Transfer Learning in Credit Default Prediction

An Exploratory Study

Anthony Bernardi[1]

[1]University of Kentucky - Department of Statistics

December 11, 2024

Overview

Introduction

- What is Transfer Learning?
  - DL tool utilizing pre-trained models on a new prediction task
- Why is it relevant and important?
  - Reduce computational costs and time
  - Helps with limited data sets
- What are its advantages?
  - Reduces the need for large data sets
  - Can be used for a variety of prediction tasks

Literature Review

- Papers of Note for this Study
  - *A Concise Review of Transfer Learning* Farahani et al, 2021
  - *A Comprehensive Survey on Transfer Learning* Zhuang et al, 2020
- Domain Literature: Credit Default Prediction
  - Kaggle: American Express Credit Default Prediction

Data and Methodology

- Primary Data Set:
  - American Express Credit Default Prediction
  - 1,000,000 + observations!
- Methodology:
  - Data Cleaning and Feature Engineering
  - Feather and Parquet Files: Data Compression

Initial Results and Future Directions

- Feature Engineering and Data Cleaning:
  - 200,000 observations after cleaning
  - 225 features
- Initial Model:
  - Feed Forward Neural Network
- Baseline Model:
  - Logistic Regression
  - Random Forest

Proposed Model Architecture

```
class NeuralNetwork(nn.Module):
  def __init__(self):
      super().__init__()
      self.flatten = nn.Flatten()
      self.linear_relu_stack = nn.Sequential(
          nn.Linear(225, 24),
          nn.ReLU(),
          nn.Linear(24, 225),
          nn.ReLU(),
          nn.Linear(225, 1),
          nn.Sigmoid()
      )
```

Proposed Model Architecture - Feed Forward

```
def forward(self, x):
    x = self.flatten(x)
    logits = self.linear_relu_stack(x)
    return logits
```

Future Directions

- Next Steps:
  - Continue to refine the base NN model
  - Compare with baseline models
  - Dropout to prevent overfitting
  - Secondary Task Credit Default Prediction
- Future Ideas:
  - Predicting other types of Default with this pre-trained model
  - Comparing with other pre-trained models (e.g. Available models on HuggingFace)