Optimization of MLP-Regressor for Predicting Student’s Grade Point Average (GPA)

|  |  |
| --- | --- |
| Ihsan Fadli Tampati Politeknik Siber dan Sandi NegaraBogor, Indonesia  [ihsan.fadli@student.poltekssn.ac.id](mailto:herlambang.rafli@student.poltekssn.ac.id) | Hermawan Setiawan  Politeknik Siber dan Sandi NegaraBogor, Indonesia  [hermawan.setiawan@student.poltekssn.ac.id](mailto:herlambang.rafli@student.poltekssn.ac.id) |

*Abstract*— The digital age brings a lot of digital data that can be turned into useful information to improve student quality and achieve sustainable education goals. To achieve these goals, educational institutions need to identify factors influencing students' academic performance and predict students' exam scores. This is useful for providing extra attention to students predicted to have lower exam scores. This process can be carried out using Educational Data Mining (EDM) techniques, which allow for in-depth analysis of students' historical data. Previous studies have shown that factors such as learning methods, study environment, parental interaction, and parental pressure significantly affect students' academic performance. Identifying these factors affecting academic performance is done through a classification approach using machine learning algorithms. In this research, the approach is different. The main idea is to train the deep learning model using the history of grade point average (GPA) from Politeknik Siber dan Sandi Negara with a regression method so that it can predict students' GPA as accurately as possible, then implement it in a web application. If there are students whose GPA is predicted to decrease in the next semester, then the lecturer is obliged to pay more attention to these students. The Multi-layer Perceptron (MLP) regressor is chosen to train, test, and evaluate the model. The research findings indicate that the MLP regressor model is more optimal when configured for the number of neurons, hidden layer’s depth, and the use of Adam optimizer and constant learning rate types. The model has been evaluated using mean squared error (MSE), root-mean squared error (RMSE), mean-absolute error (MAE), coefficient of determination (R-squared), and Breusch-Pagan Test with the best results of 0.097511, 0.312267, 0.218024, 0.118555, and 0.151559 respectively. The evaluation results show that the model is capable of reducing the MSE, RMSE, and MAE value close to zero and improving the coefficient of determination and Breusch-Pagan *p*-value. However, the coefficient of determination and Breusch-Pagan *p*-value is still far from 1. In the future, research can be continued by expanding the dataset.

Keywords—educational data mining (edm), deep learning, grade point average (gpa), regression

# Introduction

The digital age brings a lot of digital data that can be turned into useful information. This information can be analyzed in various ways to reach specific goals. To achieve sustainable education goals, we use a method called Educational Data Mining (EDM). EDM uses statistical, machine learning, and data mining algorithms to analyze educational data, helping us understand more about how students learn and their surroundings [1], [2]. Various universities have started using EDM techniques to enhance student quality, institutional quality, and institutional standards. Using EDM to improve student quality can be done by analyzing students' academic history data [1].

Researchers have analyzed that additional information such as school teaching methods, a country's educational approach, study environment, questioning intensity, participation in discussions, parent-child relationships, and parental pressure can influence students' exam scores [1], [3], [4], [5], [6], [7], [8], [9]. Analyzing the factors affecting students' exam scores and predicting students' exam scores are classification and regression problems that can be solved using machine learning algorithms [1], [3], [4], [5], [6], [7], [8], [9].

The role of machine learning is crucial in EDM techniques [1]. The use of machine learning enables the prediction of future occurrences based on analyzed data. Machine learning algorithms commonly used for prediction in regression problems include linear regression, lasso regression, decision tree regression, random forest, k-nearest neighbors (K-NN), and multi-layer perceptron (MLP) regressor. Previous researchers have demonstrated that these algorithms can define factors related to the improvement or decline of students' exam scores [1], [3], [4], [5], [6], [7], [8], [9]. However, no researcher has proven that student exam scores, in this case grade point average (GPA), can be predicted solely based on previous GPA. Therefore, this study aims to demonstrate that deep learning algorithms can be used and optimized to predict GPA based on previous GPA. The deep learning algorithm used in this study is the Multi-Layer Perceptron regressor from the scikit-learn library. The MLP-Regressor algorithm was chosen because it has been proven effective in predicting data in regression problems and has been widely applied in various fields [10], [11], [12], [13], [14], [15]. The main idea of this research is to train the MLP regressor model using the grade point average (GPA) from Politeknik Siber dan Sandi Negara with a regression method so that it can predict students' GPA as accurately as possible, then implement it in a web application. If there are students whose GPA is predicted to decrease in the next semester, then the lecturer is obliged to pay more attention to these students. This research is expected to help universities improve the quality of students by paying attention to students who have predicted low GPA in the next semester and encourage universities to utilize educational data mining using digital technology.

# Related Works

Various studies conducted by previous researchers aim to understand and predict students' academic performance through different approaches. A study in Tunisia [3] utilized two-stage analysis, namely the directional distance function Approach (DDF) and machine learning algorithms regression trees and random forest to identify key factors influencing students' academic performance. The results showed that factors such as school size, competition, class size, parental pressure, and interaction with females had significant effects, while school location had no significant influence. Another study [4] proposed a comprehensive student academic data processing model called AugmentED using various data sources and features as well as machine learning and deep learning algorithms to predict academic performance with high accuracy. A comparison was made of machine learning algorithms to predict student performance based on multi-feature data, while deep learning was used to model student performance based on big data from the virtual learning environment (VLE). A study [5] attempted to identify student and school characteristics related to student exam scores using a tree-based machine learning approach. The research was conducted on students and schools from nine different countries: Australia, Canada, France, Germany, Italy, Japan, Spain, the UK, and the USA. The results of the study showed a correlation between student and school characteristics and test scores.

Research [6] aims to identify suitable machine learning algorithms for predicting students' academic grades and categorizing final student grades. Predicting academic grades is a regression problem that the authors attempted to solve by evaluating machine learning algorithms using the mean squared error (MSE) evaluation metric, while categorizing final student grades is a classification problem that the authors attempted to solve by evaluating machine learning algorithms based on classification accuracy. The algorithms used include Decision Tree, Genetic Algorithm (GA) based Decision Tree, K-Nearest Neighbour (K-NN), and Genetic Algorithm (GA) based K-NN. The research results showed that the GA-based Decision Tree algorithm is better at categorizing final student grades compared to the other three algorithms. The obtained accuracy is 96.64%, and the GA-based Decision Tree algorithm is also better at predicting students' academic grades compared to the other three algorithms. This is evident from the smaller MSE value compared to the other three algorithms, which is 5.34.

Research [9] seeks to demonstrate the influence of using online learning resources and study duration on student performance. Performance is defined as either great or weak labels obtained from features such as learning style, family education history, non-study related work, sports, study time, use of online learning resources, and others. The relationship between these features, particularly study duration and the use of online learning resources, with the labels will be modeled using four machine learning algorithms, namely artificial neural network (ANN), naive Bayes, logistic regression, and decision trees. The results show that ANN models better than other algorithms. The best model produced by ANN interprets that there is no significant influence between study duration and the use of online learning resources on student academic performance.

Research [7] argues that now is the era where big data in education can be easily accessed, supported by increasingly advanced learning platforms, discovering problems from data with these platforms, optimizing the education environment based on the problems found, and providing recommendations for education to run smoothly. Data processing in education has been widely carried out using Educational Data Mining (EDM) techniques using machine learning algorithms. However, the use of deep learning is still rare. Researchers attempt to investigate the effectiveness of Deep Learning in modeling student performance in the form of classification. The classification labels consist of 'withdrawn-pass', 'pass-fail', 'distinction-pass', and 'distinction-fail' using datasets from VLE. The research results show that the proposed deep learning model has better accuracy compared to the logistic regression and SVM machine learning algorithms. The proposed deep learning model achieves classification accuracies from 84% to 93%, while the classification accuracies of logistic regression and SVM algorithms range from 79.82% - 85.60% and 79.95% - 89.14%, respectively.

Researcher [8] argues that in an increasingly competitive world, institutions need to be able to predict the performance of their students. This is necessary so that institutions can provide more attention to students who indeed need it. Furthermore, the success of an institution is measured by the development of its students' performance. In the researcher's case, student performance is defined as a classification problem. To predict student performance, the author uses Educational Data Mining (EDM) techniques to analyze the relationships between features of the data and student performance using Naive Bayes, ID3, C4.5, and SVM machine learning algorithms. Subsequently, the prediction accuracy among the algorithms will be compared to obtain the best modeling. Research [10] proposes a dual-input deep learning model framework using a multi-layer perceptron (MLP) and long-short term memory (LSTM) architecture to predict grade point average (GPA) data on a 4.0 scale using a regression approach. The research results show that the model achieves the best performance compared to other models by obtaining values of 0.4142 for the mean squared error (MSE), 0.418 for the mean absolute error (MAE), and 0.4879 for R-squared.

From these various approaches, these studies contribute to a better understanding of the factors influencing students' academic performance, [2], [3], [4], [5], [6], [7], [9] the development of more accurate classification models [2], [3], [4], [5], [6], [7], [9], and the development of regression models [7]. Since the regression approach is still less used, this study will use the regression approach with deep learning.

# Implementation

The implementation is carried out using the MLP-Regressor from the scikit-learn library with a dataset taken from the grade point average (GPA) data of Politeknik Siber and Sandi Negara students over the last nine years. Hyper-parameter tuning process is conducted on the number of neurons, layer’s depth in hidden layer, optimizer, and learning rate. Subsequently, the models resulting from hyper-parameter tuning will be evaluated using mean squared error (MSE), root-mean squared error (RMSE), mean-absolute error (MAE), coefficient of determination (R-squared), and Breusch-Pagan test to ensure that the model is free from heteroscedasticity and normally distributed (referensi rmse dkk dan breuch pagan). Furthermore, the model with the best statistical evaluation results will be implemented on the website application using the Streamlit framework.

## Methodology

This research methodology utilizes the method from [16] as illustrated in the following Fig. 3.1.

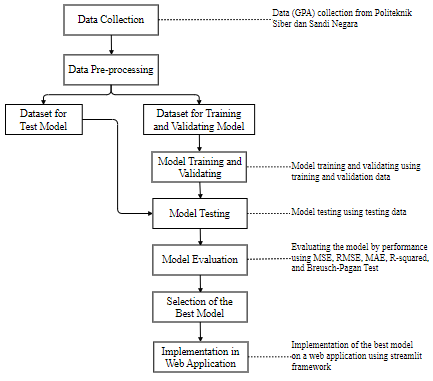


Figure III.1 Research Methodology

Based on Fig.3.1, the grade point average (GPA) data will be extracted from the GPA data of students at the Politeknik Siber dan Sandi Negara for the last nine years. The training, testing, and evaluation stages will involve hyper-parameter tuning processes for the number of neurons, layer depth within the hidden layer layer, various optimizers, and various learning rates. Then, the model will be evaluated using mean squared error (MSE), root-mean squared error (RMSE), mean-absolute error (MAE), coefficient of determination (R-squared), and Breusch-Pagan Test. The hyper-parameter tuning process is conducted to determine which parameters can be configured to reduce the MSE, RMSE, MAE and increase the coefficient of determination and Breusch-Pagan *p*-value value. Subsequently, the evaluation results between models will be compared to obtain the best model, which will be implemented on a website application using the Streamlit framework.

## Datasets

The following Fig.3.2 below provides an overview of the dataset used.

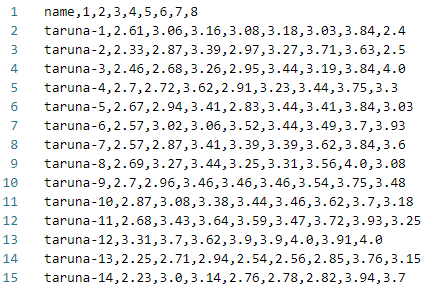


Figure III.2 GPA’s Dataset

## Neurons

Pada tahap ini, jumlah *neurons* akan dikonfigurasi dengan jumlah delapan hingga enam belas *neurons*. Tabel 3.1 dan Tabel 3.2 berikut merupakan hasil evaluasi model dengan MSE, RMSE, MAE dan koefisien determinasi (R2) pada fase validasi dan tes model.

Table III.1 Nilai MSE, RMSE, MAE, dan R2 pada Fase Validasi Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Neurons*** | **MSE** | **RMSE** | **MAE** | **R2** |
| 8 | 0.133032 | 0.364736 | 0.272875 | -0.202540 |
| 9 | 0.124096 | 0.352273 | 0.254574 | -0.121764 |
| 10 | 0.106264 | 0.325982 | 0.231283 | 0.039430 |
| 11 | 0.116069 | 0.340689 | 0.252536 | -0.049198 |
| 12 | 0.086268 | 0.293714 | 0.214810 | 0.220184 |
| 13 | 0.109435 | 0.330810 | 0.246849 | 0.010762 |
| 14 | 0.105754 | 0.325199 | 0.243254 | 0.044040 |
| 15 | 0.107648 | 0.328097 | 0.253219 | 0.026923 |
| 16 | 0.170648 | 0.413096 | 0.310840 | -0.542567 |

Table III.2 Nilai MSE, RMSE, MAE, dan R2 pada Fase Tes Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Neurons*** | **MSE** | **RMSE** | **MAE** | **R2** |
| 8 | 0.125173 | 0.353798 | 0.280785 | -0.168637 |
| 9 | 0.115388 | 0.339688 | 0.260061 | -0.077279 |
| 10 | 0.110864 | 0.332963 | 0.249052 | -0.035045 |
| 11 | 0.125576 | 0.354368 | 0.274813 | -0.172403 |
| 12 | 0.090367 | 0.300612 | 0.220819 | 0.156315 |
| 13 | 0.096716 | 0.310992 | 0.222536 | 0.097042 |
| 14 | 0.123103 | 0.350861 | 0.267037 | -0.149311 |
| 15 | 0.112331 | 0.335159 | 0.259050 | -0.048743 |
| 16 | 0.150107 | 0.387437 | 0.309738 | -0.401423 |

Berdasarkan hasil eksperimen pada Tabel 3.1 dan Tabel 3.2, ditemukan bahwa terdapat nilai negatif pada koefisien determinasi. Pada fase validasi, koefisien determinasi bernilai negatif ketika *neurons* dikonfigurasi pada jumlah delapan, sembilan, sebelas, dan enam belas *neurons*. Pada fase tes, koefisien determinasi bernilai negatif ketika *neurons* dikonfigurasi pada jumlah delapan hingga sebelas dan empat belas hingga enam belas *neurons*. Berdasarkan penelitian dari …, nilai negatif pada koefisien determinasi terjadi ketika nilai MSE dari fase tes lebih besar dibandingkan nilai MSE pada fase validasi sehingga dapat disimpulkan bahwa model pada fase tes memiliki performa prediksi yang buruk dibandingkan dengan performa prediksi pada fase validasi.

Selanjutnya, untuk model mendapatkan model regresi yang baik dalam memprediksi nilai GPA, model perlu ditinjau dari segi ketidaksamaan *variance residual* sehingga mampu mengidentifikasi model yang memiliki heteroskedastisitas menggunakan metode *Breusch-Pagan Test. Breusch-Pagan Test* dilakukan terhadap model dengan nilai koefisien determinasi yang tidak negatif. Tabel 3.3 dan Tabel 3.4 berikut merupakan hasil *Breushc-Pagan Test* pada fase validasi dan fase tes model.

Table III.3 Hasil Breusch-Pagan Test pada Fase Validasi Model

|  |  |  |
| --- | --- | --- |
| ***Neurons*** | **Breusch-Pagan Test Statistic** | **Breusch-Pagan  *p-*Value** |
| 8 | 7.533108 | 0.375555 |
| 9 | 6.981827 | 0.430774 |
| 10 | 7.864285 | 0.344713 |
| 11 | 6.687852 | 0.462087 |
| 12 | 6.834934 | 0.446266 |
| 13 | 5.398801 | 0.611416 |
| 14 | 9.296493 | 0.232064 |
| 15 | 5.342367 | 0.618259 |
| 16 | 17.215907 | 0.016055 |

Table III.4 Hasil Bresuch-Pagan Test pada Fase Test Model

|  |  |  |
| --- | --- | --- |
| ***Neurons*** | **Breusch-Pagan Test Statistic** | **Breusch-Pagan  *p-*Value** |
| 8 | 18.565359 | 0.009663 |
| 9 | 12.725760 | 0.079078 |
| 10 | 10.588772 | 0.157589 |
| 11 | 12.20278 | 0.094084 |
| 12 | 16.744093 | 0.019122 |
| 13 | 10.933424 | 0.141549 |
| 14 | 11.695947 | 0.111011 |
| 15 | 18.247843 | 0.010900 |
| 16 | 32.611364 | 0.003127 |

Heteroskedastisitas pada model diidentifikasi dari nilai Breusch-pagan test statistic yang tinggi dan nilai Breusch-pagan p-value yang lebih kecil dari 0.05. Oleh karena itu, dari hasil eksperimen pada Tabel 3.3 dan Tabel 3.4, dicari model dengan nilai Breusch-pagan test statistic yang rendah dan nilai Breusch-pagan p-value yang lebih besar dari 0.05 dan didapatkan bahwa model dengan konfigurasi tiga belas (13) *neurons* merupakan model terbaik. Pada tahap selanjutnya, konfigurasi tiga belas (13) neurons akan digunakan untuk menentukan konfigurasi terbaik dari parameter lainnya. Adapun fokus utama dari konfigurasi parameter lainnya yaitu untuk menurunkan nilai MSE, RMSE, MAE, dan Breusch-pagan test statistic serta menaikkan nilai koefisien determinasi (R2) dan Breusch-pagan p-value.

## Hidden Layer’s Depth

Pada tahap ini, kedalaman dari *hidden layer* akan dikonfigurasi dari satu hingga sembilan kedalaman. Tabel 3.5, Tabel 3.6, Tabel 3.7, dan Tabel 3.8 berikut merupakan hasil evaluasi model dengan MSE, RMSE, MAE dan koefisien determinasi (R2) pada fase validasi dan tes model.

Table III.5 Nilai MSE, RMSE, MAE, dan R2 pada Fase Validasi Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Hidden Layer’s Depth*** | **MSE** | **RMSE** | **MAE** | **R2** |
| 1 | 0.162070 | 0.402580 | 0.310083 | -0.465027 |
| 2 | 0.099629 | 0.315640 | 0.235849 | 0.099410 |
| 3 | 0.109435 | 0.330810 | 0.246849 | 0.010762 |
| 4 | 0.089254 | 0.298751 | 0.217952 | 0.193186 |
| 5 | 0.109662 | 0.331153 | 0.256150 | 0.008712 |
| 6 | 0.089938 | 0.299898 | 0.222695 | 0.187003 |
| 7 | 0.085841 | 0.292986 | 0.207943 | 0.224045 |
| 8 | 0.086710 | 0.294466 | 0.215999 | 0.216187 |
| 9 | 0.097511 | 0.312267 | 0.218024 | 0.118555 |

Table III.6 Nilai MSE, RMSE, MAE, dan R2 pada Fase Tes Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hidden Layer’s Depth** | **MSE** | **RMSE** | **MAE** | **R2** |
| 1 | 0.369468 | 0.294961 | -0.274449 | 0.136507 |
| 2 | 0.323244 | 0.242227 | 0.024492 | 0.104487 |
| 3 | 0.310992 | 0.222536 | 0.097042 | 0.096716 |
| 4 | 0.304638 | 0.217673 | 0.133561 | 0.092804 |
| 5 | 0.317712 | 0.240772 | 0.057596 | 0.100941 |
| 6 | 0.300298 | 0.216751 | 0.158075 | 0.090179 |
| 7 | 0.296095 | 0.202801 | 0.181477 | 0.087672 |
| 8 | 0.300865 | 0.213214 | 0.154891 | 0.090520 |
| 9 | 0.369468 | 0.294961 | -0.274449 | 0.136507 |

Berdasarkan hasil eksperimen pada Tabel 3.5 dan Tabel 3.6, terlihat bahwa konfigurasi kedalaman sangat berpengaruh dalam menaikkan atau menurunkan nilai-nilai MSE, RMSE, MAE, dan koefisien determninasi (R2). Terlihat bahwa semakin dangkal kedalaman *hidden layer,* model semakin buruk dalam mempelajari hubungan dari variasi variabel terikat yang ditandai dengan menurunnya nilai koefisien determinasi (R2). Selanjutnya, untuk mendapatkan konfigurasi kedalaman *hidden lauer* yang baik, model perlu ditinjau lagi menggunakan *Breusch-Pagan Test*. Tabel 3.7 dan 3.8 berikut merupakan hasil dari *Breusch-Pagan Test.*

Table III.7 Hasil Breusch-Pagan Test pada Fase Validasi Model

|  |  |  |
| --- | --- | --- |
| **Hidden Layer’s Depth** | **Breusch-Pagan Test Statistic** | **Breusch-Pagan  *p-*Value** |
| 1 | 6.487691 | 0.484091 |
| 2 | 6.766556 | 0.453584 |
| 3 | 5.398801 | 0.611416 |
| 4 | 5.422261 | 0.608576 |
| 5 | 6.032790 | 0.535925 |
| 6 | 6.348390 | 0.499708 |
| 7 | 6.850935 | 0.444563 |
| 8 | 6.024954 | 0.536838 |
| 9 | 10.714637 | 0.151559 |

Table III.8 Hasil Breusch-Pagan Test pada Fase Tes Model

|  |  |  |
| --- | --- | --- |
| **Hidden Layer’s Depth** | **Breusch-Pagan Test Statistic** | **Breusch-Pagan  *p-*Value** |
| 1 | 13.228622 | 0.066730 |
| 2 | 17.442600 | 0.014754 |
| 3 | 10.933424 | 0.141549 |
| 4 | 15.469837 | 0.030426 |
| 5 | 19.776877 | 0.006072 |
| 6 | 15.140707 | 0.0342374 |
| 7 | 12.002047 | 0.100491 |
| 8 | 13.610676 | 0.058555 |
| 9 | 11.208736 | 0.129769 |

Berdasarkan hasil Breusch-Pagan test pada tabel 3.7 dan tabel 3.8, terlihat bahwa konfigurasi tujuh kedalaman *hidden layer* merupakan konfigurasi terbaik. Selanjutnya, model dengan konfigurasi tiga belas (13) neurons dan tujuh (7) kedalaman lapisan pada *hidden layer* akan digunakan untuk menentukan model dengan *optimizers* dan *learning rate* terbaik.

## Optimizers

Pada tahap ini, jenis *optimizers* akan dikonfigurasi dengan adam, stochastic-gradient descent (sgd), dan quasi-newton (lbfgs). Tabel 3.9 dan Tabel 4.0 berikut merupakan hasil evaluasi model dengan MSE, RMSE, MAE dan koefisien determinasi (R2) pada fase validasi dan tes model.

Table III.9 Nilai MSE, RMSE, MAE, dan R2 pada Fase Validasi Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Optimizers*** | **MSE** | **RMSE** | **MAE** | **R2** |
| adam | 0.085841 | 0.292986 | 0.207943 | 0.224045 |
| sgd | 0.111789 | 0.334349 | 0.252168 | -0.010515 |
| lbfgs | 0.094406 | 0.307255 | 0.218354 | 0.146623 |

Table III.10 Nilai MSE, RMSE, MAE, dan R2 pada Fase Tes Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Optimizers*** | **MSE** | **RMSE** | **MAE** | **R2** |
| adam | 0.097511 | 0.312267 | 0.218024 | 0.118555 |
| sgd | 0.110704 | 0.332722 | 0.250531 | -0.033552 |
| lbfgs | 0.103840 | 0.322242 | 0.227131 | 0.030534 |

Berdasarkan hasil eksperimen pada tabel 3.9 dan 3.10, terlihat bahwa konfigurasi kedalaman sangat berpengaruh dalam menaikkan atau menurunkan nilai-nilai MSE, RMSE, MAE, dan koefisien determninasi (R2). Terlihat bahwa jika model dikonfigurasi dengan optimizer sgd, model memiliki nilai negatif pada koefisien determinasi (R2) sehingga dapat disimpulkan bahwa konfigurasi optimizer sgd membuat model kesulitan dalam mempelajari hubungan dari variasi variabel terikat. Selanjutnya, untuk mendapatkan konfigurasi optimizersterbaik, model perlu ditinjau lagi menggunakan *Breusch-Pagan Test*. Tabel 3.11 dan 3.12 berikut merupakan hasil dari *Breusch-Pagan Test.*

Table III.11 Hasil Breusch-Pagan Test pada Fase Validasi Model

|  |  |  |
| --- | --- | --- |
| **Optimizers** | **Breusch-Pagan Test Statistic** | **Breusch-Pagan  *p-*Value** |
| adam | 6.850935 | 0.444563 |
| sgd | 11.328200 | 0.124929 |
| lbfgs | 6.927729 | 0.436443 |

Table III.12 Hasil Breusch-Pagan Test pada Fase Tes Model

|  |  |  |
| --- | --- | --- |
| **Optimizers** | **Breusch-Pagan Test Statistic** | **Breusch-Pagan  *p-*Value** |
| adam | 10.714637 | 0.151559 |
| sgd | 9.507533 | 0.218240 |
| lbfgs | 22.565600 | 0.002028 |

Berdasarkan hasil eksperimen pada Tabel 3.11 dan Tabel 3.12, meskipun optimizer sgd memiliki nilai Breusch-Pagan test statistic dan Breusch-Pagan *p*-value yang paling baik, optimizer sgd bukan konfigurasi terbaik. Hal ini dikarenakan pada Tabel 3.9 dan Tabel 3.10 konfigurasi sgd memiliki nilai negatif pada koefisien determinasi (R2). Oleh karena itu, Optimizer adam merupakan optimizer terbaik dalam kasus ini. Oleh karena itu, pada eksperimen selanjutnya, model dengan konfigurasi tiga belas (13) neurons, tujuh (7) kedalaman lapisan pada *hidden layer*, dan optimizer adam akan digunakan untuk menentukan konfigurasi *learning rate* terbaik.

## Learning Rate

Pada tahap ini, jenis *learning rate* akan dikonfigurasi dengan constant, invscaling, dan adaptive. Tabel 3.13 dan Tabel 3.14 berikut merupakan hasil evaluasi model dengan MSE, RMSE, MAE dan koefisien determinasi (R2) pada fase validasi dan tes model.

Table III.13 Nilai MSE, RMSE, MAE, dan R2 pada Fase Validasi Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Learning***  ***Rate*** | **MSE** | **RMSE** | **MAE** | **R2** |
| constant | 0.085841 | 0.292986 | 0.207943 | 0.224045 |
| invscaling | 0.085841 | 0.292986 | 0.207943 | 0.224045 |
| adaptive | 0.085841 | 0.292986 | 0.207943 | 0.224045 |

Table III.14 Nilai MSE, RMSE, MAE, dan R2 pada Fase Tes Model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Learning***  ***Rate*** | **MSE** | **RMSE** | **MAE** | **R2** |
| constant | 0.085841 | 0.292986 | 0.207943 | 0.224045 |
| invscaling | 0.085841 | 0.292986 | 0.207943 | 0.224045 |
| adaptive | 0.085841 | 0.292986 | 0.207943 | 0.224045 |

Berdasarkan hasil eksperimen pada Tabel 3.13 dan Tabel 3.14, dapat disimpulkan bahwa konfigurasi *learning rate* tidak berpengaruh secara signifikan dalam menaikkan atau menurunkan nilai-nilai MSE, RMSE, MAE, dan koefisien determninasi (R2). Selanjutnya, untuk mendapatkan konfigurasi learning rateterbaik, model perlu ditinjau lagi menggunakan *Breusch-Pagan Test*. Tabel 3.15 dan 3.16 berikut merupakan hasil dari *Breusch-Pagan Test*.

Table III.15 Hasil Breusch-Pagan Test pada Fase Validasi Model

|  |  |  |
| --- | --- | --- |
| ***Learning Rate*** | **Breusch-Pagan Test Statistic** | **Breusch-Pagan  *p-*Value** |
| constant | 6.850935 | 0.444563 |
| invscaling | 6.850935 | 0.444563 |
| adaptive | 6.850935 | 0.444563 |

Table III.16 Hasil Breusch-Pagan Test pada Fase Tes Model

|  |  |  |
| --- | --- | --- |
| ***Learning Rate*** | **Breusch-Pagan Test Statistic** | **Breusch-Pagan  *p-*Value** |
| constant | 10.714637 | 0.151559 |
| invscaling | 12.002047 | 0.100491 |
| adaptive | 12.002047 | 0.100491 |

Berdasarkan Tabel 3.13, Tabel 3.14, Tabel 3.15, dan Tabel 3.16, diperoleh bahwa konfigurasi *learning rate* terbaik yaitu *learning rate* constant. Selanjutnya, perlu untuk memastikan bahwa model dengan konfigurasi tiga belas (13) neurons, tujuh (7) kedalaman lapisan pada *hidden layer*, optimizer adam, dan *learning rate* constant mampu mempelajari hubungan antara variabel bebas dengan variabel terikat dari *dataset* sehingga menghasilkan nilai-nilai yang terdistribusi normal dan secara visual terbebas dari heteroskedastisitas. Oleh karena itu, pada tahap selanjutnya, visualisasi hasil validasi dan tes model akan dilakukan menggunakan Q-Q plot, density plot, dan residual plot.

## Visualisasi Residual Plot, Q-Q Plot, dan Density Plot

Gambar 3.3, Gambar 3.4, dan Gambar 3.5 berikut merupakan visuliasi dari residual plot, Q-Q plot, dan density plot pada fase validasi dan tes model.

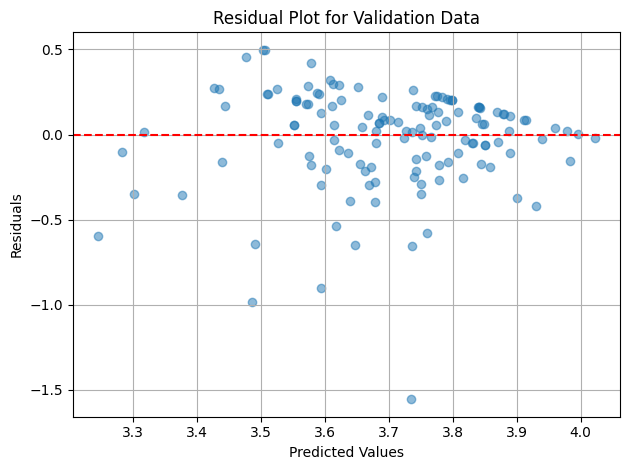
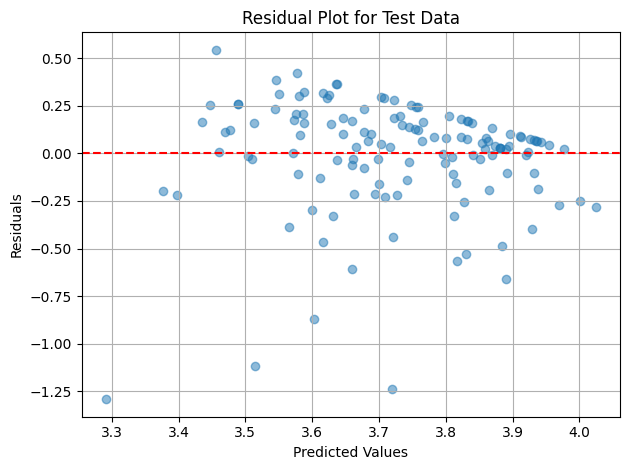
 

Figure III.3 Perbandingan Residual Plot pada Fase Validasi dengan Fase Tes Model

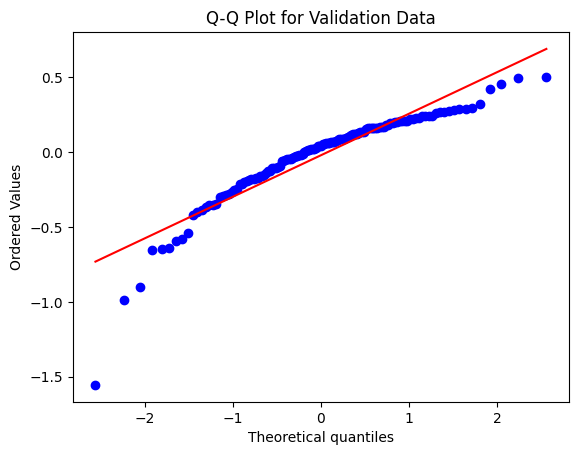
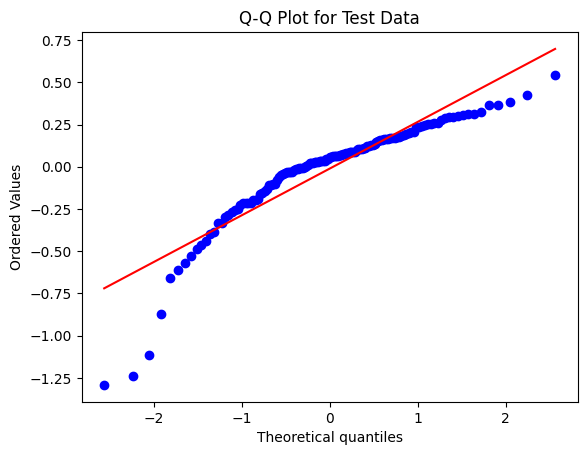
 

Figure III.4 Perbandingan Q-Q Plot pada Fase Validasi dengan Fase Tes Model

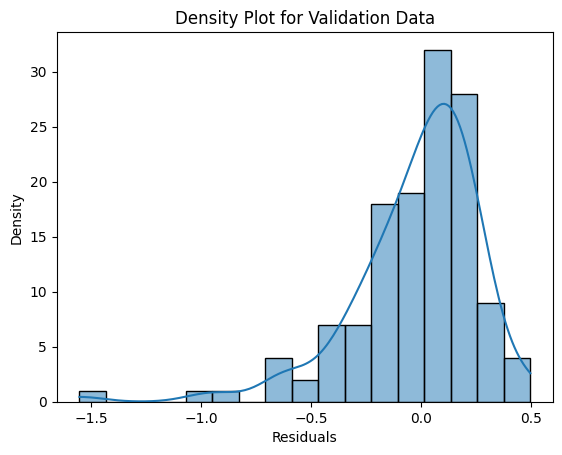
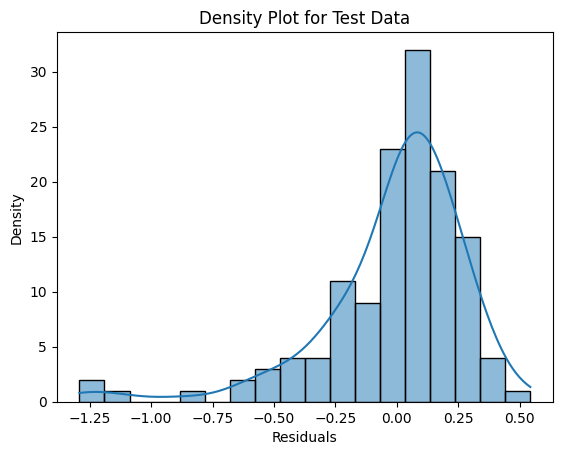
 

Figure III.5 Perbandingan Density Plot pada Fase Validasi dengan Fase Tes Model

Berdasarkan visualisasi pada Gambar 3.3, Gambar 3.4, dan Gambar 3.5, terlihat bahwa model pada fase validasi dan tes terbebas dari heteroskedastisitas, terdistribusi normal, dan memiliki sifat yang sama. Berdasarkan fakta tersebut, dapat disimpulkan bahwa model mampu menggeneralisasi nilai GPA baru dengan baik meskipun nilai koefisien determinasi (R2) dan Breusch-Pagan *p-*value model masih jauh dari nilai 1.

## Implementation in Web Application

Here is the web application interface for GPA prediction using the MLP regressor model.

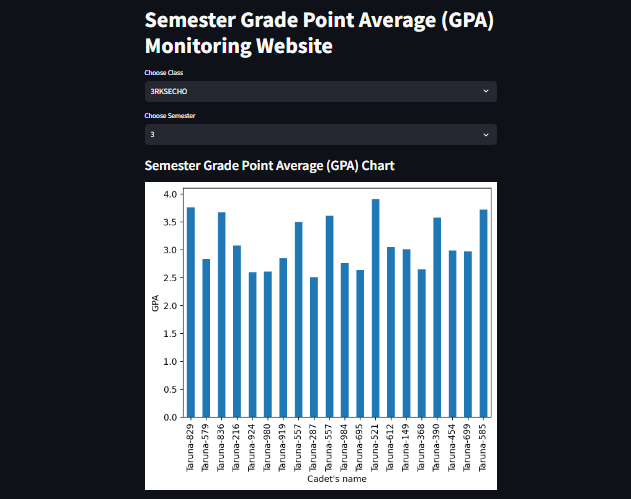


Figure III.6 Tampilan Daftar GPA

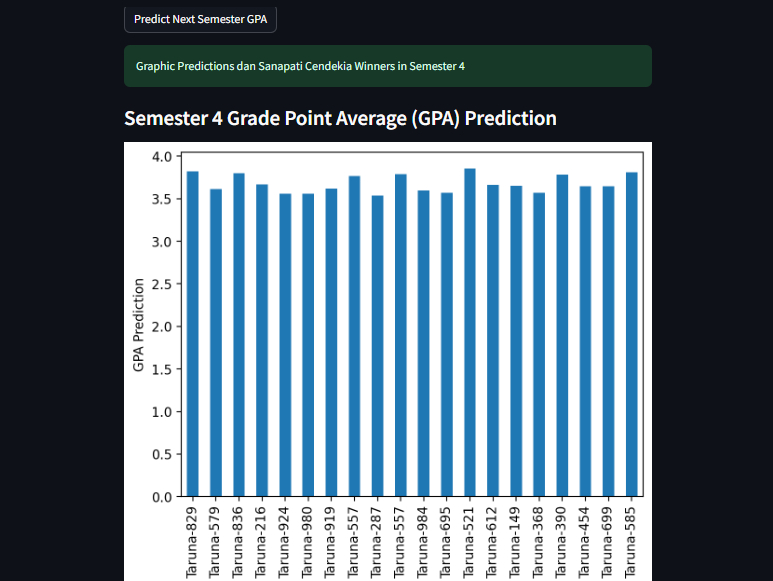


Figure III.7 Tampilan Prediksi GPA

Based on Fig.6 and Fig.7, users can select the student level and semester they want to predict. Then, the system will predict the chosen semester's GPA. If there are predicted GPA values for students that decrease in the next semester, the website will display a separate list of those students.

# Conclusion

Berdasarkan hasil eksperimen, dapat disimpulkan bahwa konfigurasi parameter learning rate justru tidak memengaruhi performa model MLP-Regressor secara signifikan. Berbeda dengan learning rate, konfigurasi parameter jumlah neurons, kedalaman layer pada hidden layer, dan jenis optimizer berpengaruh pada peningkatan atau penurunan performa regresi model. Meskipun secara statistic model memiliki performa yang baik, namun nilai koefisien determinasi (R-squared) dan Breusch-Pagan p-value masih jauh dari nilai 1. Oleh karena itu, penelitian selanjutnya dapat dikembangkan dengan memperbanyak data pada dataset guna meningkatkan nilai koefisien determinasi dan Breusch-Pagan p-value hingga mendekati nilai 1 agar model MLP-Regressor yang dihasilkan mampu secara akurat memprediksi GPA mahasiswa dan membantu institusi pendidikan memberikan perhatian lebih kepada mahasiswa yang diprediksi akan mengalami penurunan GPA.

##### References

[1] H. Pallathadka, A. Wenda, E. Ramirez-Asís, M. Asís-López, J. Flores-Albornoz, and K. Phasinam, “Classification and prediction of student performance data using various machine learning algorithms,” *Mater. Today Proc.*, vol. 80, pp. 3782–3785, 2023, doi: https://doi.org/10.1016/j.matpr.2021.07.382.

[2] C. Romero and S. Ventura, “Educational Data Mining: A Review of the State of the Art,” *IEEE Trans. Syst. Man, Cybern. Part C (Applications Rev.*, vol. 40, no. 6, pp. 601–618, 2010, doi: 10.1109/TSMCC.2010.2053532.

[3] S. Rebai, F. Ben Yahia, and H. Essid, “A graphically based machine learning approach to predict secondary schools performance in Tunisia,” *Socioecon. Plann. Sci.*, vol. 70, p. 100724, 2020, doi: https://doi.org/10.1016/j.seps.2019.06.009.

[4] L. Zhao *et al.*, “Academic Performance Prediction Based on Multisource, Multifeature Behavioral Data,” *IEEE Access*, vol. 9, pp. 5453–5465, 2021, doi: 10.1109/ACCESS.2020.3002791.

[5] C. Masci, G. Johnes, and T. Agasisti, “Student and school performance across countries: A machine learning approach,” *Eur. J. Oper. Res.*, vol. 269, no. 3, pp. 1072–1085, 2018, doi: https://doi.org/10.1016/j.ejor.2018.02.031.

[6] S. Hussain and M. Q. Khan, “Student-Performulator: Predicting Students’ Academic Performance at Secondary and Intermediate Level Using Machine Learning,” *Ann. Data Sci.*, vol. 10, no. 3, pp. 637–655, 2023, doi: 10.1007/s40745-021-00341-0.

[7] H. Waheed, S.-U. Hassan, N. R. Aljohani, J. Hardman, S. Alelyani, and R. Nawaz, “Predicting academic performance of students from VLE big data using deep learning models,” *Comput. Human Behav.*, vol. 104, p. 106189, 2020, doi: https://doi.org/10.1016/j.chb.2019.106189.

[8] H. Prabowo, A. A. Hidayat, T. W. Cenggoro, R. Rahutomo, K. Purwandari, and B. Pardamean, “Aggregating Time Series and Tabular Data in Deep Learning Model for University Students’ GPA Prediction,” *IEEE Access*, vol. 9, pp. 87370–87377, 2021, doi: 10.1109/ACCESS.2021.3088152.

[9] “Predicting Students’ Performance Using Machine Learning Techniques,” *J. Univ. BABYLON Pure Appl. Sci.*, vol. 27, no. 1 SE-Articles, pp. 194–205, Apr. 2019, doi: 10.29196/jubpas.v27i1.2108.

[10] M. I. Dutt and W. Saadeh, “A Multilayer Perceptron (MLP) Regressor Network for Monitoring the Depth of Anesthesia,” in *2022 20th IEEE Interregional NEWCAS Conference (NEWCAS)*, 2022, pp. 251–255. doi: 10.1109/NEWCAS52662.2022.9842242.

[11] H. Shimanaka, T. Kajiwara, and M. Komachi, “{RUSE}: Regressor Using Sentence Embeddings for Automatic Machine Translation Evaluation,” in *Proceedings of the Third Conference on Machine Translation: Shared Task Papers*, O. Bojar, R. Chatterjee, C. Federmann, M. Fishel, Y. Graham, B. Haddow, M. Huck, A. J. Yepes, P. Koehn, C. Monz, M. Negri, A. Névéol, M. Neves, M. Post, L. Specia, M. Turchi, and K. Verspoor, Eds., Belgium, Brussels: Association for Computational Linguistics, 2018, pp. 751–758. doi: 10.18653/v1/W18-6456.

[12] X. Xu, S. L. Keoh, C. K. Seow, Q. Cao, and S. K. B. A. Rahim, “Towards Prediction of Bus Arrival Time using Multi-layer Perceptron (MLP) and MLP Regressor,” in *2023 8th International Conference on Business and Industrial Research (ICBIR)*, 2023, pp. 669–674. doi: 10.1109/ICBIR57571.2023.10147614.

[13] E. Agirre-Basurko, G. Ibarra-Berastegi, and I. Madariaga, “Regression and multilayer perceptron-based models to forecast hourly O3 and NO2 levels in the Bilbao area,” *Environ. Model. Softw.*, vol. 21, no. 4, pp. 430–446, 2006, doi: https://doi.org/10.1016/j.envsoft.2004.07.008.

[14] X. Feng, G. Ma, S.-F. Su, C. Huang, M. K. Boswell, and P. Xue, “A multi-layer perceptron approach for accelerated wave forecasting in Lake Michigan,” *Ocean Eng.*, vol. 211, p. 107526, 2020, doi: https://doi.org/10.1016/j.oceaneng.2020.107526.

[15] Doreswamy, H. K S, Y. KM, and I. Gad, “Forecasting Air Pollution Particulate Matter (PM2.5) Using Machine Learning Regression Models,” *Procedia Comput. Sci.*, vol. 171, pp. 2057–2066, 2020, doi: https://doi.org/10.1016/j.procs.2020.04.221.

[16] J. Kamiri and G. Mariga, “Research Methods in Machine Learning: A Content Analysis,” *Int. J. Comput. Inf. Technol.*, vol. 10, no. 2, 2021, doi: 10.24203/ijcit.v10i2.79.