# 

Multilayer Perceptron

Machine Learning:

CANDIDATE NO: 277238

2024

Table of Contents

[0](#_Toc166853699)

[INTRODUCTION 2](#_Toc166853700)

[1. PERFORMANCE 2](#_Toc166853701)

[2. MULTILAYER PERCEPTRON MODEL 3](#_Toc166853702)

[I. Activation function for the output layer 3](#_Toc166853703)

[II. Loss Function 3](#_Toc166853704)

[III. Number of units in the output layer 3](#_Toc166853705)

[I. L2 Regularisation 3](#_Toc166853706)

[II. Model Complexity 4](#_Toc166853707)

[III. Batch size 4](#_Toc166853708)

[3. FEATURES AND LABELS 4](#_Toc166853709)

[I. Lag Features 5](#_Toc166853710)

[II. Exchange Rates and Land Temperatures 5](#_Toc166853711)

[4. PRE-PROCESSING 6](#_Toc166853712)

[I. Handling Missing Values 6](#_Toc166853713)

[II. Feature Scaling 6](#_Toc166853714)

[III. Encoding Variables 6](#_Toc166853715)

[CONCLUSION 7](#_Toc166853716)

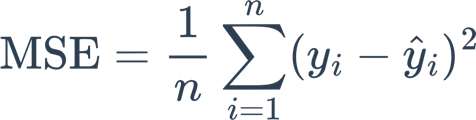
# INTRODUCTION

This report focuses on building a multilayer perceptron (MLP) model for forecasting export value of crop products, specifically three years into the future using a comprehensive dataset. By leveraging the power of machine learning techniques, we aim to capture the underlying patterns and relationships in the data to make accurate predictions. The objective is to present the development and evaluation of the MLP model. The report is structured as follows: Section 1 discusses the performance of the model. Section 2 describes the architecture of the MLP model. Section 3 focuses on the features and labels used for training and Section 4 explains the pre-processing steps taken to prepare the data for modelling.

# PERFORMANCE

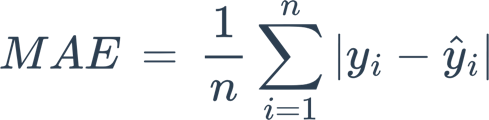
The multilayer perceptron model was evaluated using three metrics. This includes Mean Squared Error, Mean Absolute Error and R-Squared Score. These metrics offer an assessment of how the model predicts and how well it fits.

Mean Squared Error measures the average squared difference between the predicted and the actual target variables. It is calculated using the following formula:



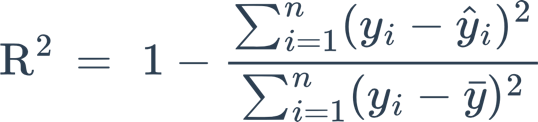
Where yi = the actual export value, ŷi = predicted export value, and n = total number of instances in the dataset. The Mean Squared Error places emphasis on errors because of the squaring process, which makes it responsive to extreme values. A lower MSE indicates better model performance as it indicates that the predicted values are closer to the true values.

Mean Absolute Error measures the average absolute difference between the predicted and the actual target variables (encord.com, n.d.).



The Mean Absolute Error (MAE) treats every error equally no matter how big or small. It offers a way to understand the error of predictions, in the same units, as the target value. Like, MAE suggests that the model is performing better.

R-Squared, commonly known as the coefficient of determination, provides information on how well the regression line approximates the actual data (Newcastle University, 2023). It quantifies the proportion of variance in the export values that can be explained by the input features. It is calculates using the following formula:



Where {"type":"$$","font":{"color":"#2c3e50","size":11.5,"family":"Roboto"},"id":"3","code":"$$\\bar{y}$$","aid":null,"backgroundColor":"#ffffff","ts":1715950478584,"cs":"VguDl0JGxVXuhg8rPriBYQ==","size":{"width":6,"height":12}} = mean of the actual export values. An R2 value close to 1 indicates that the model explains a large portion of the variance in the export values, while a value closer to 0 suggests that a model does not perform much better than a simple average baseline.

In the analysis, a total of 14,823,125 instances were utilised. The dataset was divided into two sets, training and validation, using the year 2020 as a threshold. The training set comprised 13,586,325 instances, which was used to train the model. The validation set comprised 1,236,800, was used to evaluate the performance of the model.

The MLP model achieved the following performance on the validation set:

* Mean Squared Error (MSE): 1532468379.1302736
* Mean Absolute Error (MAE): 1881.513166568991
* R-squared (R2): 0.9996998488214139

# MULTILAYER PERCEPTRON MODEL

The multilayer perceptron model, a type of artificial neural network, consists of an input layer, hidden layer, and an output layer. The specific architecture of the model is as follows:

1. Activation function for the output layer:

The output layer uses the Rectified Linear Unit (ReLU) activation function.

{"backgroundColor":"#ffffff","type":"$$","code":"$$\\text{ReLU(}x\\text{)}\\;\\text{=}\\;\\max\\left(0,x\\right)$$","font":{"color":"#000000","size":11,"family":"Arial"},"aid":null,"id":"4","ts":1715939733972,"cs":"NGESf8LOekf1YFLb0aiSpA==","size":{"width":164,"height":16}}

It returns 0 for negative input values for the input values for positive inputs. ReLU introduces non-linearity into the model, which allows to learn complex patterns and relationships in the data.

1. Loss Function:

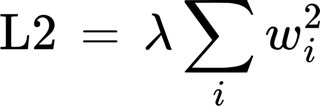
The model is trained using the Mean Squared Error loss function which is default in MLPRegressor implementation. The objective of the training process is to minimise the MSE.

1. Number of units in the output layer:   
   Since the task involves predicting the export value, a regression problem, the output layer consists of a single unit. The unit outputs the predicted export value for the given set of input features.

To prevent overfitting, which is when a model learns to fir the noise in the training data rather than the underlying patterns, multiple steps were taken during the model training process. These involved:

1. L2 Regularisation:

L2 Regularisation, also known as weight decay, adds a penalty term to the loss function that is proportional to the square of the magnitude of the model’s weights. The regularisation term is calculated as:



Where {"id":"7","backgroundColor":"#ffffff","code":"$$\\lambda$$","font":{"family":"Arial","size":12,"color":"#000000"},"type":"$$","aid":null,"ts":1715951390743,"cs":"sz51OdtSygRvsY+v6mi8eA==","size":{"width":8,"height":12}} is the regularization parameter that controls the strength of regularization, and {"font":{"family":"Arial","size":11,"color":"#000000"},"aid":null,"code":"$$w_{i}^{}$$","type":"$$","id":"5","backgroundColor":"#ffffff","ts":1715951718101,"cs":"GV1I+sZH63XIQmLfer7M5Q==","size":{"width":16,"height":12}} represents the individual weights of the model.

A regularisation parameter of 0.01 was utilised, encouraging the model to acquire weights thus simplifying the process and potentially enhancing performance.

1. Model Complexity:

The complexity of the model was kept relatively low by limiting the number of hidden layers. The first hidden layer consists of 50 units, while the second layer consists of 25. This reduces the model’s ability to memorise noise in the training data.

1. Batch size:

The batch size indicates how many instances are handled together in each round of the training process. We used a batch size of 64 during training to introduce stochasticity which helps the model generalise better.

By using methods such, as L2 regularization simplifying the model and adjusting batch size the MLP model aims to find a ground between fitting the training data and being able to make predictions on new data. These tactics are meant to reduce overfitting and enhance the model’s accuracy, for export values.

# FEATURES AND LABELS

In this analysis, 5 datasets were used to collect all the information relevant for the prediction of crop products’ export value. These datasets contained data on crop production, exchange rates, food balances, food trade indicators, and land temperature changes. The food trade dataset was used as the base, then features from the other datasets were progressively merged to form one dataset.

The target variable is ‘Export Value’, which was directly obtained from the ‘Value’ column in the Food trade dataset.

A total of 23 features were selected and used to train the MLP model. They were selected based on their potential relevance to the export value of crop products. The features listed below were used to train the model. They include:

1. Area
2. Products
3. Export Quantity
4. Crop Production
5. Exchange Rate\_x
6. Land Temperature\_x
7. Exchange Rate\_y
8. Land Temperature\_y
9. Export Value Lag 1
10. Export Value Lag 2
11. Export Value Lag 3
12. Export Quantity Lag 1
13. Export Quantity Lag 2
14. Crop Production Lag 1
15. Crop Production Lag 2
16. Exchange Rate\_x Lag 1
17. Exchange Rate\_x Lag 2
18. Land Temperature\_x Lag 1
19. Land Temperature\_x Lag 2
20. Exchange Rate\_y Lag 1
21. Exchange Rate\_y Lag 2
22. Land Temperature\_y Lag 1
23. Land Temperature\_y Lag 2

Some of the features were derived from the original dataset through different kinds of transformations:

1. Lag Features:

In each ‘Area’ and ‘Product’ category, lag features were established by shifting the related feature values by 1, 2 or 3 years. These lagged features help grasp the time related connections and past data that could be useful for forecasting export value. For example, the export value from the previous year (Lag 1) may provide valuable insights into the current year's export value. By including lag features, the model can learn from past patterns and trends in the data.

1. Exchange Rates and Land Temperatures:

The exchange rates and land temperatures were averaged per year and area to obtain a single value for each year-area combination. This helps capture the overall trends and variations in these variables. The fluctuation of exchange rates can have an impact on the competitiveness and profitability of exports whereas variations in land temperatures can directly affect crop yields and overall production. By taking into account these characteristics the model is able to factor in the economic and environmental influences that play a role in determining export values.

The inclusion of these derived characteristics is intended to furnish the MLP model with a range of data, for learning and making forecasts. The lagging features capture the time related dynamics and relationships within each sector and product category enabling the model to factor in contexts when predicting export values. The averaged currency exchange rates and land temperatures offer measures of environmental conditions allowing the model to consider their potential effects on export values.

Through selection and crafting of these features, the aim is to provide the MLP model with comprehensive information to enhance its predictive accuracy. By combining features with derived ones, a feature space is formed that covers factors affecting crop product export values aiding the model in understanding intricate patterns and making reliable long-term predictions.

# PRE-PROCESSING

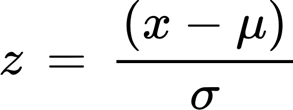
Several pre-processing steps were performed before the model was trained. This ensures data quality and improved model performance. These pre-processing techniques address common issues such as missing values, feature scaling, and encoding of categorical variables. The specific pre-processing steps employed in this analysis are as follows:

1. Handling Missing Values:

Missing numeric values were first filled with the mean value grouped by ‘Area’. This approach takes into account the variations across different areas and provides a more context-specific imputation. By considering the area-specific means, the imputed values are more representative of the local conditions. The remaining missing values after the first approach was implemented were then filled with the overall mean of the respective feature.

Handling missing values is crucial to avoid biases and information loss in the model training process

1. Feature Scaling:  
   Feature Scaling is a step that adjusts all features to a consistent scale especially aiding in the convergence and stability of the MLP model. StandardScaler from the scikit-learn library was implemented to apply standardisation to the following features:
2. Exchange Rate\_x
3. Exchange Rate\_y
4. Land Temperature\_x
5. Land Temperature\_y
6. Crop Production



1. Encoding Variables:

Categorical variables ‘Area’ and ‘Products’, cannot be directly used as input features in the MLP model as it requires numerical inputs. Label Encoding was implemented to address this, converting the categorical variables to numerical representations

These pre-processing steps—handling missing values, feature scaling, and encoding categorical variables—are essential for preparing the data for training the MLP model. They make sure that the model gets quality compatible input features in terms of numbers. Dealing with values by using area overall mean imputation reduces the impact of incomplete data allowing the model to learn from all available examples. Standardizing feature scaling brings all features to a scale enhancing the model’s convergence and stability. Encoding variables helps the model effectively process. Learn from categorical information.

By using these methods, the dataset is changed into a format suitable for training the MLP model. The processed features lay a groundwork for the model to recognize patterns, understand relationships and make predictions of export values three years ahead.

# CONCLUSION

This report details the development and evaluation of a perceptron (MLP) model that predicts the export value of crop products three years into the future. The MLP system underwent training with a dataset encompassing, past export values, crop yields, currency exchange rates, land temperatures and other relevant variables.

The model was designed with a Rectified Linear Unit (ReLU) activation function in the output layer, Mean Squared Error (MSE) loss function for training, and a single output unit for predicting the export value. Techniques such as L2 regularization, reduced model complexity, and batch size were employed to prevent overfitting and improve generalization.

A total of 23 features, including original variables and derived features like lag variables, averaged exchange rates, and land temperatures, were used to capture the potential influences on export value. Pre-processing techniques, such as handling missing values, feature scaling, and encoding categorical variables, were applied to ensure data quality and optimise model performance.

The model's performance was evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R2) score on a separate validation set. The results demonstrated the MLP model's ability to capture complex patterns and make accurate predictions for the export value of crop products three years into the future.

The validation results were saved to a CSV file named 'predictions\_results.csv', which includes the true export values and the corresponding predicted export values for the validation instances.

In conclusion, this report showcases the application of an MLP model for long-term forecasting of crop product export values. The insights gained from this analysis can assist stakeholders in the agricultural sector in making informed decisions and developing effective strategies for resource allocation and planning. However, the model's limitations should be acknowledged, and its predictions should be used as a complementary tool alongside other sources of information and domain expertise.