Nida Tazeen

Explainability

Contents

[1. Introduction to Classifier Models 2](#_Toc189152896)

[2. Decision Tree (DT) 2](#_Toc189152897)

[2.1 Train and Test Accuracy 2](#_Toc189152898)

[2.2 Feature Importance 2](#_Toc189152899)

[2.3 Discussion on Feature Importance 3](#_Toc189152900)

[3. Logistic Regression (LR) 4](#_Toc189152901)

[3.1 Train and Test Accuracy 4](#_Toc189152902)

[3.2 Feature Importance (Coefficients): 4](#_Toc189152903)

[3.3 Discussion on Feature Importance 4](#_Toc189152904)

[1.1 Image of Feature Importance (Coefficients): 5](#_Toc189152905)

[1.2 Issues or Difficulties Encountered 5](#_Toc189152906)

[2. Multi-layer Perceptron (MLP) 6](#_Toc189152907)

[2.1 Train and Test Accuracy 6](#_Toc189152908)

[2.2 Feature Importance (Permutation Importance) 6](#_Toc189152909)

[2.3 Discussion on Feature Importance 6](#_Toc189152910)

[2.4 Image of Featured Importance 7](#_Toc189152911)

[2.5 Issues or Difficulties Encountered 7](#_Toc189152912)

[AI Tools Used 7](#_Toc189152913)

# Introduction to Classifier Models

Classifier models are a cornerstone of machine learning and data analysis, enabling the automated categorization of data into distinct classes based on identified features (Classification in Machine Learning: A Guide for Beginners - DataCamp, 2022). Essentially, these models function by learning from a training dataset, which consists of instances categorized by class labels. This learning process helps the model identify patterns and relationships within the data, allowing it to predict the class of new, unseen instances accurately (IBM, 2024). The applications for classifier models are vast, extending from image recognition and email spam filtering to medical diagnoses and sentiment analysis in textual data (Classification in Machine Learning: A Guide for Beginners - DataCamp, 2022).

Various classification algorithms exist, each designed to meet different requirements and contexts. Common models include Decision Trees, Logistic Regression, Support Vector Machines (SVM), and Neural Networks, with each algorithm possessing its unique strengths and weaknesses (IBM, 2024). As businesses and organizations increasingly turn to data-driven decisions, understanding the functionality and implications of classifier models has become vital for practitioners aiming to leverage machine learning methodologies effectively to enhance their analytical capabilities (IBM, 2024).

​In sum, classifier models are integral to utilizing machine learning for practical applications, providing the ability to make informed decisions based on data-driven insights.

# Decision Tree (DT)

Decision trees are a widely used method in data mining for creating classification and regression models based on sequential data splits. They utilize a tree-like structure that divides the dataset into branches to identify patterns and relationships among variables. Each node represents a feature, while leaves denote class labels or predicted outcomes. Their interpretability and non-parametric nature make them a preferred choice for many applications, despite their susceptibility to overfitting, especially with small datasets.

## Train and Test Accuracy

| **Metric** | **Value** |
| --- | --- |
| Train Accuracy | 0.99 |
| Test Accuracy | 0.97 |

## Feature Importance

| **Feature** | **Value** |
| --- | --- |
| Salary | 0.0218 |
| Age | 0.0126 |
| Credit Score | 0.8675 |
| Debt | 0.0981 |

## Discussion on Feature Importance

* The most important feature for the Decision Tree model was **credit score**, with a feature importance score of **0.8675**. This indicates that the model is heavily relying on the credit score to make predictions, which makes sense given that credit score is typically a key determinant in approval for credit-related tasks.
* **debt** and **salary** also played important roles, though they had smaller importance values in comparison to credit score.
  1. Image of Trained Decision Tree

A diagram of a network

Description automatically generated

* 1. Issues or Difficulties Encountered
* The Decision Tree model showed high accuracy, but overfitting could be a concern as the model's training accuracy is significantly higher than test accuracy. Regularization (e.g., reducing tree depth or pruning) might help to further improve the generalization ability of the model.
* One issue could be the interpretability of the Decision Tree when the depth is large, which might lead to a complex and hard-to-understand tree.

# Logistic Regression (LR)

Logistic regression is prominent in statistical analysis, particularly for binary classification problems. It estimates the probability of a binary outcome based on one or more predictor variables (2014). Instead of assuming a linear relationship as in linear regression, logistic regression employs a logistic function to transform predicted values into probabilities bound between 0 and 1, allowing it to model categorical outcome variables effectively. While logistic regression is easy to implement and interpret, it may not perform well in cases where the relationship between predictors and the target variable is nonlinear.

## Train and Test Accuracy

| **Metric** | **Value** |
| --- | --- |
| Train Accuracy | 0.97 |
| Test Accuracy | 0.96 |

## Feature Importance (Coefficients):

| **Feature** | **Value** |
| --- | --- |
| Salary | 0.0218 |
| Age | 0.0126 |
| Credit Score | 0.8675 |
| Debt | 0.0981 |

## Discussion on Feature Importance

* Similar to the Decision Tree, **credit score** emerged as the most important feature in the Logistic Regression model, with a significant coefficient value.
* The coefficients for **debt** and **salary** also contributed to the prediction, but their effect was smaller compared to credit score. Logistic Regression models typically show how strongly each feature affects the outcome through their coefficients.

## Image of Feature Importance (Coefficients):

A graph with blue squares

Description automatically generated

## Issues or Difficulties Encountered

* Logistic Regression has the limitation of assuming a linear relationship between the features and the target variable. This could be a potential issue if the relationship is non-linear.
* Another challenge could be multicollinearity, where highly correlated features may distort the model's ability to assess their importance properly.

# Multi-layer Perceptron (MLP)

Multilayer Perceptrons (MLP) are a type of artificial neural network that consists of multiple layers of nodes, or neurons, interconnected by weighted links (Contributors to Wikimedia projects, 2005). MLPs are designed to capture complex, nonlinear relationships by employing activation functions across hidden layers to transform inputs into outputs through a feedforward architecture (Contributors to Wikimedia projects, 2005). Though they excel in handling large datasets and complex interactions, MLPs require substantial computational resources and may demand careful tuning to avoid overfitting and ensure generalizability (“Introduction to Machine Learning, Neural Networks, and Deep ...,” 2025).

## Train and Test Accuracy

| **Metric** | **Value** |
| --- | --- |
| Train Accuracy | 0.96 |
| Test Accuracy | 0.97 |

## Feature Importance (Permutation Importance)

| **Feature** | **Value** |
| --- | --- |
| Credit score | 0.3652 |
| Debt | 0.0590 |
| Salary | 0.0232 |
| Age | 0.0045 |

## Discussion on Feature Importance

* The **credit score** again proved to be the most influential feature for the MLP model, with the highest importance value of **0.3652**. This aligns with our expectations that credit score is the primary factor in credit decisions.
* **debt** also had a relatively higher importance than other features, making it a significant factor for MLP predictions as well.

## Image of Featured Importance

A graph with blue squares

Description automatically generated

## Issues or Difficulties Encountered

* Training the MLP model required a higher number of iterations (max\_iter=1100) due to the complex nature of the neural network and the dataset. This could have been optimized further with a proper early stopping criterion.
* MLP models are computationally more expensive than Decision Trees or Logistic Regression, and they might overfit if the model architecture (hidden layers) is too complex.

## AI Tools Used

ChatGPT was utilized to assist in organizing to ensuring clarity and consistency in the implementation.

**References**

*Classification in Machine Learning: A Guide for Beginners - DataCamp*. (2022). <https://www.datacamp.com/blog/classification-machine-learning>

*What are classification models? | IBM*. (2024). <https://www.ibm.com/think/topics/classification-models>

\*correspondence: yanyansongsjtu.edu.cn(Yan-yan SONG); (2025). Decision tree methods: applications for classification and prediction. In *nih.gov*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4466856/>

Introduction to Machine Learning, Neural Networks, and Deep ... (2025). In *nih.gov*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC7347027/>

*Multilayer perceptron - Wikipedia*. (2005). <https://en.wikipedia.org/wiki/Multilayer_perceptron>

Sandro Sperandei. (2014). Understanding logistic regression analysis - PMC - PubMed Central. In *Biochemia Medica*. <https://pmc.ncbi.nlm.nih.gov/articles/PMC3936971/>