

- **Yassine Sfaihi**
- www.linkedin.com/in/yassinesfaihi
- www.github.com/yassinesfaihi
- www.kaggle.com/yassinesfaihi

[Author] : Yassine Sfaihi

Fraudulent transactions detection with Keras and TensorFlow: Hyperparameter Tuning and Weight Regularization for Improved Model Performance

Introduction

This project aims to develop a binary classification model for a credit card dataset. The model is built using the deep learning framework Keras, and the numerical computation library, TensorFlow. The model is trained and evaluated on a dataset of credit card transactions, where the goal is to predict whether a transaction is fraudulent or not.

Dataset

The dataset contains information about credit card transactions, including the amount of the transaction and various features such as time, V1-V28 (anonymized features), and a target variable "Class" which indicates whether the transaction is fraudulent (1) or not (0).

Preprocessing

The dataset is loaded into a Pandas dataframe and explored to get an overview of the data, including the data types of the columns, descriptive statistics, class distribution, and distribution of the "Amount" column for each class. Missing values and outliers are also checked and removed from the dataset.

Model Development

A 2-layer neural network model is defined and trained using the Adam optimizer and binary crossentropy loss. The model is trained for 10 epochs, with checkpoints saved after each epoch. The data is split into training and validation data, where 20% of the samples are used for validation. Both the training and validation data are normalized by subtracting the mean and dividing by the standard deviation.

Hyperparameter Tuning

To improve the model performance, a grid search is performed to find the best hyperparameters for the binary classification task. The grid search uses the scikit-learn's GridSearchCV class and considers various hyperparameters and their possible values. The best hyperparameters are stored and printed.

- **Yassine Sfaihi**
- www.linkedin.com/in/yassinesfaihi
- www.github.com/yassinesfaihi
- www.kaggle.com/yassinesfaihi

Model Evaluation

The precision, recall, and F1-score of the model's predictions on the validation data are calculated, and the model loss and confusion matrix are visualized. The model performance is also visualized using loss plots, confusion matrix, ROC curve, and precision-recall curve. Class weights are computed for the training targets to handle the class imbalance, if present.

Summary

This project demonstrates the development and evaluation of a binary classification model for a credit card dataset. The model is built using Keras and TensorFlow, with hyperparameter tuning and weight regularization for improved performance. The model is evaluated using various performance metrics and visualizations, and the best hyperparameters are found using a grid search.