developed and applied to the dataset. The performance of each system was evaluated based on the accuracy of health index predictions against actual values from the dataset. The scikit-fuzzy (skfuzzy) and PYIT2FLS libraries, available in Python, were used for implementing and evaluating these fuzzy logic systems. The results obtained from the experiments demonstrated that the Interval Type 2 Fuzzy Logic Rule-Based learning System outperformed all other three systems with an impressive accuracy of 65%. This system showed a robust ability to model uncertainty and imprecision, making it highly effective in handling complex and uncertain scenarios. The second-highest accuracy was achieved by the Type 1 Fuzzy Logic Rule-Based Learning System with a 46% accuracy rate. Although not as accurate as the Interval Type 2 Fuzzy Logic Rule-Based Learning System, it still showed promising results in predicting health outcomes based on the input variables. In contrast, the Type 2 Fuzzy Logic System with auto rule generation and the Type 1 Fuzzy Logic System with auto rule generation exhibited lower accuracies of 30% and 27%, respectively. In the case of Machine Learning algorithms, all three algorithms achieved very low accuracy, and it can be clearly observed that FLS outperforms machine learning algorithms. In the thorough conclusion of our project, "Interval Type-2 Fuzzy Logic Learning Rules-based Approach for Health and Wellbeing," it was emphasized how crucial it is to manage ambiguous health data, particularly when forecasting students' health indices.

The use of learning rule-based straight from data was one noteworthy element and a significant contribution of mine. With the help of this novel concept, a more individualized and data-driven approach was provided, enabling the fuzzy logic system to adjust to the distinct nuances and trends found in the dataset. We produced outstanding results by utilizing this rule-based learning approach, particularly with the Interval Type 2 Fuzzy Logic System, demonstrating its strong capability in modeling complexity and uncertainty. Our models pay attention to adapting to, and improving as the data talks, demonstrating the dynamic nature of rule-based learning from data. When compared to other machine learning algorithms and conventional health prediction models, this approach's importance became even more clear. It is obvious that using data-driven rules enables more accuracy and adaptability, which are crucial for changing health conditions. Even though it highlights some areas for improvement, the experiment highlights the importance of direct learning from data, pointing to a positive approach for future research on student health forecasting. However, it was noted that the performance of fuzzy logic systems could be significantly influenced by various factors such as the selection of linguistic terms, membership functions, and rules, as well as the quality and size of the dataset used for testing. While these results are promising, there are avenues for further refinement and improvement. The linguistic terms, membership functions, and rules can be fine-tuned, and domain-specific knowledge can be incorporated for better predictions. The outcomes of this project reveal both its strengths and areas where enhancements are warranted. One clear limitation observed was the dataset's size.

The chosen dataset might not capture the full range of real-world scenarios, suggesting that results could vary with a different, perhaps larger, dataset. Moreover, while the auto rule generation process streamlined the application of the fuzzy logic system, it didn't always generate the most optimal rules, highlighting the value of domain expertise. The quality of data, vital for such analyses, may also have affected the reliability of our results, underscoring the importance of robust data validation processes in future studies. Furthermore, while the Interval Type 2 Fuzzy Logic System showcased promising results, it, like other systems, occasionally produced complex rules that posed interpretability challenges. Looking forward to this, there's considerable scope for improvement. Refining the rule generation algorithm by incorporating broader input variables can elevate the system's accuracy. Collaborating with experts in the health sector could lead to more grounded and realistic fuzzy rules. The system might also benefit from a dynamic adjustment mechanism that responds to changing health trends, making it more adaptable. Additionally, integrating the fuzzy logic system with other AI techniques can potentially enhance its predictive power. And lastly, an intuitive, user-friendly interface paired with a robust ethical framework can make this system more accessible and trustworthy for both professionals and the public. The aim is to continue enhancing the accuracy of health predictions for students, contributing to advancements in the field of student health and well-being. Future research could also focus on integrating the latest techniques and methodologies in fuzzy logic for improved outcomes.

5.2 Limitation and Future Work:

This study, while insightful, does encounter certain limitations. Firstly, the reliance on existing data might not fully capture the intricacies of young adults' health indices. To overcome this, a more comprehensive survey targeting secondary school students could provide a richer dataset for analysis. Additionally, while a comparative analysis of fuzzy logic systems was conducted, a more robust evaluation using real-time data could offer a more accurate depiction of the predictive capabilities of Type-1 and Interval Type-2 fuzzy logic systems. Moreover, the integration of fuzzy logic with other artificial intelligence techniques, such as neural networks and genetic algorithms, remains largely unexplored, hinting at a promising avenue for future investigation. Tailoring fuzzy logic algorithms for specific tasks like fuzzy clustering or decision-making could also enhance the precision and applicability of the systems.

Looking ahead, there are several promising directions for future research in this domain. The implications of conducting a comprehensive survey among secondary school students could significantly impact the medical industry by providing valuable insights into health prediction among young adults. Integrating survey data with fuzzy logic techniques holds the potential to unravel the complex web of factors influencing health indices. Embracing interdisciplinary collaborations and hybrid systems, combining fuzzy logic with other computational approaches, could lead to innovative solutions for addressing complex health assessment challenges. Advancing type-2 fuzzy logic systems to tackle more intricate problems is also a frontier worth exploring. Ultimately, the goal is to enhance the well-being of young individuals through accurate interventions driven by a deeper understanding of health predictors and improved fuzzy logic techniques. By addressing limitations and pursuing these avenues, researchers can contribute to the progress of health assessment and intervention strategies.

6. References

A. Torres-Iglesias and J. J. Nieta, 2006. *"Fuzzy Logic in Medicine and Bioinformatics," Journal of Biomedicine and Biotechnology, p.,* s.l.: s.n.

Ahna Ballonoff Suleiman, R. E. D., 2016. *"Leveraging Neuroscience to Inform Adolescent health: The Need for an Innovative Transdisciplinary Development Science of Adolescene," Adolescent health, pp. 240-248, 17 December 2016.,* s.l.: s.n.

Ali, J. K. R. A. N. &. M. I., 2012. *Random Forests and Decision Trees.,* s.l.: s.n.

Allawi, Z. a. A. T., 2014. *An Optimized Interval Type-2 Fuzzy Logic Control Scheme based on Optimal Defuzzification. International Journal of Computer Applications, 95(13), pp.26-30. DOI: 10.5120/16655-6633.,* s.l.: s.n.

archive.ics.uci.edu., n.d. *UCI Machine Learning Repository. [online] Available at: https://archive.ics.uci.edu/dataset/320/student+performance..* [Online].

Bedi, P. G. S. B. B. P. A. A. S. R. S., 2023. *Fusion Fuzzy Logic with Deep Learning for Identifying Depressed People Based on Their Facial Expressions. Procedia Computer Science, 218, 2795-2805. doi: 10.1016/j.procs.2023.08.,* s.l.: s.n.

Carreon-Ortiz, H. V. F. &. C. O., 2023. *Comparative Study of Type-1 and Interval Type-2 Fuzzy Logic Systems in Parameter Adaptation for the Fuzzy Discrete Mycorrhiza Optimization Algorithm. Mathematics, 11(11), 2501. https://doi.org/10.3390/,* s.l.: s.n.

Chung, J. a. T. J., 2022. *'Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges', Journal of Healthcare Engineering, 2022, pp. 1-19. doi: 10.1155/2022/6649479.,* s.l.: s.n.

Constanza Huapaya1, n.d. *Proposal of Fuzzy Logic-based Students´ Learning Assessment Model,* s.l.: s.n.

E. S. Singh, 2015. *"A Fuzzy Rule-Based Expert System to Diagnostic Mental Illness (MIDExS)," 2015.,* s.l.: s.n.

Ekong, V. E., 2013. *A fuzzy inference system for predicting depression risk levels. African Journal of Mathematics and Computer Science Research, 6(10), 197-204. doi: 10.5897/AJMCSR2013.0511,* s.l.: s.n.

Elly Robinson, B. R. &. P. B., 2008. *"Family relationships and mental illness," Australian Family Relationships Clearinghouse (AFRC) Issues, pp. 1-19, 2008.,* s.l.: s.n.

Facchinetti, G. A. T. P. T. &. M. G., 2012. *A fuzzy approach to face the multidimensional aspects of well-being. In Proceedings of the 2012 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS) (pp. 1-6),* s.l.: s.n.

Ghosh, S. B. A. S. N. N. V. N. R. &. K. A., 2017. *A Fuzzy Logic System to Analyze a Student's Lifestyle.,* s.l.: s.n.

Haghrah, A., 2023. *PyIT2FLS. [online] GitHub. Available at: https://github.com/Haghrah/PyIT2FLS [Accessed 31 Jul. 2023].,* s.l.: s.n.

Hannighofer, L. M. F. H. M. H. K. &. Z. J., 2017. *The impact of relationship status and family type on the health of mothers and their children: A 10-year longitudinal study. Frontiers in psychology, 8, 1227. doi: 10.3389/fpsyg.2017.,* s.l.: s.n.

Huapaya, C., n.d. *(., n.d. Proposal of Fuzzy Logic-based Students´ Learning Assessment Model. [online] Available at: https://core.ac.uk/download/pdf/296349722.pdf.,* s.l.: s.n.

J. Betty Jane, D., n.d. *"A Review On Big Data With Machine Learning And Fuzzy Logic For Better Decision Making." International Journal of Scientific & Technology Research, vol. 8, no. 10, October 2019, pp. 1121-1124. ISSN 2277-8616.,* s.l.: s.n.

Jiang, F. J. Y. Z. H. D. Y. L. H. M. S. W. Y. D. Q. S. H. &. W. Y., 2017. *Artificial intelligence in healthcare: past, present and future. Stroke and Vascular Neurology, 2, e000101. doi:10.1136/svn-2017-000101,* s.l.: s.n.

Julia Velten, A. B. S. S. A. W. J. M., 2018. *"Lifestyle choices and health: a longitudinal survey with German and Chinese students," BMC Public health, pp. 1-15, 2018. ,* s.l.: s.n.

Kaggle, n.d. *kaggle.com. (n.d.). Student Performance Prediction. [online] Available at: https://www.kaggle.com/competitions/1056lab-student-performance-prediction/overview..* [Online].

Kazemzadeh, A. L. S. &. N. S., 2014. *Understanding the conceptual meaning of emotion words: A computational model based on interval approach to fuzzistics. IEEE Transactions on Affective Computing, 5(1), 3-14. doi: 10.1109/TAFFC.2013.28,* s.l.: s.n.

M. C. P. de Souto, R. F. d. C. a. A. A. F., 2014. *"FCA-Based Rule Generator: A Genetic Approach for Extracting Fuzzy Classification Rules," in Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Beijing, China, July 2,* s.l.: s.n.

M., J., 2003. *Review of Uncertain rule-based fuzzy logic systems: introduction and new directions. Fuzzy Sets and Systems, 133(1), 133-135. DOI: 10.1016/S0165-0114(02)00359-7.,* s.l.: s.n.

Mendel, Q. L. a. J. M., 2000. *Fellow and IEEE, "Interval Type-2 Fuzzy LogicSystems: Theory and Design," in IEEE TRANSACTIONS ON FUZZY SYSTEMS, 2000. ,* s.l.: s.n.

Monard, M. C. M. C. H. &. C. I., 2009. *A comparative study on classic machine learning and fuzzy approaches for classification problems. [PDF] https://ieeexplore.ieee.org/document/4801906,* s.l.: s.n.

Moraga, C., 2005. *Introduction to Fuzzy Logic. FACTA UNIVERSITATIS Series Electronics and Energetics, 18(2), 319-332. DOI: 10.2298/FUEE0502319M.,* s.l.: s.n.

N. Gupta, 2014. *"Comparative Study of Type-1 and Type-2 Fuzzy Systems," International Journal of Engineering Research and General Science Volume 2, pp. 195-198, June-July 2014.,* s.l.: s.n.

Nguyen, T. K. A. C. D. &. N. S., 2015. *Medical data classification using interval type-2 fuzzy logic system and wavelets. Applied Soft Computing, 30, 812-822. DOI: 10.1016/j.asoc.2015.02.016.,* s.l.: s.n.

Nieta, A. T.-I. a. J. J., 2006. *"Fuzzy Logic in Medicine and Bioinformatics," Journal of Biomedicine and Biotechnology, p., s.l.: s.n.,* s.l.: s.n.

Setiawan, A. A. E. &. S. P., 2020. *Fuzzy membership functions analysis for usability evaluation of online credit hour form. Journal of Engineering Science and Technology, 15(5), 3189-3203.,* s.l.: s.n.

Shinde, S. a. R. P., 2018. *Intelligent health risk prediction systems using machine learning: A review. International Journal of Engineering & Technology, 7(3), pp.1019-1023. DOI: 10.14419/ijet.v7i3.12654.,* s.l.: s.n.

Smith, J., 2021. *Family Relationships and Mental Illness in Australia [PDF]. Australian Family Relationships Clearinghouse.,* s.l.: s.n.

Song YY, L. Y., 2015. *Decision tree methods: applications for classification and prediction. Shanghai Arch Psychiatry. 2015 Apr 25;27(2):130-5. doi: 10.11919/j.issn.1002-0829.215044. PMID: 26120265; PMCID: PMC4466856.,* s.l.: s.n.

Surmann, H. &. S. A., 2002. *Automatic generation of fuzzy logic rule bases: Examples I. In Proceedings of the NF2002: First International ICSC Conference on Neuro-Fuzzy Technologies (pp. 75). Cuba: ICSC Academic Press.,* s.l.: s.n.

Umoh, U. U. S. A. T., 2021. *Interval Type-2 Fuzzy Framework for Healthcare Monitoring and Prediction. In: Advances in Intelligent Systems and Computing, vol 1235. Springer, Singapore. DOI: 10.1007/978-981-16-1543-6_17.,* s.l.: s.n.

Xue Ying, 2019. *"An Overview of Overfitting and its Solutions," J. Phys.: Conf. Ser. 1168, 022022 (2019).,* s.l.: s.n.

Zadeh, L., 1965. *Fuzzy sets. Information and Control, 8(3), pp.338–353. doi:https://doi.org/10.1016/s0019-9958(65)90241-x., s.l.: s.n.,* s.l.: s.n.

7. Appendix A

This appendix section will be used to display code related to the Section's Implementation and analysis.

Data Analysis and Preprocessing:

```
In [3]: # Loading the necessary libraries
        import numpy as np
        import pandas as pd
        import matplotlib.pyplot as plt
        import seaborn as sns


        # Loading the dataset
        df = pd.read_csv('student-por.csv', sep=';')
        df1 = pd.read_csv('student-mat.csv', sep = ';')

        # Concatinate both the datasets
        df = pd.read_csv('student-por.csv', sep=';')
        df1 = pd.read_csv('student-mat.csv', sep = ';')
        data = pd.concat([df, df1], axis=0)
        data.head()

        # export the dataset
        data_new.to_csv('final_dataset.csv')
```

Fig. 7.1 Data Preprocessing

As shown in fig 7.1 is a snippet of data preprocessing code, which loads two datasets, 'student-por.csv' and 'student-mat.csv', using pandas. Both datasets are then concatenated vertically to form a unified dataset, which is subsequently saved as 'final_dataset.csv'. The complete code can be found under the filename "Data Preprocessing and Analysis" in the accompanying document.

```
plt.figure(figsize=(9,6))
ax = sns.countplot(x = 'guardian', hue = 'school', data = data)
plt.xlabel('guardian')
plt.title('guardians of students in both the school')
plt.show()
```

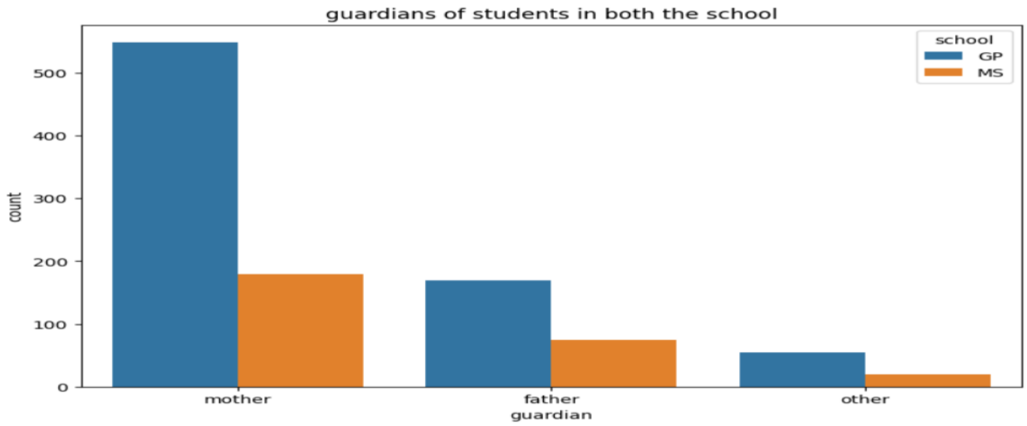

Fig 7.2 Data Analysis

Type 1 Fuzzy Logic Rule Based Learning Implementation:

This section has mentioned the code for Type-1 fuzzy logic rule-based learning system implementation. Fig 7.3 covers the part of input and output variable creation and membership function, Fig 7.4 is the code for rules of type 1 fuzzy logic rule-based learning system, and Fig 7.5 is the prediction and accuracy calculation. The complete code can be found in the supporting document under the filename "Type 1 fuzzy logic system rule-based learning."

```
import numpy as np
import skfuzzy as fuzz
from skfuzzy import control as ctrl

# Create input variables
goout = ctrl.Antecedent(np.linspace (0, 5, 100), 'goout')
famrel = ctrl.Antecedent(np.linspace(0, 5, 100), 'famrel')
freetime = ctrl.Antecedent(np.linspace(0, 5, 100), 'freetime')
studytime = ctrl.Antecedent(np.linspace(0, 5, 100), 'studytime')
Pstatus = ctrl.Antecedent(np.linspace(0, 5, 100), 'Pstatus')

# Create output variable
health = ctrl.Consequent(np.linspace(0, 5, 100), 'health')

# Define membership functions for input variables
goout['very_less_goout_time'] = fuzz.gaussmf(goout.universe, 1, 0.6)
goout['less_goout_time'] = fuzz.gaussmf(goout.universe, 2, 0.6)
goout['average_goout_time'] = fuzz.gaussmf(goout.universe, 3, 0.6)
goout['good_goout_time'] = fuzz.gaussmf(goout.universe, 4, 0.6)
goout['max_goout_time'] = fuzz.gaussmf(goout.universe, 5, 0.6)

famrel['very_bad_relationship'] = fuzz.gaussmf(famrel.universe, 1, 0.6)
famrel['bad_relationship'] = fuzz.gaussmf(famrel.universe, 2, 0.6)
famrel['average_relationship'] = fuzz.gaussmf(famrel.universe, 3, 0.6)
famrel['good_relationship'] = fuzz.gaussmf(famrel.universe, 4, 0.6)
famrel['very_good_relationship'] = fuzz.gaussmf(famrel.universe, 5, 0.6)

freetime['very_less_freetime'] = fuzz.gaussmf(freetime.universe, 1, 0.6)
freetime['less_freetime'] = fuzz.gaussmf(freetime.universe, 2, 0.6)
freetime['average_freetime'] = fuzz.gaussmf(freetime.universe, 3, 0.6)
freetime['good_freetime'] = fuzz.gaussmf(freetime.universe, 4, 0.6)
freetime['max_freetime'] = fuzz.gaussmf(freetime.universe, 5, 0.6)
```

Fig 7.3 Membership Function

```
rule1 = ctrl.Rule(goout['good_goout_time'] &
                  freetime['average_freetime'] &
                  studytime['less_studytime'] &
                  Pstatus['divorced'] &
                  famrel['good_relationship'],
                  health['good'])

rule2 = ctrl.Rule(goout['good_goout_time'] &
                  freetime['good_freetime'] &
                  studytime['less_studytime'] &
                  Pstatus['together'] &
                  famrel['bad_relationship'],
                  health['good'])

rule3 = ctrl.Rule(goout['good_goout_time'] &
                  freetime['good_freetime'] &
                  studytime['average_studytime'] &
                  Pstatus['together'] &
                  famrel['average_relationship'],
                  health['bad'])

rule4 = ctrl.Rule(goout['average_goout_time'] &
                  freetime['average_freetime'] &
                  studytime['very_less_studytime'] &
                  Pstatus['together'] &
                  famrel['average_relationship'],
                  health['good'])

rule5 = ctrl.Rule(goout['good_goout_time'] &
                  freetime['good_freetime'] &
                  studytime['average_studytime'] &
                  Pstatus['together'] &
                  famrel['average_relationship'],
                  health['average'])
```

Fig 7.4 Rules

```
health_ctrl = ctrl.ControlSystem([rule1, rule2, rule3, rule4, rule5, rule6, rule7, rule8, r


health_sim = ctrl.ControlSystemSimulation(health_ctrl)
goout_values = data['goout'].values
studytime_values = data['studytime'].values
Pstatus_values = data['Pstatus'].values
freetime_values = data['freetime'].values
famrel_values = data['famrel'].values



expected_health_values = []
predicted_health_values = []
accuracy_count = 0
for i in range(0,350):
    health_sim.input['goout'] = goout_values[i]
    health_sim.input['studytime'] = studytime_values[i]
    health_sim.input['Pstatus'] = Pstatus_values[i]
    health_sim.input['freetime'] = freetime_values[i]
    health_sim.input['famrel'] = famrel_values[i]

    data['health']
    health_sim.compute()

    health_value = health_sim.output['health']
    print("Input: goout={}, studytime={},  freetime={}, Pstatus={},famrel={}, ".format(
        goout_values[i], studytime_values[i], Pstatus_values[i],
        freetime_values[i], famrel_values[i]))
    print('Actual_value :', data['health'][i])
    expected_health_values.append(data['health'][i])
    predicted_health_values.append(health_value)
    if round(health_value) == round(data['health'][i]):
        accuracy_count += 1
    # Print the output value for 'health'
    print("Predicted value Output: health={}".format(health_value))
    print("---------------------------------------")
```

Fig 7.5 Accuracy Calculation

Interval Type 2 Fuzzy Logic Rule Based Learning Implementation:

This segment outlines the code related to the implementation of the Interval Type-2 fuzzy logic rule-based learning system. Fig 7.6 elaborates on the membership functions. In Fig 7.7, you can find the code for the rules in the Interval Type-2 fuzzy logic system. Finally, Fig 7.8 demonstrates the code for prediction and accuracy calculation in this context. The complete code is available in the supporting document under the filename "Interval Type 2 fuzzy logic system rule-based learning."

```
from pyit2fls import IT2FLS, IT2FS_Gaussian_UncertStd, IT2FS_plot, \
                     min_t_norm, max_s_norm, TR_plot, crisp
from numpy import linspace

domain = linspace(0.0, 5.0, 100)

very_bad_relationship = IT2FS_Gaussian_UncertStd(domain, [1, 0.6, 0.4, 0.8])
bad_relationship = IT2FS_Gaussian_UncertStd(domain, [2, 0.6, 0.4, 0.8])
average_relationship = IT2FS_Gaussian_UncertStd(domain, [3, 0.6, 0.4, 0.8])
good_relationship = IT2FS_Gaussian_UncertStd(domain, [4, 0.6, 0.4, 0.8])
very_good_relationship = IT2FS_Gaussian_UncertStd(domain, [5, 0.6, 0.4, 0.8])

very_less_free_time = IT2FS_Gaussian_UncertStd(domain, [1, 0.6, 0.4, 0.8])
less_free_time = IT2FS_Gaussian_UncertStd(domain, [2, 0.6, 0.4, 0.8])
average_free_time = IT2FS_Gaussian_UncertStd(domain, [3, 0.6, 0.4, 0.8])
good_free_time = IT2FS_Gaussian_UncertStd(domain, [4, 0.6, 0.4, 0.8])
max_free_time = IT2FS_Gaussian_UncertStd(domain, [5, 0.6, 0.4, 0.8])


very_less_study_time = IT2FS_Gaussian_UncertStd(domain, [1, 0.6, 0.4, 0.8])
less_study_time = IT2FS_Gaussian_UncertStd(domain, [2, 0.6, 0.4, 0.8])
average_study_time = IT2FS_Gaussian_UncertStd(domain, [3, 0.6, 0.4, 0.8])
good_study_time = IT2FS_Gaussian_UncertStd(domain, [4, 0.6, 0.4, 0.8])
max_study_time = IT2FS_Gaussian_UncertStd(domain, [5, 0.6, 0.4, 0.8])


very_less_goout_time = IT2FS_Gaussian_UncertStd(domain, [1, 0.6, 0.4, 0.8])
less_goout_time = IT2FS_Gaussian_UncertStd(domain, [2, 0.6, 0.4, 0.8])
average_goout_time = IT2FS_Gaussian_UncertStd(domain, [3, 0.6, 0.4, 0.8])
good_goout_time = IT2FS_Gaussian_UncertStd(domain, [4, 0.6, 0.4, 0.8])
max_goout_time = IT2FS_Gaussian_UncertStd(domain, [5, 0.6, 0.4, 0.8])
```

Fig 7.6 Membership Function

```
# Rules

myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", average_free_time),
                   ("study_time", less_study_time),("parents_status", divorced),
                   ("family_relationship", good_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", good_free_time),
                   ("study_time", less_study_time),("parents_status", together),
                   ("family_relationship", bad_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", good_free_time),
                   ("study_time", average_study_time),("parents_status", together),
                   ("family_relationship", average_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", good_free_time),
                   ("study_time", less_study_time),("parents_status", together),
                   ("family_relationship", good_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", average_free_time),
                   ("study_time", less_study_time),("parents_status", together),
                   ("family_relationship", good_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", good_goout_time), ("free_time", average_free_time),
                   ("study_time", good_study_time),("parents_status", together),
                   ("family_relationship", good_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", average_goout_time), ("free_time", average_free_time),
                   ("study_time", very_less_study_time),("parents_status", together),
                   ("family_relationship", average_relationship)], [("health", good)])
myIT2FLS.add_rule([("goout", average_goout_time), ("free_time", very_less_free_time),
                   ("study_time", less_study_time),("parents_status", together),
                   ("family_relationship", good_relationship)], [("health", good)])
```

Fig 7.7 Rules

```
for i in range(0,350):
    it2out, tr = myIT2FLS.evaluate({"goout": data['goout'][i], "free_time": data['freetime'][i],
                                "study_time": data['studytime'][i], "parents_status": data['Pstatus'][i],
                                "family_relationship": data['famrel'][i]
                               }, min_t_norm, max_s_norm, domain, method="Centroid", algorithm="KM")

    # Obtain the crisp output value
    crisp_output = crisp(tr["health"])
    print('Upper and Lower Limit :', tr)
    print('mental_health_of_the_student:', crisp_output)
    print('Actual_value:', data['health'][i])
    if round(crisp_output) == data['health'][i]:
        count = count + 1
    print()
print(count)
Accuracy = (count/350)*100
print('Accuracy :', round(Accuracy),'%')
```

Fig 7.8 Prediction and result

Fuzzy Logic Auto Rule Generation:

This section shows code for fuzzy logic auto rule generation for both type 1 fuzzy logic system with auto generation and type 2 fuzzy logic system with auto rule generation.

```
import skfuzzy as fuzz
from skfuzzy import control as ctrl

famrel_terms = ["very_bad_relationship", "bad_relationship", "average_relationship", "good_relationship", "very_good_relationshi
freetime_terms = ["very_less_freetime", "less_freetime", "average_freetime", "good_freetime", "max_freetime"]
studytime_terms = ["very_less_studytime", "less_studytime", "average_studytime", "good_studytime", "max_studytime"]
Pstatus_terms = ["together", "divorced"]
goout_terms = ["very_less_goout_time", "less_goout_time", "average_goout_time", "good_goout_time", "max_goout_time"]
# output
health_terms = ["bad", "average", "good"]

# Define a function to randomly generate a rule
def generate_rule():
    goout = random.choice(goout_terms)
    famrel = random.choice(famrel_terms)
    freetime = random.choice(freetime_terms)
    studytime = random.choice(studytime_terms)
    Pstatus = random.choice(Pstatus_terms)
    health = random.choice(health_terms)

    return "(goout['{}'] & freetime['{}'] & studytime['{}'] & Pstatus['{}'] & famrel['{}'], health['{}'])".format(goout, freetime

rules = [generate_rule() for _ in range()]
# print(rules)

for rule in rules:
    print(f"ctrl.Rule{rule}")
    print()
```

```
ctrl.Rule(goout['average_goout_time'] & freetime['less_freetime'] & studytime['less_studytime'] & Pstatus['together'] & famrel
['very_bad_relationship'], health['average'])

ctrl.Rule(goout['average_goout_time'] & freetime['good_freetime'] & studytime['max_studytime'] & Pstatus['together'] & famrel
['good_relationship'], health['bad'])
```

Fig 7.9 Type 1 Fuzzy Logic Auto Rule Generation

```
import random

family_relationship_terms = ["very_bad_relationship", "bad_relationship", "average_relationship", "good_relationship","very_good_
free_time_terms = ["very_less_free_time","less_free_time", "average_free_time", "good_free_time", "max_free_time"]
study_time_terms = ["very_less_study_time","less_study_time", "average_study_time", "good_study_time","max_study_time"]
parents_status_terms = ["together", "divorced"]
goout_terms = ['very_less_goout_time', 'less_goout_time','average_goout_time','good_goout_time','max_goout_time']
# output
health_terms = ["bad", "average", "good"]

# Define a function to randomly generate a rule
def generate_rule():
    goout_time = random.choice(goout_terms)
    family_relationship = random.choice(family_relationship_terms)
    free_time = random.choice(free_time_terms)
    study_time = random.choice(study_time_terms)
    parents_status = random.choice(parents_status_terms)


    health = random.choice(health_terms)

    return f"[('goout' ,{goout_time}), ('family_relationship' ,{family_relationship}), ('free_time' ,{free_time}) , ('study_time'


rules = [generate_rule() for _ in range(9)]
# print(rules)


for rule in (rules):
    print(f"myIT2FLS.add_rule({rule})")
    print()
```

Fig 7.10 Type 2 Fuzzy Logic Auto Rule Generation

Machine Learning Implementation:

This section details the implementation of the Machine Learning algorithms. Fig 7.11 represents the implementation of decision tree regression and Fig 7.12 represents the implementation of Random Forest regression and hyper parameter tuning. The complete code can be found in the supporting document under the filename "Machine Learning Implementation."

```python
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.ensemble import RandomForestRegressor
from sklearn.tree import DecisionTreeRegressor
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import mean_squared_error
from sklearn.model_selection import train_test_split,GridSearchCV, KFold
# load the dataset
data = pd.read_csv('fuzzy_use.csv')

# splitting data into train and test
X = data.drop('health', axis=1)
y = data['health']


X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Create a scaler object
scaler = StandardScaler()

# Fit the scaler to the training data
scaler.fit(X_train)

# Scale the input features
X_train_scaled = scaler.transform(X_train)
X_test_scaled = scaler.transform(X_test)

# Decision Tree Regressor
dt = DecisionTreeRegressor()
dt.fit(X_train_scaled, y_train)
```

Fig 7.11 Decision Tree Regressor

```python
# Random Forest with Grid Search CV

param_grid = {
     'n_estimators': [100, 200, 300, 400],
     'max_depth': [1, 5, 10, 15],
     'min_samples_split': [2, 5, 10]
}

grid_search = GridSearchCV(estimator=rf, param_grid=param_grid, cv=5)
grid_search.fit(X_train_scaled, y_train)

rf_with_param_tune = RandomForestRegressor(max_depth = 10,
                                           min_samples_split = 10,
                                           n_estimators = 300)

rf_with_param_tune.fit(X_train_scaled, y_train)

print('traning accuracy :', rf_with_param_tune.score(X_train_scaled, y_train)*100)
print('testing accuracy :', rf_with_param_tune.score(X_test_scaled, y_test)*100)
```

Fig 7.12 Random Forest Regressor