Reducing Overfitting Problem by Hyperparameter Optimization of Different Regularization Techniques using Linear Regression

Nawaraj Paudel[1*] and Manoj Saru[2]

[1]Central Department of Computer Science and Information Technology, Tribhuvan University

[2]Rastriya Banijye Bank Limited, Singhadarbar Plaza

**Correspondence: nawarajpaudel@cdcsit.edu.np; Ph. 9841305619*

Keywords

Overfitting, Linear regression, L1 regularization, L2 regularization, L1_L2 regularization.

Abstract

In machine learning, regularization refers to the techniques for calibrating learning models that prevent overfitting and lower the adjusted loss function. Regularization is useful for attaining the perfect fit between overfitting and underfitting. This research focuses on identifying the problem of overfitting on linear regression model and also aims to solve that problem by implementing different regularization technique like Lasso (L1 regularization), Ridge (L2 regularization) and Elastic-Net (L1_L2 regularization) which are applicable on regression model. Optimization of the hyperparameter is also carried out for all the regularization techniques in order to find the ideal value of the hyperparameter to minimize the MSE and to enhance the model performance and efficiency. The dataset used for this research is the Nepali Housing Price Dataset which has been downloaded from Kaggle repository. MSE is used as performance metrics for to analyze the performance of regularization techniques. From the experimental results, it is observed that the Elastic-Net regression outperformed both Lasso and Ridge regression with the minimum MSE of 3.5541 at lambda (λ) value of 0.9. Also, the MSE provided by Lasso and Ridge at same lambda (λ) value of 0.9 are 12.6478 and 4.6688. Thus, this research concluded that the Elastic-Net regression provides the minimum MSE for the dataset and also help to overcome the problem of overfitting by enhancing the overall performance and efficiency of the model.

Received: 12 July 2025
Revised: 2 November 2025
Accepted: 3 December 2025

ISSN: 3059 - 9687

Introduction

n machine learning, underfitting or overfitting of the data are the two main reasons for poor performance of a learning model. A model is said to be overfitted if it accurately analyzes the training data, capturing both noise and details, but affects the model's performance on fresh data. On other hand, underfitting take place when a model is unable to accurately model the training set nor generalize to new data. Since underfitting is simple to identify using fundamental performance metrics, it is rarely discussed (Brownlee, 2016).

Regularization are the methods for calibrating the learning models in machine learning which helps to avoid overfitting and decrease the adjusted loss function. Regularization is helpful to achieve the ideal fit between overfitting and underfitting. Regularization minimizes overfitting errors and helps in properly fitting the machine learning model on the provided test set (Simplilearn, 2025).

By including an additional penalty term in the error function, regularization helps to aid in fine-tuning the learning models. The extra term helps to lower the complexity of the learning model and the processing costs by controlling the overly fluctuating function so that the coefficients don't take extreme values (Trotta, 2022).

Regression in machine learning refers to mathematical techniques that allow data scientists to predict a continuous result (y) based on the values of one or more predictor variables (x). Plotting a best-fit line or curve between the data is the ultimate objective of the regression procedure. Linear regression is likely the most widely used type of regression analysis because it’s so simple to use for forecasting and prediction. Variance, bias, and error are the three primary metrics used to evaluate the regression model. A high variance results in overfitting, whereas a high bias results in underfitting. Regularization often prevents overfitting by including a penalty term in the loss function (Kurama et al., 2024).

In machine learning, Regularization is a method for reducing the overfitting problem of the coefficient in the model. Alternatively said, regularization is any adjustment we make to a learning algorithm with the goal of lowering the generalization error but not the training error. Overfitting of the model will result in decreased accuracy because the model is attempting to capture every bit of noise present in the dataset. Here noise is defined as all data points that are not accurate representations of the data but rather random chance. So, learning these kinds of data can therefore results in overfitting, which is why on the training set, the model performs well and on the test set, it performs poorly. Therefore, in order to prevent the overfitting problem, regularization is utilized to enhance the condition that prevents the learning of a more complex model (Dave & Shah, 2019).

When there are a lot of features, we use ridge, lasso and elastic-net regression to prevent overfitting. These regularization strategies are applied for regression analysis. They minimize the error between the expected and actual observations by penalizing the magnitude of the coefficients of features. Linear regression is probably one of the most popular and well-understood machine learning methods used for making prediction. Using a best-fit straight line, linear regression determines the linear connection between the dependent variable and one or more independent variables. A linear regression model typically uses the weighted sum of the input features to provide a prediction (Kurama et al., 2024).

L1 regularization is also known as Lasso regression. The overfitted models are modified by introducing a penalty equivalent to the total of the absolute values of the coefficient. L1 regularization technique is best suited for the datasets where all the attributes aren’t necessarily needed for the outcome of the final result. L1 regularization encourages weights or coefficients to be 0 if possible. L1 regularization is also worked as feature selection. Below is the cost function formula with the regularization term included. L2 regularization is also known as Ridge regression. The overfitted models are modified by introducing a penalty equivalent to the sum of the square of the magnitude of coefficients. L2 regularization technique is best suited for the datasets where there exists multi-collinearity among the attributes. Since, L2 regularization leads the weights to decay towards zero (but not exactly zero), it is often referred to as weight decay. Elastic Net regularization blends elements of both Ridge and Lasso regularization by integrating their penalty terms into the loss function. This approach aims to harness the strengths of both techniques, offering a comprehensive solution that addresses issues like multicollinearity and feature selection (Simplilearn, 2025).

Currently, there are two hyperparameters requiring adjustment: λ1 and λ2. These parameters can be optimized through cross-validation to achieve optimal outcomes. When λ1= 0, Lasso regularization takes effect, while when λ2=0, Ridge regularization is applied. Elastic Net regularization proves especially beneficial when dealing with a substantial number of attributes (DataCamp, 2019).

Literature Review

Lifferent research works had been done in the past in order to improve the efficiency of the different machine learning models and to reduce the problem of overfitting by implementing different regularization techniques. The authors of (Sagala & Cendriawan, 2022) used a linear regression model for the purpose of house price prediction on Maribisnis Housing Dataset from the Maribelajar company. They use a six-step process methodology including Business understanding, Data understanding, Data pre-processing, Data cleaning, Data Standardization and Modelling. After then, Power Business Intelligence was used to visualize the data in order to properly analyze business performance. Using the data from the experiments, the model was able to provide RMSE = 0.0334, MAE = 0.023, RSE = 0.3, RAE 0.52 and R Coefficient = 0.7. Housing prices can be somewhat predicted and evaluated by the linear regression model, according to the data analysis and testing in this study.

The authors of (Dave & Shah, 2019) compared different regularization techniques such as L1, L2, Dropout and Elastic-net to classify the class of heartbeat using ANN architecture. The dataset used was ECG heartbeat categorization dataset. The authors tested all the regularization techniques using ANN architecture. The result showed that L2 regularization provided the maximum accuracy of 0.98% where all L1, Dropout and Elastic-net regularization provided the accuracy of 0.97%.

The authors of (Kolluri et al., 2020) proposed a novel L1/4 regularization method to overcome the overfitting problem, the experimental results have shown that this novel L1/4 regularization method considerably outperforms the L1 regularization in solving the overfitting problem in the gene data analysis of flowers. This article shows L1/4 regularization is computationally more effective which may choose fewer vectorized variables and shows that the method performs well while analyzing gene data. CNN model is used on the Flavia dataset to classify the flowers containing 1200 sample images. The prediction error by L1/4 regression is 0.47% where the prediction error by Lasso regression is 0.48%.

The various data augmentation methods were employed to expand the training dataset to pre-train the specified neural network in order to increase the effectiveness of the training procedure and to reduce the problem of overfitting (Mikołajczyk & Grochowski, 2018). The authors have used different data augmentation techniques like classical image transformation (rotating, cropping, zooming, histogram-based methods), Generative Adversarial Networks, Style Transfer, Texture Transfer. Three medical datasets are used to validate the proposed technique (skin melanomas diagnosis, histopathological images and breast magnetic resonance imaging scans analysis) using a VGG16 Convolutional Neural Network for image classification in order to provide a diagnosis.

Authors in (Ghodasara, 2021) provided introduction to the Pruning of Decision trees which serves to find a sparse subnetwork with lesser parameters in a dense network that provides same overall accuracy. Different approaches like bottom-up pruning and top-down pruning and types of pruning like alpha-beta pruning had been used to help in reducing the space requirements and computational cost in operating the network. The author also mentioned the Lottery Ticket Hypothesis in order to support the advantages of pruning.

Several other researches also have been done for the improvement of the performance of machine learning models and for reducing the overfitting problems. Most of the research suggest a wide range of regularization techniques depending upon the machine learning model which can be used for the improved performance of the learning model and for reducing the overfitting problems of machine learning models. Most of the research suggest the most common L1, L2 and Elastic-net regularization for the improvement of the linear regression model and for reducing the problem of overfitting.

Methodology

A housing dataset was downloaded from Kaggle repository and all the required pre-processing of data was done. Linear regression which is one the widely used and widely known machine learning model for prediction was implemented on the pre-processed housing dataset. After evaluating the end result i.e. MSE between the training and test results, the problem of overfitting of the model was identified. Once, the overfitting problem was identified, regularization techniques like L1 regularization, L2 regularization and L1_L2 regularization which are mostly used and known to calibrate of regression model were applied in order to prevent overfitting and reduce the loss function. Also, the MSE resulted from L1 regularization, L2 regularization and L1_L2 regularization with varying values of hyperparameter were compared to get the optimal result. The regularization techniques not only reduces the problem of overfitting but also helps to reduce the generalization error.

Dataset Description

Housing price dataset (Thapa, 2020), has been used to ensure the reliability of this research work. The housing price dataset used in this research has been downloaded from Kaggle repository which is Nepali Housing Price Dataset. This dataset consists larger number of attributes and instances which is in CSV format providing more reliability to this.

Table 1. Dataset Description

Dataset Name	Dataset Size	Number of Instances	Number of Attributes	Attributes Type
Nepali Housing Price Dataset	568 kB	2211	18	Price, Distance, Rooms, Address, City, Bathroom etc.

Data Preprocessing

Preprocessing raw data which is downloaded from Kaggle as per the need and necessity is very crucial step to get the desired outcome. It helps to improve the performance and accuracy of research work by reducing noise, irrelevant information, handling missing values and also by correcting the errors in the dataset. Also, the quality and efficiency of the data input determine the overall outcome and reliability of the research work. The different steps taken during the preprocessing of the dataset are as given below.

- **Reducing Dimensionality:** Removing the unnecessary and irrelevant attributes and selecting only the most important features to improve the model performance and the end results.
- **Data Transformation and Standardization:** Data transformation helps to convert data into appropriate formats as well as helps to aligns the dataset with the model requirements. Data standardization ensures that the features have similar distribution which helps in achieving better convergence during training.
- **Errors Correction:** Correcting the errors in the dataset by finding missing and null values and replacing them by the mean value of the attribute.
- **Label Encoding:** A label encoder allows us to convert the categorical data into numerical form which in necessary as the machine learning model requires numerical data input rather than text or categorical input.

Linear Regression

One of the most basic and widely used machine learning algorithms is Linear regression. Linear regression is method of performing predictive analysis mathematically and also projections of continuous, real or mathematical variables are possible with linear regression. This method establishes linear relationships between dependent and independent variables through data modeling and analysis. Thus, this approach would simulate correlations between independent and dependent variables from analysis and learning to the present training outcomes. Linear Regression uses the range of values for the independent variable, x to estimate the value of the dependent variable, y (Maulud & Abdulazeez, 2020). If there are p independent variables, the regression equation is as given below.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

Where:

$\hat{y} = predicted\ valuey$

$\beta_0 = intercept$

$\beta_1, \beta_2, \dots, \beta_p = coefficients$

$x_1, x, \dots, x_p = predictors$

Training linear regression means learning the best values of the coefficients (β) so that the model can predict output ($\hat{y}$) as accurately as possible. Our goal is to find (β) that minimizes the cost function.

$$J(\beta) = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2$$

Where:
$m = number\ of\ training\ examples$
y_i = actual output
$\hat{y}_i$ = predicted output
$\beta = model\ parameters\ (\beta_0, \beta_1, \beta_2, \dots, \beta_p)$

Lasso Regression (L1 Regularization)

Regularization technique like lasso regression is frequently employed to enhance the performance of conventional linear regression models. The decision of Lasso regression is based on the unique characteristics of the problem as well as the desired model attributes. Lasso regression works effectively in situations where a large number of characteristics exists, but only a tiny percentage of those features are meaningful. Researchers can create dependable and effective regression models for a variety of applications with this method's strong frameworks for improving model performance and interpretability. Researchers can choose the Lasso regularization approach for their particular regression analysis by carefully assessing the benefits and drawbacks of the strategy. Lasso regression is a useful tool for managing multicollinearity and avoiding overfitting. Lasso (L1 regularization) encourages the usage of less complex models, which are beneficial for automated model selection and multicollinearity. Lasso regression is suited for the dataset containing many attributes with only few important ones (Safi et al., 2023). The cost function in lasso regression is

$$J(\beta) = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 + \lambda\sum_{j=1}^{p}|\beta_j|$$

Where:
$\lambda = regularization\ parameter$ (strength of penalty)

Ridge Regression (L2 Regularization)

In the areas of data science and machine learning, regularization technique like Ridge regression is frequently employed to enhance the performance of conventional linear regression models. The decision of Ridge regression is based on the unique characteristics of the problem as well as the desired model attributes. Ridge regression works effectively where there is presence of multicollinearity and when all features impact the model's performance. Researchers can create dependable and effective regression models for a variety of applications with this method's strong frameworks for improving model performance and interpretability. Researchers can choose the Ridge regression approach for their particular regression analysis by carefully assessing the benefits and drawbacks of the strategy. Ridge regression can help prevent overfitting and manage multicollinearity. Ridge regression (L2 regularization) is very useful for multicollinearity since it lowers estimate variability and bias. When all features have an effect on the performance of the model, ridge regression works well. Ridge regression improve the interpretability and performance of the models, allowing researchers to create dependable regression models for a range of uses (Safi et al., 2023). The cost function in ridge regression is

$$J(\beta) = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^{p} {\beta_j}^2$$

Where:

$\lambda = regularization\ parameter$ (strength of penalty)

Elastic-net Regression (L1_L2 Regularization)

Elastic net (ELNET) regression is a statistical technique that combines the best aspects of Lasso and Ridge regression. It is used to handle multicollinearity issues between predictor variables and regularize and select important predictor variables that have a significant impact on the response variable. ELNET regression is utilized to regularize and choose the key predictor variables, despite high levels of multicollinearity between the predictors resulting in a simple model with the most significant predictor variables. ELNET can eliminate or choose the predictor variables with a high correlation from the final model to improve prediction accuracy. ELNET regression combines the two best shrinkage regression techniques: Lasso regression (L1 penalty) for feature selection of regression coefficients and Ridge regression (L2 penalty) for handling multi-collinearity problems. The ELNET regression shrinks the number of predictor variables toward zero or equal to zero for the less significant variables using the sum of the absolute values of the coefficient variables (L1 norm) multiplied by the tuning parameter λ_1 where the sum of squared coefficient variables (L2 norm) multiplied by the tuning parameter λ_2 is used to address the strong correlation between the predictor variables. So, the ELNET regression principle helps to produce an interpretable fitting model by keeping the predictor variables with strong correlation within or outside of the fitted model to improve prediction accuracy (Al-Jawarneh et al., 2021). The cost function in elastic-net regression is

$$J(\beta) = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 + \lambda_1 \sum_{j=1}^{p}|\beta_j| + \lambda_2 \sum_{j=1}^{p} {\beta_j}^2$$

Where:

$\lambda_1 = L1$ penalty

$\lambda_2 = L2$ penalty

Performance Metrics

Among different performance metrics of linear regression, Mean Squared Error (MSE) is used as a performance measure in this study. The Mean Squared Error (MSE) is most common and widely used metric to determine the performance of regression model. It determines how near a regression line is to a given set of points which is accomplished by squaring the distance between the points and the regression line also referred as the “errors”. Squaring is required to eliminate the negative signs. Since the average of the set of errors is determined, it is known as Mean Squared Error. The prediction is much better when the MSE is smaller. The formula for the MSE is given by the following equation(Gregersen & Stewart, 2023).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Where,
n = number of data points
y_i = true value
$\hat{y}_i$ = predicted value

Implementation

Python programming language and associated libraries are used to implement the algorithms using Google Colab environment proved by Google. The different Python libraries used are NumPy, Pandas, Matplotlib, and sklearn. During implementation, the Nepali Housing Price dataset is loaded using panda library. After preprocessing on the dataset as mentioned in methodology, the original dataset was separated between dependent and independent variables as well as the dataset was splitted between training and testing dataset using the standard ratio of 80:20.

The dataset has been modeled using traditional Linear regression. The linear regression model also helps to analyze the relationship between the dependent variables and the independent variables. MSE is used to evaluate the performance on both the training and test dataset with the MSE of 4.2432 and 9.3807 respectively. This result also helps to identify that the model suffer from problem of overfitting as the MSE on training dataset is less as compared to the MSE on testing dataset. So, to solve this overfitting problem and to provide simpler and more cost-effective model regularization techniques like lasso, ridge and elastic net regression is employed.

Lasso regression is applied to overcome the problem of overfitting problem which occur while implementing linear regression. Lasso regression is applied along with the varying values of hyperparameter lambda (λ) for selecting the appropriate value of hyperparameter which will provides the minimum MSE. Ridge regression or L2 regularization is applied to overcome the problem of overfitting problem which occur while implementing linear regression. Ridge regression is applied along with the varying values of hyperparameter lambda (λ) for selecting the appropriate value of hyperparameter which will provides the minimum MSE. Elastic-Net regression is applied to overcome the problem of overfitting problem which occur while implementing linear regression. Elastic-Net regression is applied along with the varying values of hyperparameter lambda (λ_1 and λ_2) for selecting the appropriate value of hyperparameter which will provides the minimum MSE.

Hyperparameter Optimization

The process of determining the ideal settings for the hyperparameters in order to maximize performance metrics is known as hyperparameter optimization. For regularization techniques finding the hyperparameters is the value of lambda (λ) so that the overall performance will be optimized. The selection of the appropriate hyperparameter lambda (λ) is done using cross-validation method where the value lies between 0 and 1. The values of lambda used are [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9] to examine the influence of altering lambda values on model performance and to find out the ideal value of lambda (λ) that reduces and provides the minimum MSE through cross-validation.

System Specifications

This thesis research work is carried out in the machine with 8 GB RAM, Intel(R) Core (TM) i7-8750H CPU processor, 64-bit operating system, Windows 10 Home, Google Colab and Python 3.10.4.

Result Analysis

The experiment was conducted on the Nepali Housing Price dataset, available in Kaggle repository. Two new datasets were created from this dataset: a training dataset comprising 80% of the total and a testing dataset including the remaining 20%. Initially, the traditional linear regression was employed on both the training and testing datasets providing the MSE of 4.2432 and 9.3807 respectively. This result showed that the model suffered from problem of overfitting as the MSE on training dataset was less as compared to the MSE on testing dataset. So, to solve this overfitting problem and to provide simple and cost-effective model, regularization techniques like lasso, ridge and elastic net regression were employed. The result after employing different regularization techniques on linear regression model is shown in Table 2.

Table 2. *MSE obtained on Nepali Housing Price Dataset for L1, L2 & L1_L2 regularization.*

Hyperparameter lambda (λ)	Mean Square Error (MSE)		
	L1 regularization	L2 regularization	L1_L2 regularization
0.1	4.9038	5.3925	4.1512
0.2	4.3631	5.2515	4.0637
0.3	**4.2516**	5.1435	3.9805
0.4	4.5692	5.0471	3.9012
0.5	5.3161	4.9588	3.8254
0.6	6.4924	4.8774	3.7531
0.7	8.0974	4.8022	3.6839
0.8	10.1497	4.7328	3.6176
0.9	12.6478	**4.6688**	**3.5541**

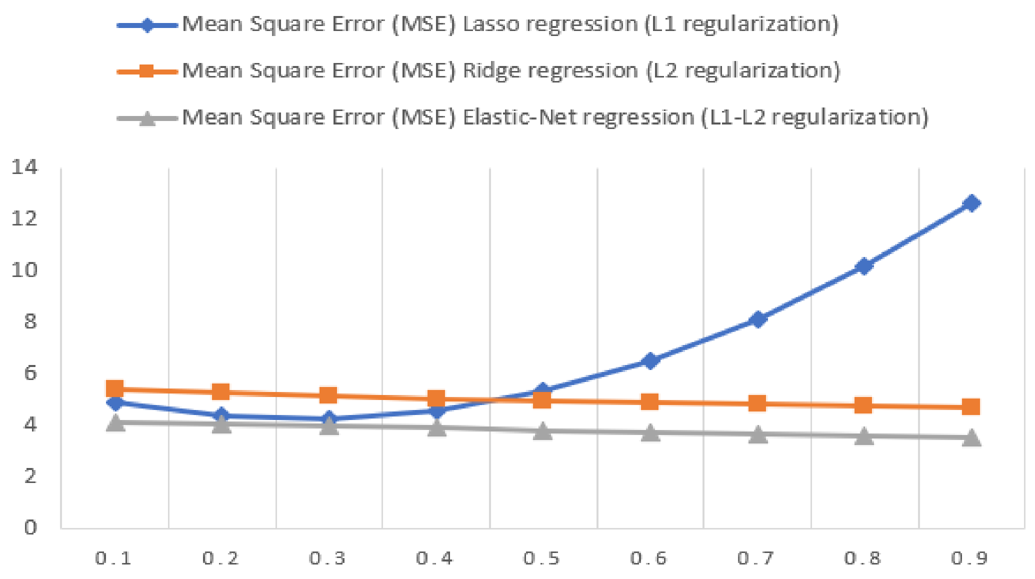

Fig. 1. *MSE obtained from different values of hyperparameter.*

As shown in Figure 1, for the Nepali housing price dataset the lowest MSE was obtained from Elastic net regression while lambda (λ) = 0.9 with MSE of 3.5541. The line graph of elastic net also shows how the MSE keep declining with the increase in the value of lambda (λ) with the ideal value of lambda (λ) represented by the lowest point on the curve. The optimal value of lambda (λ) effectively fine tunes the regression model with minimum MSE and also helps to prevent coefficients from exploding uncontrollably or being too controlled. The figure also shows that the curve representing the ridge regression exhibits similar characteristics to the Elastic net regression with initially high MSE of 5.3925 for smaller value of lambda (λ) = 0.1 which when decreases with increase in the value of lambda (λ) providing similar curve. On other hand, the curve representing lasso regression shows opposite characteristics than both the ridge and elastic net with initially low MSE of 4.9038 for smaller value of lambda (λ) = 0.1 but keep increases exponentially with higher value of lambda (λ).

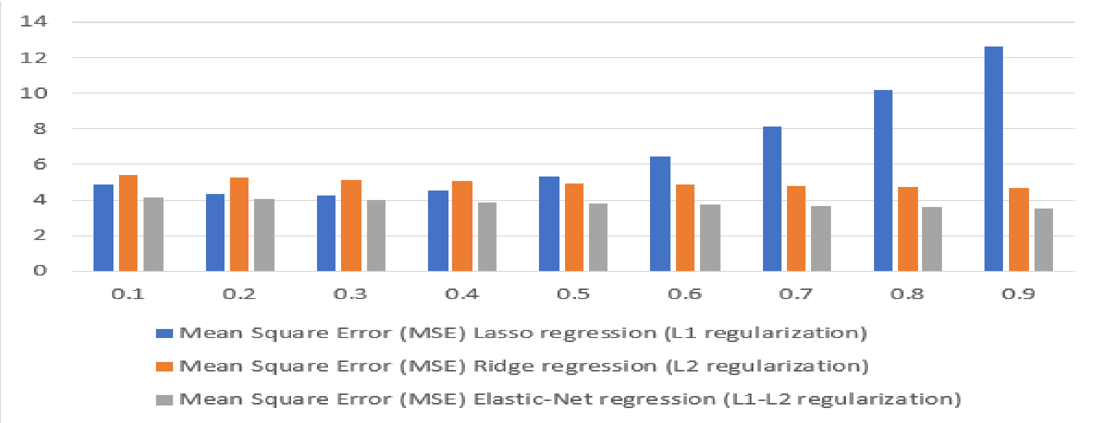

Fig. 2. MSE obtained from different values of hyperparameter for L1, L2 and L1_L2 regularization.

The bar graph in figure 2 shows the overall MSE for all L1, L2 and L1_L2 regularization with varying value of hypermeter (λ) with elastic net (L1_L2 regularization) regression providing the lowest MSE in every value of hyperparameter lambda (λ) as compared to lasso (L1 regularization) regression and ridge (L2 regularization) regression. The figure shows that elastic net regression provides the minimum MSE of 3.5541 with the ideal value of hyperparameter lambda (λ) at 0.9 as compared to ridge with MSE of 4.6688 at lambda (λ) =0.9 and lasso with MSE of 12.6478 at lambda (λ) =0.9.

Conclusion

The purpose of this research is to identify the problem of overfitting on linear regression model and to minimize the overfitting problem by using of different regularization techniques applicable on regression model. Linear regression is used to model the Nepali Housing Price dataset obtained from Kaggle repository which causes the problem of overfitting. Different regularization techniques like Lasso (L1 regularization), Ridge (L2 regularization) and Elastic-Net (L1_L2 regularization) are used to calibrate the linear regression model by adding their respective penalty terms to the loss function. MSE is used as performance metrices for this research work to analyze the performance of these regularization techniques. Optimization of the hyperparameter lambda (λ) is also carried out for all the regularization techniques to find the ideal value of lambda that provides the minimum MSE value. The elastic net regression also known as L1_L2 regularization provides the minimum MSE of 3.5541 at ideal value of lambda (λ) = 0.9 where the MSE of lasso (L1 regularization) and ridge (L2 regularization) at lambda (λ) = 0.9 are 12.6478 and 4.6688 respectively. Also, Elastic-net regression outperform both lasso and ridge regression not only at lambda value of 0.9 but at every value of lambda from 0.1 to 0.9.

Future studies may explore other machine learning models along with their respective regularization techniques, as this study only focuses on the regression model and the regularization techniques applicable on regression model. It is not certain that any particular regularization technique is applicable for all types of datasets so it depends on dataset and its attributes for determining an appropriate regularization technique. Therefore, replicating this research with different and large dataset can be an excellent idea to enhance the reliability of the results obtained in this research work.

References

Al-Jawarneh, A. S., Ismail, M. T., & Awajan, A. M. (2021). Elastic Net Regression and Empirical Mode Decomposition for Enhancing the Accuracy of the Model Selection. *International Journal of Mathematical, Engineering and Management Sciences*, *6*(2). https://doi.org/10.33889/IJMEMS.2021.6.2.034

Brownlee, J. (2016). Overfitting and Underfitting With Machine Learning Algorithms. *Machine Learning Mastery*.

DataCamp. (2019, November 12). *Regularization in R Tutorial: Ridge, Lasso & Elastic Net Regression | DataCamp*. https://www.datacamp.com/tutorial/tutorial-ridge-lasso-elastic-net

Dave, V., & Shah, K. (2019). *COMPARATIVE ANALYSIS OF REGULARIZATION TECHNIQUES IN ARTIFICIAL NEURAL NETWORKS*. https://doi.org/10.1729/Journal.22756

Ghodasara, K. (2021). Overview of Decision Tree Pruning in Machine Learning. *International Research Journal of Engineering and Technology*.

Gregersen, E., & Stewart, K. (2023, March 22). *Mean squared error (MSE) | Definition, Formula, Interpretation, & Facts | Britannica*. https://www.britannica.com/science/mean-squared-error

Kolluri, J., Kotte, V. K., Phridviraj, M. S. B., & Razia, S. (2020). Reducing Overfitting Problem in Machine Learning Using Novel L1/4 Regularization Method. *Proceedings of the 4th International Conference on Trends in Electronics and Informatics, ICOEI 2020*. https://doi.org/10.1109/ICOEI48184.2020.9142992

Kurama, V., Whitfield, B., & Oppermann, A. (2024, August 1). *Regression in Machine Learning: Definition and Examples | Built In*. https://builtin.com/data-science/regression-machine-learning

Maulud, D., & Abdulazeez, A. M. (2020). A Review on Linear Regression Comprehensive in Machine Learning. *Journal of Applied Science and Technology Trends*, *1*(2). https://doi.org/10.38094/jastt1457

Mikołajczyk, A., & Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. *2018 International Interdisciplinary PhD Workshop, IIPhDW 2018*. https://doi.org/10.1109/IIPHDW.2018.8388338

Safi, S. K., Alsheryani, M., Alrashdi, M., Suleiman, R., Awwad, D., & Abdalla, Z. N. (2023). Optimizing Linear Regression Models with Lasso and Ridge Regression: A Study on UAE Financial Behavior during COVID-19. *Migration Letters*, *20*(6). https://doi.org/10.59670/ml.v20i6.3468

Sagala, N. T. M., & Cendriawan, L. H. (2022). House Price Prediction Using Linier Regression. *2022 IEEE 8th International Conference on Computing, Engineering and Design, ICCED 2022*. https://doi.org/10.1109/ICCED56140.2022.10010684

Simplilearn. (2025, September 7). *The Best Guide to Regularization in Machine Learning | Simplilearn*. https://www.simplilearn.com/tutorials/machine-learning-tutorial/regularization-in-machine-learning

Thapa, S. (2020, April 27). *Nepali Housing Price Dataset*. https://www.kaggle.com/datasets/sagyamthapa/nepali-housing-price-dataset?resource=download

Trotta, F. (2022). *Understanding l1 and l2 Regularization | Towards Data Science*. https://towardsdatascience.com/understanding-l1-and-l2-regularization-93918a5ac8d0/