

AI-LMS: AI-Based Behavioral Modeling of User's Interaction With The Learning Management Systems

Abdulrahman Aly Awad

MSc. Thesis

May 2023

A thesis submitted to Khalifa University of Science and Technology in accordance with the requirements of the degree of MSc in Biomedical Engineering in the Department of Biomedical Engineering.



AI-LMS: AI-Based Behavioral Modeling of User's Interaction With The Learning Management Systems

by

Abdulrahman Aly Awad

A thesis submitted in partial fulfillment of the requirements for the degree of

MSc in Biomedical Engineering

at

Khalifa University

Thesis Committee

Prof. Leontios Hadjileontiadis (Main Advisor),
Khalifa University

Dr. Aamna AlShehhi (Co-Advisor),
Khalifa University

Prof. Mohamed Seghier (Co- Advisor),
Khalifa University

Dr. Asli Hassan (Co-advisor),
Khalifa University

Prof. Ahsan Khandoker (RSC Member1),
Khalifa University

Dr. Herbert Jelinek (RSC Member 2),
Khalifa University

May 2023

Abstract

Abdulrahman Aly Awad, "**AI-LMS: AI-Based Behavioral Modeling of User's Interaction With The Learning Management Systems**", M.Sc. Thesis, MSc in Biomedical Engineering, Department of Biomedical Engineering, Khalifa University of Science and Technology, United Arab Emirates, May 2023.

Learning management system became a necessary part of the pedagogical phenomenon especially after the COVID-19 pandemic. Modelling of LMS users is vital to learners, instructors, policy makers and stakeholders. QoI is an effective metric that can model the user's interaction with LMS which was proven across the different HEI levels. This work develops several deep learning models, mainly ANN to represent the expert knowledge encapsulated in the fuzzy logic system of Fuzzy-QoI allowing for generalizability and possibility of interpretation using SHAP. Deep models were interpreted, and force plots were used to provide feedback to users in order to improve their QoI. Additionally, two new datasets covering the quarantine period were collected from Moodle in Khalifa University, annotated, and prepared based on the work of Fuzzy-QoI. External validation using the new datasets provided insights into the advantages and disadvantages of the application of deep learning models in education. The feedback received from the models based on interpretation can improve both learners' and instructors' pedagogical experience while providing relevant insight to stakeholders and policymakers.

Indexing Terms: Learning Management Systems, Quality of Interaction, Fuzzy Logic, High Education Institutes, Deep Learning, SHAP.

Acknowledgement

All praise is to Allah, praise in abundance, good and blessed. It cannot be compensated for, nor can it be left, nor can it be done without.

I would like to express my gratitude to my supervisor Prof. Leontios Hadjileontiadis for his support, guidance, and patience. I am grateful for his constant support and mentorship that was instrumental when working on this thesis.

I also would like to thank Dr. Aamna AlShehhi for her guidance, support, and mentorship. Her input, feedback and tutoring provided me with needed knowledge for this thesis work and expertise that can be applied elsewhere.

I want to give my thanks to the research committee, Prof. Mohamed Seghier, Dr. Asli Hassan, Prof. Ahsan Khandoker, and Dr. Herbert for their constructive feedback that helped me improve on this work.

I would like to thank Dr. Sofia Dias for providing the original work of Fuzzy-QoI including DB1 database and MATLAB codes for the fuzzy logic system.

I would also like to thank Mr. Abdul Rafi Jaffarullah Khan for his support and help in facilitating the data acquisition process and Mr. Shaju Bharuddin & members of the IT department for providing the Moodle SQL data.

Finally, I would like to thank my parents for their unconditional support, love, and guidance.

Declaration and Copyright

Declaration

I declare that the work in this thesis was carried out in accordance with the regulations of Khalifa University of Science and Technology. The work is entirely my own except where indicated by special reference in the text. Any views expressed in the thesis are those of the author and in no way represent those of Khalifa University of Science and Technology. No part of the thesis has been presented to any other university for any degree.

Author Name: Abdulrahman Aly Awad

Author Signature: 

Date: 4 / May / 2023

Copyright ©

No part of this thesis may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, without prior written permission of the author. The thesis may be made available for consultation in Khalifa University of Science and Technology Library and for inter-library lending for use in another library and may be copied in full or in part for any bona fide library or research worker, on the understanding that users are made aware of their obligations under copyright, i.e. that no quotation and no information derived from it may be published without the author's prior consent.

Contents

Abstract iii

Acknowledgement iv

Declaration and Copyright v

Contents vi

List of Figures viii

List of Tables ix

List of Abbreviations x

CHAPTER 1 INTRODUCTION 1

1.1 Introduction 1
1.2 Research Statement 2
1.3 Objectives 3
1.4 Thesis Structure 3

CHAPTER 2 LITERATURE REVIEW 5

2.1 Introduction 5
2.2 LMS User Modeling 6
2.3 AI applications in LMS 9

CHAPTER 3 DATASETS 11

3.1 Introduction 11
3.2 DB1 Dataset 11
3.3 Development of KUQ2020 and KUQ2021 Datasets 12
3.4 Scaling of Data and Application of Fuzzy QoI 15

CHAPTER 4 DEEP MODELS FOR 12 ACTION PARAMETERS 19

4.1 Introduction 19
4.2 Methodology 19
4.2.1 Data Preprocessing 19
4.2.2 Developing Deep Learning Models (90:10 split) 20
4.2.3 Developing Deep Learning Models (75:25 split) 22
4.2.4 Evaluation of Models 22
4.3 Results 22
4.3.1 Results for 90:10 split. 24
4.3.2 Results for 75:25 split 24
4.4 Discussion 24
4.5 Conclusion 25

CHAPTER 5 INTERPRETATION OF USERS' BEHAVIOR ON LMS 27

5.1 Introduction 27
5.2 Methodology 27
5.3 Deep Model selection 28
5.4 Model Interpretation 28
5.5 Results and Discussion 29

5.6 Behavioral Modulation of LMS Users 35
5.7 Conclusion 36

CHAPTER 6 ON MODELLING QOI USING DEEP LEARNING 38

6.1 Introduction 38
6.2 Methodology 38
6.2.1 Data Splitting 39
6.2.2 Deep Model selection 39
6.2.3 External Validation 41
6.3 Results 42
6.4 Discussion 47
6.5 Implementation of AI: LMS (A proposition) 50
6.6 Conclusion 53

CHAPTER 7 EXPLORATION OF DEEP FUZZY SYSTEMS 55

7.1 Introduction 55
7.2 Literature Review 56
7.3 Methodology, Results and Discussion 59
7.3.1 Trail 1 With Deep Fuzzy Logic System 60
7.3.2 Trial 2 Using Kernel MLP 61
7.3.3 Trial 3 With Parallel Deep Fuzzy Model 62
7.4 Fuzzification of QoI Deep Models 68
7.4.1 Introduction 68
7.4.2 Methodology 69
7.4.3 Results and Discussion 71
7.5 Conclusion 71

CHAPTER 8 CONCLUSION 73

8.1 Conclusion 73
8.2 Future Work 74

References 75

Appendices 82

Appendix A. Assigned Interactions for New Datasets 82
Appendix B. MATLAB Code 91
Appendix C. Python Code: Prepare data for Fuzzy QoI 93
Appendix D. Python Code: Tunning for 14 parameters 95
Appendix E. Python Code: Cross validation 14 parameters 98

List of Figures

Figure 1 Fuzzy QoI system 18
Figure 2 Correlation heat map 23
Figure 3 Students' top 21 (from 112) factors affecting QoI results. 30
Figure 4 Professors' top 21(from 112) factors affecting QoI results. 30
Figure 5 Students' action parameters in order of affecting QoI results. 31
Figure 6 Professors' action parameters in order of affecting QoI results. 31
Figure 7 Student with low QoI. 32
Figure 8 Student with low QoI after editing. 32
Figure 9 Student with high QoI. 33
Figure 10 Student with high QoI after editing 33
Figure 11 Professor with low QoI. 33
Figure 12 Professor with low QoI after editing. 33
Figure 13 Professor with high QoI. 34
Figure 14 Professor with high QoI after editing 34
Figure 15. Scaled Exponential Linear Unit (SeLU). 40
Figure. 16. 12 Parameters models a) Tuned, b)Tapered. 43
Figure. 17. 110 Metrics models a) Tuned, b)Tapered. 44
Figure 18 Sequential deep fuzzy system. 58
Figure 19 Parallel deep fuzzy system. 58
Figure 20 a) Integrated deep fuzzy system, b) stacked generalization model. 59
Figure 21 Deep fuzzy model trial 1. 60
Figure 22 Feature map from MNIST (left) and CIFAR-10 (right). 61
Figure 23 Kernel MLP Architecture 63
Figure 24 Feature Map of Kernel MLP 65
Figure 25 Blank feature map of parallel deep fuzzy model 65
Figure 26 CIFAR 10 first 20 testing images 66
Figure 27 Fuzzy Gaussian membership functions (negative low to positive high) 66
Figure 28 Parallel Deep Fuzzy Model 67
Figure 29: Bias in model performance. 68
Figure 30: Sine wave. 69
Figure 31: Sine waves with different periods. 70
Figure 32: Sine waves with different phases. 70

List of Tables

Table 1 Dataset Periods and Breaks.....15
Table 2. Results for 90:10 split.....24
Table 3. Results for 75:25 split.....24
Table 4 Deep Model Parameters.....28
Table 5 Results for 110 parameters deep models. 29
Table 6. Bayesian optimization tested parameters.....41
Table 7 .Datasets descriptions.....42
Table 8. Datasets statistics. 42
Table 9. Model performance for tuned model (14 parameters), case 1. 44
Table 10. Model performance for tapered model (14 parameters), case 1. 45
Table 11. Model performance for tuned model (14 parameters), case 2. 45
Table 12. Model performance for tapered model (14 parameters), case 2. 45
Table 13. Model performance for tuned model (14 parameters), case 3. 45
Table 14. Model performance for tapered model (14 parameters), case 3. 45
Table 15. Model performance for tuned model (112 parameters) 46
Table 16. Model performance for tapered model (112 parameters) 46
Table 17 Results for fuzzification of deep models. 71

List of Abbreviations

1D : 1 Dimensional.

AI : Artificial Intelligence.

ANN : Artificial Neural Networks.

AS : Affective State.

CCTV : Closed Circuit Television.

CI : Confidence Interval.

CM : Cognitive Mapping.

CNN : Convolutional Neural Networks.

CTL : Center of Teaching and Learning.

ET : Engagement Time.

F2F : Face to Face.

FCM : Fuzzy Cognitive Mapping.

FIS : Fuzzy Inference System.

HEI : Higher Education Institutes.

IQR : Inter-quartile Range.

IT : Information Technology.

KU : Khalifa University.

LMS : Learning Management Systems.

LSTM : Long-short Term Memory.

MAE : Mean Absolute Error.

MLP : Multi-Layer Perceptron.

MOOC: Massive Open Online Classrooms.

MSE : Mean Squared Error.

OBS : Observations

PhD : Doctor of Philosophy

QoC : Quality of Collaboration.

QoI : Quality of Interaction.

R^2 : Coefficient of Determination.

RMSE : Root Mean Squared Error.

RNN : Recurrent Neural Networks.

SHAP : SHapley Additive exPlanations.

TCNN : Temporal Convolutional Neural Networks.

TP : Time Period.

TSK : Takeo-Sugeno-Kang.

UM : User Model.

CHAPTER 1 INTRODUCTION

1.1 Introduction

Learning management systems (LMS) are becoming increasingly integrated by higher education institutions (HEI) [1]. LMS can be defined as a system that serves several functions concurrently from course content creation, user enrollment and registration, student grading and course delivery through virtual classrooms among other functions [2]. LMS have many free and commercial brands that are available for deployment by HEI [3]. LMSs played a crucial role as main method of teaching and delivery during the COVID-19 pandemic [4] as they provided flexibility to educators and students under lockdown protocols. Though LMSs are considered relevant by today's technological standards the question of acceptance of such technologies by educators and learners is still a point of interest [3]. With different types of LMSs between those employed by HEI and by massive open online classroom (MOOCs) the need for user modeling or learner modeling becomes a focal point for all parties involved [5]. Modeling in LMS can be done on different levels pedagogical, domain or learner's model [6]. Pedagogical and domain pertain to perceived expertise or knowledge which can be assessed effectively using appropriate tests [6]. And so, the focus is on the learner's model specifically as it not just pertains to user's actions/interactions within LMS environment but also the user's affective state [6] [7]. Some of these can be evaluated separately as LMSs like Moodle and BlackBoard offer analytics that can give a general idea about user's usage and total interactions. Although, such tools are available, applying the same metrics across different LMSs (domains) provides some difficulty mainly in the domain shift and in relevance of such metrics across the different brands with the pedagogical phenomenon as the center of focus[5]. Additionally, artificial intelligence (AI) is also being integrated with LMSs to serve different purposes aiming to improve the overall experience [6,7]. This raises another concern which is the ethical

implications of deploying such systems and biases that can arise due to data collection or model development [8].

1.2 Research Statement

In this vein, this work tries to tackle the issues of applying AI models to LMS user data. As aforementioned, with the variance in brands, HEI and even types of courses like MOOCs the question of the feasibility of application of such models across different domains (whether it is different LMS brands, HEIs or MOOCs) and subsequently the question of LMS user modeling are raised. As any data collected from LMS users should aim to model the users in some manner to solve some problem or provide some level of analysis. To account for user modelling we follow on the work of Fuzzy-QoI [9] as part of a larger approach that aims to model LMS users from blended, collaborative, and affective dimensions [7,10]. Thus, we try to substitute the expert knowledge crystalized in the Fuzzy-QoI system with deep learning models that could allow for generalizability, interpretability, and ease of deployment, in addition to facilitating new forms of application like feedback and recommendation systems. With our target being to provide better educational experiences to learners and educators while also providing insights in the workings of LMS to policymakers and stakeholders. So, this thesis tries to answer the following research questions:

RQ1: Can we develop a deep model that replaces the expert knowledge system of Fuzzy-QoI?

RQ2: Can the deep model be interpreted to better understand/ modulate users' behavior?

RQ3: Is such model able to generalize across domains (universities or LMS brands)?

RQ4: How can we reduce bias/uncertainty or improve performance of the model?

1.3 Objectives

The main goal of this thesis is to develop deep learning models to substitute for the Fuzzy-QoI as part of the larger approach of AI-LMS that implements all the components of the A/B/C teach framework [7,10], and additionally the development of new datasets based on data extracted from Khalifa University (KU) LMS databases. These goals could be achieved through the following deliverables:

- Literature review of works related to LMS user modelling and applications of AI in education.
- Development of deep learning models for the dataset used in the original work of Fuzzy-QoI.
- Development and annotation of new datasets based on Khalifa University LMS user data.
- External Validation of deep learning models against the new datasets.
- Investigating the effects of deep fuzzy models on dealing with uncertainty and noise to improve overall model performance.

1.4 Thesis Structure

The thesis is structured in the following order:

Chapter 2 gives a background into the LMS user modeling, Fuzzy-QoI system and applications of AI in education.

Chapter 3 highlights the datasets used for this work and the development of two new datasets (KUQ2020, KUQ20201), covering aspects of appropriating interactions to fit the Fuzzy-QoI, discussing methods used to derive the QoI value, and justifications of those methods.

Chapter 4 discusses early trials in application of deep models to DB1 dataset, this includes methodology, results, discussion and finally a conclusion of the applied methods and its relevance to HEI education.

Chapter 5 highlights the development of deep learning models and interpretation of those models in order to modulate LMS users' behavior, similarly it includes methodology, results and discussion, recommendation of feedback system from a behavioral standpoint, and finally conclusion.

Chapter 6 discusses the development of deep learning models and external validation using KUQ2020 and KUQ2021 datasets, this chapter focuses more on improving the performance of previously developed models while also discussing effects of noise, bias and domain shift on model performance providing suggestions for improving overall outcomes.

Chapter 7 discusses the experimental development of deep fuzzy models, it includes background (literature review), methodology, a combined results and discussion and finally conclusion.

Chapter 8 is thesis contribution, conclusion, and future work.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

Learning Management Systems (LMSs) have gained an increased adoption rate by Higher Education Institutions (HEIs) because of an increased demand for new emerging technological and pedagogical approaches/practices [11]Additionally, with the advent of Covid-19 pandemic, HEIs were motived to support blended learning (hybrid) fully or partially [12,13]. This meant higher dependence on LMSs for potential sharing of knowledge, learning material creation, student engagement and evaluation/assessment, especially for HEIs switching to fully online modalities during the Covid-19 pandemic [14]. LMSs have many advantages that were further accentuated during the Covid-19 pandemic, for instance, they allow engagement of larger number of learners compared to face to face (F2F) modalities and offer several types of approaches (blended, mobile, game, etc.) [15]. The aforementioned aspects bring up the need to re-evaluate the pedagogical processes and practices under these "new" online learning environments/conditions. In this vein, Ponce et al. [16] suggests that an assessment of educational research should follow a model that allows it to generate *generalizable data*. LMSs have a plethora of brands, they are not limited to a single type of software, and they expand to all forms of digital media (videos, articles, etc.) [17], and to fulfill that condition a more generalized approach to model learner/user's Quality of Interaction in online learning environment is required. Authors in [15] suggest that for any learner model three important questions should be answered, namely:

"(i) What are the characteristics of the user we want to model? (ii) How we model them? and (iii) How we use the user model?" [15]. Thus, an indicator that can model the users' engagement with the LMS based on their interactions within the online environment and,

possibly, provide feedback to users as to act and/or gain insight about their usage/interactions, could answer all the three questions.

2.2 LMS User Modeling

Learner or user modeling while using a LMS is a point of interest not just to educators, but also to stakeholders and policymakers. The question of effectiveness of LMS in HEIs can be difficult to answer when a model to quantify the users' engagement with the system is not in place. Most LMSs collect data about the user's interaction in the form of views, access, forum discussions etc., and using those; they offer recommendations to increase system access/usage based on the average usage of same cohort. As such, this subsection will review several approaches to model learner/student interactions within an online learning environment in order to evaluate the effectiveness of knowledge acquisition or suggest a better online learning approach.

Colace et al. [18], proposed a system for adaptive learning that consists of four main components. Two of its components relate to student modeling; one related to user model (UM) and the other concerning observations (OBS). UM describes the knowledge level of the user and learning style, while the OBS is more concerned with the interactions with the system (e.g., clicks, views, time spent). The approach focused more on the evaluation aspect of learning (level of knowledge) to improve or adapt the learning experience to the user. Another approach [19]used four predictors (online learning self-efficacy, learner interactions with content, instructors, and other learners) to estimate student satisfaction and perceived knowledge level. The method used a survey to collect the data and found that these factors were relevant to the learning process outcomes. The use of survey questionnaire questions the objectivity of the method, as it is reliant on student perceived level of knowledge and satisfaction, which was noted by the author. Hussain et al. [20], predicted the student engagement in an online course

using the engagement time, clicks and student's scores to predict students with low engagement during the 1st course assessment allowing for interventions before the final evaluation (exam). The focus of the model was on evaluation, and the interactions (i.e., clicks, time) were used as predictors of engagement, which is relative to the online course environment but can be difficult to adapt in a blended learning environment. Authors in [21]; studied the student engagement in a massive open online course (MOOC) and its relation to the course completion. The study used a survey along with a model of the user adopted on data composed of the number of viewed videos and posts made. The study found that student engagement on forums has significant impact on course completion versus videos viewed. Generalization of the study approach to different online learning environments could be considered as the main limitation.

From the discussed literature it can be noted that the main factor utilized in most studies is the evaluation or assessment of the student/learner. In a physical class or F2F, the assessment of student interactions with the content, instructor, other learners and the environment could be achieved without relying on student's evaluative scores of their quality of interaction*(i.e., effective utilization of spaces/resources, and quality of engagement with the instructors*), which brings the question of which parameters are required to evaluate the user's quality of engagement with the learning system regardless of their final score or assessment. For example, an instructor using a LMS will not have a score assessment but their engagement with the learning system still needs to be evaluated. A/B/C-Teach framework suggested by [7,10]models the user based on three aspects, in particular: (i) quality of interaction (QoI), which evaluates user engagement with the LMS purely on direct interactions with the system, (ii) quality of collaboration (QoC), which investigates user interaction with other users [22], and (iii) affective state(AS) which evaluates the users' emotions during engagements/interactions within a LMS. The framework uses a multidimensional approach to model LMS users. The QoI metric was introduced with Fuzzy-QoI [9], a fuzzy logic based

expert knowledge system. The system took all user interactions on LMS of which were 110 unique interactions. They were then categorized into three main categories relevant to the pedagogical process: viewing, adding, and altering of LMS content. Each category had four subcategories forming 12 action parameters. Five fuzzy inference systems were developed with over 600 rules that receive these inputs in form of aggregated user frequencies on weekly basis to account for the sparsity of interactions by users. Two additional parameters are then added which are the time period (TP) referring to the week of the academic year and engagement time (ET) which is the aggregated sum of time spent per week. The model was first developed on HEI level (macro), then further applications across different domains and levels were done [23-29]. The QoI metric relevance is due to its ability to capture user's interactions whether they are learners or instructors while also capturing the change in trend (TP) and reflecting it in a manner that is similar to human cognition (fuzzy concept). By applying fuzzy cognitive mapping (FCM-QoI) to predict the value of QoI [29]; a proof of concept for the deployment of a machine learning system was done. Fuzzy cognitive maps (FCM) are a fuzzy graph structure [30] that links concepts and knowledge in form of fuzzy causality between fuzzy concepts. This representation using FCM-QoI allows for interpretation of the large numbers of metrics (110) when categorized into 12 action parameters or 14 counting TP and ET. The interpretation links these fuzzy concepts with the outcome of QoI metric, which is relevant to users, educators, and policy makers alike. FCM-QoI linked the 14 parameters to QoI, which leaves the question of how such concepts are linked when taken from the perspective of the 110 raw metrics. We can argue for increased generalizability and ease of integration by introducing QoI into artificial intelligence (AI) domain. But, by pushing the development of QoI into the artificial intelligence (AI) domain it compromises the interpretability while converting it from a theory-driven to a data-driven approach that could allow for creation of direct links between 110 raw metrics and the QoI. The 110 metrics are separated from the outcome of the QoI by

categorization, five fuzzy inference blocks and over 600 rules making direct correlation and linking between QoI difficult. There would be advantages to using AI to model the QoI metric, but it can also raise other concerns which are discussed in next subsection.

2.3 AI applications in LMS

AI applications in LMS vary in their target objective; efforts for adaptive learning, recommendation and feedback systems are some of the examples [31]. Adaptive learning aims to provide a user tailored materials to advance their level of knowledge based on user's own competencies [18,32]. In addition to adaptive learning, other venues like chatbots and virtual helpers are being implemented to offer a better learning experience [33]. Recommender systems [34-37]could be considered an earlier step preceding the application of an adaptive learning path. Both systems as noted would require some form of user modeling [34] before any kind of assessment to evaluate user knowledge is considered. Big data and advances in algorithms could allow for development of better data-driven user modeling AI models [38]. But that raises ethical issues related to implementation of AI in educational environment [39]. Main concerns would be privacy, bias, and fairness [39]. Another consideration for any AI application is interpretability [40,41], as compared to machine learning models that are based on expert knowledge and easily interpretable, interpretation of AI models can vary based on developed model and available tools. It is also noted that in AI models developed for education; the developers are more concerned with the technical aspects of the model than its effect on the pedagogical experience and how does it fit within such environment [31]. Thus, we can conclude that for effective AI application in education; generalizability, data privacy, elimination of bias, ensuring fairness and interpretability are sought after characteristics that need to be fulfilled. So, to account for the expert knowledge before applying AI deep learning model we followed on the work of Fuzzy-QoI which was developed using fuzzy logic and is a theory-based model that models LMS users within a blended and collaborative environment.

In summary, this work applies deep learning models to predict QoI metric across two different levels 14 and 112 parameters (including TP and ET). Three datasets from two different universities with a decade of time between are used for training and validation in order to check for effective model generalization and bias reduction. And so, this thesis main target is to develop deep learning models that replace the Fuzzy-QoI to predict the QoI metric, using data across two different HEIs and also different sources from same LMS (Moodle). The developed models are to be interpreted in order to provide feedback to users to better enhance the users learning experience and interactions with the LMS. We also discuss the effects of noise, bias, and domain shift on the performance of the models and finally we suggest several approaches to deal with those issues.

CHAPTER 3 DATASETS

3.1 Introduction

This chapter discusses the preparation of two datasets KUQ2020, KUQ2021 extracted from Moodle LMS. As we mentioned in previous chapter, in order to develop our deep models, data is an essential component in order to train our models as the QoI prediction is a supervised regression task. Thus, we used the main dataset used for development of both Fuzzy-QoI and FCM-QoI (DB1) then we developed two new datasets in order to further validate our working models in efforts to answer the first and third research questions which deal with development of AI models and their deployment across different domains. We explain the DB1 dataset and showcase the step-by-step development of the two new datasets KUQ2020 and KUQ2021 in the next subsections. It should be noted that tables for descriptions and statistics for all three datasets will be included in chapter 6 to allow for easier following of results and discussion.

3.2 DB1 Dataset

The dataset DB1 consists of raw data of the LMS use from a total of 1037 students and 75 professors, with a duration spanning of 51 weeks across two academic semesters (2009/2010) in Faculdade de Motricidade Humana, Portugal. This data was collected as the Moodle LMS was first deployed and was done retroactively so students and lecturers were not aware of a study being done that might cause any performance bias. This dataset was collected from five different subjects (HEI or macro level); Sport sciences, ergonomics, dance, sport management, and psychomotor rehabilitation, covering a range of both theoretical and practical courses. The users' weekly interactions are aggregated, then scaled to [0,1] before being given as input to Fuzzy-QoI system to output QoI values. The input features are made into two separate datasets, one taking the 110 raw metrics and the other using the 12 action parameters.

Both cases have last two features as TP and ET. These two configurations of the same dataset allow for training models that symbolize different stages of the Fuzzy-QoI system.

3.3 Development of KUQ2020 and KUQ2021 Datasets

A New dataset was acquired from the IT department at Khalifa University from the Moodle backend SQL database. The data collection for this set was started at 17th March 2020 until present, though the relevant data are until August 2021, which is the date that Khalifa University switched completely to Blackboard as the de facto learning management system.

The dataset overall (with records until data was acquired) have over 18 million interactions, with 5050 confirmed students and possibly around 406 professors. The definition of professor status was done based on combination of unique actions like creating course materials and grading, though it is still uncertain what is the exact number due to presence of teaching assistants. This caused an overlap between teaching assistants and professors' data, though some of them were confirmed after contacting the center of teaching and learning for further elaboration. This left over 1300 user entries who are likely to be graduate students. The SQL database contains a column 'edulvl' which may refer to education level, that was found to be 0 for most actions that are not related to course creation and grading allowing for easier limitation of student status. The column gives 2 for actions that are participatory, 1 for those affecting education process and 0 for any other. The main issue with using this column as identifier is that all actions on level of viewing even when done by identified professor users still give the 0 flag. As such, cooperation with IT department and CTL is paramount to correctly annotate the database. Another issue is that course identification numbers are used to identify both subject and section with no relevant connection between course code and course ID. The database contains around 9000 unique course IDs, which may not be representative of the reality, suggesting that around 9000 were taught over three major semesters.

The SQL got its own column that is similar in function to the three main categories of the Fuzzy-QoI (view, add, alter), which is referred to as ‘crud’ short for first letter of the 4 words; create, read, update, and delete. This system makes it easier to link the interaction to its major category though some unique cases that will be mentioned later may not be easy to classify.

As this is the backend of the Moodle system; matching of metric names as identified in the original Fuzzy-QoI is not straight forward and some form of domain shift to match those 110 metrics to the corresponding interactions is needed. Most of the 110 metrics were easily found as shown in Appendix and extra 130 metrics were found that were not part of the original 110(a total of 334 unique interactions were annotated). It should be noted that due to difference in Moodle software development across a decade (~ 10 years) some of the modules/plugins that were part of the Moodle in 2009/2010 are no longer available in the classic version but can be added as an extra plugin; examples of those removed from KU version are (Flashchat, Journal). To simplify the new metrics no new numbers were given beyond 110 but each were matched to one of the 12 action parameters (view, add, alter); focusing on whether the interaction allows user to view the system, add content or update and edit content. One major difference between SQL and Moodle logs is that reports generated by user (instructors) are identified as viewing and there is a subsequent action named ‘downloaded’, which refers to user downloading the report logs instead of viewing as html.

Another difference between frontend and backend is that for backend metrics a single number could refer to a group of metrics, for example viewing course module has two parts: viewing course module instance list and then viewing course module. The object of action for both is the core of the course, similar instance is present for all different modules like quizzes, scorm, hotpot…etc. As such, calling for distinct cases that match each metric was done

manually and matching with proper metric and identifying working search terms was noted as to ease the re-evaluation process if needed.

After matching each interaction with its metric and then matching rest to the 12 action parameters, certain interactions were left without allocation; majorly the action of login to the system. Another interaction that was disregarded was ‘called’ which targets object ‘meeting’, this interaction alone compromised over 3 million interactions and was assumed to be the virtual classroom given that data collection period coincides with quarantine/ lockdown period of Covid-19 in UAE. Although disregarded, the interactions still exist in the database and could be further considered for evaluation, though considering them as a single parameter could be difficult as most of these sessions had both viewing and adding in form of recording. The aspect of inter-user interaction during these sessions is also a point of interest and they might be considered for evaluation in QoC assessment.

The time period for this dataset was three major semesters; Spring 2020, Fall 2020 and Spring 2021. The dataset was divided into two parts with Spring 2020 as 1st part (KUQ2020) and the rest as the 2nd part (KUQ2021). The reason for the division is to allow for ease of application of the Fuzzy-QoI model and match the original work which was around 51 weeks. The total period for each part along with holidays and breaks is shown in Table 1. One major consideration is the need for aggregation per week for FuzzyQoI values as the user interactions are sparse when considering the full scope of the period. As such, the period was converted from timestamp format to datetime format and allocated a separate column, then every week or 604800 seconds was considered one week. Based on calendar counting started earlier than 17th March (starts at Monday 00:00) of every week, and so by adding the above value all interactions lying in between were assigned the same week number or time period in case of 17th March 2020 (week 30). Part 1(KUQ2020) was allocated numbers with 1000 added to them for example 1030, 1031 ...etc., as to help separate from the 2nd part (KUQ2021) weeks and

prevent overlap between the users across the two academic years. As the universe of discourse for the Fuzzy-QoI is from [0 1], scaling of the values is needed and division by max value for each part adjusts the column values to be within universe of discourse. The change of values across two academic years may not be representative and may differ greatly from the original Fuzzy-QoI work, hence the previous separation was considered.

Table 1 Dataset Periods and Breaks

KUQ2020		KUQ2021	
Instance	**Week**	**Instance**	**Week**
Fall 2019	W1 - W17	Fall 2020-1	W1-W14
Break	W18-W20	National Day	W15
Spring 2020(not included)	W21-W29	Fall 2020-2	W16-18
Spring 2020-1	W30-31	Break	W19-W21
Spring Break	W32	Spring 2021-1	W22 - W29
Spring 2020-2	W33-W39	Midterm	W30-31
Eid Break	W40	Spring Break	W32
Summer 2020	W40-W47	Spring 2021-2	W33-40
Break	W48-W51	Break	W41
_	_	Summer 2021	W42 - W48
_	_	Break	W49-W52

Engagement time was assigned as the difference in time between subsequent interactions done by same user. The thresholding of interaction value to 1000 per interaction was considered as some interactions had follow ups that were days apart hugely inflating the engagement time value. After the difference is calculated; some cases like login had zero time for engagement time and no correction for values were considered as aggregation of weekly interactions will accumulate engagement time and the situation where only engagement time with zero values occurs during any week was considered unlikely. Another consideration is that interactions that had time but were not assigned a value for the 12 actions parameters were dropped from the application of Fuzzy-QoI. Also, no scaling or thresholding was considered besides the initial 1000 seconds per interaction.

3.4 Scaling of Data and Application of Fuzzy QoI

In order to receive outputs for all the datasets the Fuzzy-QoI was applied to all datasets separately, the Fuzzy-QoI receives inputs as aggregated interactions frequencies (from 12

action parameters) per user in every time period. The scaling of the data needed to be done to account for the universe of discourse [0,1]. Three approaches could be considered; first is to min-max scale based on global max value for each column irrelevant of user or week which basically means that the highest possible value for that feature becomes the global maximum. Second approach is to min-max scale based on each user individually, so for each user the maximum number of interactions per column or engagement time becomes a scaler for that user, it can help offset effect of outliers across users, meaning users with low interaction against those with high interaction throughout the academic year. This can suffer effect of outliers, if the user has few instances with high value against mostly low values and vice versa. Third approach is to consider each week for all users as a collection and find the min-max scale for that grouping, in essence it would be highlighting the effect of seasonality if such effect is present as weeks with more user interactions overall will be reflected across the whole group of users and the frequnecy across low interaction periods may suffer from outlier effects. A way to account for the outlier effect is to threshold the values of each column to a fixed maximum, essentially ensuring that column will have a fixed universe of discourse across the whole period.

In this study, the approach taken is to consider the weekly interactions across groups of users and apply a double layer of scalers; the first is a robust scaler which uses the median along with a inter quartile range (IQR) [42] to adjust the values (default 25^{th} to 75^{th}) thus removing outliers during scaling while maintaining the distribution. This is followed by a min-max scaler of the adjusted values ensuring all columns lie within the universe of discourse. The values for the time period (TP) are adjusted manually as they are usually a fixed number (week of the academic year) for each group. So, after scaling all values, the original time period values are scaled based on maximum time period (51). By using this approach, the effect of outliers is minimized but such dataset suffers from information leakage if they are considered for a

deep learning model unless the splitting is done across the periods i.e., certain periods are taken as a whole for testing set. In order to apply deep models on the dataset in future work, the aggregation and scaling by user maybe considered before the scaling is applied, by removing certain users from the training set, this would allow the training set to cover all academic year period while preventing information leakage from scaling. At the end of this step two unique datasets are formed one for Spring 2020, and the other for the 2020/2021 academic year containing 77670 and 172227 weekly entries, across 3377 student users. The lack of confidence in assigning the flag to professors and differentiating them from teaching assistants/ PhD students is the main reason for disregarding those users until further affirmation is acquired.

Fuzzy-QoI model is applied to the scaled data and the received QoI output is used as the main outcome or ground truth for the deep learning model training and testing, i.e., the model predicts the value of QoI per weekly interaction input. The data transformation was applied using Python library [43] and scikit learn[44], while the Fuzzy QoI was applied using MATLAB [45]. Equations (1, 2) are for robust scaling and min-max scaling respectively.

$$x^{\circ} = \frac{x_i - x_{median}}{IQR} \quad (1)$$

$$x^{\circ} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (2)$$

As aforementioned, the 110 metrics are categorized into 12 action parameters for dimensionality reduction, these 12 parameters are then aggregated on weekly basis, and the scaled output becomes the input of the first 3 FIS blocks. FIS1 for view, FIS2 for add, and FIS3

for alter. The output of these three blocks is the input to 4th Block FIS4 which produces the action output. It should be mentioned that all four blocks use three trapezoid membership functions. The next step is to combine action with TP and ET as input to the last block FIS5 (five membership functions) which produces the QoI metric output. It should be noted that the system accepts single input instances (1D vectors), then the output is stacked into a matrix (table). Figure 1 shows the Fuzzy-QoI system blocks.

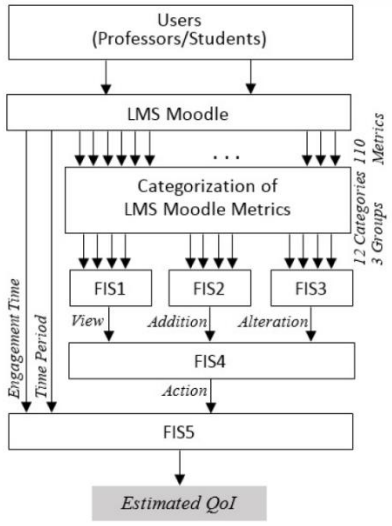


Figure 1 Fuzzy QoI system [9]

CHAPTER 4 DEEP MODELS FOR 12 ACTION PARAMETERS

4.1 Introduction

As mentioned in the literature review, the issue of adapting AI models into any LMS requires certain precautions in addition to the need of a LMS user model. We circumvent the model requirement by relying on the previous work of Fuzzy-QoI and focus on the development of deep models that would substitute the working components of the system at different levels (12 parameters and 110 metrics). In an effort to partly answer the first research question which aims to replace the expert knowledge with a deep learning model and also explore if the model can be somewhat be interpreted based on mechanism of output calculation (depending on model architecture to draw analysis), we developed several deep models that take the 12 action parameters with addition to TP and ET in order to predict the QoI value. We try to understand the behavior of the model and the effect of hyper parameterization on the output from the angle of inter-feature and intra-feautre relation/ correlation to derive analysis that could be of use to learners and educators alike.

4.2 Methodology

In order to model the QoI using a deep neural network; the dataset DB1 was used (see subsection 3.2). Overall, this section covers the methodology of the development of the proposed deep learning models upon DB1.

4.2.1 Data Preprocessing

The Fuzzy QoI, derived 12 action parameters based on the 110 metrics extracted from the Moodle LMS. These 12 action parameters are divided into three main categories (4 each); view, add and alter. The three categories group the 110 parameters based on users' interactions

with the LMS. For example, the view action parameters group any user interaction that pertains to viewing LMS content, whether it is a course module, quiz, post and/or grades. Similarly, the add and alter action parameters are concerned with user interactions that add new content (quizzes, posts, and course material) and edit existing content (editing, updating, and deleting) respectively. As such, the 110 parameters are grouped into 12 action parameters resulting in dimensionality reduction. The proposed models used the 12 parameters along with time period (which refers to the week number in the academic year) and total engagement time. To avoid bias from the engagement time; for example, users leaving the screen open while away; a threshold of around 2 hours per day or 14 hours per week was put. The daily users' action parameters and engagement time were aggregated as the QoI output was derived on a weekly basis. The data splitting was done in 90:10 and 75:25, the latter to match the FCM-QoI split. In both instances, the data was scaled to range (0 -1) without robust scaling. The input data is represented as a 1D vector having the 12 action parameters (frequency/week), time period and total engagement time per week. The output consists of the weekly QoI indicator value which is between (0-1).

4.2.2 Developing Deep Learning Models (90:10 split)

The performance of four different models was tested; the parameters of loss function (mean squared error (MSE)), epochs (50), optimizer (ADAM), batch size (32), learning rate (0.001), dropout (0.15) and output of a sigmoid activation function (Dense =1, activation = 'sigmoid') were fixed for all four models. Such choice could affect the models' performance negatively, but it was made while considering this step (90:10 split) as a preliminary check for the feasibility of the proposed models in predicting the QoI indicator. The models were artificial neural network (ANN), long short-term memory recurrent neural network (LSTM), one dimensional convolutional neural network (1D-CNN) and temporal convolutional neural network (T-CNN). ANNs or multilayer perceptron (MLPs) [46], are formed of input layer,

hidden layer(s), and output layer. ANNs are used to learn real world data representations, and they can be used to model the QoI which is a regression problem. The ANN developed consisted of 3 hidden layers containing 64, 32 and 16 hidden nodes. LSTM is a type of recurrent neural network (RNN) that was developed to overcome the vanishing gradient problem with regular RNN [47] and, similar to ANN they can be used for regression. The LSTM developed consisted of one layer containing 50 nodes followed by three hidden layers similar to ANN. CNNs are the standard model for computer vision and image processing applications [48]. To generalize application to 1D signals (not images), 1D-CNNs were proposed. In this study the 1D-CNN consisted of two layers; 64 filters with kernel of size 9 and 32 filters of kernel size 3 respectively, followed by fully connected layer of similar parameters to ANN.

T-CNNs or TCN are a modification of CNN and can be considered an evolution of both CNN and RNN [49,50]. The T-CNN uses dilated convolution instead of regular convolution, which applies convolution across input points separated by fixed step, and thus by extension increasing the history of the data being covered. So, for example, a kernel size 3 will always apply convolution for current data point and two points before it, but with dilated convolution of (2) it will apply to current point and two points at (2) and (4) steps behind it, respectively. In this way, by adjusting the size of step (dilation) and kernel size the extent of history modeled can be controlled. In this study, the parameters used to infer QoI are a total of 14 (4 for view, 4 for add, 4 for alter, time period (TP) and engagement time (ET)) and adjusting the kernel size and dilation would effectively correlate these factors at different points. For data splitting of 90:10; two levels of dilation are used (2) (one step in between), (4) (three steps in between) in the T-CNN layer which is followed by a fully connected layer of similar parameters to ANN.

4.2.3 Developing Deep Learning Models (75:25 split)

Five different T-CNN models were developed with fixed single hidden layers of 64 neurons and dropout of 0.15. The five models had different dilations, namely, T-CNN248: (2,4,8), T-CNN24: (2,4), T-CNN2: (2), T-CNN4: (4), and T-CNN8: (8). Note that the numbers in brackets indicate the levels of dilation. All models were trained for 100 epochs and the results are for 4-fold cross validation.

4.2.4 Evaluation of Models

Models developed were compared to the results of FCM-QoI model, using parameters of root mean squared error (RMSE), mean absolute error (MAE)[51] and correlation of determination (R^2) [52]. RMSE (equation 3) and MAE (equation 4) could give a good idea about the error in prediction, while R^2 (equation 5) will provide a good indicator of how well the predicted QoI follows the True QoI.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}{N}} \quad (3)$$

$$MAE = \frac{\sum_{i=1}^{N}|\hat{x}_i - x_i|}{N} \quad (4)$$

$$R^2 = \left(\frac{\mathrm{N}(\sum x\hat{x}) - (\sum x)(\sum \hat{x})}{\sqrt{(N\sum x^2 - (\sum x)^2)(N\sum \hat{x}^2 - (\sum \hat{x})^2)}}\right)^2 \quad (5)$$

N: Total number of data points.

x: True value.

$\hat{x}$: Predicted value.

i: Variable i.

4.3 Results

The correlation between the 14 parameters and the QoI is shown in Figure 2. Pearson's correlation coefficient [53]was used to find correlation between the numerical variables, which

includes the action parameters, engagement time and QoI indicator value. The time period is an ordinal value (categorical) as it represents which week the aggregated data refers to in the academic year (week 1- week 51). Spearman's rank correlation coefficient [54]was used for the time period correlation with all the other variables. The final correlation matrix represents the combined Pearson's and Spearman's correlation coefficients. It can be noticed that the time period has a low correlation (0.35) with the QoI. Moreover, the highest correlation was between view parameter 4, engagement time and QoI. This can indicate that the most common action between users that affects both time and QoI is viewing content.

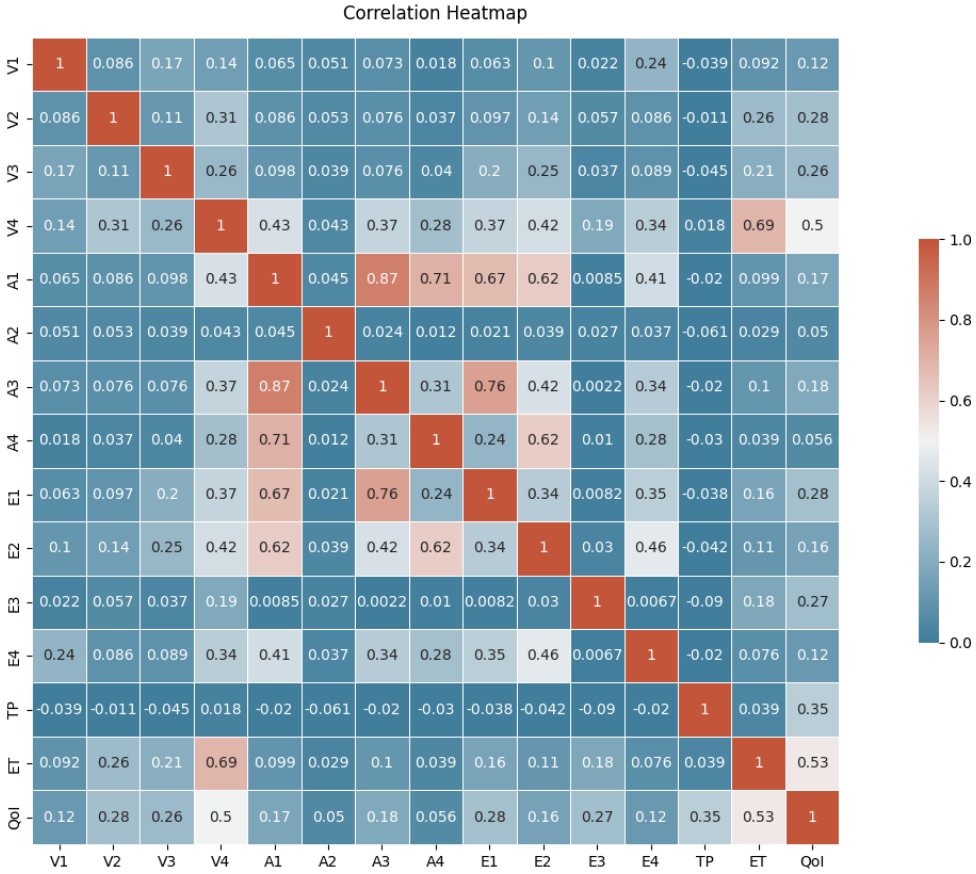


Figure 2 Correlation heat map; V = view, A = add, E = alter, TP = time period, ET = engagement time, QoI = quality of interaction.

4.3.1 Results for 90:10 split.

Table 2 shows the results for the ANN, LSTM, 1D-CNN, and T-CNN models compared to the FCM-QoI model, which are for 10-fold cross validation. The lowest MAE was for T-CNN at (0.029), FCM-QoI had the lowest (RMSE) better than T-CNN. T-CNN also had the best R^2 score at (0.685).

Table 2. Results for 90:10 split.

Model	MAE	RMSE	R2
ANN	.038	.069	.649
LSTM	.055	.09	.437
1D-CNN	.048	.084	.479
T-CNN	.029	.066	.685
FCM-QoI	__	.026	.536

Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Correlation of Determination (R2).

4.3.2 Results for 75:25 split

Table 3 shows the results for the 5 T-CNN models. Overall, the best performance was for T-CNN248 (three levels of dilation), while the worst was T-CNN8 with (8 step dilation). T-CNN248 performed better than T-CNN with 90:10 split though RMSE is still higher than FCM-QoI.

Table 3. Results for 75:25 split.

Model	MAE	RMSE	R2
T-CNN2	0.029	0.067	0.664
T-CNN4	0.029	0.067	0.678
T-CNN8	0.068	0.096	0.37
T-CNN24	0.027	0.066	0.685
T-CNN248	**0.027***	**0.066***	**0.698**
FCM-QoI	___	0.026	0.536

*: rounded value, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Correlation of Determination (R2).

4.4 Discussion

Several models were developed to represent the Fuzzy QoI model and derive its output value. The 90:10 split utilized more of the data for training and tested several models. Though

parameters for LSTM and 1D-CNN can be tweaked for better performance, T-CNN outperformed both, though RMSE was still higher than FCM-QoI model, while T-CNN had a better R^2 score. The performance of ANN highlighted that the data can be interpreted as separate non-temporal parameters as compared to a time series signal, which might explain the drop in performance of LSTM and 1D-CNN. The T-CNN developed models with different dilations can be used to interpret the relations between the parameters and their impact on the value of QoI. In addition, through the correlation heat map (see Figure. 2), the relations between the parameters and value of QoI are demonstrated. On the other hand, the T-CNN models investigate the effect of the parameters when convoluted at different intervals. Moreover, the worst performance was for TCNN8; this can be explained by how the process of convolution is applied, as the kernel (size = 3) fills the missing values (padding) with zeroes and since the input has 14 total parameters, the convolution is mostly applied between two parameters at a time with the 3rd always padded by zero effectively reducing the kernel to size 2. For single level dilations, the T-CNN4 had better performance than both T-CNN8 and T-CNN2, which shows that inter-parameter relations (i.e., view, add, alter) are more relevant as features, as the steps usually include parameters across two categories. Using highest level dilation (2, 4, 8) correlated the data across all parameters and thus had highest performance, though at cost of increased training time. It seems that the parameters in each category affect the QoI derivation and, although some seasonality might be present through the time period variable, its correlation is low and thus has modest effect on QoI inference.

4.5 Conclusion

The present study presented several deep learning models in order to replace the fuzzy QoI model, and the best performance was for T-CNN248. The proposed model can be deployed in a LMS after a preprocessing step dividing the metrics into three categories of view, add and alter. The present model can be considered the first step in fully replacing the fuzzy QoI model,

where the best outcome is to utilize the LMS metrics with minimal preprocessing to predict the QoI. Additionally, the QoI can be considered a good indicator that models the user's engagement with the LMS, giving users, policymakers, and stakeholders insights on how LMS affects the pedagogical process and feedback to improve users' experience. Finally, the next steps would be the utilization of the full 110 metrics from Moodle raw interactions in development of the deep learning model and interpretation of deep models to provide users with relevant feedback which are discussed in the next chapter.

CHAPTER 5 INTERPRETATION OF USERS' BEHAVIOR ON LMS

5.1 Introduction

In the previous chapter we were able to substitute the 12 action parameters with deep learning models, proving the feasibility of AI application of the QoI prediction model. Though the models didn't exceed the FCM-QoI in performance (RMSE); they can be considered an early proof of concept. So, to further tackle the 1st research question and develop a model that can substitute the full Fuzzy-QoI model; we focus on the starting point which is the 110 metrics instead of the 12 parameters. Additionally, we also investigate the 2nd research question which deals with the interpretability of the developed AI models and the process in which a feedback system that can modulate the user's behavior can be developed. So, in this vein, this chapter undertakes these tasks by first developing deep learning models for the 110 metrics (no categorization) and then investigates the interpretability using SHAP "SHapley Additive exPlanations" [55,56] through various available tools in that library.

5.2 Methodology

In order to model the QoI using a deep neural network; the dataset DB1 was used (see subsection 3.2). Overall, this section covers the methodology of the development of the proposed deep learning models upon DB1. The difference in methodology from the previous chapter is using 112 features as input (110, TP and ET) as opposed to 14. This would in turn increase the data sparsity, though it should be noted that unlike the previous methodology the students and professors were separated as such we expect a difference in the performance between the two approaches especially with the use of Bayesian hyperparameter optimization [57].

5.3 Deep Model selection

Two different models were developed for the 110 Moodle parameters, ANN, and 1D-CNN. Unlike the 12 parameters models, Bayesian optimization was used to select the best hyperparameters through 'Keras Tuner' from Keras python library. The optimization was run through 50 trials each spanning 15 epochs, the lowest validation (MSE) was used to select the best model, the batch size was fixed to 32. A 4-fold cross validation was then applied to the best models using scikit-learn KFold method (n_splits= 4, random_state = None, shuffle = False), four different models in total were developed as the data was separated into students and professors in order to gain a better interpretation as some of the interactions are unique to either user case. Models for students have letter 'S' as last letter, while professors' models have the letter 'P' to distinguish between the models. The parameters for each model are listed in Table 4. Model evaluation was done using three metrics, which are (RMSE), (MAE) and (R^2).

Table 4 Deep Model Parameters.

Model	Layer 1	Layer 2	Output	Optimizer (Learning rate)
ANN_S	Dense (1024, ReLu)	Dense (800, ReLu)	Dense (1, Sigmoid)	RMSprop (0.001)
ANN_P	Dense (1024, ReLu)	Dense (1024, ReLu)	Dense (1, Sigmoid)	RMSprop (0.001)
CNN_S	Conv1D (128, padding= same, Kernel= 3, Maxpooling1D = 2, ReLu)	Flatten, Dense (1024, ReLu)	Dense (1, Sigmoid)	Adam (0.001)
CNN_P	Conv1D (32, padding= valid, Kernel = 3, Maxpooling1D = 2, ReLu)	Dropout = 0.2, Flatten, Dense (32, ReLu)	Dense (1, Sigmoid)	Adam (0.001)

5.4 Model Interpretation

Interpretation of the deep learning models was done using SHAP "SHapley Additive exPlanations" [55,56]. The library offers explainers for tree and deep models based on game theory proposing SHAP values to determine feature importance through different methods and explainers. Tree explainer and Deep explainer were used to derive the SHAP values for gradient boosting model XGBoost (extreme Gradient Boosting) [58], which is an ensemble of

additive logistic decision trees[59] and ANN models for 110 parameters were used for interpretation force plots respectively. Similarly, The XGBoost was applied to both 12 parameters and 110 parameters. The XGBoost parameters for 12 and 110 were set to (n_estimators=2000, max_depth=7, eta=0.1, subsample=0.7, colsample_bytree=0.8).

5.5 Results and Discussion

Table 5 Results for 110 parameters deep models.

Model	RMSE	MAE	R2
ANN_S	0.061 ± 0.0122	0.024 ± 0.0138	0.717 ± 0.0247
ANN_P	0.082 ± 0.024	0.035 ± 0.0166	0.672 ± 0.0871
CNN_S	0.062 ± 0.0113	0.025 ± 0.0137	0.706 ± 0.0388
CNN_P	0.084 ± 0.0209	0.036 ± 0.0167	0.6557 ± 0.0871

The results for the XGBoost RMSE were as follows:

- Students 110 parameters = 0.0662
- Professors 110 parameters = 0.0826
- Students 14 parameters = 0.0678
- Professors 14 parameters = 0.0829

The results from both models for students and professors are considered reasonable when compared to the results from the 12 action parameters (Table 3). The models have an improvement in RMSE as result of applying Bayesian optimization algorithm to select the best hyperparameters as shown in Table 4. This allowed the cross validated data to be trained while ensuring the lowest validation loss. Another improvement with the application of early stopping and reduction of learning rate on non-change of validation loss for 10 and 5 iterations, respectively which allowed faster training and ensuring that best models to be saved for reference. This can explain to some extent the improvement in performance when compared to the results of the 12 action parameters models although they suffer from less sparsity and have a lower number of features, that in addition to the separation of students and professor classes.

The extent to which it can be improved using optimization and tuners is a point of interest. The separation between students and professors could be also a determining point as professors' interactions are considered high when compared to students during same periods. By using SHAP library, the four XGBoost models were interpreted through a bee-swarm plot as shown in Figure (3-6):

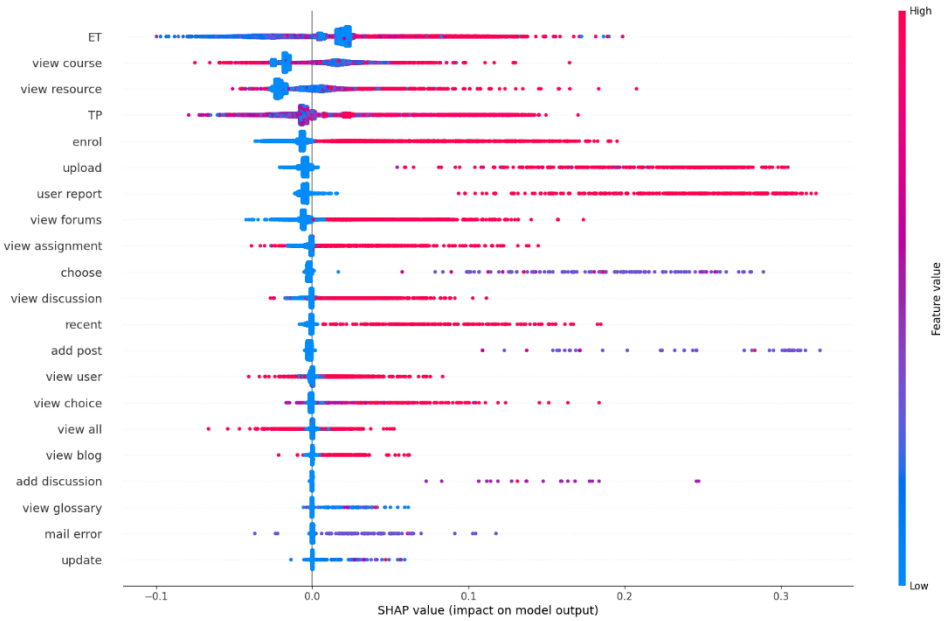


Figure 3 Students' top 21 (from 112) factors affecting QoI results.

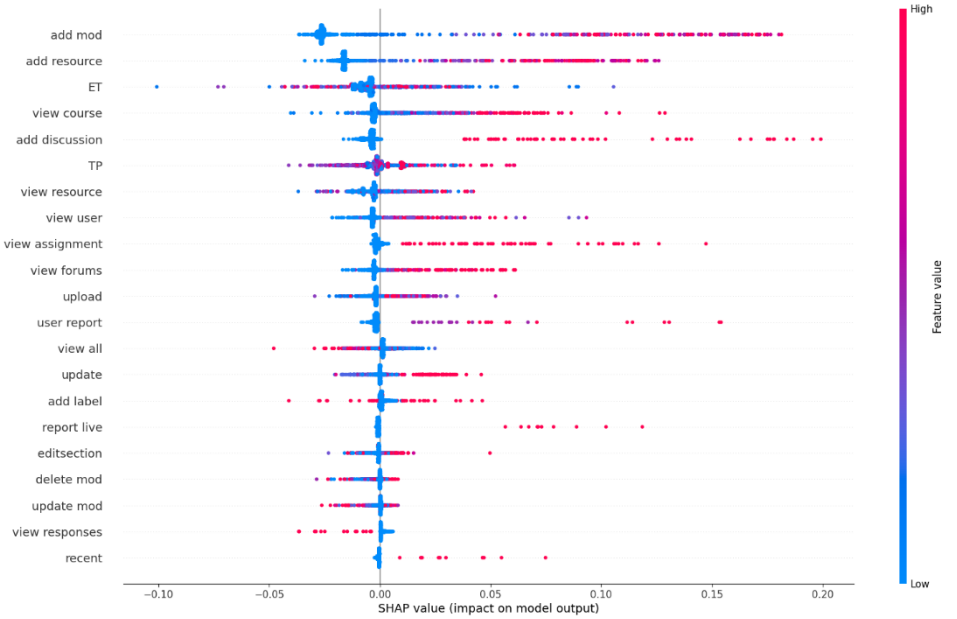


Figure 4 Professors' top 21(from 112) factors affecting QoI results.

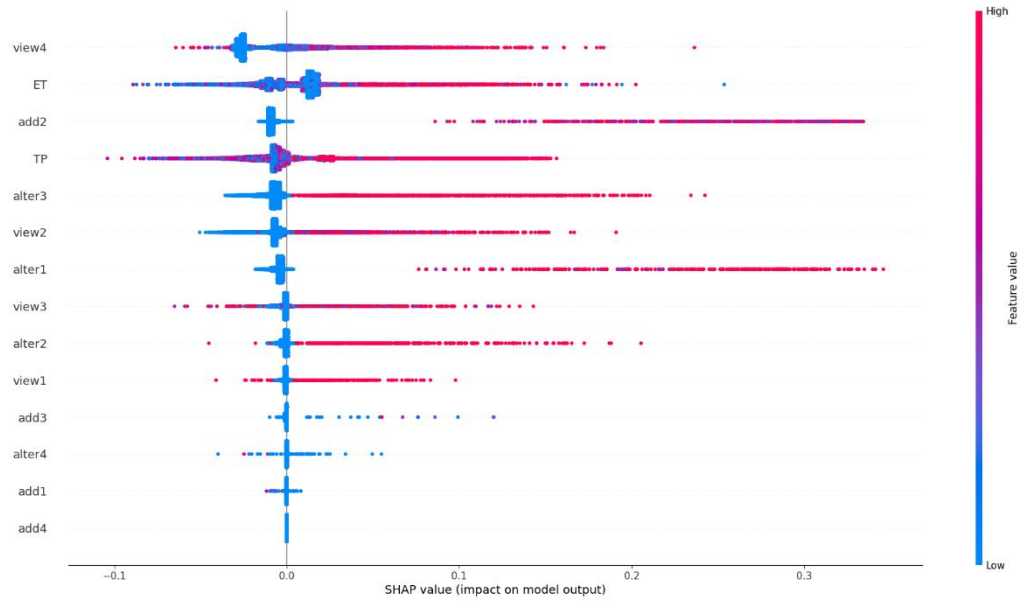


Figure 5 Students' action parameters in order of affecting QoI results.

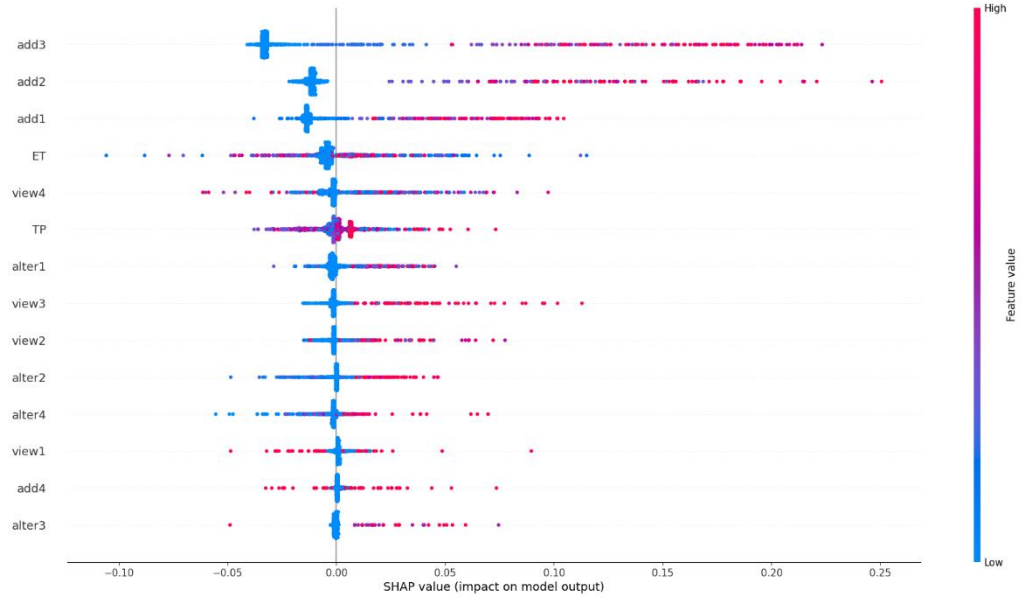


Figure 6 Professors' action parameters in order of affecting QoI results.

By examining the results of the bee-swarm plots; it can be noticed that the professors' top features affecting QoI value are those concerned with adding content, which fits their role. Engagement Time (ET) also affects it to a reasonable degree, TP on the other hand comes

fourth (students) or sixth (professors) in order of importance. Given that the academic year has some form of seasonality in the pattern of midterms, finals, and project submissions; the effect of TP fluctuates across the academic year causing its increase to show both an increase and a decrease in the QoI value. When examining the students' SHAP values in the bee-swarm plot ET comes first and second for 110 and 12 parameters respectively with having the top factors affecting the output are viewing and uploading along with TP. Finally, using Deep Explainer from the SHAP library to understand the effect of individual features on the output or value of the QoI metric through the use of force plots. Two samples of each group were used (110 parameters), one from both students and professors having both low and high QoI value. Figures (7, 9, 11, 13) show the force plots for each case. Deep explainer was applied to models trained on same parameters as best 110 models for ANN, the explainer used 1000 samples from training set to initiate the explainer method. Then it was used to explain one instance of the testing set. Two instances of low and high QoI were chosen randomly for each; students and professors and the force plots were plotted.

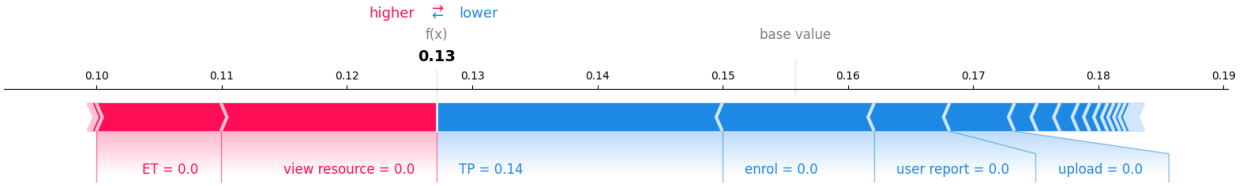


Figure 7 Student with low QoI.

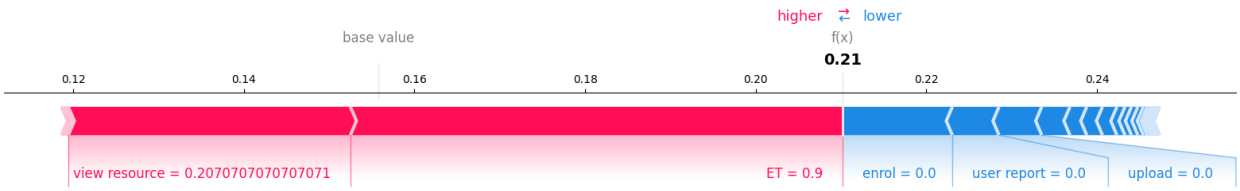


Figure 8 Student with low QoI after editing ET and TP (+ 0.08).

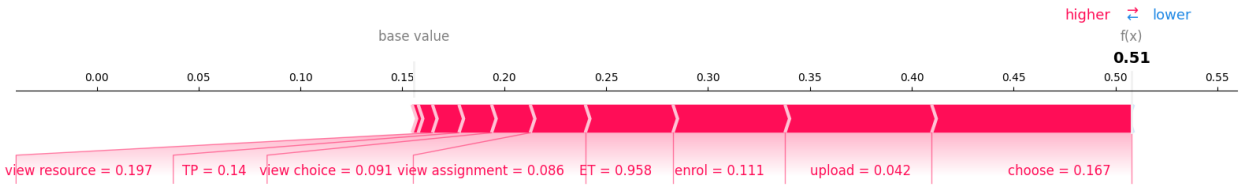


Figure 9 Student with high QoI.

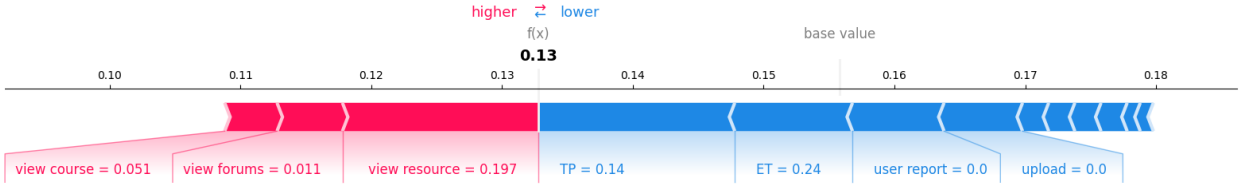


Figure 10 Student with high QoI after editing choose, enroll, upload and ET (- 0.38).

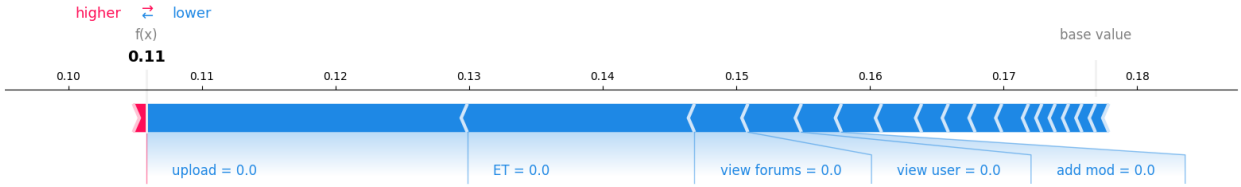


Figure 11 Professor with low QoI.

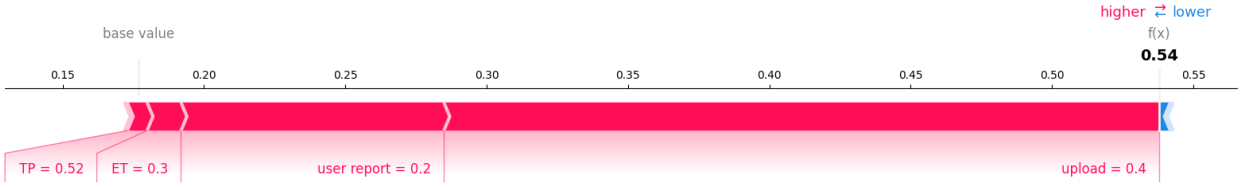


Figure 12 Professor with low QoI after editing upload, user report, info, and ET (+ 0.43).

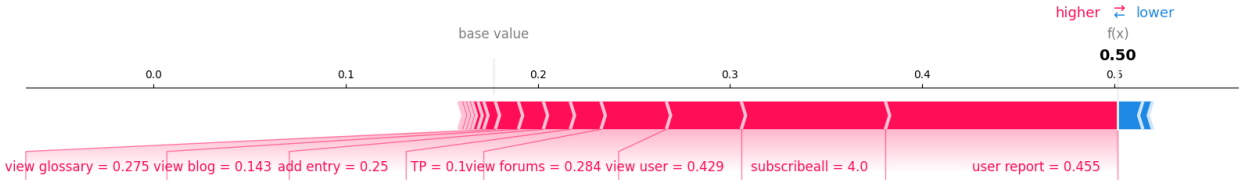


Figure 13 Professor with high QoI.

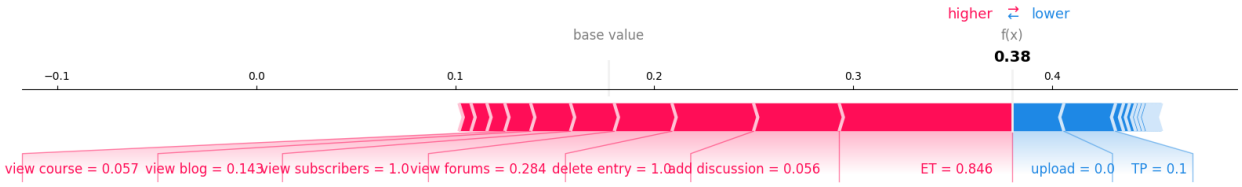


Figure 14 Professor with high QoI after editing view user, add entry, subscribe all and user-report (- 0.12).

Looking at Figures. (7,9) we can notice the factors affecting the results of both students. By modifying the values of ET and TP for low QoI student and choose, upload, ET and enroll for student with high QoI we can modulate the results of the prediction of the QoI based on the model's interpretation of the features governing or explaining the prediction. Figure. (8,10) show the modulated QoI after changing the values. On the other hand, the professors QoI are affected more by actions that pertain to other users like viewing forums, viewing users, uploading, and adding mods. Similar to the experiment with students we can modulate the QoI of the professor through changing values of view user, add entry, subscribe all and user-report leading to a reduction in value (-0.12), when compared to Figure 14 while upload, user report, info and ET increased the QoI of from 0.11 to 0.54 (+ 0.43) as in Figure 12. This small experiment can indicate that increasing the value of QoI based on interpretation is more feasible than decreasing it, though further testing with more values can provide a better understanding of the extent of such change. An explanation of such phenomenon can be attributed to masking effect of high impact values on features with less importance. The second point is that with a

low QoI increasing any value should increase the QoI to some extent if specific features like TP are not considered. Based on this small experiment an application utilizing the QoI metric to increase user's participation and engagement is feasible. The following section discusses the possible approaches to apply such a system.

5.6 Behavioral Modulation of LMS Users

Influencing student's participation and access to the LMS is not a new idea, several approaches like gamification [60], use of multimedia [61] and use of other technological [62]aspects in education to affect student participation, satisfaction, and pedagogical experience respectively were explored. Modern LMSs can be said to combine several of these approaches, though some instructors may use other modules and apps to introduce a game like experience to the education process. For example, Kahoot can be used in place of quizzes, and in one study it was found to be a good predictor of success, as those with good outcomes performed well in quiz games [63]. This brings the question of what aspects of motivation does gamification affect? whether intrinsic, extrinsic or both. Authors in [64]found that in cases of extrinsic motivation and competition between students, especially on social level could have negative effects while intrinsic motivation provided more positive effects overall on the pedagogical outcomes of the students. The use of gamification although has a wide range to define what constitutes gamification, its main aspects could be considered the novelty and rewards (both intrinsic and extrinsic). In non-educational settings, it was found that intrinsic motivation in form of a progress bar can help those receiving treatment from drug abuse by utilizing the concept of 'need to complete' [65]. Also, it was found that extrinsic motivation is more relevant and effective for tasks that are not interesting to the group while self-determination and intrinsic motivation are usually associated with interesting topics [66]. And so, from the perspective of education it can be said that an intrinsic motivation approach that induces self-efficacy and determination would probably have better outcomes.

So, to modulate student behavior on LMS using the QoI metric and the interpretation from the force plots and SHAP values, we for example introduce a progress bar that represents the student's percentage of QoI for this week and suggest relevant actions that can improve it to 100%. As a precautionary measure against abuse, (i.e., making it foolproof) by randomly doing actions on LMS to improve QoI value; weights could be given to actions that conform to suggestions (higher effect on QoI) versus other random actions. Another game like concept is to introduce milestones, missions or aesthetic rewards like badges and borders to students. Milestones and missions could be tailored based on the students' available pool of actions with relation to available course materials, recent content uploaded by instructor or deadline of assignments and projects. In a sense, it will ensure relevance of actions with regards to the learning process. Aesthetics like badges and borders could be awarded per semester or year if student's QoI was maintained above a certain level, though it would be better if such rewards remain as an individual accessed reward and not shared on social level as [63]noted it could affect students negatively if they are in a competitive social environment for rewards. Later on, a study to investigate the effects of QoI based motivation on student behavior whether successful or not could be done. It can also investigate its effect on students' grades and outcomes.

5.7 Conclusion

The development of deep learning models utilizing the 12 and 110 parameters of the Moodle user's interaction was completed. The models provided some insight into the behavior of both students and instructors when engaging with LMS, which can be relevant to users, policy makers and stakeholders. It also provided a metric (QoI) that can effectively through model interpretation modulate user's engagement through personalized feedback. The proposed model of game like intrinsic motivation to influence the behavior of the user could be a potential tool that may improve user's overall pedagogical experience and satisfaction.

The model is purely theoretical requiring to be applied through a pilot study that could include a small handful of students in order to verify its potential short- and long-term effects on learning experience. The study also is lacking in magnitude of data size as it used only 1037 students and 75 professors from one HEI, which can be augmented in follow up studies to have a more comprehensive representation of the various user interactions, i.e., utilize more interactions for the model interpretation as opposed to 21 out 112 and from different HEIs. This could be the main subsequent follow up study, while applying the other parts of A / B/ C Teach Framework is also another venue for exploration, mainly the QoC and AS components development and application. Finally, this chapter concludes the answers to the 1st and 2nd research questions, though the performance of the deep learning models leaves more to be desired, the next section focuses mainly on 3rd and 4th research questions dealing with domain shift, model performance improvement, and issues with regards to AI application in education.

CHAPTER 6 ON MODELLING QOI USING DEEP LEARNING

6.1 Introduction

The previous chapter concluded the 1st two research questions, though the issues of domain shift and improving model performance remain to be solved or investigated. In summary, we need a performance comparable to the FCM-QoI while also allowing for model deployment to other LMS brands or other HEIs. So, we focus on the aspects related to AI application across domains and problems that could rise from such application using two main approaches; 1st is to apply to data extracted from Moodle backend as opposed to frontend data of DB1, and the 2nd is validate the data against new data from Khalifa University. This allows for in-depth investigation for the pros and cons of such application while also providing insights into methods that could help improve performance and reduce bias. In summary, this chapter focuses mainly on answering the 3rd research question and partly the 4th research question.

6.2 Methodology

In this chapter, the methodology focuses on replicating the same scaling approach applied in the development of the two new datasets KUQ2020, and KUQ2021. Robust scaling followed by min-max scaling is applied across periods, then same distribution is applied to all testing and external validation sets. The model selection is done using two approaches a Bayesian optimized model and pyramid (wide base -input, small top – output) models. The approach focuses more on dealing with the data sparsity through careful selection of hyperparameters that are known to deal with this problem. One minor consideration is that TP is considered a numerical value following the Fuzzy-QoI convention.

6.2.1 Data Splitting

The previous work of FCM-QoI [29], used a 75:25 split rate between training and testing. Matching that rate; the data was split into training and testing sets based on users (75% of users in training and 25% in testing). This was done to prevent information leakage [67] between both sets as the scaling is done based on time period (TP) as mentioned in previous subsections. Additionally, to ensure that all possible outcomes of QoI are included in training and testing (both contain all ranges of QoI metric output); the output ranges were set to three thresholds (0.2, 0.4, 0.6). These ranges were selected based on maximum value of QoI within DB1(for students) which was 0.57. The Fuzzy-QoI system had five membership functions for the last fuzzy inference block going from very low to very high. Following this convention, the three ranges were defined as very low, low, and mid respectively. As such, the splitting was done to stratify these three ranges considering them as classes and hence users that exhibited for example 'mid' QoI output were split into 75 :25 between training and testing. Similarly, the other two classes were also split, and a 4-fold stratified cross validation [68] between training and testing splits based on users was done to further investigate the effect of different splits on the performance of the deep learning model.

6.2.2 Deep Model selection

DB1 data was formatted into two sets; the 1st with 12 action parameters, and the 2nd with 110 raw metrics as the input data in form of aggregated weekly frequencies per user (both not scaled). Both cases had two extra features TP and ET as last two columns for total features 14, 112 for 1st and 2nd respectively. Given that the data is tabular; multilayer perceptron [46] is compatible with tabular data, though the sparsity of user interactions requires adjusting for model parameters to allow model to learn effectively. Two considerations were taken, 1st is to use an activation function that can allow the model to continuously learn, and the 2nd is that the

optimizer should have adaptive learning rate in order to adjust the weights based on the rare positive interactions of users when compared to the total interactions. The activation function selected was scaled exponential linear unit (SeLU) [69] which is characterized by a linear scale for values greater than zero and exponential (exp(x) -1) for those lower than zero. The declining negative exponential slope ensures that the model never stops learning unless the output is zero which for a sparse data input can help train the rare occurrences. The activation function of the last dense layer was sigmoid function to ensure that output is between [0,1] matching the QoI metric output. The optimizer selected was Adamax [70] which is a variant of the Adam optimizer that can adjust learning rate based on input data and has the ability to learn temporal correlations within data, which might help reflect the effect of changing time period on the QoI metric output. Figure 15 shows the SeLU function.

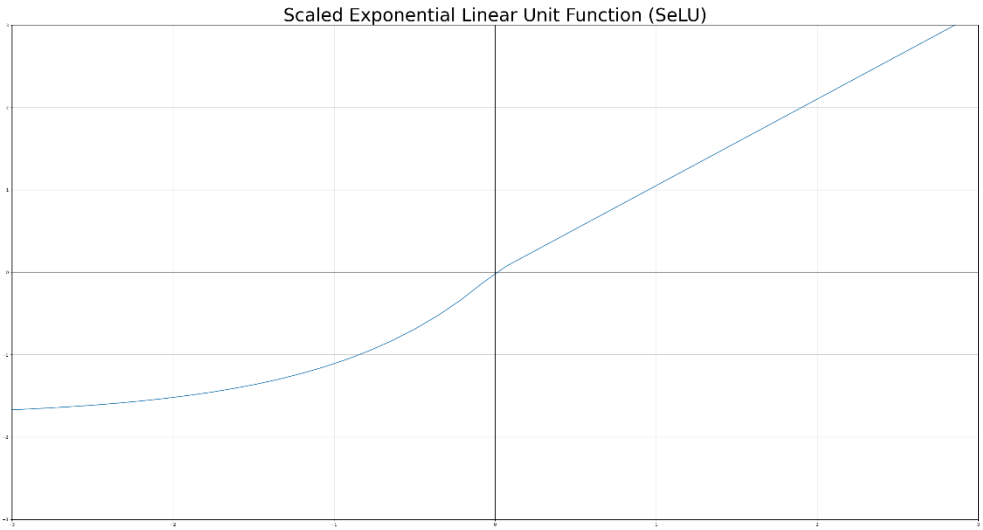


Figure 15. Scaled Exponential Linear Unit (SeLU).

Optimization of the dropout [71], the number of layers and the number of units in each layer was done using Bayesian optimization [57] through the keras tuner as part of keras library

[72]. The tested parameters are shown in Table 6, the number of layers was considered up to 7 layers not including the last layer with single unit. This model would be compared to a tapering model with 6 layers (not including last layer), which starts with large number of units and reduces gradually towards final output. Parameters of both models are shown in the results section. The best models for both 12 action parameters and 110 raw metrics were saved for training. One minor consideration for batch size was done in order to improve training speed and was fixed to 16384, as the number of epochs was set to 10000 to allow for the model to learn from the sparse input and dropout was used as regularization to prevent overfitting.

Mean squared error (MSE) [73], was used as loss function as it is sensitive to large outliers which would allow the relevant values of QoI to be trained as compared to the whole data since they are fewer in numbers and can be considered outliers. Metrics for evaluation were root mean squared error (RMSE), mean absolute error (MAE) [51], and coefficient of determination (R^2) [52]. The deep learning model performance for 12 actions parameters was compared against the FCM-QoI [29] model.

Table 6. Bayesian optimization tested parameters.

Hyperparameter	Minimum	Maximum	Step
Dense layer number	1	6	1
Dense layer units	32	1024	32
Dropout layer number	1	6	1
Dropout	0	0.2	0.1

6.2.3 External Validation

The results of the deep model training were evaluated mainly using the testing set as mentioned in previous subsection, then 12 action parameter model was externally validated against the KUQ2020 and KUQ2021 datasets in order to further test the model's ability to generalize. No external validation was done for 110 raw metrics model as the number of unique interactions for both datasets exceed the number of interactions used for the Fuzzy-QoI (334

versus 110), thus the categorization of said parameters was done to match the main 12 action parameters. Results for both variants of models including tapering and optimized models is included in the following section. Table 7 shows the description of each dataset, Table 8 shows the total number of users, interactions and time spent for each dataset (hourly).

It should be noted that for 14 action parameters three experiments were done:

- Case 1: robust scaling with IQR (25-75) followed by min-max [0,1] as preprocessing step.
- Case 2: robust scaling with IQR (25-75) without min-max scaling.
- Case 3: robust scaling with IQR (10-90) followed by min-max [0,1].

Table 7 .Datasets descriptions.

Name	Start	End	Students	COVID-19	University	Country
DB1	26/8/2009	18/8/2010	1037	Pre	Faculdade de Motricidade Humana	Portugal
KUQ2020	16/03/2020	16/8/2020	3377	During	Khalifa University	UAE
KUQ2021	24/08/2020	14/8/2021	3377	During	Khalifa University	UAE

Table 8. Datasets statistics.

Parameter	DB1	KUQ2020	KUQ2021
View	499328	7427385	64090
Add	3940	280391	8751
Alter	13219	246031	11515
Weeks	51	22	51
Total time (hours)	200868.4	213315.1	1428.5

6.3 Results

The developed models' architecture for both tuned and tapering models (pyramid shape) is shown in Figure. (16a, 16b) for 14 parameters and Figure. (17a, 17b) for 112 parameters. It should be noted that dropout was included for all dense layers of the tapering models for both 14 and 112 and was set at 0.2, though it is not included in figures to reduce image space. The tuned model shown below in Figure**.** 16a is has the best model performance for MSE when iterated for 200 different combinations each spanning 200 epochs, with one trial

per execution. Layers that had dropout at zero value were removed as they are redundant and don't affect model performance.

The performance of the models was evaluated using training, testing, KUQ2020 and KUQ2021 with RMSE, MSE, MAE and R^2 as evaluations metrics. The average of the 4-fold cross validation is reported with 95% confidence interval (CI). The performance for 14 action parameters is compared with FCM-QoI. Results from both types of models for 14 and 112 parameters are shown in Tables (9-16).

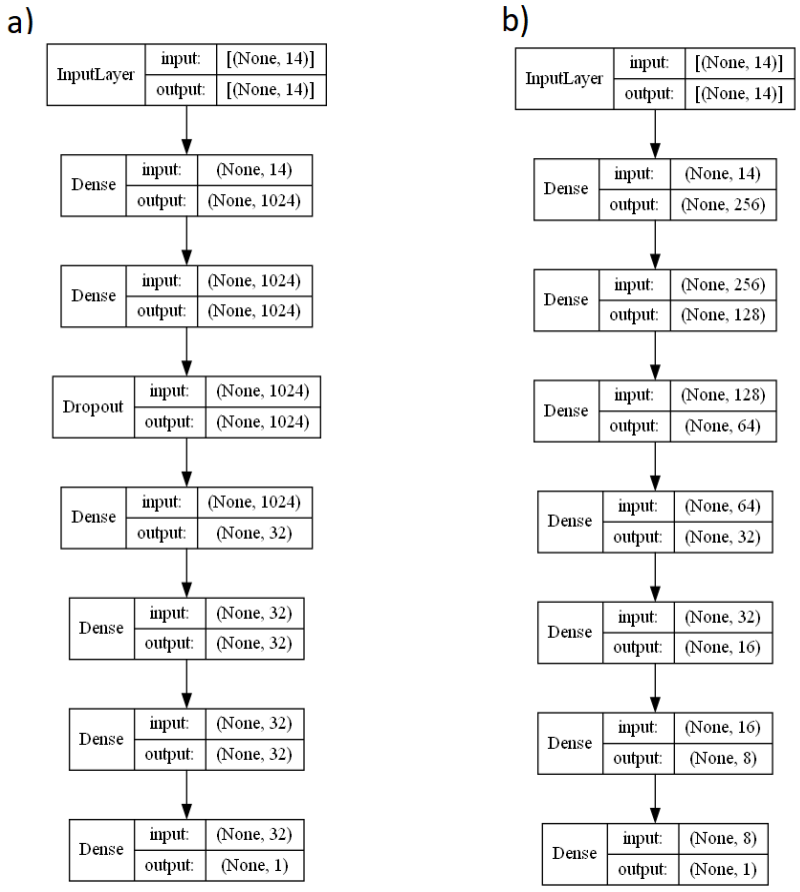


Figure. 16. a) Bayesian optimized model for 14 parameters input, the dropout rate for all dropout layers is 0.2, b) Tapering model (pyramid) with dropout layer after every dense layer except last, value of dropout is also set to 0.2.

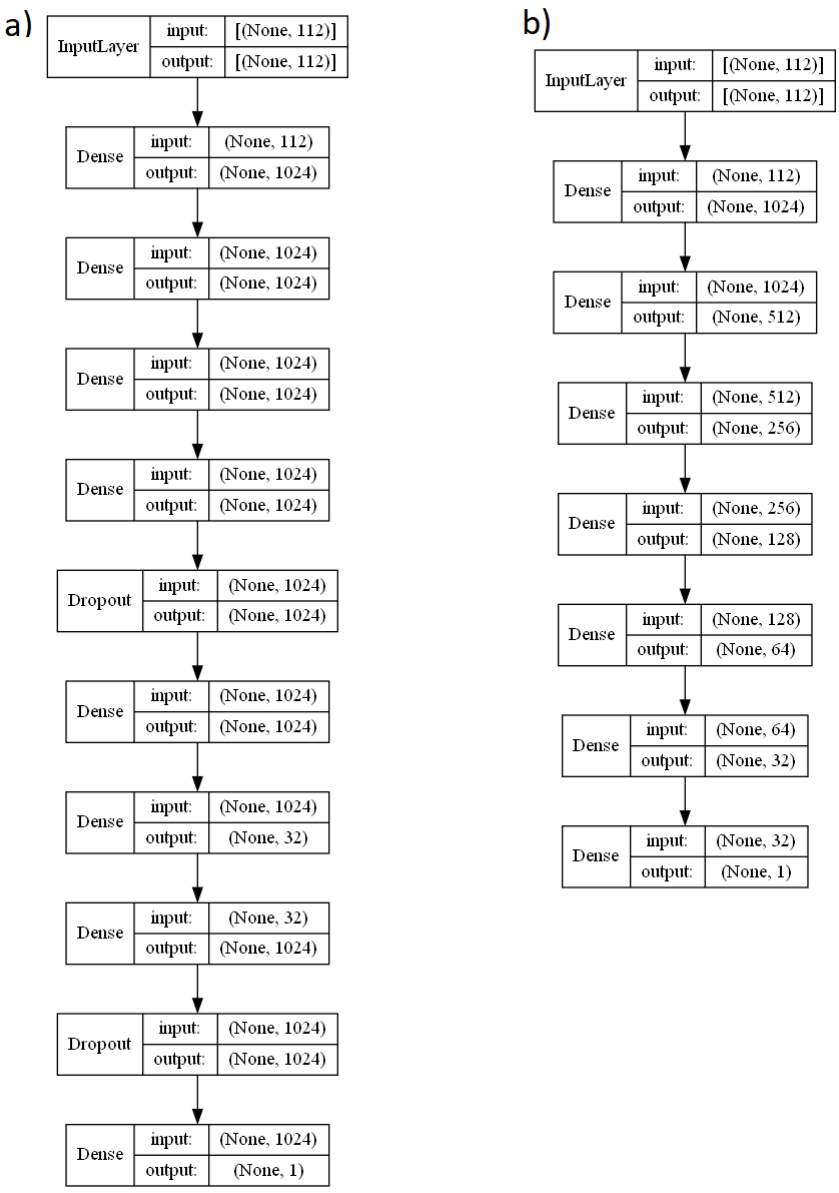


Figure. 17. a) Bayesian optimized model for 112 parameters input, the dropout rate for all dropout layers is 0.2, b) Tapering model (pyramid) with dropout layer after every dense layer except last, value of dropout is also set to 0.2.

Table 9. Model performance for tuned model (14 parameters), case 1.

Parameter	RMSE	CI (95%)	MSE	CI (95%)	MAE	CI (95%)	R^2	CI (95%)
Training	0.00519	0.00157	0.00003	0.00001	0.00197	0.00038	0.99545	0.00242
Testing	**0.02601**	**0.00585**	**0.00069**	**0.00030**	**0.00464**	**0.00103**	**0.88749**	**0.04513**
KUQ2020	0.26483	0.05835	0.07115	0.02976	0.14817	0.03921	-5.32861	2.64758
KUQ2021	0.03350	0.00517	0.00113	0.00033	0.00394	0.00041	0.09565	0.26412
FCM-QoI	0.0264	0.003	N/A	N/A	N/A	N/A	0.5363	0.01

Table 10. Model performance for tapered model (14 parameters), case 1.

Parameter	RMSE	CI (95%)	MSE	CI (95%)	MAE	CI (95%)	R^2	CI (95%)
Training	0.01985	0.00539	0.00040	0.00021	0.00490	0.00148	0.93351	0.03466
Testing	**0.02351**	**0.00495**	**0.00056**	**0.00023**	**0.00541**	**0.00158**	**0.90850**	**0.03095**
KUQ2020	0.19624	0.00997	0.03854	0.00391	0.11162	0.00857	-2.42811	0.34820
KUQ2021	**0.02401**	**0.00832**	**0.00060**	**0.00043**	**0.00353**	**0.00054**	**0.52246**	**0.34788**
FCM-QoI	0.0264	0.003	N/A	N/A	N/A	N/A	0.5363	0.01

Table 11. Model performance for tuned model (14 parameters), case 2.

Parameter	RMSE	CI (95%)	MSE	CI (95%)	MAE	CI (95%)	R^2	CI (95%)
Training	0.02987	0.00305	0.00089	0.00019	0.00722	0.00148	0.85280	0.02847
Testing	0.05436	0.00275	0.00296	0.00030	0.01420	0.00145	0.51321	0.04327
KUQ2020	0.21897	0.06758	0.04930	0.02947	0.12896	0.04625	-3.38558	2.62132
KUQ2021	0.02888	0.00407	0.00084	0.00023	0.00406	0.00038	0.32900	0.18764
FCM-QoI	**0.0264**	**0.003**	**N/A**	**N/A**	**N/A**	**N/A**	**0.5363**	**0.01**

Table 12. Model performance for tapered model (14 parameters), case 2.

Parameter	RMSE	CI (95%)	MSE	CI (95%)	MAE	CI (95%)	R^2	CI (95%)
Training	0.03128	0.00031	0.00098	0.00002	0.00758	0.00019	0.83892	0.00545
Testing	0.04252	0.00253	0.00181	0.00022	0.01013	0.00065	0.70064	0.05720
KUQ2020	0.19879	0.01012	0.03955	0.00401	0.11519	0.00755	-2.51775	0.35669
KUQ2021	0.03077	0.00160	0.00095	0.00010	0.00423	0.00017	0.24194	0.07917
FCM-QoI	**0.0264**	**0.003**	**N/A**	**N/A**	**N/A**	**N/A**	**0.5363**	**0.01**

Table 13. Model performance for tuned model (14 parameters), case 3.

Parameter	RMSE	CI (95%)	MSE	CI (95%)	MAE	CI (95%)	R^2	CI (95%)
Training	0.00576	0.00065	0.00003	0.00001	0.00210	0.00034	0.99451	0.00124
Testing	**0.02534**	**0.00285**	**0.00064**	**0.00015**	**0.00463**	**0.00102**	**0.89403**	**0.01495**
KUQ2020	0.22419	0.01183	0.05030	0.00531	0.12556	0.00963	-3.47467	0.47193
KUQ2021	0.02868	0.00225	0.00082	0.00013	0.00350	0.00054	0.34078	0.10264
FCM-QoI	0.0264	0.003	N/A	N/A	N/A	N/A	0.5363	0.01

Table 14. Model performance for tapered model (14 parameters), case 3.

Parameter	RMSE	CI (95%)	MSE	CI (95%)	MAE	CI (95%)	R^2	CI (95%)
Training	0.01932	0.00530	0.00038	0.00022	0.00463	0.00126	0.93689	0.03821
Testing	**0.02337**	**0.00631**	**0.00056**	**0.00031**	**0.00525**	**0.00164**	**0.90870**	**0.04345**
KUQ2020	0.20338	0.02462	0.04154	0.01026	0.11740	0.01799	-2.69545	0.91253
KUQ2021	0.02650	0.00996	0.00073	0.00052	0.00381	0.00093	0.41466	0.41805
FCM-QoI	0.0264	0.003	N/A	N/A	N/A	N/A	0.5363	0.01

Table 15. Model performance for tuned model (112 parameters)

Parameter	RMSE	CI (95%)	MSE	CI (95%)	MAE	CI (95%)	R^2	CI (95%)
Training	0.00357	0.00571	0.00002	0.00006	0.00163	0.00190	0.99618	0.01001
Testing	0.03992	0.00409	0.00160	0.00033	0.00702	0.00232	0.73651	0.03961

Table 16. Model performance for tapered model (112 parameters)

Parameter	RMSE	CI (95%)	MSE	CI (95%)	MAE	CI (95%)	R^2	CI (95%)
Training	0.00750	0.00076	0.00006	0.00001	0.00177	0.00037	0.99070	0.00179
Testing	0.03408	0.00293	0.00116	0.00020	0.00552	0.00117	0.80876	0.01308

Figure. (16 ,17) show the developed models both from Bayesian optimization (16a, 17a) and tapering models (16b, 17b). The use of dropout as a form of regularization to reduce overfitting was evident in tapering model as we expected the model to run for large number of epochs due to data sparsity, where during the hyperparameter optimization they only ran for 200 epochs per trial which may not show the effect of overfitting for a slower optimizer like Adamax. The performance for the models across 14 actions parameters was best for case 1 as testing set had best performance (RMSE = 0.02351) for tapering model and KUQ2021 exceeded the FCM-QoI benchmark using same model (RMSE = 0.02401). Case 2 showed worst performance overall as the effect of min-max scaling is removed. Case 3 had a slightly better performance but the increased range of IQR gave rise to more outliers reducing the performance. The most notable result is performance of KUQ2020 in all 14 parameters models having high RMSE (> 0.19) and negative R^2 value.

The 112 parameter model performance for testing set was (RMSE = 0.0399, R^2 = 0.7365) and (RMSE = 0.0341, R^2 = 0.8088) for tuned and tapering models, respectively. It can be noticed that for both models some degree of overfitting can be observed (more for tuned) models, which may require further regularization methods to reduce. The performance of testing set for 112 parameters is relatively close to performance of FCM-QoI under 14 action parameters which indicates that the model was able to learn some underlying correlations between raw input and QoI output.

6.4 Discussion

The descriptions shown in Table (5,6) highlight the main difference between the two datasets (considering both Khalifa University datasets as one) which is the COVID-19 lockdown status. Both datasets were considered a point where both universities implemented Moodle in some sense, for Faculdade de Motricidade Humana it was the first-time implementation for Moodle, while for Khalifa University it was the switching to online classrooms as the modus operandi. When looking at the three datasets; KUQ2020 shows a higher intensity in interactions even though it covers a smaller duration (22 weeks), while KUQ2021 shows lower interactions than both sets even though the lockdown protocols were still ongoing. This could be attributed to the attraction of system due to novelty which may have subsided as the situation became accepted as the norm, and also due to usage of BlackBoard in addition to Moodle during the 2nd year which reduced the total interactions captured on Moodle (currently only BlackBoard is being used in Khalifa University). Another point to take into consideration is that for both KUQ2020 and KUQ2021 interactions that represent the online classroom (virtual classroom) were not considered for the interactions accounting for either 12 action parameters or ET. This is mainly to account for the fact that Fuzzy-QoI was developed to be utilized for a blended learning environment, where the pandemic shift to online modality created a singular mode of learning (online) assuming the definition of blended learning to be the combination of both face to face(F2F) and through LMS [74-76].

From the above results, we can deduce that the developed deep models are able, to a great extent, to learn the data representations of the Fuzzy-QoI system based on given input and output data. The 14 parameters tapering model (case 1) had better performance when compared to the 112 tapering model mainly due to lower number of features allowing the model

to capture the representation of data despite the sparsity of interactions. In essence, it captured the expert knowledge represented in the Fuzzy-QoI system.

The low performance of external validation dataset KUQ2020 raises a point of concern, though when looking over the number of interactions overall we noticed that they were greater in magnitude than the other two datasets, which may be relevant when considering that the preprocessing steps of robust scaling and mix-max scaling are affected by the number of interactions across every time period. One other explanation for the worsening performance is the difference between the arrangement of TP across the two datasets (DB1 and KUQ2020) as for example the starting week (week 30 for KUQ2020) is the easter holiday for DB1 and is expected to have low interactions while for KUQ2020 it is the start of full online modality showing possibly the highest level of interactions. Moreover, the datasets come from two different populations with different behaviors, culture, and expectations. This difference brings the issue of bias and fairness to the discussion. Depending on how the data is scaled; the minimum acceptable interactions of users will vary, for example when we consider the interactions from a fuzzy perspective; a student studying during the whole semester regularly for two hours a day could be considered ‘good’ while same student during the final exams period studying for the same duration could be considered ‘bad’. The Fuzzy-QoI is able to reflect such relativity through the time period, aggregation and scaling across the week for users which shows some level of seasonality of the interactions across the academic year. For our model that was trained on the DB1, weeks that were considered holidays during which students had low interactions and when we scale across the period then apply same distribution to testing sets, we notice the worsening performance due to disparity in interactions between the two periods. One possible solution is adding a label (class) to identify common time periods between HEIs, for example class 0 for beginning of semester, class 1 for mid-term, class 2 for spring break and so on. These labels could help the model generalize based on context of the

week rather than its value, thus by extension help reduce the bias and errors, though our initial experimentation with this method showed no improvement on the performance.

Another approach is to set a fixed acceptable number of interactions per study week for each category, built using expert knowledge to bolster the model ability to generalize, this could be positive as long as it is updated regularly, and no special situations/interventions arise (COVID-19 pandemic) which might make it less favorable as a permanent solution. We can also consider the number of interactions per student, switching it to a subjective perspective where the user's maximum until present sets the scale for all his other interactions. This approach could allow for tailored experience where QoI is evaluated relative to student's overall interactions rather than the average of his peers. Main issue with this approach, is that for students with extremely high or extremely low interactions the value of QoI will not be able to represent their interactions correctly as a single week with few interactions for a student with no previous interactions will be the highest on that week (high QoI value) even though it is objectively low, and any subsequent interaction will be considered lower as long that benchmark is not exceeded.

This necessitates a robust and fair/just scaling system that is resistant to noise (for example students leaving screen open or random clicking with no purpose on LMS content) to reduce error and eliminate bias. The measures of using robust scaling and min-max scaling could offset some of the noise effects but as seen with special situations some robust methods that could help model learn better are required. On a final note, in KUQ2021 dataset, the small number of interactions could be attributed to the gradual migration of Khalifa University's LMS users to BlackBoard, making it the main LMS by August 2021, which means that a substantial portion of interactions is not accounted for due to different LMS integration.

6.5 Implementation of AI: LMS (A proposition)

The results and discussion from the current and previous chapters showcase the potential of application of AI, at least to predict QoI, as part of the A/ B / C Teach framework [7,10]. The framework encapsulates the pedagogical phenomenon using an affective (A), blended (B), and collaborative (C) approaches. The components of each of these approaches rely on thoughtfully and meticulously developed metrics (QoI, QoC) to effectively model the interactions of students / teachers within a blended (hybrid) learning environment. Both metrics were developed using fuzzy logic which represents expert knowledge. Moreover, this framework allows the evaluation of leaners' actions and interests using qualitative measures which could allow it to be utilized to provide meaningful and effective feedback. Furthermore, the framework takes into account several considerations (some are mentioned below) in order for it to be fully utilized:

- Adaptability of blended and collaborative learning scenarios, by providing constant updates to already proposed approaches to both keep up and provide a tailored experience based on HEIs' and learners' requirements.

- Online learning environment techno-pedagogical sustainability, through continuous development, training and support.

- Utilization of advanced and innovative techniques in education, similar to the framework's introduction of fuzzy logic, other modelling techniques could be appropriated in order to further improve the user modeling.

- Worldwide application of collaborative and blended learning contexts, QoI and/or QoC were applied across three different institutes (Faculdade de Motricidade Humana, Portugal, Aristotle University of Thessaloniki, Greece, and Khalifa University, UAE), giving it the potential for expansion across different institutes and LMSs [7,10].

So, by introducing AI into the framework, certain precautions must be taken; first is to deal with issues of privacy due to the data collection, 2nd is the effect of bias due to generalization across different populations and lastly to deal with other data collection issues like surveillance, data sharing, anonymity [77]. As it is known, for any research requiring data collection from participants, ethical approval from supervising institute and consent from participants is required (may not be needed for retrospective data). The data collected after fulfilling both requirements is expected to be under secure protocols ensuring no leakage and in case that data is to be made public or shared, any identifying characteristics should be removed. Within this framework, student's data (name, age, gender, or ID) are removed from collected data to ensure privacy. Though it should be noted that for some LMS brands the data may not be on the HEI premise, but instead collected from the servers of the service provider based on agreements between both parties. As such, cooperation between institutes and service providers is needed to collect the user data for implementation of both QoI and QoC. The issue of privacy at this point will rely mostly on efforts of HEIs in de-identifying this data whether they are sole holder of the data or acquiring it from a 3rd party service provider. Additionally, users should be made aware of what data is being collected as they use LMSs. Other than interaction data, user's videos are also included as part of virtual classrooms or online presentations. The security of this data could be within the responsibility of 3rd parties, but also the HEIs if such data is shared (viewed by students) which may be subject to dissemination that is difficult to control (example: students sharing lecture recordings across same or different years). Although there are efforts in ensuring data privacy, education data doesn't enjoy the same treatment given to more sensitive data (health care records), as such although there are efforts by 3rd party providers and HEIs in ensuring that their systems are secure, the level of security will vary based on measurements taken. Thus, it can be concluded that privacy is and

will be a main focus during implementation of any AI-based LMSs including the A/B/C Teach framework.

The 2nd issue of bias mainly due to difference in populations, can be addressed using several approaches. When considering the A/B/C Teach framework, the application of fuzzy logic system across population allows for expert knowledge to be incorporated based on population requirements. The system on its own is robust, but introducing AI into the mix adds a complication of bias as the model, different from fuzzy logic, is trained on population data and tries to match those characteristics when applied to other populations resulting in the obvious bias as observed from our results. One of the techniques used to overcome this problem, is federated learning [78] the technique trains models for each population separately then combine them in another model that can aggregate their weights to develop a more robust model that can deal with heterogenous data. It also ensures privacy as data doesn't need to be shared, rather the model parameters are the ones being shared. The approach has requirements in hardware, software resources and agreement with all participating parties though it offers a better outcome in reducing bias.

The final issue of surveillance is a concern for users whether through LMSs or any other platform that collects data. By surveillance we mean monitoring of user's actions while using certain services or platforms. Although most of us accept being surveilled by CCTVs during our day-to-day interactions, there is some sense of foreboding when it comes to online presence. [77] mentions that due to surveillance, student's participation in online environment could be hindered if they are aware that their actions are monitored (performance bias). Though, we can also argue the same for all other online platforms, which makes the focus of the argument shift from the act of surveilling to how such data is being employed. Is the data being used to provide predictions on student's outcomes? Will their behavior determine their treatment in the institute and so on. So, we can argue that the repercussions of the actions taken

based on that surveillance are what truly matters. In that vein, the A/B/C Teach framework avoids those effects by removing the student outcome(grades) out of the equation, and the metrics measure qualitative indictors of learner/instructor interactions within the environment that do not rely on specific personal information or take student's success as focus of prediction. Affective state could require more data in form of physiological signals (EEG, EDA, video feed) but the expected effect of their inclusion is to provide the instructor with same feedback that could be inferred from interacting with the students in F2F environment.

These issues are of major concern in any field using AI, but as education has far reaching consequences in shaping the learner's future; HEIs and LMS service providers should be more open about their privacy policies, data collection/sharing and data security. This would ensure users are aware of what data being collected and how it is being used, providing understanding and easing the concerns related to privacy and surveillance. On a final note, the deployment of the AI:LMS following the A/B/C Teach framework would yield better results if users have an understanding of the potential uses of its qualitative metrics and how it can improve their learning experience.

6.6 Conclusion

In conclusion, several deep models were developed to replace the Fuzzy-QoI expert knowledge, the models outperformed the FCM-QoI benchmark model for 14 parameters for DB1 testing set and KUQ2021 external validation set. The low performance across KUQ2020 could be attributed to difference in weeks between two institutions when it comes holidays and the high interactions associated with relatively new implementation due to COVID-19 lockdown protocols. This issues of difference in distribution or interactions could also arise when considering applying deep learning models to different levels of HEIs (meso or micro levels), applying across different LMSs, or different domains of expertise (theory-based

education versus practical and technical education). The Fuzzy-QoI as an expert system will not suffer such issues as it applied on weekly basis but may require further evaluation if transferred to another domain (new LMS) or if new interactions are to be considered (virtual classrooms). This brings the issue of interpretability of both models, when considering the 14 parameters it might be easier to interpret the Fuzzy-QoI model, while when considering the system as a whole; explainable AI [41] could offer a better alternative. The value of such model from a pedagogical perspective is the ease of deployment, flexibility, and interpretability if relevant tools are considered. This could allow the models to provide feedback to the user to increase their interactions with the system through relevant recommended interactions. At the same time, it can offer policymakers and educators insights on the dynamics of LMS and can be used as a predictive tool to anticipate the effect of adding new materials on students' interactions. Finally, the discussion of AI implementation of the A/B/C Teach framework highlights some of the precautions needed to ensure transparency, fairness, justice and upholding of user's privacy and autonomy. The chapter concludes the 3rd research question and part of the 4th research question. This leaves the issue of noise, uncertainty, or bias to be further investigated. The next chapter investigates the possibility of applying a deep fuzzy model to utilize its robustness in dealing with noise and uncertainty in data as compared to a deep learning model.

CHAPTER 7 EXPLORATION OF DEEP FUZZY SYSTEMS

7.1 Introduction

From the previous chapter, we noticed that noise, bias, or some degree of uncertainty exists in our models as results of domain shift and presence of intervention in form of switching to online learning as a result of the COVID-19 pandemic. This raises the need to investigate a solution that could possibly deal with such noise and uncertainty in data. As such, the focus of this chapter is mainly on the 4th research question in trying to deal with data uncertainty using a combined approach between deep and fuzzy models. Although a prominent approach like federated learning [78] could be a valid solution, it is not considered due to the small number of datasets, resulting in development of only two models to tackle the issue, where the performance of individual models is expected to be high, as we noticed during implementation of the fuzzy logic system in order to create the ground truth labels for both populations, the fuzzy logic system produced relevant outputs while being applied separately for each population. Therefore, combining the robustness of the fuzzy logic system could help deal with uncertainty. Additionally, when it comes to our data, it is tabular, as previously discussed with no definitive identifying features of what constitutes an input. Meaning, that any combination of values for features could constitute a valid input, as such to evaluate our experimental deep fuzzy models we use image data as they are easier to interpret, and we can also investigate effects of fuzzification across different levels. Another reason for these experiments is that the KU data was not available until late January 2023, thus no definite problem of domain shift was encountered until the external datasets were annotated and models were applied as mentioned previously. So, in this chapter we will start with a literature review into the deep fuzzy models, followed by our experimental methods in form of a combined methodology, results and discussion section for each experiment, and finally application of two models in order to deal with our model bias.

7.2 Literature Review

Deep learning models can learn real world representations of data through simplifying complex features. Deep learning models usually consist of an input layer, several hidden layers, and an output layer [79] . The arrangement of layers, number of neurons and activation functions determine the type of the model. Uncertainty can arise when handling large or big data as deep models require well defined input criteria for training and testing [80]. In order to reduce this uncertainty, fuzzy inference systems are being used in conjunction with deep models. Several approaches to combine the benefits of fuzzy systems into deep learning models were developed, the models are divided into three categories based on approach into [79]; sequential, parallel and integrated. Sequential deep fuzzy models apply the fuzzification process either before, after and/or in between deep model operations. Several techniques utilizing this architecture were developed, one approach [81] used an ensemble of deep learning models followed by a fuzzy inference system in order to evaluate cancer mortality risk in breast cancer. [82]combined autoencoders with fuzzy clustering in order to deal with high dimensional data, the approach is considered a double layered clustering where the autoencoder (encoder part) reduces the dimensionality of the data, followed by further clustering using fuzzy logic. [83] aimed to improve control of nonlinear system by adding a Gaussian fuzzy membership function before the fully connected layer of the deep model, the membership functions act as another layer of non-linearities similar to activation function increasing the model robustness in dealing with non-linear data. Finally, [84] focused on deep model interpretation through passing of the convolutional layer feature maps through a fuzzy clustering from which a new representation could be drawn and data interpretation could be derived.

The 2nd approach is parallel deep fuzzy systems where both systems run parallel to each other, then combine in later steps through another deep model (neural network). [80]

developed a combined deep learning and fuzzy model utilizing the expert knowledge from the fuzzy inference system which combines with the fully connected layer of the deep model as concatenated input to solve image classification problems. [85]used a similar approach but the deep learning model was a CNN instead of MLP to predict traffic flow (image regression problem). Similarly, [86] applied MLP with Pythagorean fuzzy numbers and intuitionistic fuzzy sets in order to predict cloud resources needed to be purchased (from user side). Finally, [87] used restricted Boltzmann machines as deep learning model and FCM followed by fuzzy memberships then combines both at a fully connected layer that act as the rule layer to predict user's movement in indoor living environment.

The 3rd and last approach is the integrated deep fuzzy models. In this approach the model structure is modified to take advantage of either deep or fuzzy systems. [88] considered the filter output of convolution as a fuzzy set which is given to a normalization layer to combine all branches of output filters of convolution. The method aims to find the best kernel parameters for image classification tasks. [89] used fuzzy clustering with dictionary deep learning in unsupervised learning method to learn changes in satellite or drone images. [90] used a unique approach of connecting inputs to ensemble of fuzzy systems and deep learning models that are also interconnected, meaning inputs to ensemble in parallel and outputs of first membership group goes to rule layer that also receives input, then to hidden layer that also receives input and finally a summing layer to predict lung tumor movements based on breathing parameters. [91]utilized the stacked generalization scheme to a develop a deep fuzzy model from Takeo-Sugeno-Kang (TSK) inference system. The model stacks several TSKs in layers and uses adversarial learning to derive the IF THEN rules against data labels using five fuzzy Gaussian membership functions. Each derivation of TSK layer is combined with original input as input to next layer which also are pitted against labels and so on. The model adapts the deep models learning capabilities into the fuzzy logic allowing it to derive data representations which are

limited by the membership functions. As seen, integrated approaches modify the model structure to take advantage of the strength of fuzzy system in order to reduce uncertainty in data. Figure (18 – 20) show sequential, parallel and integrated models including the stacked generalization. From the above literature review, we can reach the conclusion that although deep fuzzy systems offer some utility in handling uncertainty, they are often designed in order to deal with specifically designed problems. And so, in order to deal with noise and uncertainty in our data, we need to test the feasibility of application onto a quantifiable system which in our case is the image classification task.

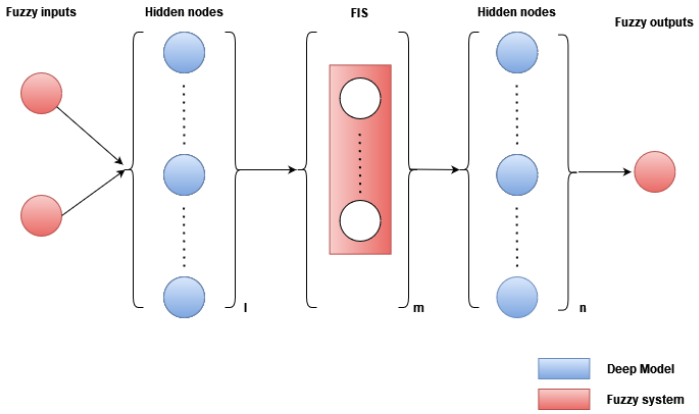


Figure 18 Sequential deep fuzzy system.

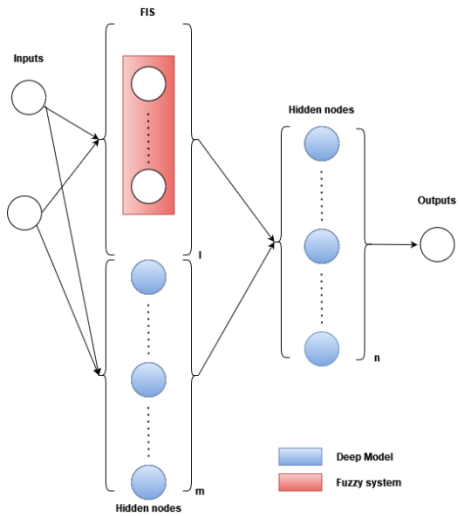


Figure 19 Parallel deep fuzzy system.

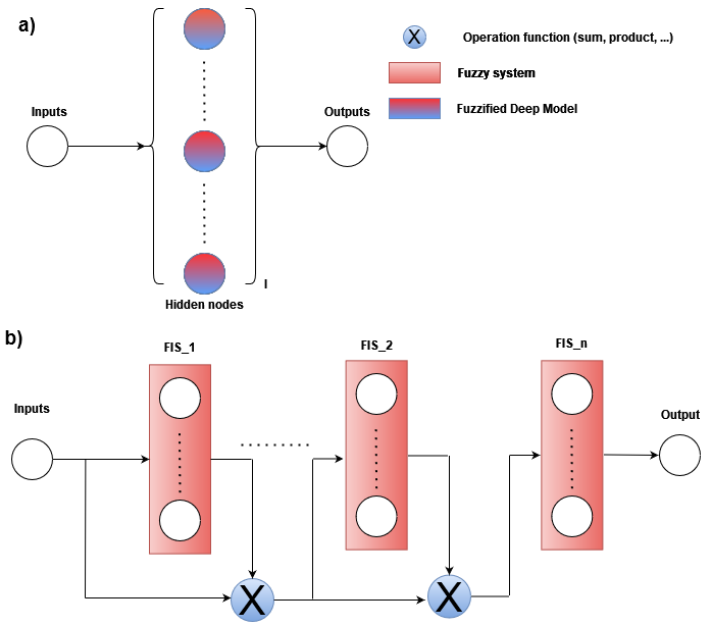


Figure 20 a) Integrated deep fuzzy system, b) stacked generalization model.

7.3 Methodology, Results and Discussion

As mentioned in the introduction, we will discuss the three experiments for deep fuzzy systems, using two main dataset MNSIT and CIFAR10 both are available through the keras python library. MNIST dataset contains grey scaled images of handwritten numbers from 0 to 9, while CIFAR-10 contains images of airplanes, birds, cars, cats, deer, dogs, frogs, horses, ships, and trucks. The next subsections will discuss trials applied mostly on CIFAR-10 as the task for MNIST classification achieves high performance even using simple MLP. These trials are part of investigative effort in order to answer the 4th research question which deals with bias and noise in our data (external validation data).

7.3.1 Trail 1 With Deep Fuzzy Logic System

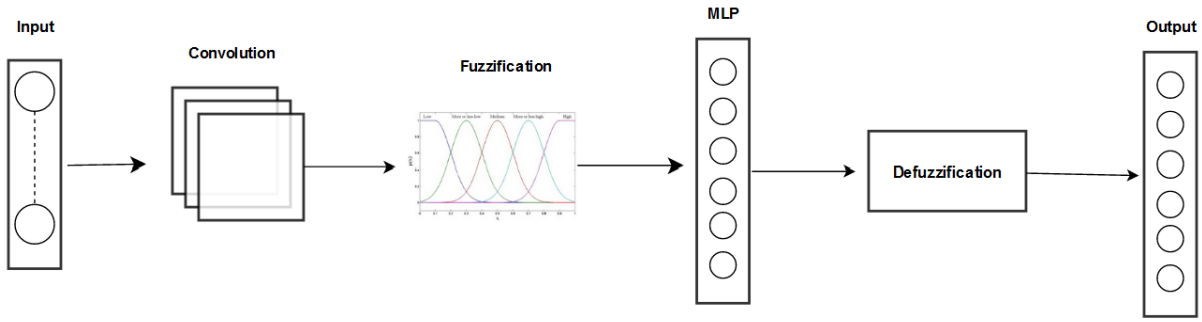


Figure 21 Deep fuzzy model trial 1.

Trials for development of a deep fuzzy model were done. The model in Figure. 21 consists of input image, convolutional layer, fuzzification using five Gaussian membership functions, flattening followed by MLP layer for rule inference, defuzzification using centroid method and finally MLP for output classification. The method excluded activation functions and max pooling from convolution layers, aiming to create a fuzzy system that can replace those functions. The main challenges are the increased number of parameters at bottleneck (flatten layer), and loss of spatial correlation at MLP layer. The MNIST and CIFAR10 are being used to develop the models. Fig 22 shows some of the outputs of the convolution layer (feature maps). The issue of the bottleneck layer causes heavy memory consumption, and to overcome such problem; development of a customized MLP layer that functions as the rule layer in fuzzy logic system is undertaken. In such approach, the MLP is initiated but the weight multiplication is done across slices that are formed of filters x fuzzy sets, so if the number of filters is 10 and we constructed three output membership functions, the weight connections that will be used are (5x10x3) for each pixel (32x32) for a total of (153,600) as compared to (157,491,200) without this modification.

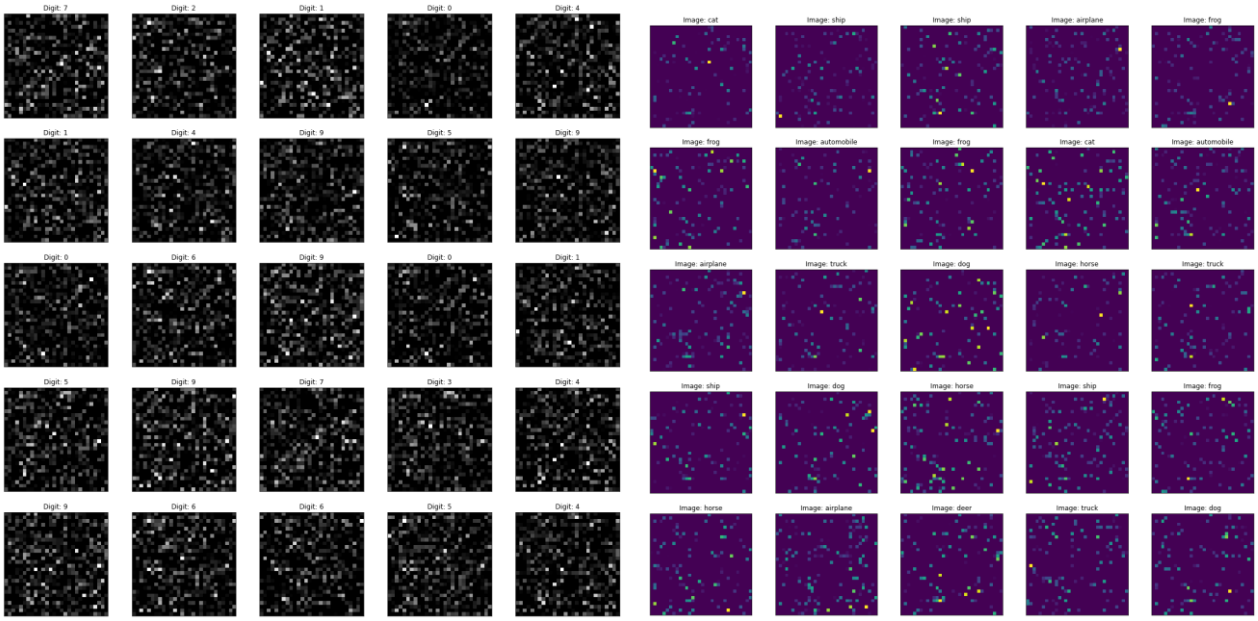


Figure 22 Feature map from MNIST (left) and CIFAR-10 (right).

7.3.2 Trial 2 Using Kernel MLP

So, to follow up on the development of the modified MLP layer with custom connections and after actually forming the layer. By using it in a model it was confirmed that although the weights are not being updated (zero value), they nonetheless were initialized in memory leading to the same conclusion as the previous trial. Figure 23 shows a plot of the used neural network, Example of the Feature map output from the defuzzification of the modified kernel MLP replacing the MLP in Figure 21 is shown in Figure 24. The network had 0.643 accuracy for training and 0.605 for testing after 5 epochs. Increasing the number of blocks caused the performance of the model to be reduced reaching ~40% accuracy for testing set when using three consecutive blocks. This could be explained by the loss of the filters during the center of sum method application. Which restores the image extracted features into a single channel of defuzzified output. In a sense it is removing the features we extracted from the feature maps. Several trials to overcome the problem were done, including passing the

convolution unchanged forward, while having the fuzzy system in parallel so we can maintain the features extracted from convolution which are then concatenated at the fully connected layer. The discussion of the trial is the next subsection.

7.3.3 Trial 3 With Parallel Deep Fuzzy Model

To overcome the previous problem, a new approach to overcome this hurdle was implemented which through a concept similar to Xception transfer model architecture [92]but in this case we use the fuzzy logic output as to run parallel to the convolution model thus retaining the convolutional feature maps and adding extra features that will be concatenated with final feature map along with other fuzzy blocks before the flattening at MLP layer. The model reduced the memory usage from the previously created rule layer (kernel MLP) through utilizing 3D convolution; the 3D convolution is applied in two steps, the 1^{st} using a (1,1,3) kernel pushing the model to learn some features across the fuzzified channel as the extra 3D filters. Then another (2,2,3) kernel with (2,2,4) strides is used. The model achieved 0.1 for both training and testing and the main reason could be the increase in number of defuzzy memberships before applying center of sum which is equal to the number of filters. Reducing the number of filters produced similar results to kernel MLP but the model runs faster, around (1 minute compared to 10 minutes/epoch). This allowed for increasing of batch size which was limited to less than 20 and filters which constrained to maximum of 9 in order for device to not run out of memory.

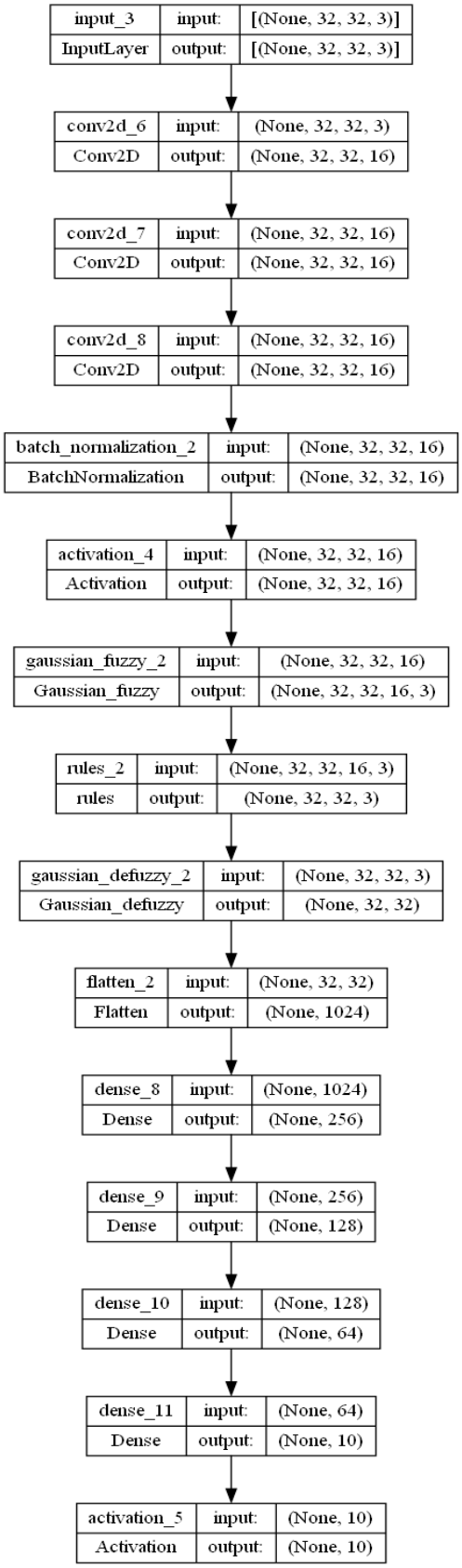


Figure 23 Kernel MLP Architecture

So, this configuration overall has better computational performance though further tweaking is required. Figures (25, 28) show the output of the 1st reshape layer (feature map) as the defuzzy layer output is flattened to make it easier for concatenation afterwards and the parallel deep fuzzy model. Figure 26 shows the original CIFAR 10 output for the first 20 images of the testing set. The filter shows a blank image which means that it is not learning any representations, hence the low performance. Figure 27 shows the gaussian membership functions used for most of the models. These trials show that although fuzzification has some effect on the output of the convolution, to replace the functionality of any of its components a similar approach that maintains the validity of the extracted features and also provides higher dimensionality in form of fuzzy memberships is needed. The computational cost should also be a part of the equation as high number of parameters similar to transfer learning model would be inefficient if similar performance is not achieved. A Final note to consider is to focus the fuzzification where it is needed to deal with problems like uncertainty and noise, which is being applied and discussed in the next subsection.

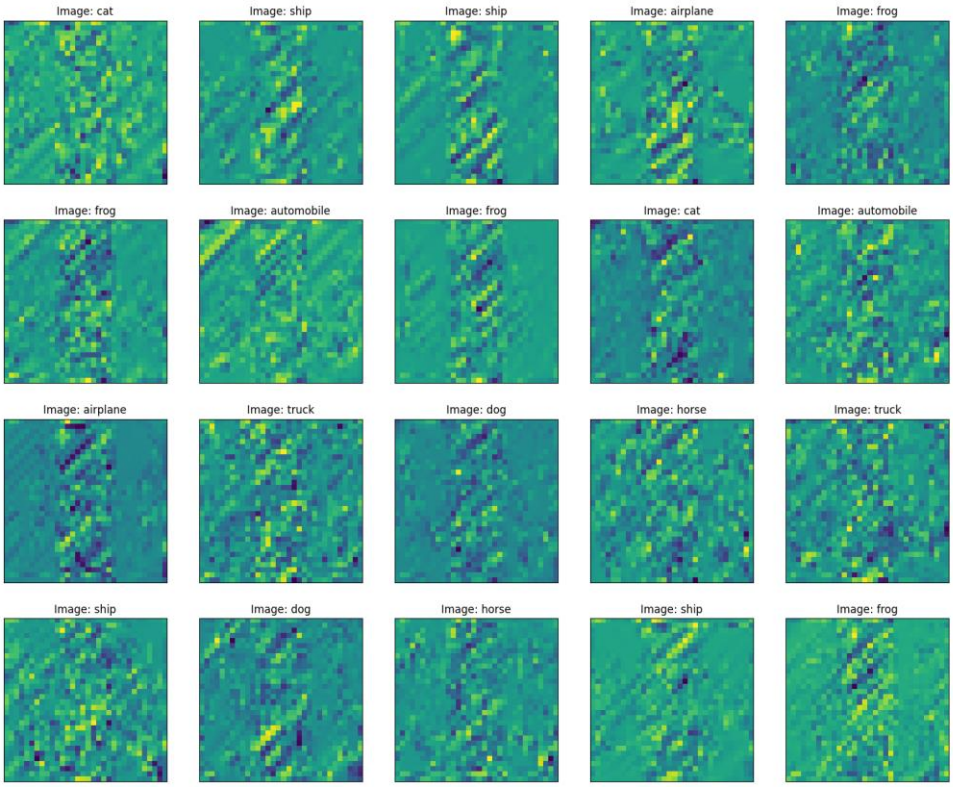


Figure 24 Feature Map of Kernel MLP

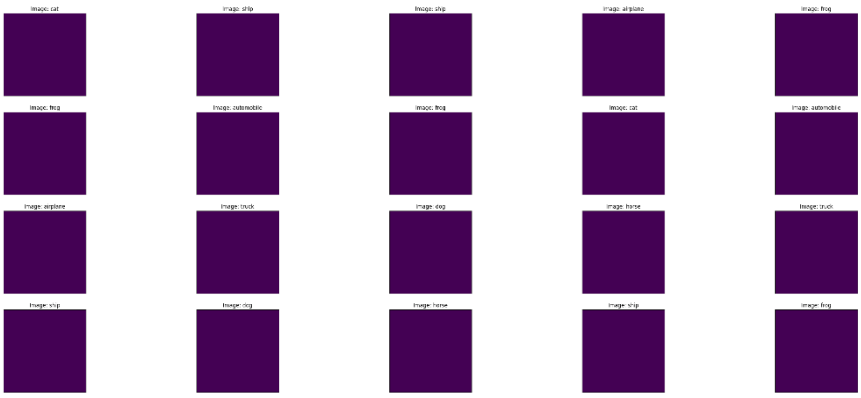


Figure 25 Blank feature map of parallel deep fuzzy model

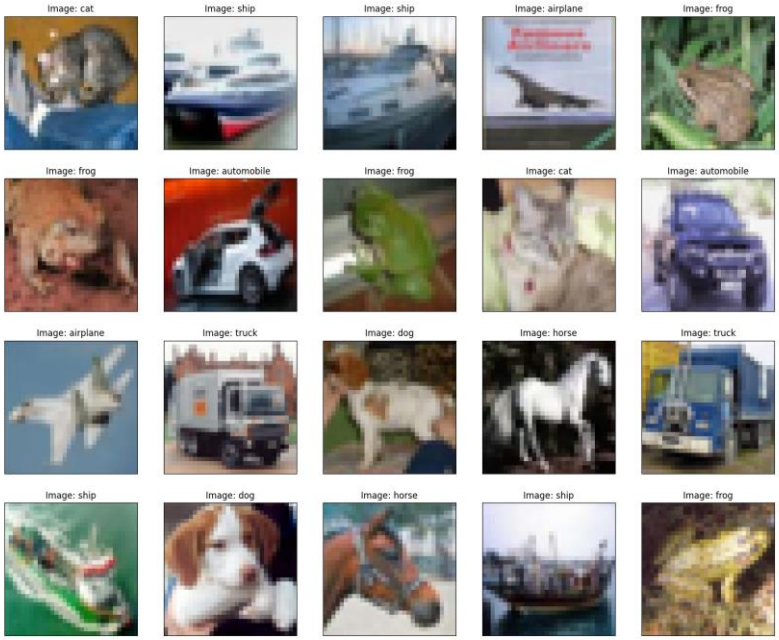


Figure 26 CIFAR 10 first 20 testing images

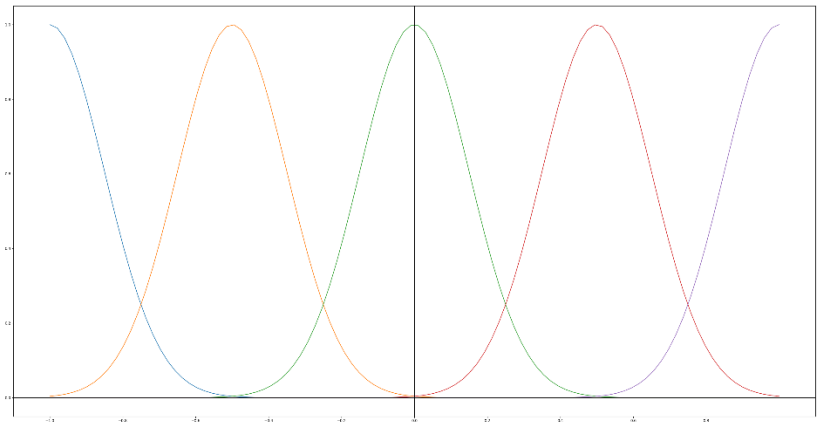


Figure 27 Fuzzy Gaussian membership functions (negative low to positive high)

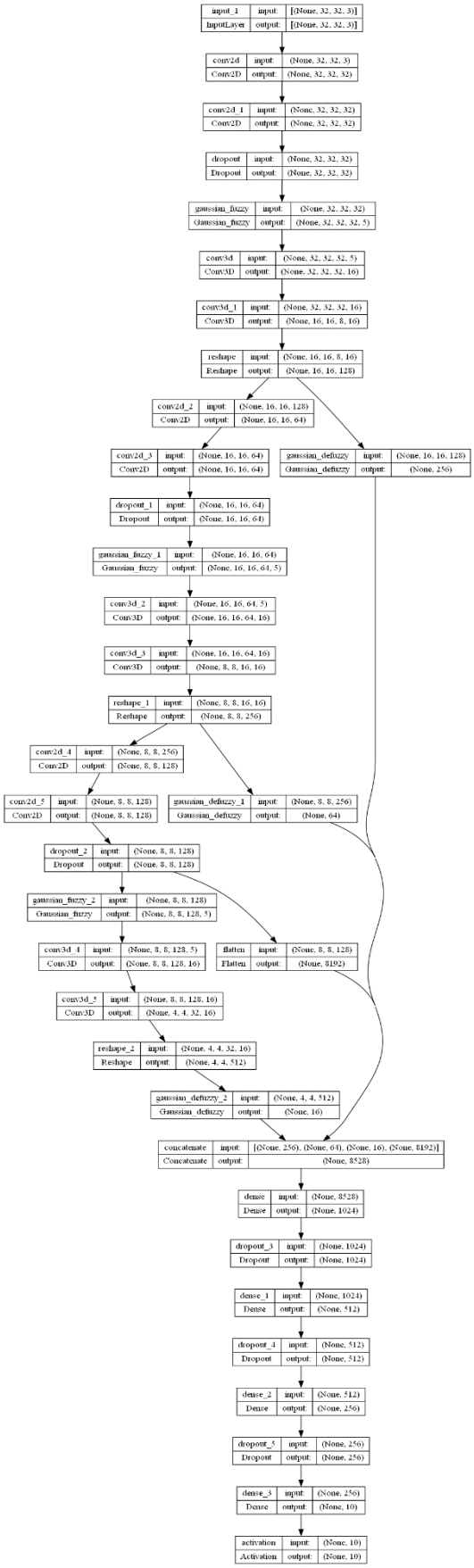


Figure 28 Parallel Deep Fuzzy Model

7.4 Fuzzification of QoI Deep Models

7.4.1 Introduction

Before applying any fuzzification to our deep models, we need to identify the exact effect of bias from the perspective of model performance. The model performed poorly on the KUQ2020 dataset, while having good performance (low RMSE) for KUQ2021. The main explanation is the difference in magnitude of interactions across KUQ2020 and DB1. KUQ2021 had low interactions within the maximum limit of DB1. The application of robust scaling and mix-max scaling reduces the effect of outliers while ensuring all values are within [0 -1]. If we project the maximum interactions of DB1 to KUQ2020 we notice that some features end up with values greater than 1. Figure 29 shows an example of what the model is trained on versus what it receives. So, our main focus should be to allow the model to understand values that it hasn't seen (greater than 1) while still providing valid output for those features.

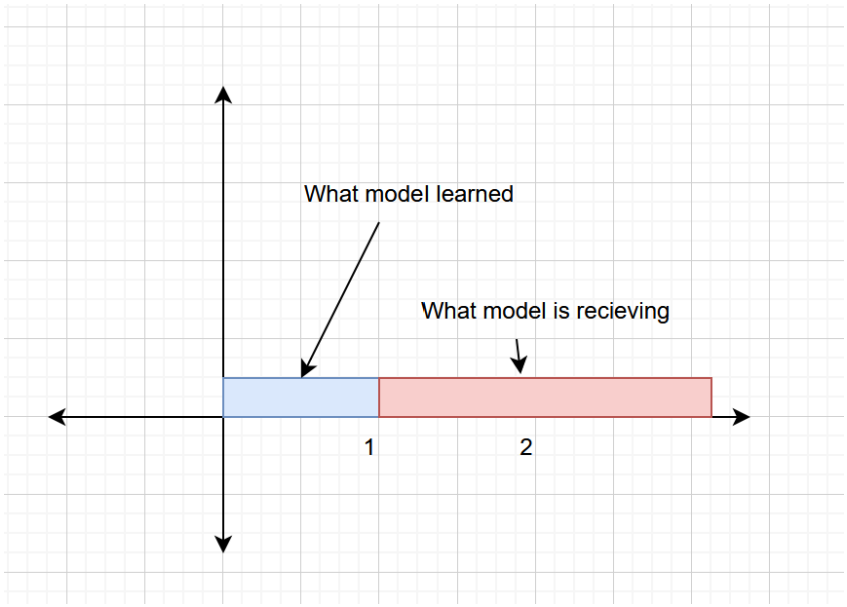


Figure 29: Bias in model performance.

7.4.2 Methodology

To allow the model to encounter values it never saw before but at the same time produce meaningful output, the introduction of a fuzzy membership to the input was explored. The best outcomes were from utilizing periodic signals (sine wave). The sine waves as shown in Figure 30, modify the values of the features to be within [-1, 1]. Though the repetition of the signal causes certain values (0, 0.5, 1) to have same output (0). As such, we introduced two main approaches to allow the model to learn significance of these values. The first was to use different periods in a way simulating different membership functions. And the second was to introduce signals with different phases. Figures (31,32) show examples of these signals. Though both figures showcase 3 signals per figure, while the model utilized five different signals. The models also applied MSE and log cosh as loss functions as opposed to MSE in original models.

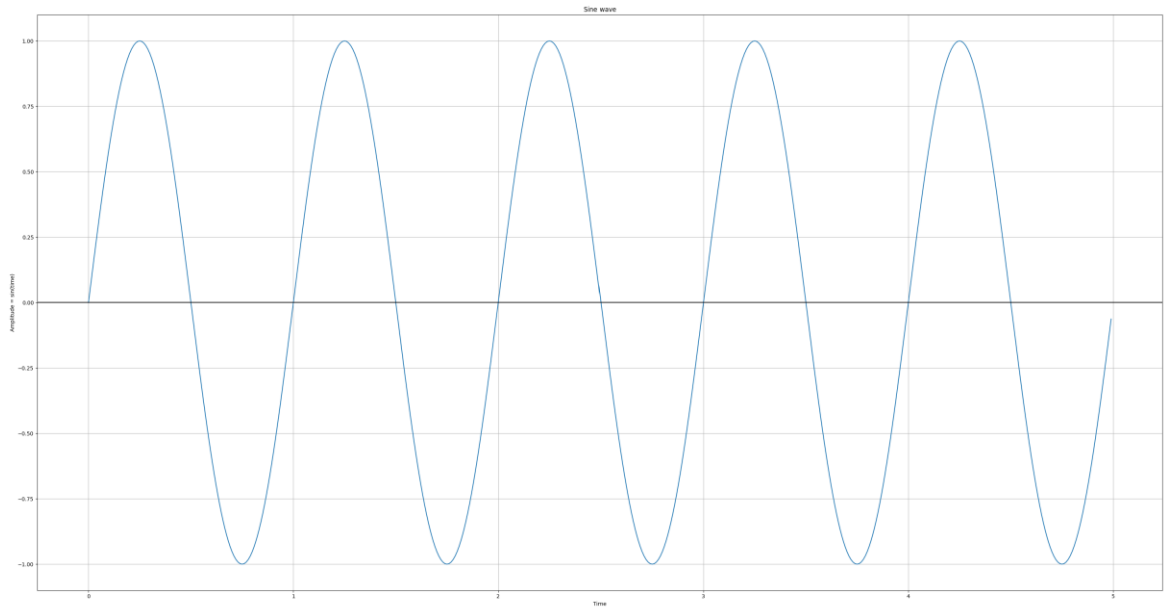


Figure 30: Sine wave.

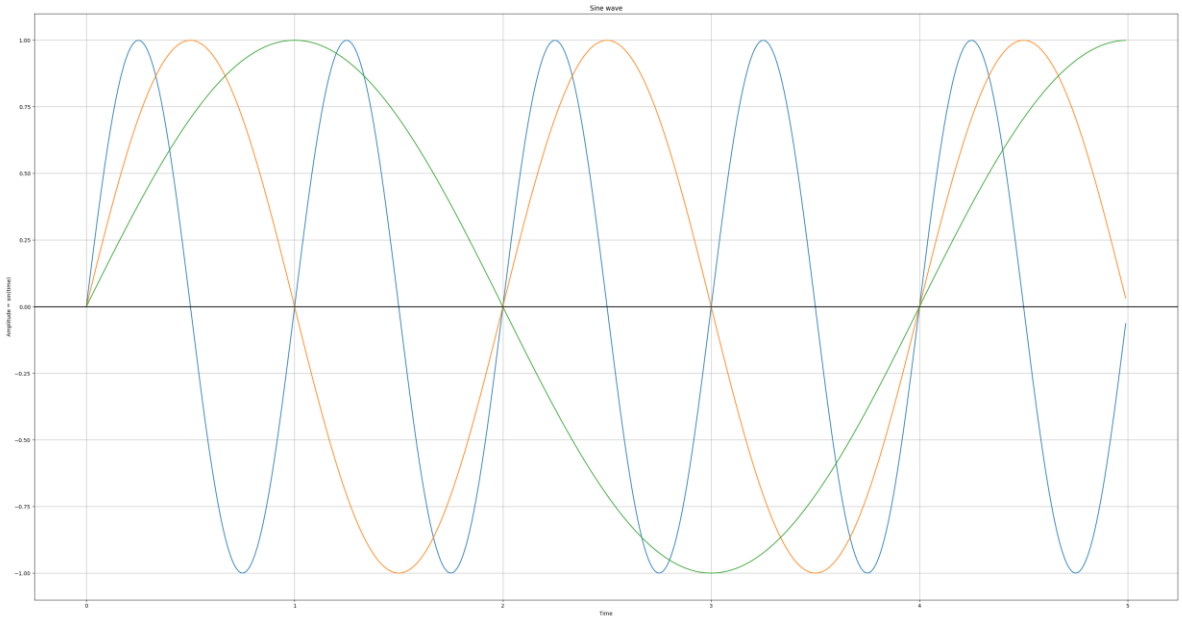


Figure 31: Sine waves with different periods.

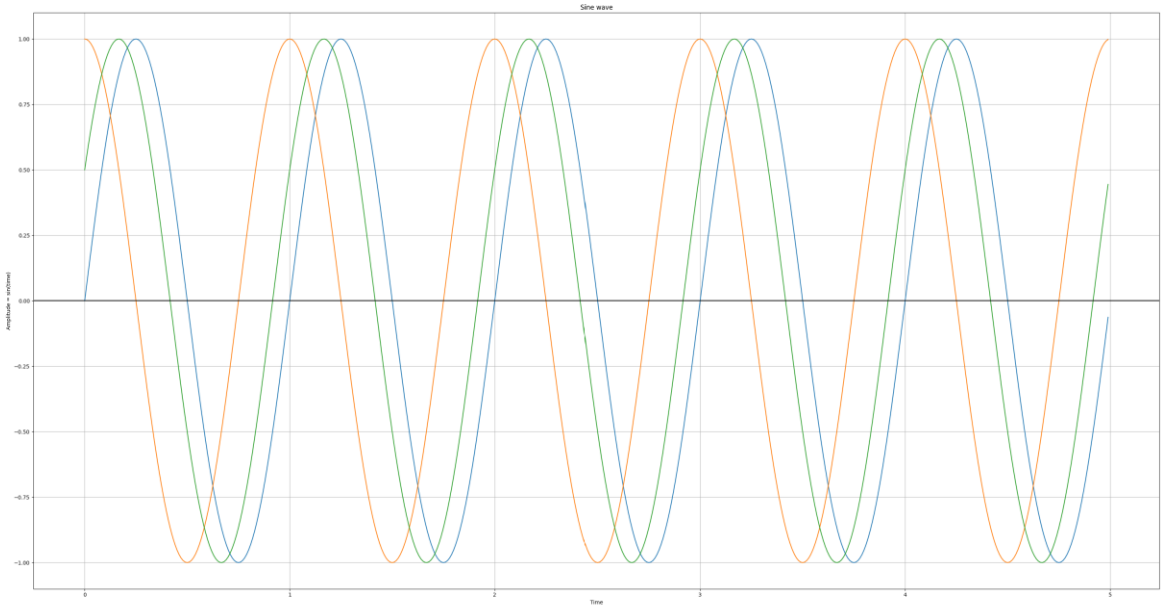


Figure 32: Sine waves with different phases.

7.4.3 Results and Discussion

Table 17 Results for fuzzification of deep models.

	RMSE	MSE	MAE	R2
Phase (Cosh)	0.200279	0.040147	0.118143	-2.571254
Period (Cosh)	**0.166168**	**0.027764**	**0.082366**	-1.469737
Phase (MSE)	0.20012	0.040063	0.118246	-2.563776
Period (MSE)	0.175028	0.030758	0.091324	-1.736001
Case 1 (Pyramid)	0.19624	0.03854	0.11162	-2.42811

The results in Table 17 show ~15% improvement as compared to the base deep model for the memberships consisting of different periods using log cosh as loss function. The solution still had a high R^2 negative value, though RMSE error was reduced. Although the results indicate a minor success in the approach it doesn't solve the issue of model bias but rather alleviates it to a certain degree. This may suggest that a more in detail fuzzy memberships developed specifically to deal with this problem based on expert knowledge (accounts for the differences between the two populations) is required. On the other hand, applying and training individual models through federated learning approach may provide a more robust solution that can deal with the heterogeneous data, though for evaluation a wider range of datasets from different populations is required.

7.5 Conclusion

Several experimental trials were done to investigate the effect of deep fuzzy models on image classification task using mainly CIFAR-10 dataset. These experiments were done in order to answer the 4^{th} research question and the problems highlighted due to domain shift and external validation in the previous chapter. The developed models although shown difference in performance due to position of the fuzzy logic component within the deep model, they also highlight some of the problems that would be encountered when increasing the dimensionality of the data using the fuzzy memberships (curse of dimensionality). The suggested solutions tried to reduce the dimensionality while trying to pass through as much information after the

defuzzification process. As follow up work, the target of fuzzification should aim to replace specific components of the convolution or deal with narrow problems in image classification, for example foggy image classification. From our perspective with regards to educational data, the difficulty lies in defining what constitutes valid input and how to track its changes while developing the deep learning model. The application of deep fuzzy system to KUQ2020 data helped reduce the effect of bias, though federated learning which is used in structured healthcare data could be the best solution. With this we answered the 4th research question to a great extent, and the next chapter concludes the thesis work, its contribution and future work.

CHAPTER 8 CONCLUSION

8.1 Conclusion

In conclusion, we developed two new datasets based on the backend data extracted from Khalifa University's Moodle LMS for the period of the COVID-19 lockdown. The development of these datasets allowed for a proof of concept of domain shift across LMS frontend and backend and also across HEIs. Additionally, several deep learning models were developed that successfully learned the expert knowledge representation of the Fuzzy-QoI models for both 14 and 112 parameters. These models can be used to predict the value of QoI in order to provide feedback to the user (learner or instructor) and also provide insight into the workings of LMS for policymakers and stakeholders. The interpretation based on these models if combined with sound psychological and pedagogical approaches can possibly allow for improved learning experience. We noted the effect of different motivators and effect of gamification as a recommendation or feedback system to the users and how to possibly optimize the effect of deployment of such system. Finally, we investigated several deep fuzzy models in order to deal with uncertainties and noise encountered during the external validation.

The limitations of this work are highlighted by the previously mentioned point, which is the uncertainty raised due to migration of the deep model to another domain (in this case another university during an intervention period), the suggested approaches could help reduce its effects, though complete elimination requires a more robust system of which a deep fuzzy model and federated learning could be considered for target solutions. Also, in order to annotate all interactions included in other domains, it has to be done manually at the categorization step, requiring some sort of expert intervention in order to include those parameters as part of the 12 action parameters input of the model. Of course, only categorization into 12 actions parameters is required to derive the Fuzzy-QoI model, but for a deep learning model, the model needs to

be retrained if the raw metrics are to be considered (every time the metrics are increased the model needs to be retrained with new features as extra columns). As such, although the system allows for generalizability, adaptability, ease of deployment and interpretation it also requires careful categorization of all discovered interactions. Finally, we discussed the implications of AI:LMS application under the A/B/C Teach framework within the concepts of privacy, bias, fairness, surveillance and autonomy.

8.2 Future Work

Our future work will first include acquisition of data from both Moodle and BlackBoard to further improve model performance and investigate domain shift issues, this could be done by cooperation with CTL and IT department to procure more data. Second focus will be on the investigation into federated learning after acquisition of more data to allow training of several independent models before aggregation. Third focus would be the development of study protocols to investigate the effects motivators and gamification as part of a QoI feedback system. Finally, would be the development of study protocols for both QoI and AS as part of the A/B/C teach framework.

References

[1] D Weaver, C Spratt, CS Nair. Academic and student use of a learning management system: Implications for quality, Australasian journal of educational technology. 24 (2008).

[2] W Watson, SL Watson. An argument for clarity: What are learning management systems, what are they not, and what should they become. (2007).

[3] N Fathema, D Shannon, M Ross. Expanding the Technology Acceptance Model (TAM) to examine faculty use of Learning Management Systems (LMSs) in higher education institutions. Journal of Online Learning & Teaching. 11 (2015).

[4] MN AL-Nuaimi, OS Al Sawafi, SI Malik, M Al-Emran, YF Selim. Evaluating the actual use of learning management systems during the covid-19 pandemic: an integrated theoretical model, Interactive Learning Environments. (2022) 1-26.

[5] A Awad, A AlShehhi, SB Dias, SJ Hadjileontiadis, LJ Hadjileontiadis, On Modeling LMS Users' Quality of Interaction Using Temporal Convolutional Neural Networks, (2023) 145-154.

[6] R Luckin, W Holmes, M Griffiths, LB Forcier. Intelligence unleashed: An argument for AI in education, (2016).

[7] SB Dias, J Bidarra. A/B/C-teach framework: the application of AI in blended learning, The Envisioning Report for Empowering Universities. (2020) 22-24.

[8] A Jobin, M Ienca, E Vayena. The global landscape of AI ethics guidelines, Nature Machine Intelligence. 1 (2019) 389-399.

[9] SB Dias, JA Diniz. FuzzyQoI model: A fuzzy logic-based modelling of users' quality of interaction with a learning management system under blended learning, Comput.Educ. 69 (2013) 38-59.

[10] SB Dias, S Hadjileontiadou, JA Diniz, L Hadjileontiadis, Towards an intelligent learning management system: The A/B/C-TEACH approach, (2019) 397-411.

[11] DR Garrison, H Kanuka. Blended learning: Uncovering its transformative potential in higher education, The internet and higher education. 7 (2004) 95-105.

[12] Pelletier, K., Brown, M., Brooks, D.C., McCormack, M., Reeves, J., Arbino, N., Bozkurt, A., Crawford, S., Czerniewicz, L., Gibson, R., Linder, K., Mason, J. & Mondelli, V., EDUCAUSE Horizon Report Teaching and Learning Edition., Boulder, CO: EDU. 2021.

[13] K Pelletier, M McCormack, J Reeves, J Robert, N Arbino, C Dickson-Deane, et al. No title, 2022 EDUCAUSE Horizon Report Teaching and Learning Edition. (2022).

[14] A Shahzad, R Hassan, AY Aremu, A Hussain, RN Lodhi. Effects of COVID-19 in E-learning on higher education institution students: the group comparison between male and female, Quality & quantity. 55 (2021) 805-826.

[15] K Chrysafiadi, M Virvou. Student modeling approaches: A literature review for the last decade, Expert Syst.Appl. 40 (2013) 4715-4729.

[16] OA Ponce, J Gómez, N Pagán. Current scientific research in the humanities and social sciences: central issues in educational research, Eur.J.Sci.Theol. 15 (2019) 81-95.

[17] PCd Oliveira, CJCdA Cunha, MK Nakayama. Learning Management Systems (LMS) and e-learning management: an integrative review and research agenda, JISTEM-Journal of Information Systems and Technology Management. 13 (2016) 157-180.

[18] F Colace, M De Santo, L Greco. E-learning and personalized learning path: A proposal based on the adaptive educational hypermedia system, International Journal of Emerging Technologies in Learning (Online). 9 (2014) 9.

[19] E Alqurashi. Predicting student satisfaction and perceived learning within online learning environments, Distance education. 40 (2019) 133-148.

[20] M Hussain, W Zhu, W Zhang, SMR Abidi. Student engagement predictions in an e-learning system and their impact on student course assessment scores, Computational intelligence and neuroscience. 2018 (2018).

[21] F Bonafini, C Chae, E Park, K Jablokow. How much does student engagement with videos and forums in a MOOC affect their achievement? Online Learning Journal. 21 (2017).

[22] SB Dias, SJ Hadjileontiadou, JA Diniz, LJ Hadjileontiadis. Computer-based concept mapping combined with learning management system use: An explorative study under the self-and collaborative-mode, Comput.Educ. 107 (2017) 127-146.

[23] SB Dias, JA Diniz, LJ Hadjileontiadis, SB Dias, JA Diniz, LJ Hadjileontiadis. On Modeling Users' Quality of Interaction with LMS Using Fuzzy Logic, Towards an Intelligent Learning Management System Under Blended Learning: Trends, Profiles and Modeling Perspectives. (2014) 151-168.

[24] SB Dias, SJ Hadjileontiadou, J Alves Diniz, LJ Hadjileontiadis, FuzzyQoI-Based Estimation of the Quality of Interaction in Online Learning Amid Covid-19: A Greek Case-Study, (2021) 249-262.

[25] SB Dias, JA Diniz, LJ Hadjileontiadis, Exploring B-learning scenarios using fuzzy logic-based modeling of users' LMS quality of interaction in ergonomics and psychomotor rehabilitation academic courses, (2014) 233-243.

[26] SB Dias, JA Diniz, LJ Hadjileontiadis, Dynamic Fuzzy Logic-Based Quality of Interaction within Blended-Learning: The Rare and Contemporary Dance Cases., ERIC 2014.

[27] SB Dias, SJ Hadjileontiadou, J Diniz, LJ Hadjileontiadis. DeepLMS: a deep learning predictive model for supporting online learning in the Covid-19 era, Scientific reports. 10 (2020) 1-17.

[28] SB Dias, LJ Hadjileontiadis, JA Diniz, On enhancing blended-learning scenarios through fuzzy logic-based modeling of users' LMS quality of interaction the rare & contemporary dance paradigms, 2 (2014) 765-772.

[29] SB Dias, SJ Hadjileontiadou, LJ Hadjileontiadis, JA Diniz. Fuzzy cognitive mapping of LMS users' Quality of Interaction within higher education blended-learning environment, Expert Syst.Appl. 42 (2015) 7399-7423.

[30] B Kosko. Fuzzy cognitive maps, International journal of man-machine studies. 24 (1986) 65-75.

[31] X Zhai, X Chu, CS Chai, MSY Jong, A Istenic, M Spector, et al. A Review of Artificial Intelligence (AI) in Education from 2010 to 2020, Complexity. 2021 (2021) 1-18.

[32] D Kem. Personalised and adaptive learning: Emerging learning platforms in the era of digital and smart learning, International Journal of Social Science and Human Research. 5 (2022) 385-391.

[33] A Muniasamy, A Alasiry. Deep learning: The impact on future eLearning, International Journal of Emerging Technologies in Learning (Online). 15 (2020) 188.

[34] A Klašnja-Milićević, M Ivanović, A Nanopoulos. Recommender systems in e-learning environments: a survey of the state-of-the-art and possible extensions, Artif.Intell.Rev. 44 (2015) 571-604.

[35] G George, AM Lal. Review of ontology-based recommender systems in e-learning, Comput.Educ. 142 (2019) 103642.

[36] JK Tarus, Z Niu, G Mustafa. Knowledge-based recommendation: a review of ontology-based recommender systems for e-learning, Artif.Intell.Rev. 50 (2018) 21-48.

[37] NW Rahayu, R Ferdiana, SS Kusumawardani. A systematic review of ontology use in E-Learning recommender system, Computers and Education: Artificial Intelligence. (2022) 100047.

[38] N Rubens, D Kaplan, T Okamoto, E-Learning 3.0: anyone, anywhere, anytime, and AI, (2014) 171-180.

[39] W Holmes, K Porayska-Pomsta, K Holstein, E Sutherland, T Baker, SB Shum, et al. Ethics of AI in education: Towards a community-wide framework, International Journal of Artificial Intelligence in Education. (2021) 1-23.

[40] LH Gilpin, D Bau, BZ Yuan, A Bajwa, M Specter, L Kagal, Explaining explanations: An overview of interpretability of machine learning, (2018) 80-89.

[41] P Linardatos, V Papastefanopoulos, S Kotsiantis. Explainable ai: A review of machine learning interpretability methods, Entropy. 23 (2020) 18.

[42] DL Whaley III. No title, The interquartile range: Theory and estimation. (2005).

[43] W Python. Python, Python Releases Wind. 24 (2021).

[44] O Kramer, O Kramer. Scikit-learn, Machine learning for evolution strategies. (2016) 45-53.

[45] S Matlab. Matlab, The MathWorks, Natick, MA. (2012).

[46] MW Gardner, SR Dorling. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences, Atmos.Environ. 32 (1998) 2627-2636.

[47] Y Yu, X Si, C Hu, J Zhang. A review of recurrent neural networks: LSTM cells and network architectures, Neural Comput. 31 (2019) 1235-1270.

[48] S Kiranyaz, O Avci, O Abdeljaber, T Ince, M Gabbouj, DJ Inman. 1D convolutional neural networks and applications: A survey, Mechanical systems and signal processing. 151 (2021) 107398.

[49] C Pelletier, GI Webb, F Petitjean. Temporal convolutional neural network for the classification of satellite image time series, Remote Sensing. 11 (2019) 523.

[50] S Bai, JZ Kolter, V Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling, arXiv preprint arXiv:1803.01271. (2018).

[51] T Chai, RR Draxler. Root mean square error (RMSE) or mean absolute error (MAE), Geoscientific model development discussions. 7 (2014) 1525-1534.

[52] DJ Ozer. Correlation and the coefficient of determination. Psychol.Bull. 97 (1985) 307.

[53] I Cohen, Y Huang, J Chen, J Benesty, J Benesty, J Chen, et al. Pearson correlation coefficient, Noise reduction in speech processing. (2009) 1-4.

[54] R Artusi, P Verderio, E Marubini. Bravais-Pearson and Spearman correlation coefficients: meaning, test of hypothesis and confidence interval, Int.J.Biol.Markers. 17 (2002) 148-151.

[55] SM Lundberg, S Lee. A unified approach to interpreting model predictions, Advances in neural information processing systems. 30 (2017).

[56] SM Lundberg, G Erion, H Chen, A DeGrave, JM Prutkin, B Nair, et al. From local explanations to global understanding with explainable AI for trees, Nature machine intelligence. 2 (2020) 56-67.

[57] J Snoek, O Rippel, K Swersky, R Kiros, N Satish, N Sundaram, et al., Scalable bayesian optimization using deep neural networks, (2015) 2171-2180.

[58] T Chen, T He, M Benesty, V Khotilovich, Y Tang, H Cho, et al. Xgboost: extreme gradient boosting, R package version 0.4-2. 1 (2015) 1-4.

[59] J Friedman, T Hastie, R Tibshirani. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors), The annals of statistics. 28 (2000) 337-407.

[60] A Mazarakis, P Bräuer, Gamification of an open access quiz with badges and progress bars: An experimental study with scientists. (2020) 62-71.

[61] H Cigdem, M Ozturk. Factors affecting students' behavioral intention to use LMS at a Turkish post-secondary vocational school, International Review of Research in Open and Distributed Learning. 17 (2019) 276-295.

[62] AM Lavin, L Korte, TL Davies. The impact of classroom technology on student behavior, Journal of Technology Research. 2 (2011) 1-13.

[63] K Aleksic-Maslac, B Sinkovic, P Vranesic. Influence of gamification on student engagement in education, International Journal of Education and Learning Systems. 2 (2017).

[64] F Faiella, M Ricciardi. Gamification and learning: a review of issues and research, Journal of e-learning and knowledge society. 11 (2015).

[65] I Chen, G Teng, C Chen, T Lan, H Liu. The Autonomic Progress Bar Motivates Treatment Completion for Patients of Stimulant Use Disorder and Cannabis Use Disorder, Frontiers in psychiatry. 10 (2020) 944.

[66] AM Isen, J Reeve. The influence of positive affect on intrinsic and extrinsic motivation: Facilitating enjoyment of play, responsible work behavior, and self-control, Motiv.Emotion. 29 (2005) 295-323.

[67] A Elangovan, J He, K Verspoor. Memorization vs. generalization: quantifying data leakage in NLP performance evaluation, arXiv preprint arXiv:2102.01818. (2021).

[68] D Berrar. No title, Cross-Validation. (2019).

[69] G Klambauer, T Unterthiner, A Mayr, S Hochreiter. Self-normalizing neural networks, Advances in neural information processing systems. 30 (2017).

[70] DP Kingma, J Ba. Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980. (2014).

[71] P Baldi, PJ Sadowski. Understanding dropout, Advances in neural information processing systems. 26 (2013).

[72] F Chollet. No title, keras. (2015).

[73] DM Allen. Mean square error of prediction as a criterion for selecting variables, Technometrics. 13 (1971) 469-475.

[74] P Sharma. Blended learning, ELT journal. 64 (2010) 456-458.

[75] M Tayebinik, M Puteh. Blended Learning or E-learning? arXiv preprint arXiv:1306.4085. (2013).

[76] S Hrastinski. What do we mean by blended learning? TechTrends. 63 (2019) 564-569.

[77] S Akgun, C Greenhow. Artificial intelligence in education: Addressing ethical challenges in K-12 settings, AI and Ethics. (2021) 1-10.

[78] L Li, Y Fan, M Tse, K Lin. A review of applications in federated learning, Comput.Ind.Eng. 149 (2020) 106854.

[79] R Das, S Sen, U Maulik. A survey on fuzzy deep neural networks, ACM Computing Surveys (CSUR). 53 (2020) 1-25.

[80] Y Deng, Z Ren, Y Kong, F Bao, Q Dai. A hierarchical fused fuzzy deep neural network for data classification, IEEE Trans.Fuzzy Syst. 25 (2016) 1006-1012.

[81] C Yang, S Moi, M Hou, L Chuang, Y Lin. Applications of deep learning and fuzzy systems to detect cancer mortality in next-generation genomic data, IEEE Trans.Fuzzy Syst. 29 (2020) 3833-3844.

[82] Q Feng, L Chen, CP Chen, L Guo. Deep fuzzy clustering—a representation learning approach, IEEE Trans.Fuzzy Syst. 28 (2020) 1420-1433.

[83] A Sarabakha, E Kayacan. Online deep fuzzy learning for control of nonlinear systems using expert knowledge, IEEE Trans.Fuzzy Syst. 28 (2019) 1492-1503.

[84] M Yeganejou, S Dick, J Miller. Interpretable deep convolutional fuzzy classifier, IEEE Trans.Fuzzy Syst. 28 (2019) 1407-1419.

[85] W Chen, J An, R Li, G Xie. Tensor-train fuzzy deep computation model for citywide traffic flow prediction, IEEE Access. 7 (2019) 120581-120593.

[86] D Chen, X Zhang, L Wang, Z Han. Prediction of cloud resources demand based on hierarchical pythagorean fuzzy deep neural network, IEEE Transactions on Services Computing. 14 (2019) 1890-1901.

[87] X Zhang, F Chung, S Wang. An interpretable fuzzy DBN-based classifier for indoor user movement prediction in ambient assisted living applications, IEEE Transactions on Industrial Informatics. 16 (2019) 42-53.

[88] F Wang, Y Wang, H Wang, C Tang. T–S fuzzy model based multi-branch deep network architecture, IEEE Access. 8 (2020) 155039-155046.

[89] G Yang, H Li, W Wang, W Yang, WJ Emery. Unsupervised change detection based on a unified framework for weighted collaborative representation with RDDL and fuzzy clustering, IEEE Trans.Geosci.Remote Sens. 57 (2019) 8890-8903.

[90] S Park, SJ Lee, E Weiss, Y Motai. Intra-and inter-fractional variation prediction of lung tumors using fuzzy deep learning, IEEE journal of translational engineering in health and medicine. 4 (2016) 1-12.

[91] S Gu, F Chung, S Wang. A novel deep fuzzy classifier by stacking adversarial interpretable TSK fuzzy sub-classifiers with smooth gradient information, IEEE Trans.Fuzzy Syst. 28 (2019) 1369-1382.

[92] F Chollet, Xception: Deep learning with depthwise separable convolutions, (2017) 1251-1258.

Appendices

Appendix A. Assigned Interactions for New Datasets

No.	eventname	component	action	target	objecttable	ac12
1	\core\event\user_enrolment_created	core	created	user_enrolment	user_enrolments	11
2	\core\event\user_enrolment_updated	core	updated	user_enrolment	user_enrolments	10
3	\core\event\message_contact_added	core	added	message_contact	message_contacts	6
4	\mod_assign\event\workflow_state_updated	mod_assign	updated	workflow_state	assign	10
5	\mod_forum\event\discussion_subscription_deleted	mod_forum	deleted	discussion_subscription	forum_discussion_subs	12
6	\core\event\message_deleted	core	deleted	message	message_user_actions	12
7	\mod_forum\event\discussion_viewed	mod_forum	viewed	discussion	forum_discussions	2
8	\assignsubmission_file\event\assessable_uploaded	assignsubmission_file	uploaded	assessable	assign_submission	9
9	\mod_survey\event\response_submitted	mod_survey	submitted	response	NULL	5
10	\mod_assign\event\submission_duplicated	mod_assign	duplicated	submission	assign_submission	7
11	\core\event\role_assigned	core	assigned	role	role	11
12	\core\event\message_viewed	core	viewed	message	message_user_actions	2
13	\mod_forum\event\discussion_deleted	mod_forum	deleted	discussion	forum_discussions	12
14	\mod_wiki\event\page_history_viewed	mod_wiki	viewed	page_history	wiki_pages	1
15	\mod_glossary\event\entry_viewed	mod_glossary	viewed	entry	glossary_entries	1
16	\mod_workshop\event\course_module_viewed	mod_workshop	viewed	course_module	workshop	4
17	\core\event\badge_created	core	created	badge	badge	8
18	\core\event\calendar_event_created	core	created	calendar_event	event	6
19	\mod_assign\event\submission_status_viewed	mod_assign	viewed	submission_status	NULL	3
20	\mod_questionnaire\event\attempt_submitted	mod_questionnaire	submitted	attempt	NULL	6
21	\mod_glossary\event\comment_created	mod_glossary	created	comment	comments	6
22	\forumreport_summary\event\report_viewed	forumreport_summary	viewed	report	NULL	3
23	\core\event\calendar_event_deleted	core	deleted	calendar_event	event	12
24	\mod_oublog\event\post_viewed	mod_oublog	viewed	post	oublog_posts	1
25	\mod_attendance\event\report_viewed	mod_attendance	viewed	report	attendance	3
26	\core\event\question_category_created	core	created	question_category	question_categories	7
27	\mod_assign\event\submission_unlocked	mod_assign	unlocked	submission	assign	10
28	\report_questioninstances\event\report_viewed	report_questioninstances	viewed	report	NULL	3
29	\core\event\grade_item_updated	core	updated	grade_item	grade_items	10
30	\core\event\blog_association_deleted	core	deleted	blog_association	blog_association	12
31	\mod_facetoface\event\course_module_viewed	mod_facetoface	viewed	course_module	facetoface	4
32	\mod_forum\event\discussion_unpinned	mod_forum	unpinned	discussion	forum_discussions	12
33	\mod_questionnaire\event\questionnaire_previewed	mod_questionnaire	previewed	questionnaire	questionnaire	4
34	\mod_quiz\event\attempt_reviewed	mod_quiz	reviewed	attempt	quiz_attempts	12
35	\mod_hvp\event\course_module_viewed	mod_hvp	viewed	course_module	hvp	4
36	\mod_quiz\event\user_override_deleted	mod_quiz	deleted	user_override	quiz_overrides	12

37	\mod_lesson\event\course_module_instance_list_viewed	mod_lesson	viewed	course_module_instance_list	NULL	1
38	\mod_quiz\event\group_override_deleted	mod_quiz	deleted	group_override	quiz_overrides	12
39	\mod_quiz\event\attempt_summary_viewed	mod_quiz	viewed	attempt_summary	quiz_attempts	3
40	\mod_scheduler\event\booking_form_viewed	mod_scheduler	viewed	booking_form	scheduler	1
41	\mod_questionnaire\event\all_responses_saved_as_text	mod_questionnaire	text	all_responses_saved_as	questionnaire	6
42	\mod_facetoface\event\add_session	mod_facetoface	session	add	facetoface_sessions	6
43	\mod_forum\event\course_searched	mod_forum	searched	course	NULL	4
44	\mod_questionnaire\event\attempt_resumed	mod_questionnaire	resumed	attempt	questionnaire	6
45	\core\event\role_allow_assign_updated	core	updated	role_allow_assign	role_allow_assign	10
46	\tool_capability\event\report_viewed	tool_capability	viewed	report	NULL	3
47	\mod_wiki\event\page_map_viewed	mod_wiki	viewed	page_map	wiki_pages	1
48	\core\event\question_category_deleted	core	deleted	question_category	question_categories	12
49	\tool_usertours\event\tour_ended	tool_usertours	ended	tour	tool_usertours_tours	5
50	\mod_bigbluebuttonbn\event\recording_edited	mod_bigbluebuttonbn	edited	recording	bigbluebuttonbn	12
51	\mod_tab\event\course_module_viewed	mod_tab	viewed	course_module	tab	4
52	\report_log\event\report_viewed	report_log	viewed	report	NULL	3
53	\core\event\course_deleted	core	deleted	course	course	12
54	\mod_lesson\event\lesson_resumed	mod_lesson	resumed	lesson	lesson	5
55	\mod_hvp\event\attempt_submitted	mod_hvp	submitted	attempt	NULL	6
56	\core\event\message_sent	core	sent	message	messages	6
57	\core\event\message_user_unblocked	core	unblocked	message_user	message_users_blocked	10
58	\mod_data\event\record_created	mod_data	created	record	data_records	5
59	\core\event\grouping_deleted	core	deleted	grouping	groupings	12
60	\mod_glossary\event\course_module_viewed	mod_glossary	viewed	course_module	glossary	1
61	\mod_lesson\event\course_module_viewed	mod_lesson	viewed	course_module	lesson	1
62	\assignsubmission_onlinetext\event\submission_created	assignsubmission_onlinetext	created	submission	assignsubmission_onlinetext	7
63	\mod_oublog\event\course_module_viewed	mod_oublog	viewed	course_module	oublog	1
64	\core\event\grade_letter_updated	core	updated	grade_letter	grade_letters	10
65	\mod_quiz\event\attempt_viewed	mod_quiz	viewed	attempt	quiz_attempts	3
66	\mod_scorm\event\tracks_viewed	mod_scorm	viewed	tracks	NULL	3
67	\mod_questionnaire\event\non_respondents_viewed	mod_questionnaire	viewed	non_respondents	questionnaire	3
68	\mod_glossary\event\entry_created	mod_glossary	created	entry	glossary_entries	6
69	\core\event\role_allow_switch_updated	core	updated	role_allow_switch	role_allow_switch	10
70	\mod_data\event\field_updated	mod_data	updated	field	data_fields	10
71	\core\event\group_member_removed	core	removed	group_member	groups	12
72	\mod_questionnaire\event\question_deleted	mod_questionnaire	deleted	question	NULL	12
73	\mod_data\event\template_viewed	mod_data	viewed	template	NULL	4
74	\core\event\blog_association_created	core	created	blog_association	blog_association	6
75	\core\event\course_section_updated	core	updated	course_section	course_sections	12
76	\mod_oublog\event\comment_deleted	mod_oublog	deleted	comment	oublog_comments	12
77	\mod_assign\event\submission_locked	mod_assign	locked	submission	assign	10

78	\mod_folder\event\course_module_viewed	mod_folder	viewed	course_module	folder	4
79	\mod_scheduler\event\slot_deleted	mod_scheduler	deleted	slot	scheduler_slots	12
80	\core\event\user_updated	core	updated	user	user	10
81	\core\event\user_password_updated	core	updated	user_password	NULL	12
82	\mod_attendance\event\attendance_taken_by_student	mod_attendance	student	attendance_taken_by	attendance_log	10
83	\mod_forum\event\post_updated	mod_forum	updated	post	forum_posts	10
84	\mod_scheduler\event\appointment_list_viewed	mod_scheduler	viewed	appointment_list	scheduler	4
85	\mod_hvp\event\course_module_instance_list_viewed	mod_hvp	viewed	course_module_instance_list	NULL	4
86	\mod_glossary\event\entry_deleted	mod_glossary	deleted	entry	glossary_entries	12
87	\core\event\course_module_updated	core	updated	course_module	course_modules	10
88	\core\event\course_category_updated	core	updated	course_category	course_categories	10
89	\mod_oublog\event\post_updated	mod_oublog	updated	post	oublog_posts	10
90	\core\event\role_allow_view_updated	core	updated	role_allow_view	role_allow_view	10
91	\tool_usertours\event\tour_started	tool_usertours	started	tour	tool_usertours_tours	5
92	\core\event\completion_defaults_updated	core	updated	completion_defaults	course_completion_defaults	10
93	\assignsubmission_file\event\submission_created	assignsubmission_file	created	submission	assignsubmission_file	6
94	\mod_assign\event\user_override_updated	mod_assign	updated	user_override	assign_overrides	10
95	\core\event\blog_entry_deleted	core	deleted	blog_entry	post	12
96	\gradereport_grader\event\grade_report_viewed	gradereport_grader	viewed	grade_report	NULL	3
97	\core\event\tag_created	core	created	tag	tag	8
98	\core\event\tag_removed	core	removed	tag	tag_instance	12
99	\gradereport_history\event\grade_report_viewed	gradereport_history	viewed	grade_report	NULL	3
100	\core\event\user_enrolment_deleted	core	deleted	user_enrolment	user_enrolments	12
101	\mod_lesson\event\essay_assessed	mod_lesson	assessed	essay	lesson_grades	11
102	\core\event\blog_entries_viewed	core	viewed	blog_entries	NULL	1
103	\mod_assign\event\remove_submission_form_viewed	mod_assign	viewed	remove_submission_form	NULL	1
104	\core\event\badge_updated	core	updated	badge	badge	10
105	\core\event\enrol_instance_created	core	created	enrol_instance	enrol	11
106	\mod_wiki\event\comments_viewed	mod_wiki	viewed	comments	wiki_pages	1
107	\mod_scheduler\event\slot_added	mod_scheduler	added	slot	scheduler_slots	8
108	\core\event\user_graded	core	graded	user	grade_grades	11
109	\mod_bigbluebuttonbn\event\recording_deleted	mod_bigbluebuttonbn	deleted	recording	bigbluebuttonbn	12
110	\core\event\scale_updated	core	updated	scale	scale	10
111	\mod_book\event\chapter_updated	mod_book	updated	chapter	book_chapters	10
112	\mod_data\event\course_module_viewed	mod_data	viewed	course_module	data	4
113	\core\event\grouping_group_assigned	core	assigned	grouping_group	groupings	11
114	\mod_hotpot\event\course_module_viewed	mod_hotpot	viewed	course_module	hotpot	3
115	\mod_forum\event\discussion_subscription_created	mod_forum	created	discussion_subscription	forum_discussion_subs	6
116	\core\event\prediction_action_started	core	started	prediction_action	analytics_predictions	5
117	\mod_quiz\event\attempt_preview_started	mod_quiz	started	attempt_preview	quiz_attempts	6
118	\mod_lesson\event\page_created	mod_lesson	created	page	lesson_pages	5
119	\assignsubmission_comments\event\comment_deleted	assignsubmission_comments	deleted	comment	comments	12

120	\assignsubmission_onlinetext\event\submission_updated	assignsubmission_onlinetext	updated	submission	assignsubmission_onlinetext	10
121	\mod_questionnaire\event\course_module_viewed	mod_questionnaire	viewed	course_module	questionnaire	3
122	\mod_choice\event\answer_deleted	mod_choice	deleted	answer	choice_answers	12
123	\mod_chat\event\course_module_viewed	mod_chat	viewed	course_module	chat	2
124	\mod_quiz\event\user_override_updated	mod_quiz	updated	user_override	quiz_overrides	10
125	\mod_attendance\event\status_updated	mod_attendance	updated	status	attendance_statuses	10
126	\mod_questionnaire\event\attempt_saved	mod_questionnaire	saved	attempt	NULL	6
127	\mod_quiz\event\user_override_created	mod_quiz	created	user_override	quiz_overrides	7
128	\mod_lti\event\course_module_instance_list_viewed	mod_lti	viewed	course_module_instance_list	NULL	4
129	\mod_turnitintooltwo\event\delete_submission	mod_turnitintooltwo	submission	delete	turnitintooltwo	12
130	\mod_attendance\event\session_duration_updated	mod_attendance	updated	session_duration	attendance_sessions	10
131	\core\event\role_updated	core	updated	role	role	10
132	\mod_data\event\template_updated	mod_data	updated	template	NULL	10
133	\mod_assign\event\user_override_deleted	mod_assign	deleted	user_override	assign_overrides	12
134	\core\event\course_module_deleted	core	deleted	course_module	course_modules	12
135	\mod_forum\event\discussion_pinned	mod_forum	pinned	discussion	forum_discussions	10
136	\core\event\notification_viewed	core	viewed	notification	notifications	4
137	\core\event\course_completed	core	completed	course	course_completions	5
138	\core\event\blog_entry_created	core	created	blog_entry	post	7
139	\core\event\question_created	core	created	question	question	7
140	\mod_questionnaire\event\all_responses_viewed	mod_questionnaire	viewed	all_responses	questionnaire	2
141	\mod_attendance\event\session_added	mod_attendance	added	session	attendance_sessions	6
142	\mod_scorm\event\report_viewed	mod_scorm	viewed	report	NULL	3
143	\mod_scorm\event\interactions_viewed	mod_scorm	viewed	interactions	NULL	3
144	\mod_lti\event\course_module_viewed	mod_lti	viewed	course_module	lti	4
145	\mod_forum\event\discussion_created	mod_forum	created	discussion	forum_discussions	6
146	\core\event\question_viewed	core	viewed	question	question	3
147	\mod_lesson\event\content_page_viewed	mod_lesson	viewed	content_page	lesson_pages	1
148	\mod_forum\event\post_deleted	mod_forum	deleted	post	forum_posts	12
149	\mod_quiz\event\question_manually_graded	mod_quiz	graded	question_manually	question	10
150	\core\event\course_category_created	core	created	course_category	course_categories	8
151	\core\event\user_deleted	core	deleted	user	user	12
152	\mod_quiz\event\report_viewed	mod_quiz	viewed	report	NULL	3
153	\mod_assign\event\submission_confirmation_form_viewed	mod_assign	viewed	submission_confirmation_form	NULL	1
154	\core\event\question_moved	core	moved	question	question	10
155	\block_comments\event\comment_created	block_comments	created	comment	comments	6
156	\core\event\course_created	core	created	course	course	5
157	\core\event\question_category_updated	core	updated	question_category	question_categories	12
158	\mod_quiz\event\course_module_viewed	mod_quiz	viewed	course_module	quiz	3
159	\core\event\course_module_completion_updated	core	updated	course_module_completion	course_modules_completion	10
160	\core\event\user_list_viewed	core	viewed	user_list	course	4

161	\mod_questionnaire\event\response_viewed	mod_questionnaire	viewed	response	questionnaire	2
162	\mod_chat\event\sessions_viewed	mod_chat	viewed	sessions	chat	2
163	\core\event\webservice_service_created	core	created	webservice_service	external_services	8
164	\mod_lesson\event\page_updated	mod_lesson	updated	page	lesson_pages	10
165	\mod_assign\event\feedback_viewed	mod_assign	viewed	feedback	assign_grades	4
166	\mod_attendance\event\session_updated	mod_attendance	updated	session	attendance_sessions	10
167	\mod_groupselect\event\course_module_viewed	mod_groupselect	viewed	course_module	groupselect	2
168	\core\event\recent_activity_viewed	core	viewed	recent_activity	NULL	10
169	\core\event\course_content_deleted	core	deleted	course_content	course	12
170	\core\event\role_allow_override_updated	core	updated	role_allow_override	role_allow_override	10
171	\core\event\grade_deleted	core	deleted	grade	grade_grades	12
172	\core\event\group_created	core	created	group	groups	6
173	\mod_choice\event\report_viewed	mod_choice	viewed	report	choice	3
174	\mod_assign\event\identities_revealed	mod_assign	revealed	identities	assign	10
175	\core\event\tag_added	core	added	tag	tag_instance	8
176	\core\event\tag_deleted	core	deleted	tag	tag	12
177	\mod_assign\event\assessable_submitted	mod_assign	submitted	assessable	assign_submission	6
178	\gradereport_outcomes\event\grade_report_viewed	gradereport_outcomes	viewed	grade_report	NULL	3
179	\core\event\question_category_viewed	core	viewed	question_category	question_categories	3
180	\tool_recyclebin\event\category_bin_item_created	tool_recyclebin	created	category_bin_item	tool_recyclebin_category	8
181	\mod_url\event\course_module_viewed	mod_url	viewed	course_module	url	1
182	\mod_scheduler\event\booking_removed	mod_scheduler	removed	booking	scheduler_slots	12
183	\mod_lesson\event\question_answered	mod_lesson	answered	question	lesson_pages	7
184	\core\event\course_resources_list_viewed	core	viewed	course_resources_list	NULL	4
185	\core\event\blog_external_viewed	core	viewed	blog_external	NULL	1
186	\core\event\course_section_created	core	created	course_section	course_sections	5
187	\core\event\question_updated	core	updated	question	question	12
188	\mod_quiz\event\group_override_created	mod_quiz	created	group_override	quiz_overrides	7
189	\core\event\webservice_token_created	core	created	webservice_token	external_tokens	8
190	\mod_choice\event\course_module_instance_list_viewed	mod_choice	viewed	course_module_instance_list	NULL	3
191	\tool_usertours\event\tour_reset	tool_usertours	reset	tour	tool_usertours_tours	10
192	\mod_lesson\event\question_viewed	mod_lesson	viewed	question	lesson_pages	1
193	\core\event\capability_assigned	core	assigned	capability	role_capabilities	11
194	\core\event\scale_created	core	created	scale	scale	8
195	\mod_scorm\event\course_module_viewed	mod_scorm	viewed	course_module	scorm	3
196	\core\event\group_member_added	core	added	group_member	groups	6
197	\core\event\grade_letter_deleted	core	deleted	grade_letter	grade_letters	12
198	\mod_wiki\event\page_created	mod_wiki	created	page	wiki_pages	6
199	\core\event\notification_sent	core	sent	notification	notifications	6
200	\core\event\capability_unassigned	core	unassigned	capability	role_capabilities	12
201	\mod_bigbluebuttonbn\event\recording_published	mod_bigbluebuttonbn	published	recording	bigbluebuttonbn	5
202	\mod_forum\event\subscription_created	mod_forum	created	subscription	forum_subscriptions	6

203	\mod_turnitintooltwo\event\list_submissions	mod_turnitintooltwo	submissions	list	turnitintooltwo	7
204	\mod_attendance\event\attendance_taken	mod_attendance	taken	attendance	attendance_log	10
205	\report_log\event\user_report_viewed	report_log	viewed	user_report	NULL	3
206	\core\event\course_completion_updated	core	updated	course_completion	NULL	10
207	\core\event\course_restored	core	restored	course	course	5
208	\mod_lesson\event\lesson_ended	mod_lesson	ended	lesson	lesson	5
209	\core\event\grouping_updated	core	updated	grouping	groupings	10
210	\mod_lesson\event\lesson_restarted	mod_lesson	restarted	lesson	lesson	5
211	\mod_lesson\event\essay_attempt_viewed	mod_lesson	viewed	essay_attempt	lesson_attempts	1
212	\mod_forum\event\course_module_viewed	mod_forum	viewed	course_module	forum	2
213	\mod_lesson\event\lesson_started	mod_lesson	started	lesson	lesson	5
214	\mod_scheduler\event\booking_added	mod_scheduler	added	booking	scheduler_slots	6
215	\mod_lesson\event\page_deleted	mod_lesson	deleted	page	lesson_pages	12
216	\core\event\role_unassigned	core	unassigned	role	role	12
217	\mod_forum\event\user_report_viewed	mod_forum	viewed	user_report	NULL	3
218	\mod_folder\event\folder_updated	mod_folder	updated	folder	folder	10
219	\report_participation\event\report_viewed	report_participation	viewed	report	NULL	3
220	\mod_assign\event\user_override_created	mod_assign	created	user_override	assign_overrides	7
221	\mod_book\event\chapter_deleted	mod_book	deleted	chapter	book_chapters	12
222	\mod_glossary\event\entry_updated	mod_glossary	updated	entry	glossary_entries	10
223	\core\event\message_contact_removed	core	removed	message_contact	message_contacts	12
224	\tool_recyclebin\event\course_bin_item_restored	tool_recyclebin	restored	course_bin_item	tool_recyclebin_course	5
225	\mod_assign\event\extension_granted	mod_assign	granted	extension	assign	11
226	\core\event\course_information_viewed	core	viewed	course_information	course	1
227	\mod_quiz\event\attempt_abandoned	mod_quiz	abandoned	attempt	quiz_attempts	6
228	\core\event\user_created	core	created	user	user	6
229	\core\event\course_category_deleted	core	deleted	course_category	course_categories	12
230	\mod_quiz\event\attempt_deleted	mod_quiz	deleted	attempt	quiz_attempts	12
231	\mod_chat\event\course_module_instance_list_viewed	mod_chat	viewed	course_module_instance_list	NULL	2
232	\mod_turnitintooltwo\event\add_submission	mod_turnitintooltwo	submission	add	turnitintooltwo	9
233	\mod_forum\event\assessable_uploaded	mod_forum	uploaded	assessable	forum_posts	6
234	\core\event\course_section_deleted	core	deleted	course_section	course_sections	12
235	\core\event\user_profile_viewed	core	viewed	user_profile	user	4
236	\core\event\calendar_event_updated	core	updated	calendar_event	event	10
237	\mod_quiz\event\attempt_started	mod_quiz	started	attempt	quiz_attempts	6
238	\mod_bigbluebuttonbn\event\recording_viewed	mod_bigbluebuttonbn	viewed	recording	bigbluebuttonbn	1
239	\mod_choice\event\answer_created	mod_choice	created	answer	choice_answers	6
240	\mod_assign\event\group_override_created	mod_assign	created	group_override	assign_overrides	7
241	\mod_data\event\record_deleted	mod_data	deleted	record	data_records	12
242	\mod_scorm\event\sco_launched	mod_scorm	launched	sco	scorm_scoes	5
243	\mod_oublog\event\comment_created	mod_oublog	created	comment	oublog_comments	6
244	\mod_survey\event\course_module_viewed	mod_survey	viewed	course_module	survey	3
245	\mod_quiz\event\attempt_submitted	mod_quiz	submitted	attempt	quiz_attempts	6

246	\assignsubmission_file\event\submission_updated	assignsubmission_file	updated	submission	assignsubmission_file	10
247	\mod_quiz\event\course_module_instance_list_viewed	mod_quiz	viewed	course_module_instance_list	NULL	3
248	\mod_oublog\event\participation_viewed	mod_oublog	viewed	participation	oublog	1
249	\core\event\email_failed	core	failed	email	NULL	2
250	\mod_assign\event\course_module_viewed	mod_assign	viewed	course_module	assign	3
251	\mod_wiki\event\course_module_viewed	mod_wiki	viewed	course_module	wiki	1
252	\mod_data\event\course_module_instance_list_viewed	mod_data	viewed	course_module_instance_list	NULL	4
253	\tool_usertours\event\step_shown	tool_usertours	shown	step	tool_usertours_steps	4
254	\core\event\dashboard_viewed	core	viewed	dashboard	NULL	4
255	\report_stats\event\report_viewed	report_stats	viewed	report	NULL	3
256	\mod_wiki\event\page_viewed	mod_wiki	viewed	page	wiki_pages	1
257	\core\event\grouping_group_unassigned	core	unassigned	grouping_group	groupings	12
258	\core\event\course_category_viewed	core	viewed	course_category	course_categories	4
259	\mod_assign\event\submission_form_viewed	mod_assign	viewed	submission_form	NULL	1
260	\mod_choice\event\course_module_viewed	mod_choice	viewed	course_module	choice	3
261	\mod_assign\event\grading_table_viewed	mod_assign	viewed	grading_table	NULL	3
262	\core\event\enrol_instance_deleted	core	deleted	enrol_instance	enrol	12
263	\mod_data\event\field_created	mod_data	created	field	data_fields	5
264	\core\event\course_updated	core	updated	course	course	10
265	\mod_hotpot\event\attempt_started	mod_hotpot	started	attempt	hotpot	6
266	\core\event\notes_viewed	core	viewed	notes	NULL	1
267	\core\event\blog_entry_updated	core	updated	blog_entry	post	10
268	\report_outline\event\report_viewed	report_outline	viewed	report	NULL	3
269	\core\event\message_deleted	core	deleted	message	NULL	12
270	\core\event\courses_searched	core	searched	courses	NULL	4
271	\core\event\grouping_created	core	created	grouping	groupings	6
272	\mod_page\event\course_module_viewed	mod_page	viewed	course_module	page	4
273	\mod_lesson\event\page_moved	mod_lesson	moved	page	lesson_pages	10
274	\core\event\enrol_instance_updated	core	updated	enrol_instance	enrol	10
275	\mod_chat\event\message_sent	mod_chat	sent	message	chat_messages	6
276	\mod_forum\event\subscription_deleted	mod_forum	deleted	subscription	forum_subscriptions	12
277	\tool_recyclebin\event\course_bin_item_deleted	tool_recyclebin	deleted	course_bin_item	tool_recyclebin_course	12
278	\core\event\grade_letter_created	core	created	grade_letter	grade_letters	8
279	\core\event\group_deleted	core	deleted	group	groups	12
280	\core\event\course_reset_ended	core	ended	course_reset	NULL	5
281	\mod_oublog\event\post_created	mod_oublog	created	post	oublog_posts	6
282	\mod_assign\event\submission_viewed	mod_assign	viewed	submission	assign_submission	3
283	\mod_book\event\course_module_viewed	mod_book	viewed	course_module	book	1
284	\mod_scorm\event\user_report_viewed	mod_scorm	viewed	user_report	NULL	3
285	\mod_forum\event\course_module_instance_list_viewed	mod_forum	viewed	course_module_instance_list	NULL	2
286	\assignsubmission_onlinetext\event\assessable_uploaded	assignsubmission_onlinetext	uploaded	assessable	assign_submission	9

287	\core\event\webservice_service_user_added	core	added	webservice_service_user	external_services	6
288	\gradereport_user\event\grade_report_viewed	gradereport_user	viewed	grade_report	NULL	3
289	\mod_quiz\event\edit_page_viewed	mod_quiz	viewed	edit_page	NULL	3
290	\mod_questionnaire\event\all_responses_deleted	mod_questionnaire	deleted	all_responses	questionnaire	12
291	\mod_assign\event\submission_status_updated	mod_assign	updated	submission_status	assign_submission	10
292	\mod_wiki\event\page_locks_deleted	mod_wiki	deleted	page_locks	wiki_pages	12
293	\mod_resource\event\course_module_viewed	mod_resource	viewed	course_module	resource	4
294	\mod_book\event\chapter_created	mod_book	created	chapter	book_chapters	5
295	\core\event\question_category_moved	core	moved	question_category	question_categories	10
296	\mod_scorm\event\scoreraw_submitted	mod_scorm	submitted	scoreraw	scorm_scoes_track	6
297	\core\event\course_viewed	core	viewed	course	NULL	4
298	\report_loglive\event\report_viewed	report_loglive	viewed	report	NULL	3
299	\gradereport_singleview\event\grade_report_viewed	gradereport_singleview	viewed	grade_report	NULL	3
300	\core\event\webservice_service_updated	core	updated	webservice_service	external_services	10
301	\mod_forum\event\subscribers_viewed	mod_forum	viewed	subscribers	NULL	1
302	\core\event\group_updated	core	updated	group	groups	10
303	\mod_wiki\event\page_diff_viewed	mod_wiki	viewed	page_diff	wiki_pages	1
304	\block_comments\event\comment_deleted	block_comments	deleted	comment	comments	12
305	\mod_assign\event\submission_graded	mod_assign	graded	submission	assign_grades	10
306	\core\event\note_created	core	created	note	post	6
307	\tool_recyclebin\event\category_bin_item_deleted	tool_recyclebin	deleted	category_bin_item	tool_recyclebin_category	12
308	\mod_mediasite\event\course_module_viewed	mod_mediasite	viewed	course_module	mediasite	4
309	\core\event\course_module_created	core	created	course_module	course_modules	5
310	\mod_attendance\event\session_deleted	mod_attendance	deleted	session	attendance_sessions	12
311	\mod_questionnaire\event\response_deleted	mod_questionnaire	deleted	response	questionnaire	12
312	\tool_recyclebin\event\course_bin_item_created	tool_recyclebin	created	course_bin_item	tool_recyclebin_course	8
313	\mod_assign\event\course_module_instance_list_viewed	mod_assign	viewed	course_module_instance_list	NULL	3
314	\core\event\question_deleted	core	deleted	question	question	12
315	\assignsubmission_comments\event\comment_created	assignsubmission_comments	created	comment	comments	6
316	\mod_data\event\field_deleted	mod_data	deleted	field	data_fields	10
317	\mod_assign\event\grading_form_viewed	mod_assign	viewed	grading_form	NULL	1
318	\mod_forum\event\post_created	mod_forum	created	post	forum_posts	6
319	\mod_bigbluebuttonbn\event\recording_unpublished	mod_bigbluebuttonbn	unpublished	recording	bigbluebuttonbn	12
320	\core\event\config_log_created	core	created	config_log	config_log	6
321	\gradereport_overview\event\grade_report_viewed	gradereport_overview	viewed	grade_report	NULL	3
322	\core\event\message_viewed	core	viewed	message	message_read	2
323	\mod_scorm\event\status_submitted	mod_scorm	submitted	status	scorm_scoes_track	6
324	\core\event\grade_item_created	core	created	grade_item	grade_items	6
325	\report_stats\event\user_report_viewed	report_stats	viewed	user_report	NULL	3
326	\core\event\course_backup_created	core	created	course_backup	course	5
327	\mod_quiz\event\attempt_becameoverdue	mod_quiz	becameoverdue	attempt	quiz_attempts	6

328	\mod_bigbluebuttonbn\event\activity_viewed	mod_bigbluebuttonbn	viewed	activity	bigbluebuttonbn	2
329	\core\event\webservice_token_sent	core	sent	webservice_token	external_tokens	6
330	\core\event\message_user_blocked	core	blocked	message_user	message_users_blocked	12
331	\mod_book\event\chapter_viewed	mod_book	viewed	chapter	book_chapters	1
332	\mod_data\event\record_updated	mod_data	updated	record	data_records	10
333	\mod_wiki\event\page_updated	mod_wiki	updated	page	wiki_pages	10
334	\mod_assign\event\reveal_identities_confirmation_page_viewed	mod_assign	viewed	reveal_identities_confirmation_page	NULL	4

Appendix B. MATLAB Code

```
fis1 = readfis('C:\Users\awada\Downloads\AbdulrahmanFUZZYQoI(1)\FS1.fis');
fis2 = readfis('C:\Users\awada\Downloads\AbdulrahmanFUZZYQoI(1)\FS2.fis');
fis3 = readfis('C:\Users\awada\Downloads\AbdulrahmanFUZZYQoI(1)\FS3.fis');
fis4 = readfis('C:\Users\awada\Downloads\AbdulrahmanFUZZYQoI(1)\FS4.fis');
fis5 = readfis('C:\Users\awada\Downloads\AbdulrahmanFUZZYQoI(1)\FS5a.fis');
%%data2020 = load('C:\Users\awada\Desktop\data2020_2.mat'); %77671
data2020 =
load('C:\Users\awada\Desktop\\Thesis\ThesisProject\MATLAB\data2020_scaled3.mat');
data2021 =
load('C:\Users\awada\Desktop\\Thesis\ThesisProject\MATLAB\data2021_scaled3.mat');
%%data2021 = load('C:\Users\awada\Desktop\data2021.mat'); %172227

%create variables
userid = data2020.struct.userid;
view1 = data2020.struct.view1;
view2 = data2020.struct.view2;
view3 = data2020.struct.view3;
view4 = data2020.struct.view4;
add1 = data2020.struct.add1;
add2 = data2020.struct.add2;
add3 = data2020.struct.add3;
add4 = data2020.struct.add4;
edit1 = data2020.struct.edit1;
edit2 = data2020.struct.edit2;
edit3 = data2020.struct.edit3;
edit4 = data2020.struct.edit4;
tp = data2020.struct.TP;
eng_time = data2020.struct.eng_time;
input1 = vertcat(view1,view2,view3,view4);
input2 = vertcat(add1,add2,add3,add4);
input3 = vertcat(edit1,edit2,edit3,edit4);
View = [];
Add = [];
Alter = [];
Action = [];
QoI = [];
for i = 1 : 74294
    a = input1(:,i).';
    b = input2(:,i).';
    c = input3(:,i).';
    d = tp(:,i);
    e = eng_time(:,i);
    V = evalfis(fis1,a);
    A = evalfis(fis2,b);
    E = evalfis(fis3,c);
    comb1 = [V,A,E];
    AC = evalfis(fis4,comb1);
    comb2 = [AC,d,e];
    Q = evalfis(fis5,comb2);
    View = cat(1,View,V);
    Add = cat(1,Add,A);
    Alter = cat(1,Alter,E);
    Action = cat(1,Action,AC);
    QoI = cat(1,QoI,Q);
end
%%
all_5 =
[userid',view1',view2',view3',view4',add1',add2',add3',add4',edit1',edit2',edit3',edit4',tp
',eng_time',View, Add, Alter, Action, QoI];
filename = 'C:\Users\awada\Desktop\Thesis\ThesisProject\kuq2020.xlsx';
writematrix(all_5,filename,'Sheet',1)
    %%
```

```
%create variables
userid = data2021.struct.userid;
view1 = data2021.struct.view1;
view2 = data2021.struct.view2;
view3 = data2021.struct.view3;
view4 = data2021.struct.view4;
add1 = data2021.struct.add1;
add2 = data2021.struct.add2;
add3 = data2021.struct.add3;
add4 = data2021.struct.add4;
edit1 = data2021.struct.edit1;
edit2 = data2021.struct.edit2;
edit3 = data2021.struct.edit3;
edit4 = data2021.struct.edit4;
tp = data2021.struct.TP;
eng_time = data2021.struct.eng_time;
input1 = vertcat(view1,view2,view3,view4);
input2 = vertcat(add1,add2,add3,add4);
input3 = vertcat(edit1,edit2,edit3,edit4);
View = [];
Add = [];
Alter = [];
Action = [];
QoI = [];
for i = 1 : 172227
    a = input1(:,i).';
    b = input2(:,i).';
    c = input3(:,i).';
    d = tp(:,i);
    e = eng_time(:,i);
    V = evalfis(fis1,a);
    A = evalfis(fis2,b);
    E = evalfis(fis3,c);
    comb1 = [V,A,E];
    AC = evalfis(fis4,comb1);
    comb2 = [AC,d,e];
    Q = evalfis(fis5,comb2);
    View = cat(1,View,V);
    Add = cat(1,Add,A);
    Alter = cat(1,Alter,E);
    Action = cat(1,Action,AC);
    QoI = cat(1,QoI,Q);
end
all_5 =
[userid',view1',view2',view3',view4',add1',add2',add3',add4',edit1',edit2',edit3',edit4',tp
',eng_time',View, Add, Alter, Action, QoI];
filename = 'C:\Users\awada\Desktop\\Thesis\ThesisProject\kuq2021.xlsx';
writematrix(all_5,filename,'Sheet',1)
```

Appendix C. Python Code: Prepare data for Fuzzy QoI

```
import sys
import math
from scipy.io import savemat
import scipy.io
from scipy import stats
import numpy as np
import matplotlib.pyplot as plt
import itertools
import string
import pandas as pd
import tensorflow as tf
from sklearn.model_selection import train_test_split

# if time_steps > tcn_layer.receptive_field, then we should not
# be able to solve this task.
# Load data
dir = r'C:\\Users\\awada\\Desktop\\Thesis\\DATA\\Data\\'
data_p = np.load(dir + 'db1_sem_p_sorted.npy')
data_s = np.load(dir + 'db1_sem_s_sorted.npy')
weeks = np.unique(data_s[:,13])
colums =
['userid','view1','view2','view3','view4','add1','add2','add3','add4','edit1','edit
2','edit3','edit4','TP','eng_time']

s_14 = pd.DataFrame(data_s,columns=colums)
p_14 = pd.DataFrame(data_p,columns=colums)
acc = np.zeros(shape=(0, 15), dtype=int)
for i in weeks:
    x = np.array(s_14.iloc[:,1:][s_14['TP'] == i])
    x1 = np.array(s_14.iloc[:,0][s_14['TP'] == i]).reshape(-1,1)
    rc, mc = RobustScaler(),MinMaxScaler()
    o = rc.fit_transform(x)
    p = mc.fit_transform(o)
    p[:, 12] = (i+1) / 51
    p = np.hstack((x1,p))
    acc = np.vstack((acc,p))
scaled = np.asarray(acc)

s14_scaled = pd.DataFrame(scaled, columns=colums)
s14_scaled.to_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\s14_scaled.xlsx')



a_dict = {col_name : s14_scaled[col_name].values for col_name in
s14_scaled.columns.values}
savemat(r'C:\Users\awada\Desktop\Thesis\ThesisProject\MATLAB\s14.mat',
{'struct':a_dict})
print('Done')



acc = np.zeros(shape=(0, 15), dtype=int)
for i in weeks:
    x = np.array(p_14.iloc[:,1:][p_14['TP'] == i])
    x1 = np.array(p_14.iloc[:,0][p_14['TP'] == i]).reshape(-1,1)
    rc, mc = RobustScaler(),MinMaxScaler()
    o = rc.fit_transform(x)
    p = mc.fit_transform(o)
    p[:, 12] = (i+1) / 51
    p = np.hstack((x1,p))
    acc = np.vstack((acc,p))
scaled = np.asarray(acc)

p14_scaled = pd.DataFrame(scaled, columns=colums)
p14_scaled.to_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\p14_scaled.xlsx')
```

```
a_dict = {col_name : p14_scaled[col_name].values for col_name in
p14_scaled.columns.values}
savemat(r'C:\Users\awada\Desktop\Thesis\ThesisProject\MATLAB\p14.mat',
{'struct':a_dict})
print('Done')
```

Appendix D. Python Code: Tunning for 14 parameters

```
import scipy.io
from scipy import stats
import numpy as np
import matplotlib.pyplot as plt
import itertools
import string
import pandas as pd
from sklearn.metrics import r2_score
# import tensorflow_transform as tft
from sklearn.preprocessing import StandardScaler, OneHotEncoder, OrdinalEncoder,
Normalizer, LabelEncoder, \
    LabelBinarizer, MultiLabelBinarizer, RobustScaler, MinMaxScaler
from sklearn.decomposition import PCA
from sklearn.model_selection import train_test_split, cross_val_score,
cross_validate
from sklearn.metrics import confusion_matrix, roc_auc_score, accuracy_score,
precision_score, recall_score, \
    mean_squared_error, mean_absolute_error, r2_score
from keras.models import Sequential
from keras.wrappers.scikit_learn import KerasClassifier
from sklearn.pipeline import Pipeline
from keras.layers import Dense, BatchNormalization, Activation, Conv1D,Conv2D,
MaxPooling1D, Flatten, Dropout, Lambda, LSTM
from keras import backend as K
from tensorflow.keras import regularizers
import tensorflow as tf
from sklearn.model_selection import train_test_split
import statistics
from scipy.stats import kruskal, friedmanchisquare, wilcoxon, spearmanr, f_oneway
import torch
from shap import DeepExplainer

# if time_steps > tcn_layer.receptive_field, then we should not
# be able to solve this task.
# Load data
dir = r'C:\\Users\\awada\\Desktop\\Thesis\\DATA\\Data\\'
data_p = np.load(dir + 'db1_sem_p_sorted.npy', allow_pickle=True)
data_s = np.load(dir + 'db1_sem_s_sorted.npy', allow_pickle=True)
colums =
['userid','view1','view2','view3','view4','add1','add2','add3','add4','edit1','edit
2','edit3','edit4','TP','eng_time','View','Add','Alter','Action','QoI']
students =
pd.read_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\s14_qoi.xlsx', header =
None,names = colums)
professors =
pd.read_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\p14_qoi.xlsx', header =
None,names = colums)
y_s = np.array(students.loc[:,['userid','QoI']]).reshape(-1,2)
y_p = np.array(professors.loc[:,['userid','QoI']]).reshape(-1,2)

s_u = np.unique(data_s[:,0]).reshape(-1,1)
p_u = np.unique(data_s[:,0]).reshape(-1,1)
weeks = np.unique(data_s[:,13]).reshape(-1,1)

#Adjust QoI interaction Classes
range_qoi,max_qoi,min_qoi = np.unique(y_p[:,1]),np.max(y_p[:,1]),np.min(y_p[:,1])
sub = max_qoi - min_qoi
#adjust ranges 0.2,0.4, 0.6
users_2 = np.unique(y_s[(y_s[:,1] >= 0.4),0])
users_4 = np.unique(y_s[(y_s[:,1] >= 0.2),0])
users_6 = np.unique(y_s[(y_s[:,1] <= 0.4),0])
cls = np.zeros(shape=(y_s.shape[0],1),dtype= int)
y = np.hstack((y_s,cls))
y[(y[:,1] <= 0.2),2] = 0
y[(y[:,1] > 0.2),2] = 1
```

```
y[(y[:,1] > 0.4),2] = 2

cls_users = np.zeros(shape=(0,2),dtype= int)
for i in s_u:
    print(i)
    a = np.max(y[(y[:,0] == i),2])
    a = np.hstack((i,a))
    cls_users= np.vstack((cls_users,a))


# split based on users:
su_train, su_test= train_test_split(s_u, test_size=0.25,
random_state=0,stratify=cls_users[:,1])
X_train, X_test, y_train, y_test = data_s,data_s,y_s,y_s
# Prepare training and testing sets
for i in su_test:
    X_train = X_train[~(X_train[:,0] == i),:]
    y_train = y_train[~(y_train[:,0] == i),:]

for i in su_train:
    X_test = X_test[~(X_test[:,0] == i),:]
    y_test = y_test[~(y_test[:, 0] == i), :]

X_train, X_test, y_train,y_test= X_train[:,1:],
X_test[:,1:],y_train[:,1],y_test[:,1]

# scaling per period
x_train = np.zeros(shape=(0, 14), dtype=int)
x_test = np.zeros(shape=(0, 14), dtype=int)

for i in weeks:
    print(i)
    x1 = X_train[(X_train[:,12] == i),:]
    x2 = X_test[(X_test[:,12] == i),:]
    rc, mc = RobustScaler(),MinMaxScaler()
    o1 = rc.fit_transform(x1)
    o2 = rc.transform(x2)
    p1 = mc.fit_transform(o1)
    p2 = mc.transform(o2)
    p1[:, 12] = (i+1) / 51
    p2[:, 12] = (i + 1) / 51
    x_train,x_test = np.vstack((x_train,p1)),np.vstack((x_test,p2))


import gc
def reset_tensorflow_keras_backend():
    # to be further investigated, but this seems to be enough
    tf.keras.backend.clear_session()
    tf.compat.v1.reset_default_graph()
    _ = gc.collect()
    return
reset_tensorflow_keras_backend()

print('Done before Tunning')
# Tuning
# Baysian optimization
from keras_tuner.tuners import BayesianOptimization


def build_model(hp):
    model = Sequential()

    model.add(
        Dense(units=hp.Int('dense-bot', min_value=32, max_value=1024, step=32),
input_shape=(X_train.shape[1],), activation='selu'))
    for i in range(hp.Int('num_dense_layers', 1, 6)):
        # model.add(LSTM(units=hp.Int('lstm_' + str(i), min_value=32,
max_value=1024, step=32), activation='relu'))
```

```
        model.add(Dense(units=hp.Int('dense_' + str(i), min_value=32,
max_value=1024, step=32), activation='selu'))
        model.add(Dropout(hp.Choice('dropout_' + str(i), values=[0.0, 0.1, 0.2])))

    model.add(Dense(1, activation="sigmoid"))


    model.compile(loss='mse', metrics='mae', optimizer='adamax')

    return model


print(tf.config.list_physical_devices('GPU'))

n_epochs = 200

bayesian_opt_tuner = BayesianOptimization(
    build_model,
    objective='val_loss',
    max_trials=200,
    executions_per_trial=1,
    directory=dir,
    project_name='Tunning_14',
    overwrite=True)

bayesian_opt_tuner.search(x_train, y_train, epochs=n_epochs,batch_size = 16384,
                          validation_split=0.25,
                          verbose=0)

bayes_opt_model_best_model = bayesian_opt_tuner.get_best_models(num_models=1)
model = bayes_opt_model_best_model[0]
model.summary()
import pydot
model.save(dir+'\Tunning_14\model_12.h5')
tf.keras.utils.plot_model(model, to_file=dir+'\Tunning_14\model_14.png',
show_shapes=True,
                          show_layer_names=True)
```

Appendix E. Python Code: Cross validation 14 parameters

```
import scipy.io
from scipy import stats
import numpy as np
import matplotlib.pyplot as plt
import itertools
import string
import pickle
import math
import pandas as pd
from sklearn.metrics import r2_score
# import tensorflow_transform as tft
from sklearn.preprocessing import StandardScaler, OneHotEncoder, OrdinalEncoder,
Normalizer, LabelEncoder, \
    LabelBinarizer, MultiLabelBinarizer, RobustScaler, MinMaxScaler
from sklearn.decomposition import PCA
from sklearn.model_selection import train_test_split, cross_val_score,
cross_validate
from sklearn.metrics import confusion_matrix, roc_auc_score, accuracy_score,
precision_score, recall_score, \
    mean_squared_error, mean_absolute_error, r2_score
from keras.models import Sequential
from keras.wrappers.scikit_learn import KerasClassifier
from sklearn.pipeline import Pipeline
from sklearn.model_selection import StratifiedKFold
import keras
from keras.layers import Dense, BatchNormalization, Activation, Conv1D,Conv2D,
MaxPooling1D, Flatten, Dropout, Lambda, LSTM
from keras import backend as K
from tensorflow.keras import regularizers
import tensorflow as tf
from sklearn.model_selection import train_test_split
import statistics
from scipy.stats import kruskal, friedmanchisquare, wilcoxon, spearmanr, f_oneway
import torch
from shap import DeepExplainer

# if time_steps > tcn_layer.receptive_field, then we should not
# be able to solve this task.
# Load data
dir = r'C:\\Users\\awada\\Desktop\\Thesis\\DATA\\Data\\'
data_p = np.load(dir + 'db1_sem_p_sorted.npy', allow_pickle=True)
data_s = np.load(dir + 'db1_sem_s_sorted.npy', allow_pickle=True)
colums =
['userid','view1','view2','view3','view4','add1','add2','add3','add4','edit1','edit
2','edit3','edit4','TP','eng_time','View','Add','Alter','Action','QoI']
students =
pd.read_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\s14_qoi.xlsx', header =
None,names = colums)
professors =
pd.read_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\p14_qoi.xlsx', header =
None,names = colums)
y_s = np.array(students.loc[:,['userid','QoI']]).reshape(-1,2)
y_p = np.array(professors.loc[:,['userid','QoI']]).reshape(-1,2)

s_u = np.unique(data_s[:,0]).reshape(-1,1)
p_u = np.unique(data_s[:,0]).reshape(-1,1)
weeks = np.unique(data_s[:,13]).reshape(-1,1)

#Adjust QoI interaction Classes
range_qoi,max_qoi,min_qoi = np.unique(y_p[:,1]),np.max(y_p[:,1]),np.min(y_p[:,1])
sub = max_qoi - min_qoi
#adjust ranges 0.2,0.4, 0.6
users_2 = np.unique(y_s[(y_s[:,1] >= 0.4),0])
users_4 = np.unique(y_s[(y_s[:,1] >= 0.2),0])
users_6 = np.unique(y_s[(y_s[:,1] <= 0.4),0])
```

```
cls = np.zeros(shape=(y_s.shape[0],1),dtype= int)
y = np.hstack((y_s,cls))
y[(y[:,1] <= 0.2),2] = 0
y[(y[:,1] > 0.2),2] = 1
y[(y[:,1] > 0.4),2] = 2

cls_users = np.zeros(shape=(0,2),dtype= int)
for i in s_u:
    print(i)
    a = np.max(y[(y[:,0] == i),2])
    a = np.hstack((i,a))
    cls_users= np.vstack((cls_users,a))
cls_users = cls_users.astype(np.int64)


import gc
def reset_tensorflow_keras_backend():
    # to be further investigated, but this seems to be enough
    tf.keras.backend.clear_session()
    tf.compat.v1.reset_default_graph()
    _ = gc.collect()
    return
reset_tensorflow_keras_backend()

print('Done before cross validation')

#models
def model_tuned(a):
    inputs = keras.Input(a.shape[1],)
    x = Dense(1024, activation=tf.keras.activations.selu)(inputs)
    x = Dense(1024, activation=tf.keras.activations.selu)(x)
    x = Dropout(0.2)(x)
    x = Dense(32, activation=tf.keras.activations.selu)(x)
    # x = Dropout(0.0)(x)
    x = Dense(32, activation=tf.keras.activations.selu)(x)
    # x = Dropout(0.0)(x)
    x = Dense(32, activation=tf.keras.activations.selu)(x)
    outputs = Dense(1, activation=tf.nn.sigmoid)(x)
    model = keras.Model(inputs=inputs, outputs=outputs)
    model.compile(optimizer='adamax', loss='mse', metrics=['mae'])
    return model

def model_taper(a):
    neurons = 256
    dr = 0.2
    inputs = keras.Input(a.shape[1])
    l1 = Dense(neurons, activation=tf.keras.activations.selu)(inputs)
    d1 = Dropout(dr)(l1)
    l2 = Dense(int(neurons / 2), activation=tf.keras.activations.selu)(d1)
    d2 = Dropout(dr)(l2)
    l3 = Dense(int(neurons / 4), activation=tf.keras.activations.selu)(d2)
    d3 = Dropout(dr)(l3)
    l3 = Dense(int(neurons / 8), activation=tf.keras.activations.selu)(d3)
    d3 = Dropout(dr)(l3)
    l3 = Dense(int(neurons / 16), activation=tf.keras.activations.selu)(d3)
    d3 = Dropout(dr)(l3)
    l3 = Dense(int(neurons / 32), activation=tf.keras.activations.selu)(d3)
    d3 = Dropout(dr)(l3)
    outputs = Dense(1, activation=tf.nn.sigmoid)(d3)
    model = keras.Model(inputs=inputs, outputs=outputs)
    model.compile(optimizer='adamax', loss='mse', metrics=['mae'])
    return model

kuq2020 =
pd.read_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\data2020_2.xlsx')
kuq2021 =
pd.read_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\data2021.xlsx')
y2 =pd.read_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\kuq2020.xlsx',
```

```
header = None)
y3 =pd.read_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\kuq2021.xlsx',
header = None)
y2, y3 = np.array(y2.iloc[:,-1]).reshape(-1,1),np.array(y3.iloc[:,-1]).reshape(-
1,1)
n2020 = np.array(kuq2020)
n2020 = n2020[:,2:]
n2021 = np.array(kuq2021)
n2021 = n2021[:,2:]
# split based on users:
skf = StratifiedKFold(n_splits=4, random_state=None, shuffle=True)
models = [model_tuned, model_taper]
model_list = ['tuned', 'taper']
metric_list = []
for j,k in zip (models,model_list):
    fold = 1
    for train_index, test_index in skf.split(s_u,cls_users[:,1]):
        X_train, X_test, y_train, y_test = data_s, data_s, y_s, y_s
        su_train, su_test = s_u[train_index], s_u[test_index]
        # Prepare training and testing sets
        for i in su_test:
            X_train = X_train[~(X_train[:, 0] == i), :]
            y_train = y_train[~(y_train[:, 0] == i), :]

        for i in su_train:
            X_test = X_test[~(X_test[:, 0] == i), :]
            y_test = y_test[~(y_test[:, 0] == i), :]

        X_train, X_test, y_train, y_test = X_train[:, 1:], X_test[:, 1:],
y_train[:, 1], y_test[:, 1]
        # scaling per period
        x_train = np.zeros(shape=(0, 14), dtype=int)
        x_test = np.zeros(shape=(0, 14), dtype=int)
        kuq1 = np.zeros(shape=(0, 14), dtype=int)
        kuq2 = np.zeros(shape=(0, 14), dtype=int)

        for i in weeks:
            print(i)
            x1 = X_train[(X_train[:, 12] == i), :]
            x2 = X_test[(X_test[:, 12] == i), :]
            q2 = n2021[(n2021[:, 12] == i+1), :]
            rc, mc = RobustScaler(quantile_range=(10,90)), MinMaxScaler()
            o1 = rc.fit_transform(x1)
            o2 = rc.transform(x2)
            o4 = rc.transform(q2)
            p1 = mc.fit_transform(o1)
            p2 = mc.transform(o2)
            p4 = mc.transform(o4)
            p1[:, 12] = (i + 1) / 51
            p2[:, 12] = (i + 1) / 51
            p4[:, 12] = (i + 1) / 51
            if i+1 > 29:
                q1 = n2020[(n2020[:, 12] == i+1), :]

                o3 = rc.transform(q1)

                p3 = mc.transform(o3)

                p3[:, 12] = (i + 1) / 51

                kuq1 = np.vstack((kuq1, p3))
            x_train, x_test = np.vstack((x_train, p1)), np.vstack((x_test, p2))
            kuq2 = np.vstack((kuq2, p4))
        model = j(x_train)
        history = model.fit(x_train, y_train, batch_size=16384 , epochs=10000,
verbose = 0)
        predicted_QoI0,predicted_QoI, predicted_QoI2,predicted_QoI3 =
model.predict(x_train),model.predict(x_test),model.predict(kuq1),model.predict(kuq2
```

```
)
        MSE,MSE1,MSE2,MSE3 = mean_squared_error(y_train,
predicted_QoI0),mean_squared_error(y_test, predicted_QoI),mean_squared_error(y2,
predicted_QoI2),mean_squared_error(y3, predicted_QoI3)
        RMSE,RMSE1,RMSE2,RMSE3 =
math.sqrt(MSE),math.sqrt(MSE1),math.sqrt(MSE2),math.sqrt(MSE3)
        MAE,MAE1,MAE2,MAE3 =mean_absolute_error(y_train, predicted_QoI0),
mean_absolute_error(y_test, predicted_QoI),mean_absolute_error(y2,
predicted_QoI2),mean_absolute_error(y3, predicted_QoI3)
        R2,R21,R22,R23 = r2_score(y_train, predicted_QoI0,
multioutput='variance_weighted'),r2_score(y_test, predicted_QoI,
multioutput='variance_weighted'),r2_score(y2, predicted_QoI2,
multioutput='variance_weighted'),r2_score(y3, predicted_QoI3,
multioutput='variance_weighted')
        metric_list.append([[RMSE, MSE,MAE,R2],[RMSE1, MSE1,MAE1,R21],[RMSE2,
MSE2,MAE2,R22],[RMSE3, MSE3,MAE3,R23]])
        file_pi = open(dir +'\cv14_'+k+str(fold)+'10k3.obj', 'wb')
        pickle.dump(history.history, file_pi)
        file_pi.close()
        # if c == 3:
        #     model.save(dir +'/crorssval/'+k+str(fold)+'.h5',
include_optimizer=False, save_traces = False)
        fold += 1
metrics = np.array(metric_list).reshape(-1,16)
columns =
['RMSE','MSE','MAE','R2','RMSE1','MSE1','MAE1','R21','RMSE2','MSE2','MAE2','R22','R
MSE3','MSE3','MAE3','R23']
cv14 = pd.DataFrame(metrics, columns=columns)
cv14.to_excel(r'C:\Users\awada\Desktop\Thesis\ThesisProject\cv14_metrics3.xlsx')
```