Credit Risk Prediction using Machine Learning and Deep Learning Models

Harshini J

School Of Computer Science and Engineering Vellore Institute Of Technology(VIT) Chennai **Chennai, India**
harshini.j2021@vitstudent.ac.in

Gerina Mary C

School Of Computer Science and Engineering Vellore Institute Of Technology(VIT) Chennai **Chennai, India**
gerinamary.c2021@vitstudent.ac.in

Abstract:

This study deals with the application of deep learning models in credit risk prediction. Credit risk assessment is a key task in the banking and finance industry when assessing a borrower's likelihood of defaulting on a loan. Traditional methods are often based on statistical techniques and machine learning algorithms. However, the increasing availability of data and the complexity of credit risk profiles require more sophisticated models. This paper explores the use of various deep learning architectures including GRU, LSTM, CNN and Multilayer Perceptron (MLP) for credit risk prediction. Models are trained and evaluated on a real dataset using Python libraries such as Keras, Scikit-learn and Pandas. Experimental results show the effectiveness of deep learning models in improving the accuracy of credit risk forecasting compared to traditional machine learning algorithms.

1. Introduction:

Credit risk forecasting is a critical task in the financial industry because it aims to estimate the probability of a borrower defaulting on a loan. Accurate risk assessment helps financial institutions reduce potential losses and make informed lending decisions. Traditional credit risk models are often based on statistical analysis and traditional machine learning algorithms. However, these models can struggle to capture complex patterns in credit risk data. Deep learning techniques offer a promising way to improve the accuracy of credit risk prediction by automatically learning complex patterns and relationships from large data sets.

2. Data preprocessing:

The dataset used in this study originates from a real credit risk dataset. The data goes through pre-processing steps such as handling missing values, coding categorical variables and scaling numerical functions. Duplicate records are removed and missing values are imputed using appropriate techniques. Categorical variables are coded with a label code and an ordinal code to convert them into numbers suitable for deep learning models.

3. Model development:

The study investigates four deep learning architectures for credit risk prediction: GRU, LSTM, CNN and Multilayer Perceptron (MLP). These architectures are implemented using Python's Keras library. Models are trained on a preprocessed dataset with appropriate hyperparameters and evaluated using standard metrics such as accuracy, F1 score, and confusion matrix.

4. Test results:

The experimental results obtained from the deep learning models—GRU, LSTM, CNN, and MLP—provide valuable insights into their performance for credit risk prediction.

Both the GRU and LSTM models achieved a high test accuracy of 91%, indicating their ability to make accurate predictions based on experimental data. In addition, both models showed relatively low test loss values, with GRU at 0.26 and LSTM at 0.27. These results suggest that both models effectively learned patterns in the data and made accurate predictions.

GRU and LSTM are both recurrent neural networks (RNNs) designed to capture sequential dependencies in data. They are suitable for tasks where the order of the data is important, so they are suitable for time series data, such as credit risk forecasting. The similar performance of both models suggests that they are equally effective for this task.

The CNN model achieved a lower test accuracy of 78% compared to the GRU and LSTM models. However, its test loss value is comparable to the LSTM model at 0.52.

CNNs are typically used for tasks involving spatial information, such as image recognition, but can also be applied to sequential data such as time series. By predicting credit risk, CNNs could capture spatial patterns in the data that other models could not, affecting their performance.

The MLP model achieved the same high test accuracy as GRU and LSTM, 91%. However, its sampling loss is significantly higher at 3.3, indicating that the model had difficulty generalizing to unseen data. A large test loss indicates that the MLP model may have overfitted the training data, resulting in poor performance on the test set.

MLPs are versatile neural networks commonly used for various machine learning tasks. However, they may not be as efficient as specialized architectures such as RNNs and CNNs for tasks involving sequential or spatial information. The large test loss of the MLP model in this context shows its limitations in capturing complex patterns in credit risk data.

In conclusion, although RNN models (GRU and LSTM) and CNN model achieved similar accuracy, they showed different robustness in terms of test loss. The CNN model showed competitive performance despite its lower accuracy.front. In contrast, the MLP model found it difficult to generalize and perform this task well, indicating its limitations in dealing with sequential data. These findings highlight the importance of choosing the right model architecture for specific tasks and highlight the potential benefits of using deep learning techniques for credit risk prediction.

Model	**Test Accuracy**	**Test Loss**
GRU	0.91	0.26
LSTM	0.91	0.27
CNN	0.78	0.52
MLP	0.91	3.3

Table 1 talks about the test accuracy and test loss of all 4 models

5. Discussion:

The discussion section analyzes the strengths and limitations of deep learning models for credit risk forecasting. Although deep learning models offer excellent results in capturing complex patterns, they can suffer from overfitting, especially when data is limited. Strategies such as stop tuning and hyperparameter tuning can help reduce overfitting and improve model generalization. In addition, the interpretability of deep learning models remains a challenge that can be answered with model interpretation techniques and feature importance analysis.

6. Conclusion:

In conclusion, this study demonstrates the effectiveness of deep learning models to improve the accuracy of credit risk forecasting. GRU, LSTM, CNN, and MLP architectures show promising results, outperforming traditional machine learning algorithms in terms of predictive performance. Future research may focus on further optimizing deep learning models, exploring ensemble techniques, and incorporating industry knowledge to improve credit risk assessment in real-world scenarios.