Development of an Academic Performance Monitoring System Using Least Squares Regression for Predicting Student Academic Performance Status on Professional Courses

Cheene Rose C. Pilarca[1], Defaje Marie O. Opiña[2], Jhon Simon B. Leuterio[3], Fretzl Ann S. Toledo[4], Marjorie M. Bellones[5], Rexes S. Gerona[6], Jason P. Sermeno[7*]

Abstract: This study deals with the development of an academic performance monitoring system using least squares regression to predict a student's academic performance status on professional courses. The proposed system monitors the student's academic performance to promote education and achieve learning outcomes. The study aims to design and create a Graphical User Interface (GUI) and its functions, to utilize MATLAB for forecasting and a least squares regression model to show and predict the relationship between two factors, and to evaluate the system's predicting accuracy using Percentage Forecast Error (PFE) and Mean Absolute Percent Error (MAPE) to compare actual and forecasted data for every semester. The sample size includes 182 students of BS Information Technology, of whom thirteen (13) were successfully chosen using simple random sampling with their GPAs for the purpose of predicting future academic performance.

Keywords: Academic Performance, Least Squares Regression, Forecasting Method, Prediction, Percentage Forecast Error

[1] College of Computer Studies, University of Antique, Sibalom, Antique, Philippines
Email: singkit003@gmail.com
[2] College of Computer Studies, University of Antique, Sibalom, Antique, Philippines
Email: defayemjez@gmail.com
[3] College of Computer Studies, University of Antique, Sibalom, Antique, Philippines
Email: simonleuterio@gmail.com
[4] College of Computer Studies, University of Antique, Sibalom, Antique, Philippines
Email: toledofretzlann@gmail.com
[5] College of Computer Studies, University of Antique, Sibalom, Antique, Philippines
Email: marjbellones@gmail.com
[6] College of Computer Studies, University of Antique, Sibalom, Antique, Philippines
Email: geronarexes@gmail.com
[7*] College of Computer Studies, University of Antique, Sibalom, Antique, Philippines
Email: jasonpsermeno@gmail.com (Corresponding Author)

Received [December 26, 2022]; Revised [March 20, 2023]; Accepted [April 11, 2023]

Published by InnoCon Publishing
ISSN 2704-4440

1. Introduction

Tracking student progress in a class plays an important role in measuring the effectiveness of a teaching method, a course, and the teacher [1]. Students must be aware of their learning growth and review how it relates to bigger goals, which increases their investment. In addition, it helps teachers stay invested and continue to increase their own effectiveness by seeing trends in student data. Tracking is necessary to ensure you meet state standards, but checking off whether you've covered a curriculum requirement doesn't necessarily mean that your students are learning the required skills. Tracking student growth lets you know how and when to adjust information as you teach to guarantee students' learning. The teacher can move forward in their lessons whenever the students are performing well to keep them engaged and immersed. However, the teacher can always slow down with the learning process and supplement necessary lessons whenever the students aren't mastering the skills required for the lesson [2][3]. Keeping track of performance data gives you the ability to respond quickly and optimize your students' education. Tracking of the student's progress also indicates when to recognize a student's successes, improvements, and failures [4][5].

Universities in South Korea are using a curriculum to track and assess their students' academic performance [6]. The National Assessment of Educational Accomplishment (NAEA) is a comprehensive assessment system that not only maintains school education quality but also examines the current state and trends in Korean students' academic achievement using the national curriculum [7]. The NAEA assesses each student's academic performance. The assessment's findings help to clarify students' overall performance levels by school and region, and they can be used to hold educational institutions accountable.

In addition, some Korean universities have employed the Assessment of Higher Education Learning Outcomes (AHELO) [8]. The AHELO aims to complement the institution-based assessments by providing a direct evaluation of student learning outcomes at the global level and to enable institutions to benchmark the performance of their students against their peers as part of their improvement efforts. Given AHELO's global scope, it is essential that measures of learning outcomes are valid across diverse cultures and languages as well as different types of higher education institutions (HEIs) [9].

Similarly, in universities in the Philippines, they are using a formative approach to assessing their students' academic performance [10]. In Western Visayas, universities use the students' Grade Point Average (GPA) to track the students' academic performance. A student's GPA is typically measured on a scale of zero to four, with higher GPAs representing higher grades in the classroom [11][12]. Somehow, at the University of Antique, the institution uses this kind of strategy, which involves the traditional paper and pen in assessing, tracking, and monitoring the academic performance of the students. Its current system focuses more on academic tracking and monitoring the academic status of a student. With the current issues, the university's system is not capable of predicting the students' performance status for the next two succeeding semesters. They don't have the capability to forecast the students' performance for the next semester based on the trends. In this case, prediction is an important aspect because it can also produce or generate data that could be used later for enrolment purposes [13][14]. And also, it will enable the heads of the colleges in every university to decide what other courses they are going to offer for the next semester or summer for those students who will be flanking [15].

This research is now geared toward the development of an academic performance monitoring system using least squares regression [16][17] for predicting academic performance status on professional courses of students at the University of Antique. The application is simple to use and navigate; by entering the grades of students, the faculty can examine the students' performance.

The rest of this paper is organized as follows: Section 2 outlines the methodology used in conducting the study and developing the LSR prediction system; Section 3 details the results and discussion; and Section 4 discusses the conclusions and recommendations for further studies.

2. Methodology

The developmental method was used to develop an academic performance monitoring system using least squares regression to predict the academic performance status of professional courses that are taken by the students in the next two succeeding semesters at the University of Antique.

The Mean Absolute Percentage Error (MAPE) measures the accuracy of the forecasting that was used in the academic performance monitoring system. It represents the average of the absolute percentage errors of each entry in a dataset, showing, on average, how accurate the forecasted quantities were in comparison with the actual quantities as indicated in equation (1).

$$MAPE = \frac{(\text{ABS (Actual}-\text{Forecast))}}{\text{Actual}} \text{ X } 100 \quad (1)$$

The linear regressions are used to show or predict the relationship between two variables or factors, as indicated in equation (2). The factor that is being predicted is called the dependent variable. The factors that are used to predict the value of the dependent variable are called the independent variables.

$$y = a + bX \quad (2)$$

A sample of random sampling is a randomly selected subset of a population. In this sampling method, each member of the population has an exactly equal chance of being selected. This method was chosen because it uses randomization; any research performed on this sample should have high internal and external validity.

In addition, the prototype model was used for the development of the software. The software prototype is created prior to the actual software to get valuable feedback. Feedback is implemented, and the prototype is again reviewed for any changes. This process goes on until the desired outcome is met. Figure 1 represents the sequence and duration of the software development.

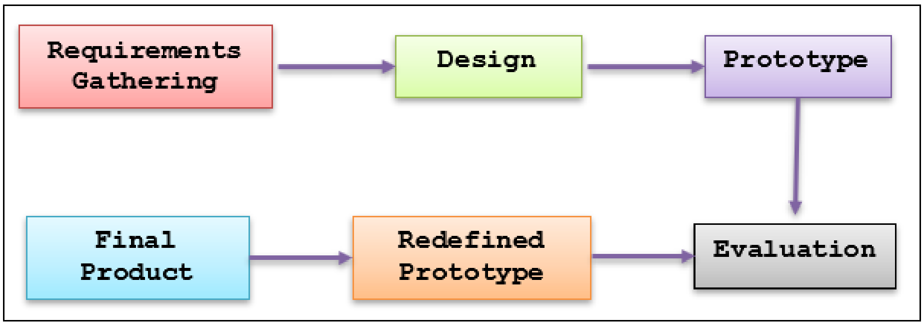

Figure 1. Prototyping Model

In this study, the Linear Regression Forecasting method was used to predict the students' academic performance by Cumulative Grade Point Average (CGPA) and Course Completion Rate on professional courses at the University of Antique. In this study, the linear regression forecasting method was used, which uses the data of students to make informed estimates that are predictive in determining the direction of future trends.

The system's core functions are limited to student information recording, predicting academic performance, collecting, and gathering academic reports. The study was conducted during the academic year 2015-2019. The first semester of AY 2018-2019 and the second semester of AY 2018-2019 have been forecasted by the proponents. The result of the forecasted students' performance status was compared to the actual students' performance to determine the percentage forecast error. This study is applicable to all departments and courses at the University of Antique, but in the meantime, this study will only be conducted on the records of the Bachelor of Science in Information Technology graduates at the University of Antique as test subjects.

3. Results and Discussion

3.1 The Interface and Its Function

3.1.1 Extracting and Managing Data from CSV Files

An interface for extracting and managing data that is taken from a CSV file was designed and created. By extracting the CSV dataset file into the program, the program is able to manage and sort the list of IT students with their data. Figure 2 depicts the graphical user interface for extracting the data.

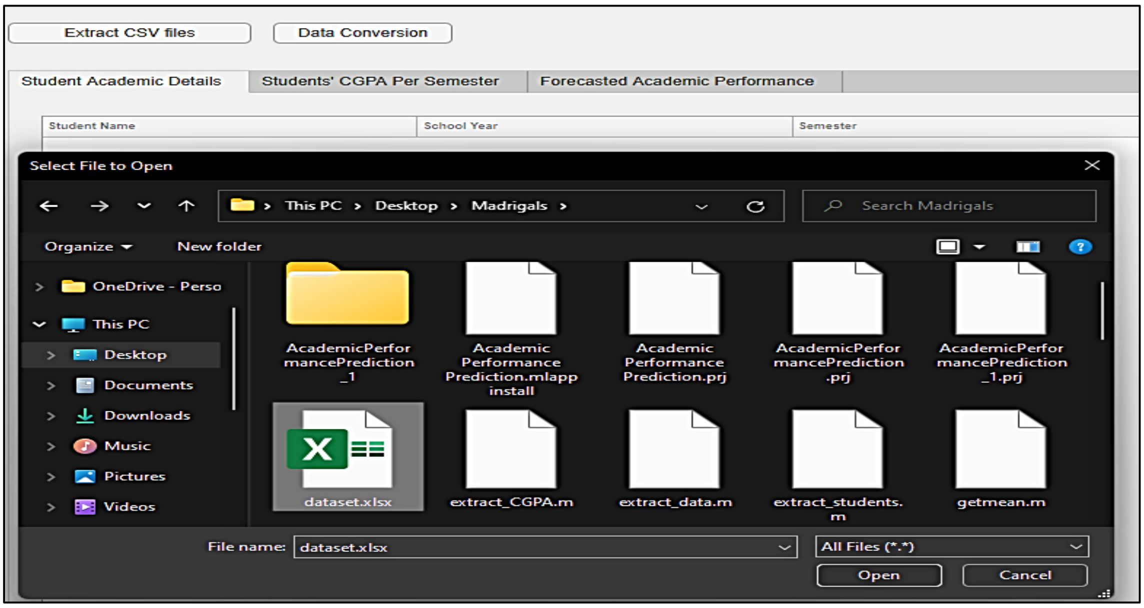

Figure 2. Data Extraction

3.1.2 Generate Summary

The proponents were able to design and create an interface for generating a summary of the cumulative grade point average of each student for every semester they are enrolled, including the number of courses they took and the number of courses they conditionally passed. This interface performs grade transmutation into a two-digit point system in order for the program to analyze and plot the prediction values in a graph. The data conversion of the grade summary is presented in Figure 3.

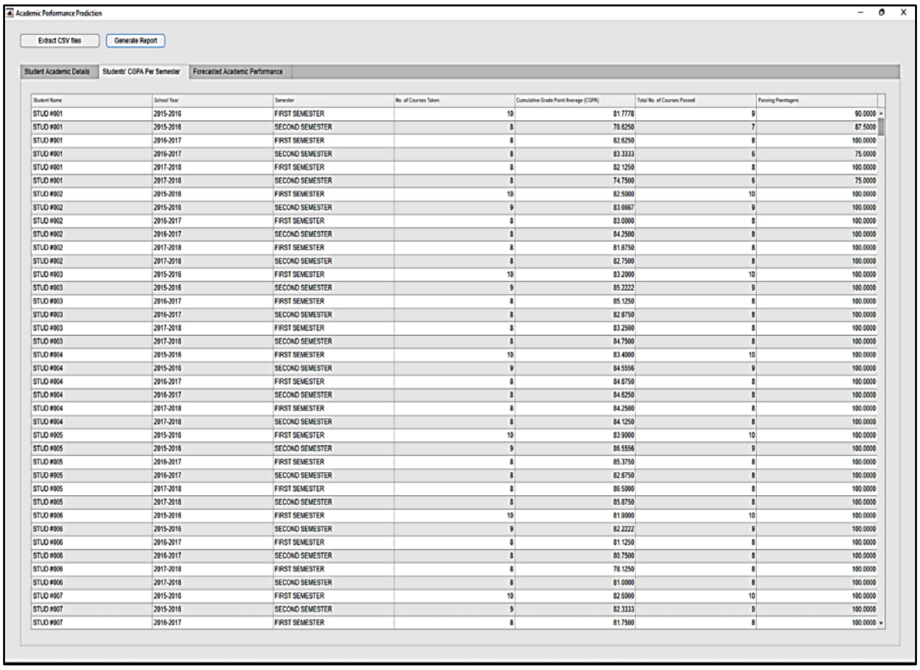

Figure 3. Data Conversion

3.2 Data Prediction Analysis Interface

3.2.1 Predicting the CGPA

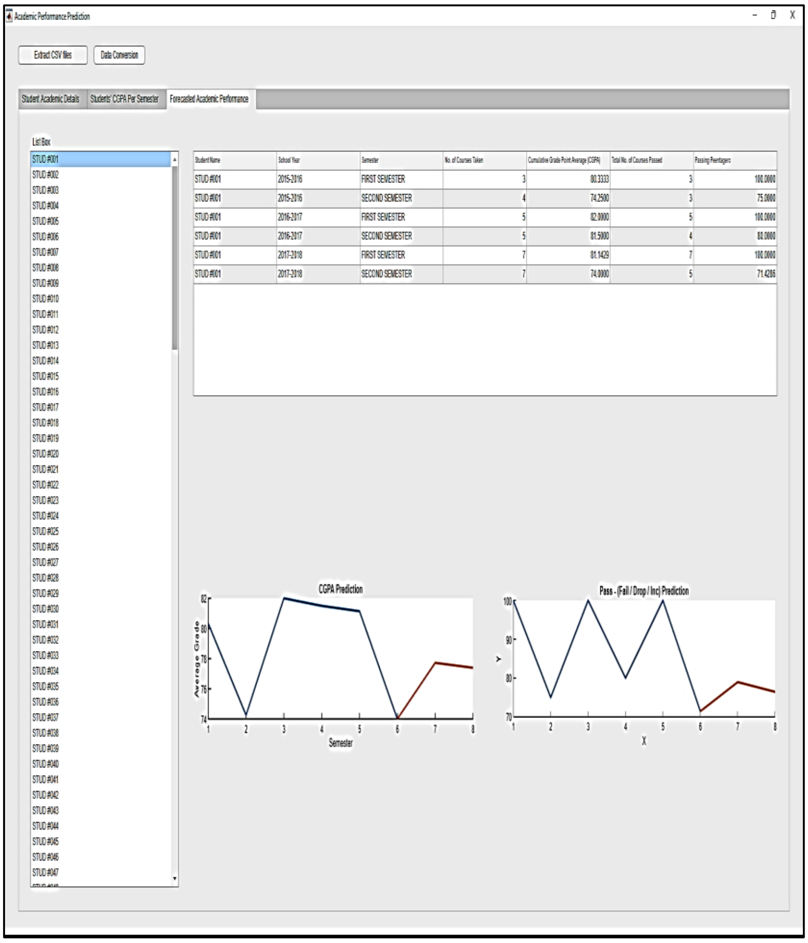

Figure 4. CGPA Interface

Figure 4 illustrates the students' Forecasted Academic Performance. The system takes in six existing CGPA values from the academic years 2015-2016 to 2017-2018 in order to predict the selected students' CGPA for the next two semesters of AY 2018-2019.

The red line in the CGPA Prediction Pass/Fail graph represents the student's forecasted CGPA and CCR for the next two semesters.

3.2.2 Predicting the CCR

Figure 5 illustrates the students' Forecasted Academic Performance. The system takes in six existing course completion rate (CCR) values from the academic years 2015-2016 to 2017-2018 in order to predict the selected students' CCR for the next two semesters of AY 2018-2019.

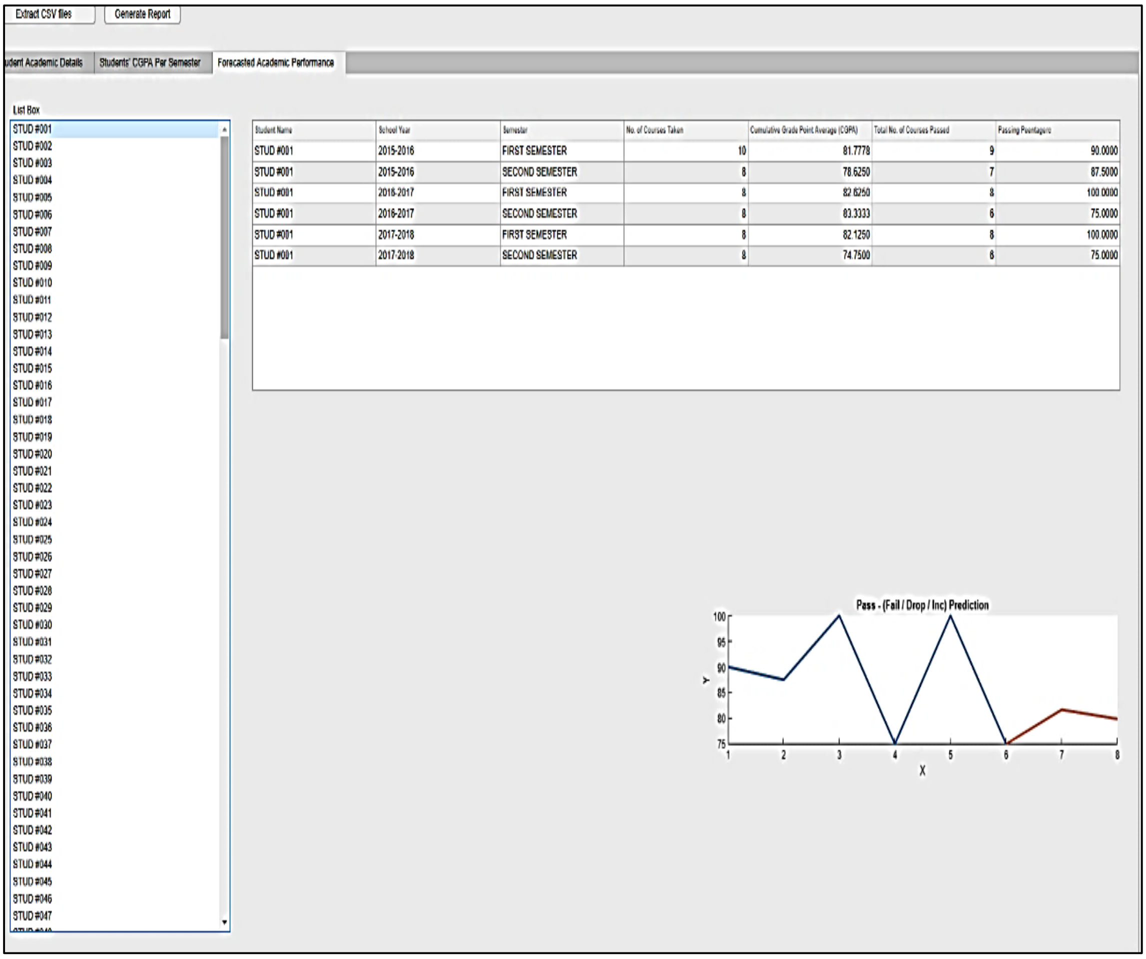

Figure 5. CCR Prediction

3.3 System Evaluation Using PFE and MAPE

3.3.1 Result in Predicting the CGPA

The results of the CGPA prediction are shown in Table 1. It shows the students' Forecasted Academic Performance. The proponents used the students' academic data from the first and second semesters' CGPAs from 2015-2018 to predict the next two semesters of AY 2018-2019. The table shows the list of the students with their forecasted and actual CGPA for the first and second semesters, and the PFE-CGPA of each student. The table also presents the MAPE of the 13 students picked for testing.

Table 1. Predicting CGPA for First and Second Semesters of School Year 2018-2019

Randomly Selected Students	1st Semester, 2018-2019			2nd Semester, 2018-2019		
	F-CGPA	A-CGPA	PFE-CGPA	F-CGPA	A-CGPA	PFE-CGPA
Student 14	86.17	85.67	0.58	86.34	86	0.40
Student 20	81.94	81.17	0.95	81.89	93	11.95
Student 31	78.03	71.67	8.87	77.67	70	10.96
Student 47	80.10	81	1.11	79.72	90	11.41
Student 67	83.77	83.33	0.53	84.08	89	5.53
Student 76	79.87	77	3.73	80.37	82	1.99
Student 94	84.50	81.50	3.68	85.83	91	5.68
Student 115	81.23	81.57	0.42	82.20	79.5	3.40
Student 118	81.92	82.33	0.50	82.20	85	3.29
Student 129	81.20	81.50	0.37	81.17	88	7.76
Student 134	84.87	83.71	1.39	84.90	76.5	10.98
Student 165	82.84	80.83	2.49	82.73	94	11.99
Student 180	81.52	82.33	0.98	81.46	88	7.43
MAPE-CGPA	1.97			7.14		

3.3.2 Result in Predicting the CCR

Table 2. Predicting CCR for First and Second Semesters of School Year 2018-2019

Randomly Selected Students	1st Semester, 2018-2019			2nd Semester, 2018-2019		
	F-CCR	A-CCR	PFE-CCR	F-CCR	A-CCR	PFE-CCR
Student 14	100	100	0.00	100	100	0.00
Student 20	100	100	0.00	100	100	0.00
Student 31	100	50	100.00	42.86	25	71.44

Student 47	100	100	0.00	100	100	0.00
Student 67	100	100	0.00	100	100	0.00
Student 76	100	100	0.00	100	100	0.00
Student 94	100	100	0.00	81.43	100	18.57
Student 115	100	100	0.00	100	100	0.00
Student 118	100	100	0.00	100	100	0.00
Student 129	100	100	0.00	100	100	0.00
Student 134	100	100	0.00	99.68	100	0.32
Student 165	100	100	0.00	100	100	0.00
Student 180	100	100	0.00	100	100	0.00
MAPE-CCR	7.69			6.95		

Table 2 exhibits the results of predicting CCR. The graphs show the students' academic status over the past years and semesters. The table shows the list of students with their forecasted and actual CCR for the first and second semesters, as well as the PFE-CCR of each student. The table also presents the MAPE-CCR of the 13 students picked for testing.

3.4 Evaluation Summary

Figure 6 represents the MAPE in predicting the CGPA and the CCR of the students used in testing. The proponents test 13 students and come up with a result of 1.97% for the first semester and 7.14% for the second semester, and for the CCR, 7.69% for the first semester and 6.95% for the second semester.

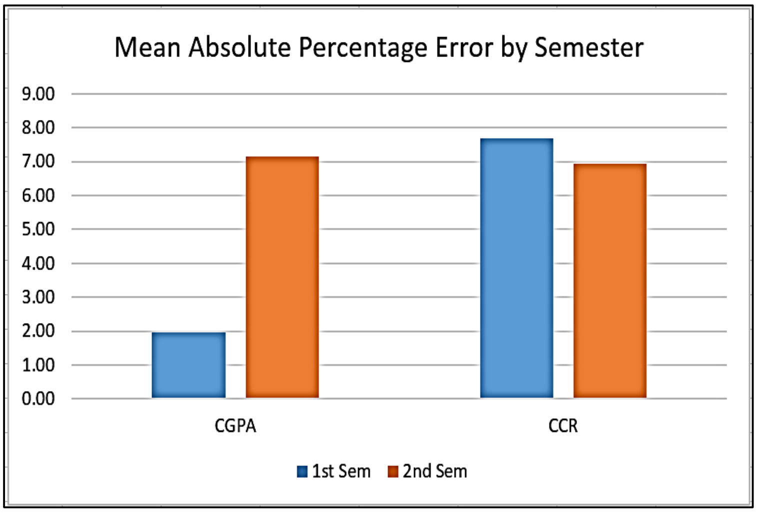

Figure 6. Mean Absolute Percentage Error

Figure 7 shows the overall MAPE for CGPA and CCR for both the first and second semesters. Results show 4.55% for the CGPA and 7.32% for the CCR.

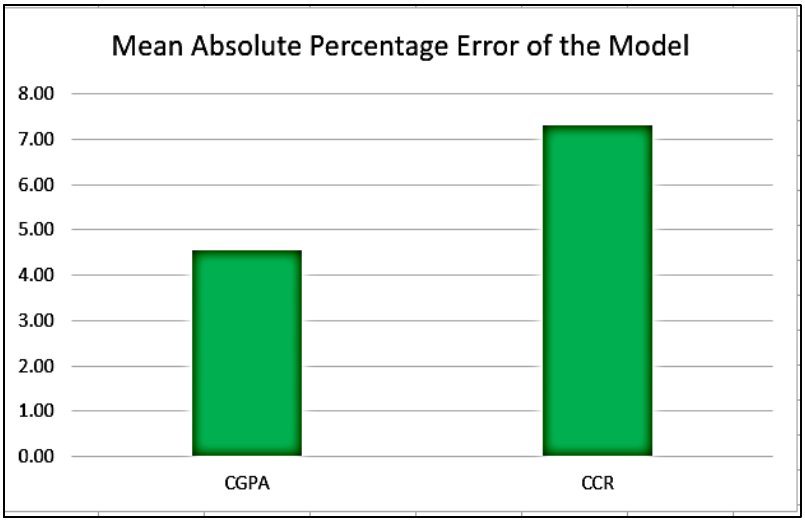

Figure 7. Mean Absolute Percentage Error of the Model

4. Conclusions and Recommendations

The following conclusions and suggestions were made based on the aforementioned results:

1. Monitoring students' academic performance is needed to promote education and achieve the learning outcome of every student. The designed and created interface for data extraction, management, and conversion is simple, has fewer buttons, and is easy to manipulate, excluding the fact that running time happens to be slower when extracting the data.
2. Another interface was designed and created. The data prediction analysis interface makes it easier to identify and label the numerical values and necessary data for each student needed for prediction. Yet, the interface data prediction of CGPA can be dragged anywhere with its value.
3. The proposed method for developing the system to predict the academic performance of CCS students is relatively accurate in terms of evaluating the students actual CGPA and forecasted CGPA for the first semester. The result of MAPE for the first semester coincides with the baseline, which is 5%. The system also considers the remarks of the students with incomplete (INC), failed, and dropped in the course, which is effective for determining the total percentage of pass or fail and the number of courses they passed for the first semester.

The 2nd semester average result exceeds the baseline at 5%, which means the forecasted CGPA and number of courses passed were not relatively accurate. The proposed system provides a smarter and more transparent way to predict student academic performance. It is also easy to manipulate, as shown in every figure of the GUI interface. The system is implemented using the Least Squares Regression/Linear Regression Toolbox in MATLAB.

On the basis of the basic findings made from the study and the conclusions drawn, the following are recommended:

1. The proponents recommend improving the systems' data extraction to avoid the delay in extracting the data from the CSV file. Otherwise, reduce the number of students on the list to enhance the performance and systems' running time.

2. With regards to the system's interface, in order to make it more presentable, it is recommended improving the system's GUI by adding icons and a homepage.
3. The proponents recommend coming up with other methods to prove the accuracy of predicting students' academic performance.
4. The proponents recommend that future researchers use or explore different prediction models or integrate other forecasting models. Some of the models that can be used are: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Educational Data Mining (EDUNET), and Deep Neural approaches for predicting the CGPA of the students. The proponents should not only focus on professional courses but should also consider the minor subjects; just like in other related studies, they used both major and minor subjects in predicting the CGPA. Future researchers can also use another dataset to feed the program and test its effectiveness.

References

[1] The TFA Editorial Team, "*Early Childhood Education Stories*", www.oese.ed.gov/files/2015/06/tfanarrative.pdf (Accessed November 2, 2022).

[2] S. Chand, K. Chaudhary, A. Prasad, and V. Chand, "*Perceived Causes of Students' Poor Performance in Mathematics: A Case Study at Ba and Tavua Secondary Schools*", Frontiers in Applied Mathematics and Statistics, vol. 7, April 2021, pp. 1-13, doi: 10.3389/fams.2021.614408.

[x] C. C. Kiu, "*Data Mining Analysis on Student's Academic Performance through Exploration of Student's Background and Social Activities*," 2018 Fourth International Conference on Advances in Computing, Communication & Automation (ICACCA), Subang Jaya, Malaysia, October 26-28, 2018, pp. 1-5, doi: 10.1109/ICACCAF.2018.8776809.

[y] Z. Zakaria, R. Abu Kassim, A. Mohamad, and N. Buniyamin, "*The impact of environment on engineering students' academic performance: A pilot study*," 2011 3rd International Congress on Engineering Education (ICEED), Kuala Lumpur, Malaysia, December 7-8, 2011, pp. 113-118, doi: 10.1109/ICEED.2011.6235371.

[5] Y. K. Mekonena and R. A. Fitiavana, "*Assessment of Learning Outcomes in Higher Education: Review of Literature*", International Journal of Research Publications, vol. 71, no. 1, 2021, pp. 69-76, doi: 10.47119/IJRP100711220211766.

[6] S. Ra, S. Kim, and K. J. Rhee, "*Developing National Student Assessment Systems for Quality Education: Lessons from the Republic of Korea*", Asian Development Bank, December 2019, doi: 10.22617/TCS190597-2.

[7] Timms 2015 Encyclopedia, "*Monitoring Student Progress in Mathematics and Science*", www.timssandpirls.bc.edu/timss2015/encyclopedia/countries/korea/monitoring-student-progress-in-mathematics-and-science/ (Accessed November 2, 2022).

[8] M. S. Jeong, "*New Mission for New Time in Korean Higher Education*", Journal of Ccomparative and International Higher Education, vol. 10, 2018, pp. 32-38, doi: 10.32674/jcihe.v10iWinter.687.

[9] K, Tremblay, D. Lalancette, and D. Roseveare, "*Assessment of Higher Education Learning Outcomes*", Feasibility Study Report, OECD 2012, vol. 1, www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume1.pdf (Accessed November 2, 2022).

[10] P. Black and D. William, "*Assessment and classroom learning*", Assessment in Education: Principles, Policy & Practice, vol. 5, no. 1, 1998, pp7-74, doi: 10.1080/0969595980050102.

[11] T. N. T. Yaakub, W. R. W. Ahmad, Y. Husaini, and N. Burham, "*Influence Factors in Academic Performance among Electronics Engineering Student: Geographic Background, Mathematics Grade and Psycographic*

Characteristics," 2018 IEEE 10th International Conference on Engineering Education (ICEED), Kuala Lumpur, Malaysia, November 8-9, 2018, pp. 30-33, doi: 10.1109/ICEED.2018.8626963.

[12] S. Verma and R. K. Yadav, "*Effect of Different Attributes on the Academic Performance of Engineering Students*," 2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI), Buldhana, India, December 30, 2020, pp. 1-4, doi: 10.1109/ICATMRI51801.2020.9398442.

[13] S. M. Ajibade, N. B. Ahmad, and S. M. Shamsuddin, "*A Heuristic Feature Selection Algorithm to Evaluate Academic Performance of Students*," 2019 IEEE 10th Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia, August 2-3, 2019, pp. 110-114, doi: 10.1109/ICSGRC.2019.8837067.

[14] J. Jamesmanoharan, S. H. Ganesh, M. L. P. Felciah, and A. K. Shafreenbanu, "*Discovering Students' Academic Performance Based on GPA Using K-Means Clustering Algorithm*," 2014 World Congress on Computing and Communication Technologies, Trichirappalli, India, February 27-March 1, 2014, pp. 200-202, doi: 10.1109/WCCCT.2014.75.

[15] M. R. Islam Rifat, A. Al Imran, and A. S. M. Badrudduza, "*EduNet: A Deep Neural Network Approach for Predicting CGPA of Undergraduate Students*," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), Dhaka, Bangladesh, May 3-5, 2019, pp. 1-6, doi: 10.1109/ICASERT.2019.8934616.

[16] X. Zhang and Q. Zhang, "*An Analysis of Students' Failing in University Based on Least Square Method and a New arctan – exp Logistic Regression Function*", Mathematical Problems in Engineering, vol. 2022, Article ID 6940855, February 2022, doi: 10.1155/2022/6940855.

[17] S. Sarwat, N. Ullah, S. Sadiq, R. Saleem, M. Umer, A. A. Eshmawi, A. Mohamed, and I. Ashraf, "*Predicting Students' Academic Performance with Conditional Generative Adversarial Network and Deep SVM*," Sensors (Basel), vol. 22, no. 13, June 2022, pp. 1-18, doi: 10.3390/s22134834.