TRAINING CNN&DNN MODELS FOR LOAN APPROVAL PREDICTION WITH MULTI MODAL INPUT AND EXPLAINABILITY

AUTHOR

Diya Sounoqrot - 2286340
AIN3002 : Deep Learning - Spring 2025
Bahcesehir University

INTRODUCTION

In this project, I trained two deep learning models a DNN and a CNN for loan approval prediction using multi-modal inputs. The input combines structured financial data and unstructured loan descriptions, where the text is processed through a frozen BERT encoder and then fused with the tabular features. I evaluated both models under identical conditions and applied SHAP explainability techniques to interpret their behavior.

MOTIVATION & CHALLENGES

CNNs are rarely applied to tabular data, and I was motivated to explore whether they could learn meaningful patterns from financial inputs. I also wanted to compare how both models perform under the same multi-modal pipeline.

In financial settings, models must be not only accurate but also explainable. Traditional systems often Ignore borrower provided text, limiting insight into real decision factors.

Key challenges included ensuring fair evaluation and handling imbalanced real-world data.

MY CONTRIBUTIONS

- DESIGNED TWO DISTINCT ARCHITECTURES (DNN, CNN) FOR MULTI-MODAL LOAN PREDICTION USING BERT.
- INTRODUCED CONV1D-BASED TABULAR PROCESSING IN THE CNN MODEL
- APPLIED SHAP FOR GLOBAL AND LOCAL INTERPRETABILITY
- COMPARED DNN VS CNN ON PERFORMANCE, COMPLEXITY, AND FEATURE ATTRIBUTION

METHODOLOGY

Dataset Used

I used a filtered subset of the Lending Club Loan Dataset (2015-2016) from Kaggle. After cleaning and feature selection, a balanced sample of 20,000 loans (10k approved, 10k rejected) was used for modeling.

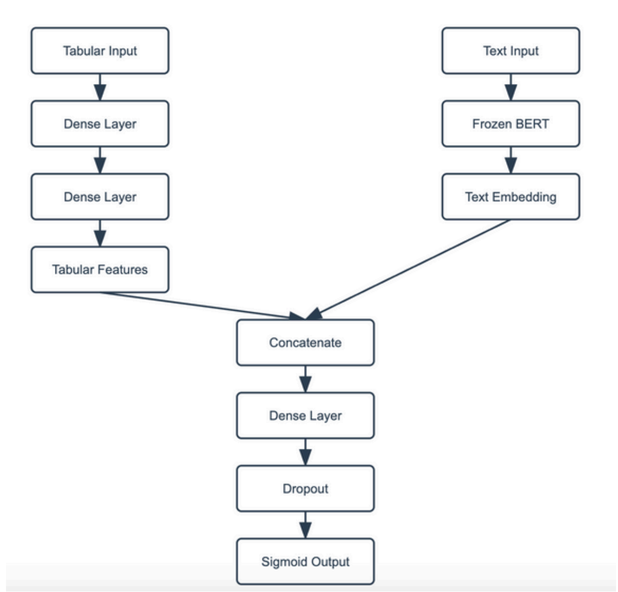
Multi-Modal Pipeline Overview

The model processes two input types:

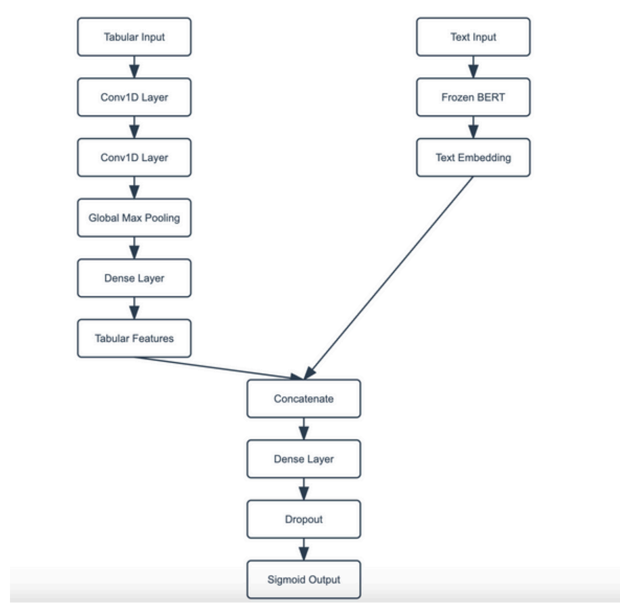
- Tabular data is preprocessed using StandardScaler (numerical) and OneHotEncoder (categorical).
- Text data (loan purpose + title) is tokenized and passed through a frozen BERT encoder.

The encoded tabular and text features are concatenated and fed into a shared dense layer for final binary classification.

DNN Model Architecture



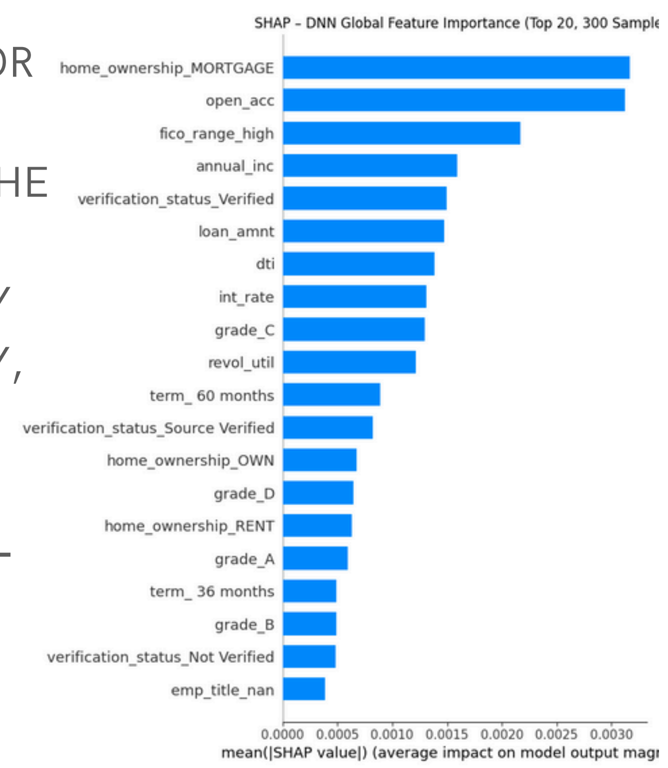
CNN Model Architecture

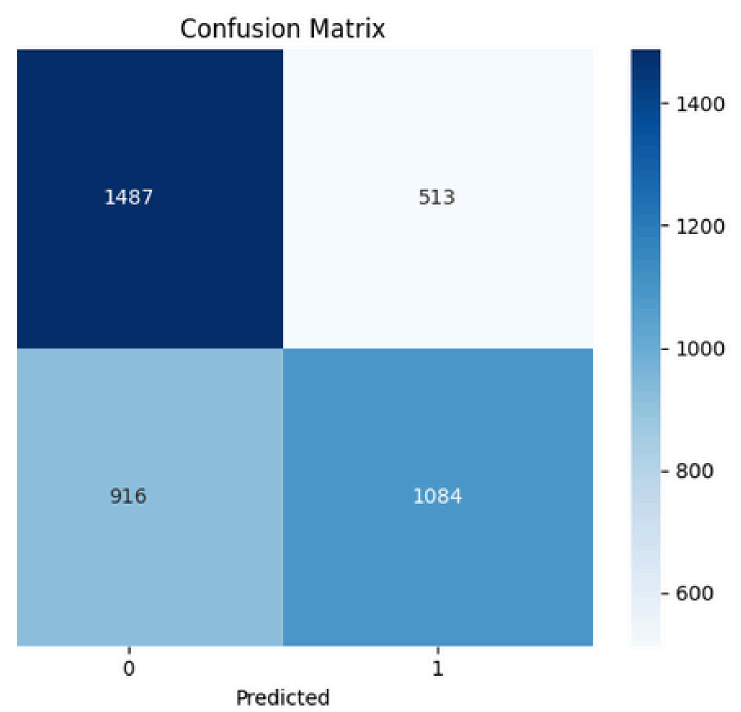


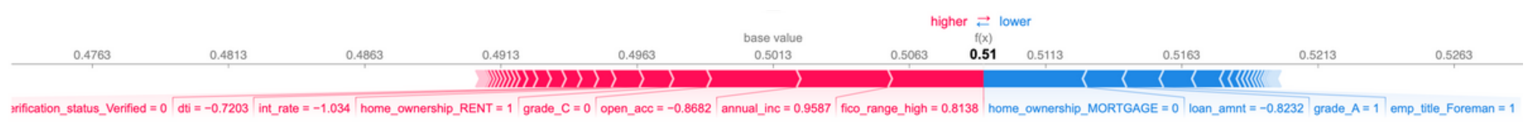
RESULTS & FINDINGS

I used SHAP summary plots for global feature importance and force plots for local prediction explanations.

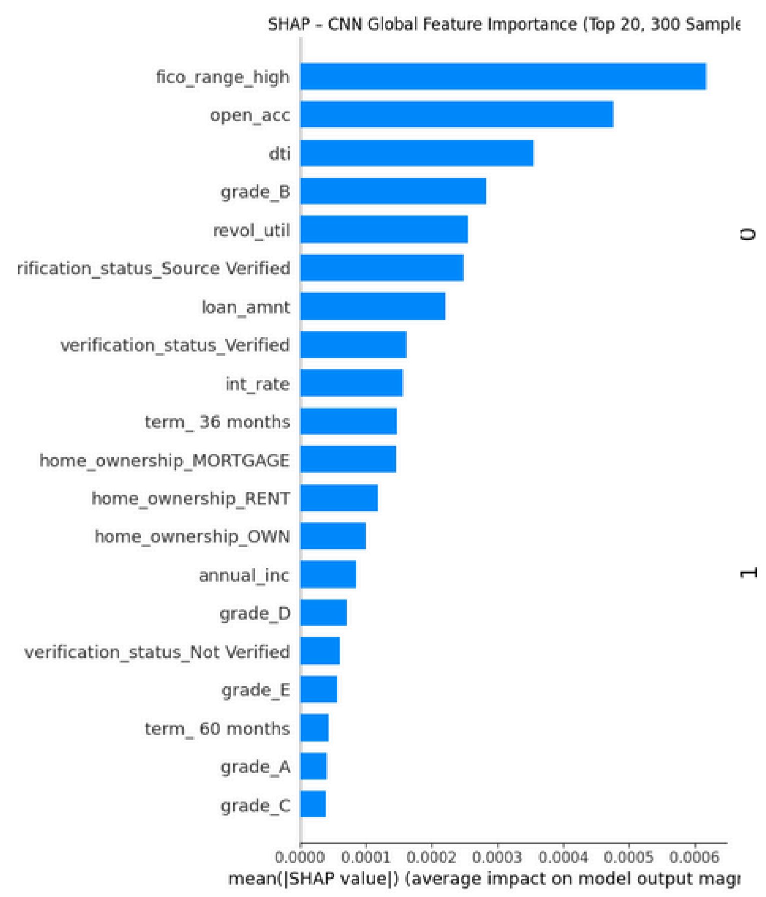
DNN Model

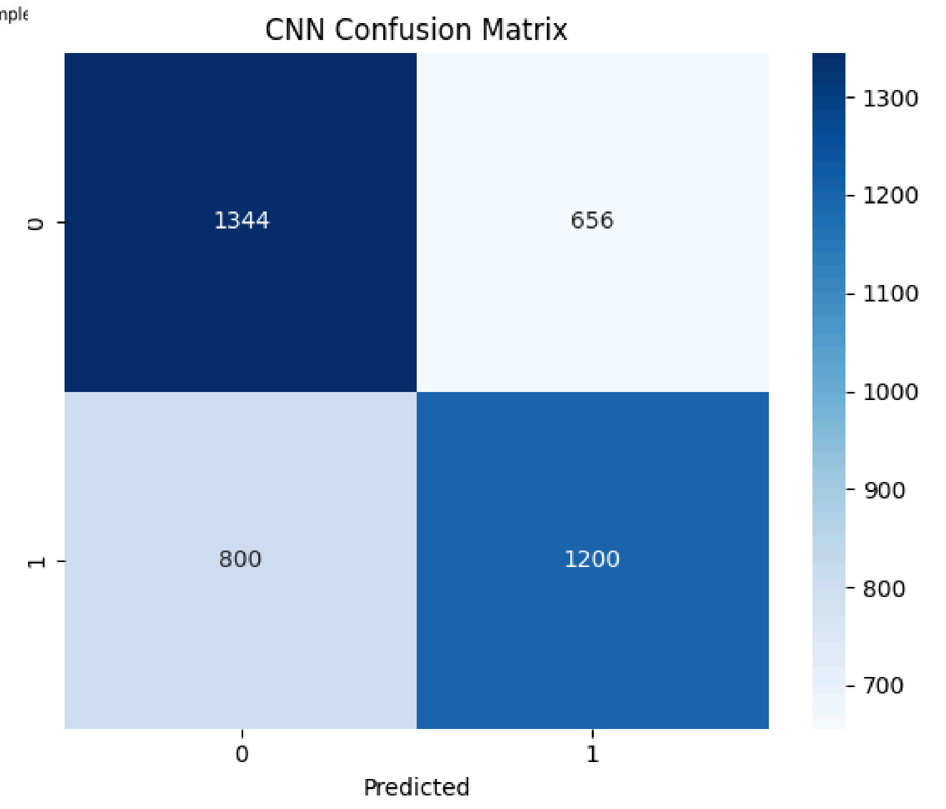


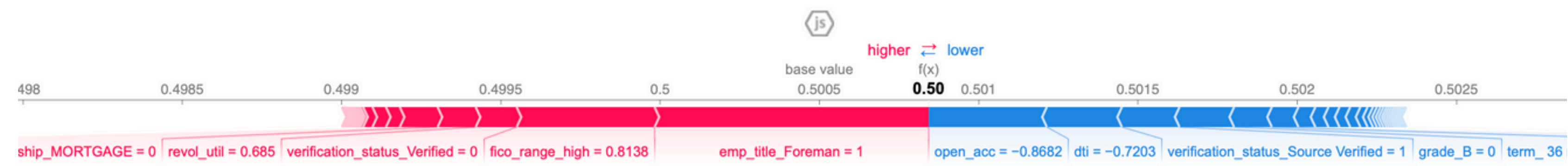




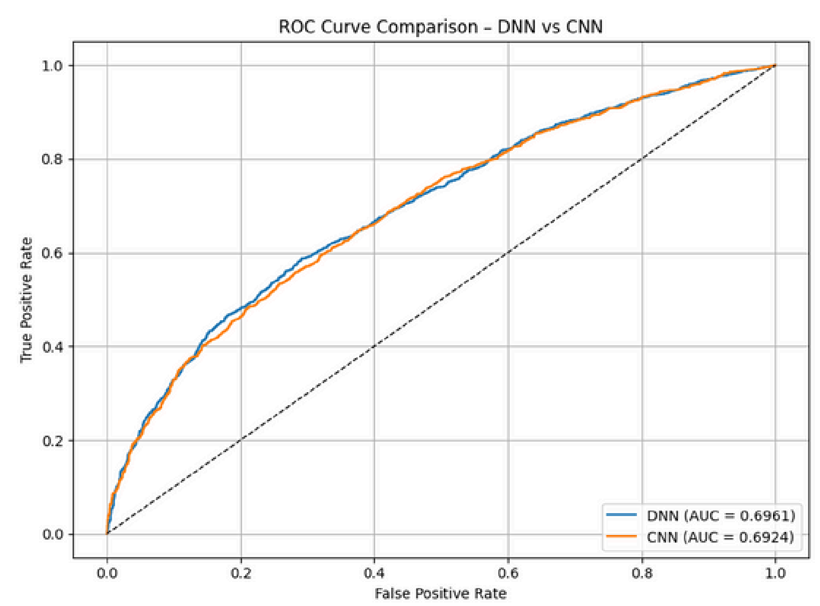
CNN Model







CONCLUSION



Metric	DNN	CNN
Accuracy	64%	64%
Precision	0.68	0.65
Recall	0.54	0.60
F1-Score	0.60	0.62
ROC AUC	0.696	0.692
Parameters	465K	58K

This project demonstrated that both DNN and CNN architectures can effectively predict loan approvals using multi-modal data that fuses structured financial features with free-text descriptions.

Despite having 8× fewer parameters, the CNN model achieved comparable accuracy and ROC-AUC to the DNN, while offering slightly higher recall and F1-score making it a strong candidate for deployment in recall-critical scenarios.

On the other hand, the DNN model produced more distributed feature attributions, which may be preferred when explainability and transparency are critical for stakeholders.