AI and applications Report

Amirali Famili

March 10, 2024

Introduction

This report is trying to highlight the design choices and performance of the model, we begin with the problem, this model aims to predict the default payments made by the consumers based on the factors inside the CCD dataset, finally the model can evaluate it's performance with the knowledge of the final column which is indicating if the payment was a default one or not. Therefore, the model has a binary shape, it's trying to answer a yes or no question. The python code is divided into 4 main files excluding the read-me and dataset, they are properly commented and are easy to understand.
This report is divided into 3 sections, Dataset part is trying to clarify the process of cleaning the data and feeding the dataset into the model, Architecture and Performance section is justifying the design choices for the neural network as well as compiling the method and it's hyper-parameters, Training and Evaluation part is trying to express the reason behind the values for hyper-parameters through the experimental results made.

Dataset

Data Extraction

to begin with CCD dataset is a normal standard dataset the first row is an object which holds both indexes being Xs and lables for the data, to clean the data these headers being first and second rows had to be removed to be able to split the remaining data into training sets and test sets by removing the Y axis from Xs and by assigning it to Y, to then be able to train the model with training sets and validate the model performance of the model based on the Y values. After we have our clean data we can feed it to the Neural Network.

Architecture and Performance

Neural Network

This section is trying to justify the design choices made and how they improve the performance of the model, the main model recommended is NN inside the models file, it consists of 6 layers they all use relu activation function because it's a non-linear algorithm that accelerates the convergence of gradient descent towards the global minimum and since only

a certain number of neurons are activated and are statistically important neurons. Furthermore, the model starts with a input layer that has 512 neurons, the reason for this huge number is that we are trying to decrease the number of neurons within the hidden layers to be able to get a final binary result on the final layer with only 1 neuron using sigmoid activation function, in between the hidden layers we have a few Dropout layers, these layers are improving the performance of the model by stooping it from over-fitting which is a statistical error it significantly drops the performance of the model when the model has a limited data resource, the functions are too closely aligned, this will cause the model to produce false positive result.

Compiling the model

Currently, the model is producing a accuracy of around %82 and a loss around %13. the model is using Adam optimiser since it's a modern and proper algorithm for gradient problem of the data, accuracy matrices is used for determining the accuracy of the model, for the loss function of the model binary cross entropy is chosen because it makes more sense when we are looking for a binary result however the mse loss function is the closest alternative, they both can quickly reduce the loss by approximately %40, since the model has a high number of epochs as well as neurons and batch sizes, it has more time and chances to compare it's result with ground truth. The validation and training loss graph is plotted after the model is trained.

Training and Evaluation

Experimental Results

To begin with the model has a relatively small dataset and it can be the main reason why the model is not at it's peak performance however many efforts have been made to improve the efficiency of the model, when the model was first created it had 50 epochs and 23 batches however after a few run it became obvious that this model performs best with high number of epochs and batch size since the dataset is relatively small, although increasing the number of epochs was an improvement, the model acts strangely with high batch sises and the loss significantly rises. After many models trained and many results produced, Many efforts have been made to try to improve the models efficiency currently it has a accuracy of around %81 and a loss near %17