Research Article

Prediction of MUET Results Based on K-Nearest Neighbour Algorithm

Norlina Mohd Sabri* and Siti Fatimah Azzahra Hamrizan

Universiti Teknologi MARA Cawangan Terengganu, Kuala Terengganu, Malaysia
norli097@uitm.edu.my; sitiazza1411@gmail.com
*Correspondence: norli097@uitm.edu.my

Received: 1st September 2022; Accepted: 5th May 2023; Published: 5th October 2023

Abstract: The machine learning based prediction has been applied in various fields to solve different kind of problems. In education, the research on the predictions of examination results is gaining more attentions among the researchers. The adaptation of machine learning for the prediction of students' achievement enables the educational institutions to identify the high failure rate, learning problems, and reasons for low student performance. This research is proposing the prediction of the Malaysian University English Test (MUET) results based on the K-Nearest Neighbour Algorithm (KNN). KNN is a powerful algorithm that has been applied in various prediction problems. The prediction of the MUET results would help the students and lecturers to be more well prepared and could improve the required English language skills accordingly before the actual examination. The MUET result prediction is based on the student's English courses grades and there are 516 data of students' results that have been collected from Universiti Teknologi MARA (UiTM) Dungun campus. The performance measurement that has been used are the mean accuracy, percentage error and mean squared error (MSE). In this research, the KNN prediction model has generated an acceptable performance with 65.29% accuracy. For future work, KNN could be modified or hybridized to further improve its performance. Furthermore, other algorithms could also be explored into this problem to further validate the best predictive model for the prediction of the MUET results.

Keywords: ***English Language Examination; K-Nearest Neighbour; MUET; Prediction***

1. Introduction

Malaysian University English Test (MUET) is a set of examinations which are conducted in Malaysia to measure the students' English proficiency levels before enrolling to universities. MUET is made compulsory as a pre-requisite requirement to the Bachelor degree programmes in Malaysian universities. MUET is consisting of four language skills which are speaking, listening, reading and writing. In MUET, the aggregated scores which could be obtained by students from 0 to 300 are classified on a Band scale ranging from 1 to 6. In the examination, the Band 1 represents the lowest band, while the Band 6 represents the highest band [1]. The students who are eligible for MUET are those from the Diploma, foundation, matriculation and form six students and they need to pass the MUET to be able to enrol to universities. Since the examination comes with a cost, the students must be well prepared to avoid taking the examination repeatedly. If they fail the MUET, redo will cause another expenditure and would also lose the chance to enrol to the universities. Therefore, for students to be well prepared for the MUET, a lot of efforts must be made by the students and the lecturers as well. Among of the efforts are to detect the students' weaknesses in any of the language skills earlier. After many efforts have been made, it is also important to predict the examination result as it could help in further detecting any weaknesses that need to be more addressed by the students and lecturers.

Currently, the research on the predictions of examination results which are based on the machine learning are gaining more attentions among the educators. One of the research which has adapted machine learning algorithms has shown that early prediction of students' grades might help to improve

Norlina Mohd Sabri and Siti Fatimah Azzahra Hamrizan, "Prediction of MUET Results Based on K-Nearest Neighbour Algorithm", *Annals of Emerging Technologies in Computing (AETiC)*, Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 50-59, Vol. 7, No. 5, 5th October 2023, Published by International Association for Educators and Researchers (IAER), DOI: 10.33166/AETiC.2023.05.005, Available: http://aetic.theiaer.org/archive/v7/v7n5/p5.html.

students' success rate, result and course selection. Gamulin *et al.* [2] predicts students' performance for the final examination based on the machine learning and it has successfully improved the learning process. In the research, the predictions have given the chance for students and teachers to improve based on their predicted results. The prediction on the final examination performance would also help the teachers, school authorities or other relevant parties to plan for the appropriate actions on the subject. The research done by Ma and Zhou [3] states that the adaptation of machine learning for the prediction of students' achievement or behaviour is very important for the educational institutions. The early prediction could help to identify the reasons for poor student achievement, poor learning ability and the high failure rate. Based on the predictions, students could also be classified into categories of low, medium, and high based on their academic scores. This could help in identifying the slow learners, where early actions could be taken to focus more on the weaker students [4].

Based on the significance of the prediction of the examination results, this research has proposed the prediction of MUET results for the Diploma students in Universiti Teknologi MARA (UiTM) based on the K-Nearest Neighbour (KNN) algorithm. KNN is a simple and easier to implement, but powerful algorithm that has been proven to perform acceptably in various classification problems [5-9]. According to Adeniyi *et al.* [8], KNN classifier has been consistent, simple to understand and easier to implement compared to other machine learning techniques. It has been reported that the KNN algorithm has been able to generate high accuracy in an examination performance prediction [10]. Therefore, for this research, KNN has been chosen to be explored for the prediction of the MUET results of the UiTM Diploma students.

2. Literature Review

This section provides brief reviews on the predictions of the examination results using machine learning algorithms and reviews on the K-Nearest Neighbour (KNN) algorithm. The summaries on the previous prediction of the examination results and the description on the KNN algorithm are presented in this section.

2.1. Predictions of Examination Results

There are many research that have been done in the machine learning based prediction of examination results. Ma and Zhou [3] have adapted machine learning algorithms for the prediction of the student pass rates from Portuguese school students. The aim of the research is to help students to be able to graduate on time and to help in handling the difficulties in their studies through data mining techniques. Gunawan *et al.* [11] have conducted a study on the prediction on on-time graduation for students and to discover what attributes impact the prediction based on decision tree algorithm. Research on predicting student's learning results has also been conducted in Indonesia. In the research, the algorithms that have been adapted are the Decision tree and Naïve Bayes [12]. Amra and Maghari [13] have applied KNN and Naïve Bayesian for the prediction of results for the secondary school students in Gaza. The teachers could then classify the performance of new students based on the classification model. KNN and SVM have also been adapted in the examination performance prediction in Iran [10]. Based on the previous research, the machine learning based predictions on the examination results have proven to be beneficial in the institutions. The prediction is important for educators in order to obtain early feedback and to take appropriate actions to improve the students' performance. The prediction could help to solve many students' performance related problems earlier and thus could help improving the institution's performance. Table 1 represents the summarization of the recent research on the examination result predictions. Different algorithms have been adapted in solving various examination results prediction problems. Most of the research have shown that the algorithms could solve the different examination results prediction with good accuracies. In the previous research, although KNN has not achieved the best performance in all of the predictions, the algorithm was still capable to generate acceptable and good performance in most of the examination result predictions. In this research, KNN has been selected to solve the MUET result prediction based on the capability of the algorithm in solving various other problems with good performance**.** The performance of KNN is to be explored and expected to be acceptable in solving this examination result prediction**.**

Table 1. Summary of the Recent Examination Result Predictions

No.	Technique/ Algorithm	Objective	Problem	Result	Reference
1.	Naïve Bayes, Bayesian Network (BN), SVM, Multilayer Perceptron (MLP), Decision Trees, KNN	To determine the best algorithm in predicting students' final performance	Complexity in the computer course lessons	Naïve Bayes generated the highest accuracy	[14]
2.	Logistic regression, linear discriminant analysis, KNN, Naive Bayes	To forecast students results	Monitoring progress of learners	Linear discrimination analysis performed the best in this problem	[15]
3.	ANN, Decision Tree, SVM and KNN	To find most suitable classifier for the student performance prediction	Need to predict the student performance and handle weaker students	ANN and Decision Tree performed better than other classifiers.	[16]
4.	SVM, Decision Trees, Feed forward feed-forward Multi-Layer Perceptron (MLP), KNN, Random Forest, Linear Discriminant Analysis and Logistic Regression	To find the most reliable prediction model for student performance	Need to analyse and determine the hidden and useful information from very large data	MLP performed better than other algorithms	[17]
5.	Decision tree classification with C4.5 algorithm	To forecast the students on-time graduation and to discover what attributes impact the prediction.	Many universities or colleges have not fully utilized the academic database with data mining	The accuracy of the algorithm is acceptable.	[11]
6.	SVM and DT	To help those who have difficulties in their studies.	Need to improve the passing rates of students and reduce the dropouts	Both algorithms have achieved good results.	[3]
7.	Naïve Bayes and DT	To predict results based on academic databases.	Need to explore student dataset from various academic databases	Naïve Bayes performed better than Decision Tree.	[12]
8.	KNN, NB	Classify new students according to their performance.	To forecast results for education ministry	Accuracy for Naïve Bayesian was better in this problem.	[13]
9.	SVM, KNN	To forecast the student's grade and compare the algorithms' performance accuracies.	Educators need to forecast student performance	Both SVM and KNN generated good accuracies.	[10]

2.2. K-Nearest Neighbour (KNN)

The simple KNN algorithm is an algorithm that could produce good results and has been widely adapted for classification and regression problems [18]. KNN algorithm has also depicted as a lazy learning due to its characteristic in classifying a new object based on the training samples and attributes [19]. KNN works by learning the training dataset and afterwards determine the label of the new object in the training dataset based on the labels of its closest neighbours. The technique considers that the nearby objects could have the same label. There are few methods available for choosing the k value. The common method is the trial and error, where the small odd numbers are tested one by one to obtain the desirable results. Another method is to add the number of classes in the research with 1. There are also methods to apply other algorithms to determine the k value. The common way of identifying the object's closest neighbours is to compare the distances of the individual object to other objects in the neighbourhood. In KNN, mathematical calculation has been used to measure the distances between objects. The commonly used distance metrics are Euclidean Distance, Minkowski Distance, Hamming Distance and Manhattan Distance. The basic procedures of the KNN algorithm are shown as the following [18]:

Algorithm 1. KNN Algorithm

1: Set the number of nearest neighbours, k
2: Evaluate the length for each object a and b in the training dataset $d(a,b)$
3: Select class according to majority of neighbourhood (the k neighbours)
4: Repeat steps 2–3 until the maximum iteration has been reached

Based on the procedures of KNN, the first step is to initialise the value of k, which is the number of nearest neighbours. The next step is to evaluate the length for each object a and b $d(a,b)$, with the selected distance formula. Then, the object will be classified according to the most of the k neighbours in the neighbourhood. The step 2 to step 3 processes will be repeated until all of the objects in the neigbourhood have been classified. The process will stop after the maximum iteration has been reached. Figure 1 shows that the algorithm would start with the initial value of k, then the processes of calculating the distance and selecting class would be repeated until the end of the iteration.

Due to the simplicity of implementation and good performance of KNN, the original KNN algorithm has been chosen to be implemented in solving the MUET result prediction problem. This is to investigate the capability of the original KNN in solving the MUET result prediction problem, while further improvement of the performance would be addressed in the future work of the research.

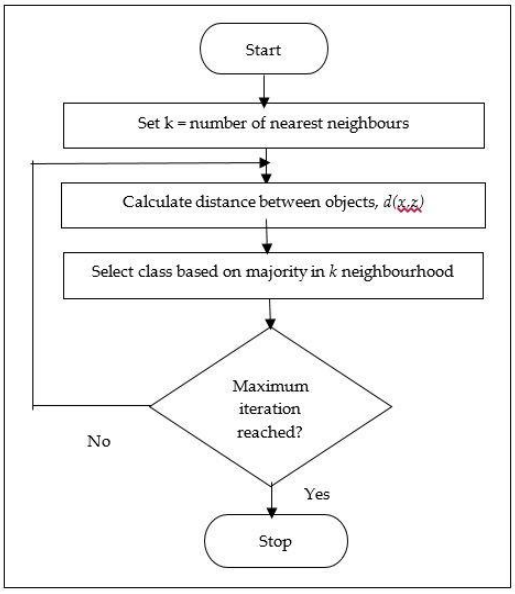

Figure 1. Flowchart of KNN Algorithm

3. Methodology

3.1. Experimental Data

In this initial research, the main data source was obtained from the Academic Affairs Division of Universiti Teknologi MARA (UiTM) Dungun campus. The collected data were consisting of the results of the English language courses from the 3 Diploma programmes in the campus. The programmes are the Diploma in Hotel Management, Diploma in Accountancy and Diploma in Electrical Engineering. The prediction of the MUET results is based on the 3 grades of the English language proficiencies from the Diploma courses which are the Integrated Language Skills I (ELC121/ELC120), Integrated Language Skills II (ELC151/ELC150) and Integrated Language Skills III (ELC231/ELC230). The grades from the courses represent the attributes for the prediction system. The data that had been collected for this research were the results of the students' English language proficiencies from the first semester of 2015 to second semester of 2016. There were 516 data of students' results that had been collected and the dataset was converted to numerical values from the grades. The grades were from A+ to C and had been converted to continuous numerical values as the data representation. These values were used in the calculation of the Euclidean distance in the KNN. The prediction is expected to be able to classify the band of the MUET based on the 3 attributes to the system. The classification is consisting of 6 levels of bands which are from Band 1 to Band 6. Table 2 shows the attributes for the prediction of the MUET results and the classification output.

Table 2. Attributes of the prediction and the classification output

Prediction Attributes	Integrated Language Skills I (ELC121/ELC120) grade
	Integrated Language Skills II (ELC151/ELC150) grade
	Integrated Language Skills III (ELC231/ELC230) grade
Classification Output	Level 1 to Level 6

3.2. Prediction System Architecture

System architecture represents the conceptual model which contains the details of the system structure with views of the system. Figure 2 depicts the prediction system architecture that was used in this project. It consists of three major phases which are user interface, data collection and preparation and the engine part. The prediction system architecture starts with the data collection and preparation phase. In this phase, the data are collected from the Academic Affairs Division of UiTM Dungun campus. Data pre-processing is where the data is cleaned and transformed into a suitable form to be used for the prediction process. Then the data would be trained using the engine from the KNN algorithm. From the user interface phase, the prediction of the MUET result begins with user enter the data of the grade of results for each of English proficiency language. The input data will be processed by the engine which is based on the KNN classification technique for the prediction process. The back end process starts with calculation of the length between input data and training data. After calculating the distance, the algorithm sorts the distance to the closest input data among the training data and lastly it selects the class from most of the closest data. As a result, the output of the prediction is the MUET result in Band level from 1 to 6, to be displayed to the user.

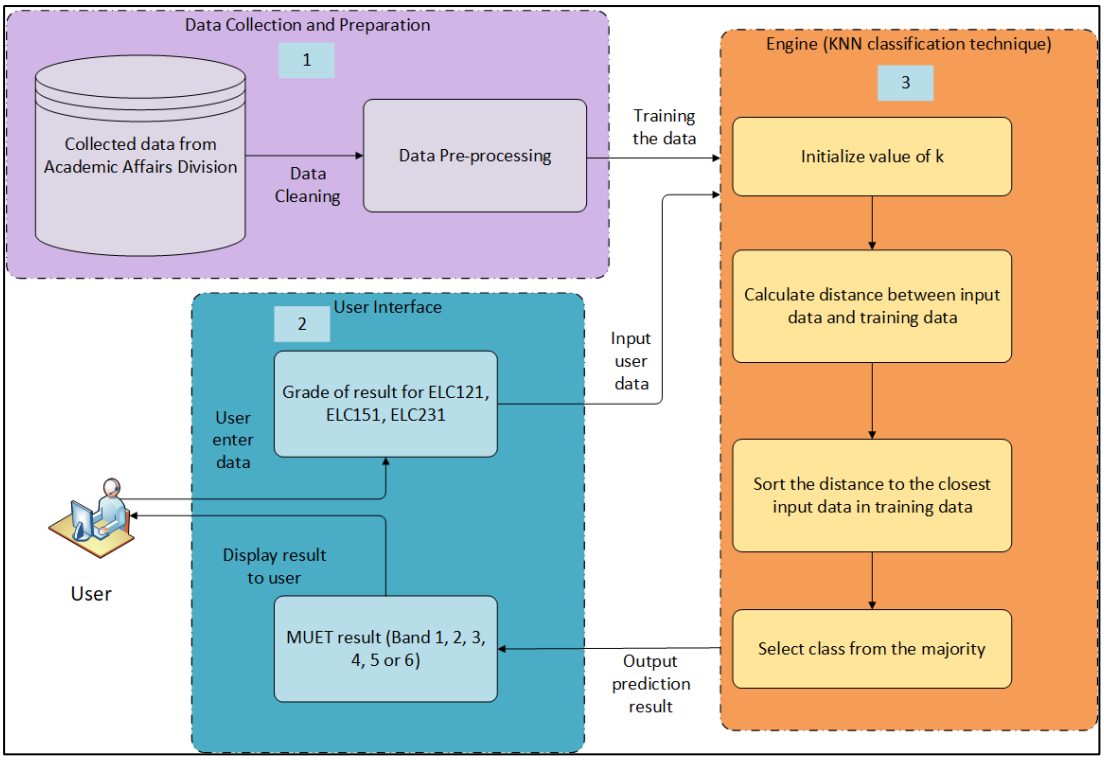

Figure 2. Prediction System Architecture

3.3. Prediction Model Process Flow

The flowchart of the KNN based prediction model is shown in Figure 3. The flowchart represents a workflow or process of the implementation of the prediction model. Based on Figure 3, the process starts with the data acquisition and data pre-processing. Pre-processed data has been transformed into the suitable form that could be used by the algorithm. The data will be trained for the prediction of the results after the pre-processing phase. Next, using the graphical user interface, the user will input their English course grades which are the ELC121, ELC151 and ELC231 results. The back-end engine will predict MUET result from the data using the KNN algorithm. Euclidean Distance has been used to calculate the length

between the input data and the neighborhood data. After that, it will sort the distance to the nearby input data among the neighbourhoods. Then, it will select the class of the majority of the neighbourhood, and it will display to the user the output of which band level that has been predicted based on the English course results. The band level is ranged from level 1 to level 6. Level 1 represents the lowermost band, while level 6 represents the uppermost band.

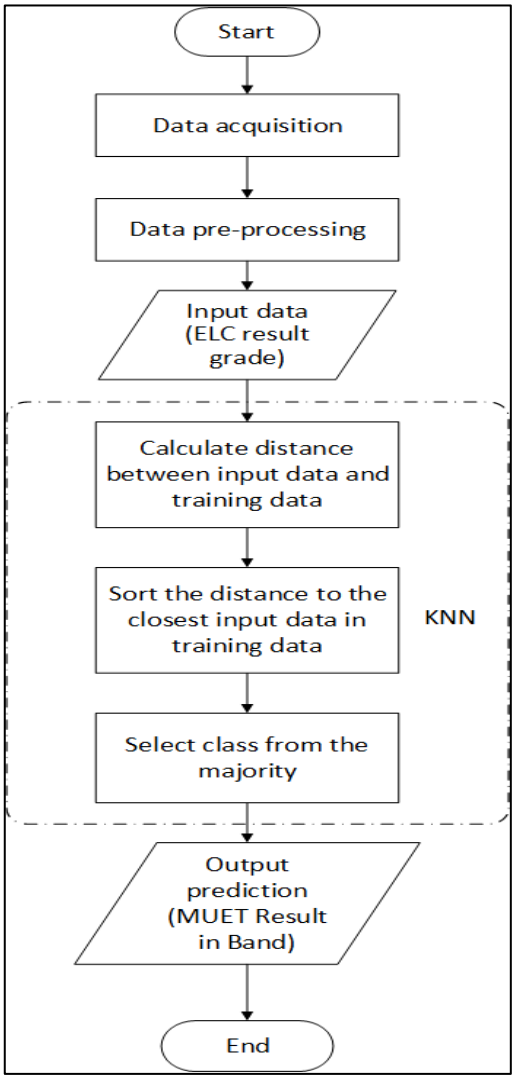

Figure 3. Flowchart of the MUET results prediction process

4. Results and Finding

In the performance evaluation, there are several values of k that has been tested for the KNN algorithm to get the best results. The values of k that have been tested were the small, odd numbers because generally an odd number is used if the number of classes is 2 or more. In this research, the value of k is related to the error rate of the model. In order to prevent overfitting or underfitting, various values of k has been tested. The results of the various k that have been tested are presented in the mean accuracy, percentage error and the mean squared error.

4.1. Mean Accuracy

The mean accuracy is calculated from the average of the accurate results over the overall dataset. The results of the mean accuracies of the prediction with different values of k starting from the odd number of 1 until 9 are shown in Table 3. In the testing, the lowest accuracy is obtained when the k is 3 with the mean accuracy of 57.45% and the highest accuracy is achieved when the k is 9 with the mean accuracy of 65.29%.

Table 3. Mean Accuracy of KNN Prediction

Value of *k*	Mean Accuracy (%)
k = 1	59.02
k = 3	57.45
k = 5	60.39
k = 7	63.14
k = 9	65.29

To summarize the results that have been shown in Table 3, the trend of the mean accuracy of the prediction model is shown in Figure 4. The graph shows that the mean accuracy has increased from the value of *k*=3 up until *k*=9. Based on the graph, the mean accuracies of the prediction have been improving with the increased value of *k*. For the testing phase, the value of *k*=9 has been chosen based on the best accuracy results that has been achieved.

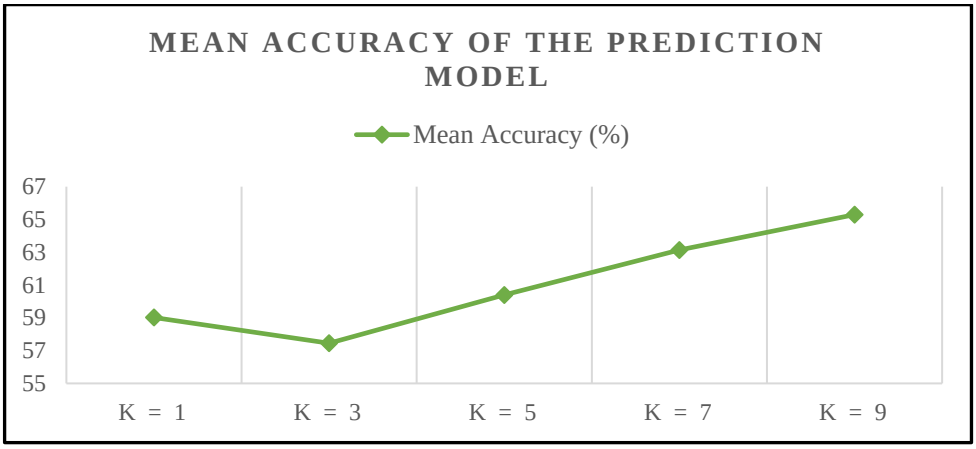

Figure 4. Graph of Mean Accuracy

4.2. Percentage Error

The percentage error is the opposite of the percentage accuracy. The lower percentage error represents the lower error of the prediction. The results in Table 4 shows that the lowest error is when the *k*=9 with 25.49% error and the highest error is when the value of *k* is 1 with 35.29% of error.

Table 4. Percentage Error

Value of k	Percentage Error (%)
k = 1	35.29
k = 3	31.37
k = 5	29.41
k = 7	27.45
k = 9	25.49

To summarize the results that have been shown in Table 4, Figure 5 shows the trend of the percentage error of the prediction model. The graph shows that the percentage error decreases from the value of *k* = 1 up to *k* = 9. This shows that the error is continually decreasing with the increasing value of *k*.

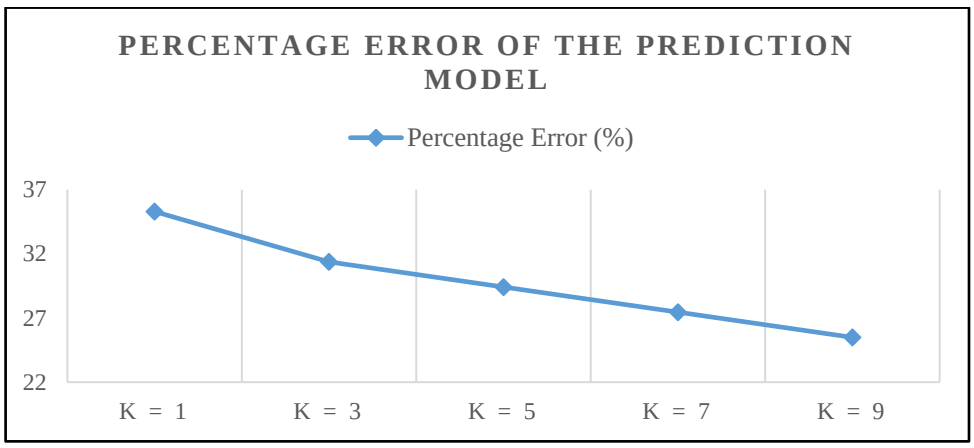

Figure 5. Graph of Percentage Error

4.3. Mean Squared Error

Mean Squared Error (MSE) is measured by the square root of the average square difference between the correct value and predicted value. The value that is closer to zero shows the better the prediction model is. The results of the MSE for each of the k value are shwon in Table 5. In this research, the lowest MSE value is 0.63 when the value of *k* is 9 and the highest MSE value is 0.77 when the *k* value is 1. Based on the results, the value of *k*=9 could yield better prediction.

Table 5. Mean Squared Error (MSE)

Value of *k* for KNN	Mean Squared Error
k = 1	0.77
k = 3	0.66
k = 5	0.71
k = 7	0.67
k = 9	0.63

Figure 6 shows the trend of the MSE with the different k values. The graph shows the highest MSE value is at *k* = 1 and the lowest is at *k* = 9. The MSE has decreased with the bigger number of *k*. In this research, the larger value of *k* is more significant as it produces smaller error in the prediction.

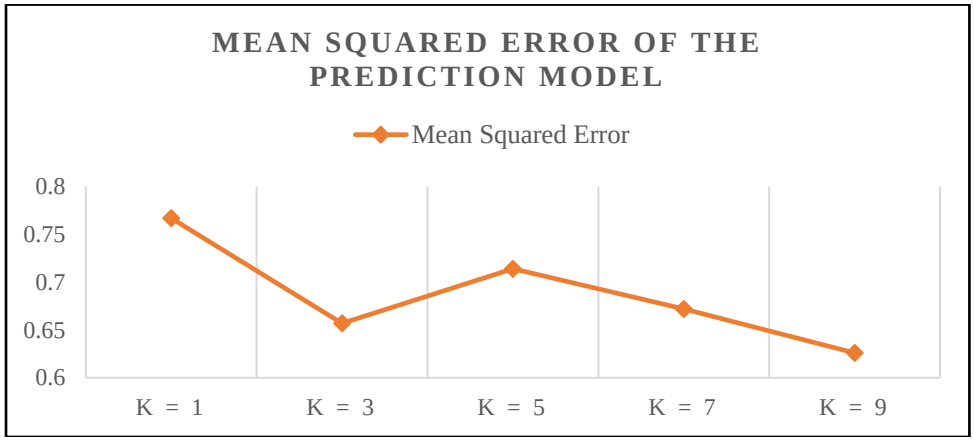

Figure 6. Graph of Mean Squared Error (MSE)

4.4. Prediction Model Interface

The prediction model interface has been developed from the built-in library of Streamlit. Streamlit was used to deploy the prototype into the web-app. Streamlit is a library based in Python open-source that enables a simple personalised web app for machine learning and data science to be easily constructed and distributed. The prediction model starts with the user entering the ELC121, ELC151 and ELC231 grades and the prediction of the MUET result is obtained by pressing the Enter key. The result would be displayed as shown in Figure 7.

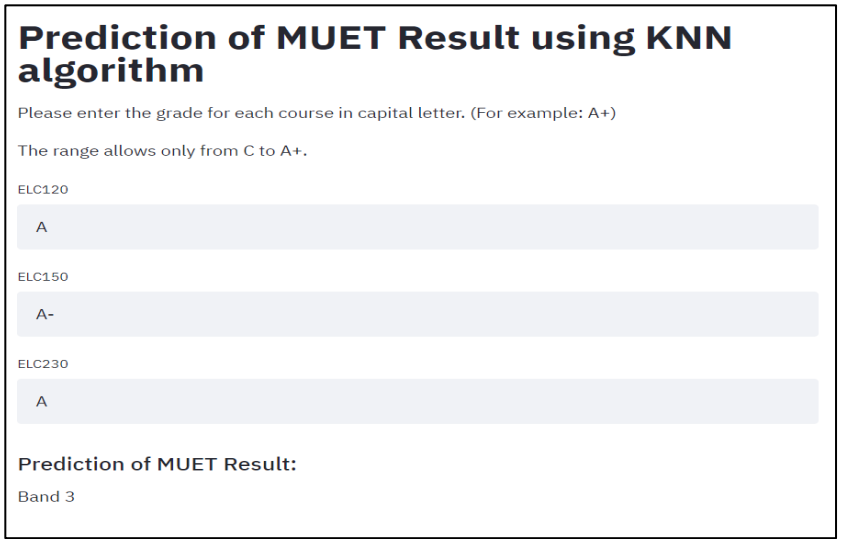

Figure 7. Prediction Model User Interface

5. Conclusion and Recommendation

This research has implemented the K-Nearest Neighbour (KNN) algorithm to predict the Malaysian University English Test (MUET) results based on the Diploma English courses grades. In this research, KNN has been initially explored to test its performance to solve the MUET prediction problem. Based on the performance evaluation, KNN has shown an acceptable performance with the accuracy of 65.29% in the prediction. The prediction model has been developed as a tool to help the Diploma students of Universiti Teknologi MARA (UiTM) to predict their MUET band level to apply for the Bachelor programmes in Malaysian universities. The results of the prediction could enable students to prepare for the actual MUET and improve their weakness earlier in the English language skills. As for the lecturers, they could figure out which language skills are the most problematic to students and could later focus on the specific language skills to help students to pass the MUET. For the future prediction model improvement, there are also other relevant English examination results that could be considered as the attributes to the system such as the English placement test results, Sijil Pelajaran Malaysia (SPM) English result and the student's CGPA. These results might further improve the accuracy of the prediction since they are related to the students' English performance. Moreover, the accuracy of the prediction might also be higher if larger number of datasets have been used in the training and testing phases. Modified or hybrid techniques could also be considered to further improve the performance of KNN in solving this particular problem. For future work, KNN could be modified or hybridized to further improve or enhance its performance. Furthermore, other algorithms such as Naïve Bayes, ANN and SVM could also be tested and compared with the KNN algorithm in the prediction of the MUET results. The exploration of various algorithms into this problem could further validate the best predictive model for the prediction of the MUET results.

Acknowledgement

Special gratitude goes to Universiti Teknologi MARA Cawangan Terengganu for the support given to this research.

References

[1] Yarshini Thamilarasan and Raja Rina Raja Ikram, "MyMUET : A Design of a Mobile based Crowdsourced Assessment for Malaysia University English Test with Non-Synchronous Participant Interaction", *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, ISSN: 2278-3075, pp. 2971–2977, Vol. 8, No. 11, September 2019, Published by Blue Eyes Intelligence Engineering & Sciences Publication, DOI: 10.35940/ijitee.K2277.0981119, Available: https://www.ijitee.org/portfolio-item/K22770981119/.

[2] Jasna Gamulin, Ozren Gamulin and Dragutin Kermek, "Data mining in hybrid learning: Possibility to predict the final exam result", in *Proceedings of the 2013 36th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 20-24 May 2013, Opatija, Croatia, ISBN: 978-953-233-073-1, pp. 591–596, Published by IEEE, Available: https://ieeexplore.ieee.org/document/6596327.

[3] Xiaofeng Ma and Zhurong Zhou, "Student pass rates prediction using optimized support vector machine and decision tree", in *Proceedings of the 2018 IEEE 8th Annual Computing and Communication Workshop and Conference*, 27 February 2018, Las Vegas, NV, USA, ISBN: 978-1-5386-4649-6, DOI: 10.1109/CCWC.2018.8301756, pp. 209–215, Published by IEEE, Available: https://ieeexplore.ieee.org/document/8301756.

[4] Iti Burman and Subhranil Som, "Predicting Students Academic Performance Using Support Vector Machine", in *2019 Amity International Conference on Artificial Intelligence (AICAI)*, 4-6 Feb. 2019, Dubai, United Arab Emirates, ISBN: 978-1-5386-9346-9, DOI: 10.1109/AICAI.2019.8701260, pp. 756–759, Published by IEEE, Available: https://ieeexplore.ieee.org/document/8701260.

[5] Gegap Putri Gempita, Diyanti Wilasari, Prisma Kristalina and Sritusta Sukaridhoto, "Implementation of K-NN Fingerprint Method on Receiving Server for Indoor Mobile Object Tracking", in *Proceedings of the 2019 International Electronics Symposium (IES)*, 27-28 Sept. 2019, Surabaya, Indonesia, pp. 411–416, ISBN: 978-1-7281-4449-8, DOI: 10.1109/ELECSYM.2019.8901570, Available: https://ieeexplore.ieee.org/document/8901570.

[6] Okfalisa, Ikbal Gazalba, Mustakim and Nurul Gayatri Indah Reza, "Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification", in *Proceedings of the 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 1-2 November 2017, Yogyakarta, Indonesia, ISBN: 978-1-5386-0658-2, DOI: 10.1109/ICITISEE.2017.8285514, pp. 294–298, Published by IEEE, Available: https://ieeexplore.ieee.org/document/8285514.

[7] Syed Hasan Adil, Mansor Ebrahim, Kamran Raza and Syed Saad Azhar Ali, "Prediction of Eye State Using KNN Algorithm", in *Proceedings of the 2018 International Conference on Intelligent and Advanced System (ICIAS)*, 13-14 August 2018, Kuala Lumpur, Malaysia, ISBN: 978-1-5386-7269-3, DOI: 10.1109/ICIAS.2018.8540596, pp. 1–5, Published by IEEE, Available: https://ieeexplore.ieee.org/document/8540596.

[8] Adedayo D. Adeniyi, Zune Wei and Y. Yongquan, "Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method", *Applied Computing and Informatics*, pp. 90–108, Vol. 12, No. 1, January 2016, Published by Elsevier B.V., DOI: 10.1016/j.aci.2014.10.001, Available: https://www.sciencedirect.com/science/article/pii/S221083271400026X.

[9] Wan Rosemehah Wan Omar, Norfaiza Fuad, Mohd Nasir Taib, Rozita Jailani, Roshakimah Mohd Isa *et al.*, "Brainwave classification for acute ischemic stroke group level using k-NN technique", in *Proceedings of the 2014 5th International Conference on Intelligent Systems, Modelling and Simulation*, 27-29 January 2014, Langkawi, Malaysia, Electronic ISBN: 978-1-4799-3858-2, Published by IEEE, DOI: 10.1109/ISMS.2014.26, pp. 117–120, Available: https://ieeexplore.ieee.org/document/7280890.

[10] Huda Al-shehri, Amani Al-Qarni, Leena Al-Saati, Arwa Batoaq, Haifa Badukhen *et al.*, "Student Performance Prediction Using Support Vector Machine and K-Nearest Neighbor", in *Proceedings of the 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, 30 April-3 May 2017, Windsor, ON, Canada, ISBN: 978-1-5090-5538-8, pp. 17–20, Published by IEEE, DOI: 10.1109/CCECE.2017.7946847, Available: https://ieeexplore.ieee.org/document/7946847.

[11] Gunawan, Hanes, and Catherine, "Information Systems Students' Study Performance Prediction Using Data Mining Approach", in *Proceedings of the 2019 Fourth International Conference on Informatics and Computing (ICIC)*, 16-17 October 2019, Semarang, Indonesia, ISBN: 978-1-7281-2207-6, DOI: 10.1109/ICIC47613.2019.8985718, Published by IEEE, Available: https://ieeexplore.ieee.org/document/8985718.

[12] Masna Wati, Wahyu Indrawan, Joan Angelina Widians and Novianti Puspitasari, "Data mining for predicting students' learning result", in *Proceedings of the 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT)*, 8-10 August 2017, Kuta Bali, Indonesia, ISBN: 978-1-5386-0600-1, DOI: 10.1109/CAIPT.2017.8320666, pp. 1–4, Available: https://ieeexplore.ieee.org/document/8320666/.

[13] Ihsan A. Abu Amra and Ashraf Y. A. Maghari, "Students Performance Prediction Using KNN and Naïve Bayesian", in *Proceedings of the* 2017 *8th International Conference on Information Technology (ICIT)*, 17-18 May 2017, Amman, Jordan, pp. 909–913, ISBN: 978-1-5090-6332-1, Published by IEEE, DOI: 10.1109/ICITECH.2017.8079967, Available: https://ieeexplore.ieee.org/document/8079967.

[14] Chia Yin Ko and Fang Yie Leu, "Examining Successful Attributes for Undergraduate Students by Applying Machine Learning Techniques", *IEEE Transactions on Education*, ISSN: 1557-9638, pp. 50–57, Vol. 64, No. 1, February 2021, DOI: 10.1109/TE.2020.3004596, Available: https://ieeexplore.ieee.org/document/9144429/.

[15] Hina Gull, Madeeha Saqib, Sardar Zafar Iqbal and Saqib Saeed, "Improving Learning Experience of Students by Early Prediction of Student Performance using Machine Learning", in *Proceedings of the 2020 IEEE International Conference for Innovation in Technology (INOCON)*, 6-8 November 2020, Bangluru, India, pp. 1–4, ISBN: 978-1-7281-9744-9, DOI: 10.1109/INOCON50539.2020.9298266, Available: https://ieeexplore.ieee.org/document/9298266/.

[16] Eman Al Nagi and Nailah Al-Madi, "Predicting Students Performance in Online Courses using Classification Techniques", in *Proceedings of the 2020 International Conference on Intelligent Data Science Technologies and Applications (IDSTA)*, 19-22 October 2020, Valencia, Spain, pp. 58–65, ISBN: 978-1-7281-8376-3, DOI: 10.1109/IDSTA50958.2020.9264113, Available: https://ieeexplore.ieee.org/document/9264113/.

[17] Thaer Thaher and Rashid Jayousi, "Prediction of Student's Academic Performance using Feedforward Neural Network Augmented with Stochastic Trainers", in *Proceedings of the 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT)*, 7-9 October 2020, Tashkent, Uzbekistan, ISBN: 978-1-7281-7386-3, DOI: 10.1109/AICT50176.2020.9368820, Published by IEEE, Available: https://ieeexplore.ieee.org/document/9368820/.

[18] Mark Wickham, *Practical Java Machine Learning*, 1st ed. Berkley, United States: Apress, 2018.

[19] Mohammed KH Zoroub and Ashraf Y. Maghari, "Candidate teacher performance prediction using classification techniques: A case study of high schools in Gaza-strip", in *Proceedings of the 2017 International Conference on Promising Electronic Technologies (ICPET)*, 16-17 Oct. 2017, Deir El-Balah, Palestine, ISBN: 978-1-5386-2269-8, pp. 129–134, DOI: 10.1109/ICPET.2017.30, Available: https://ieeexplore.ieee.org/document/8109052/.