Part a

Group 46

1. Satya Prakash Pandit - 2022AC05040
2. Pushkar Kumar Verma – 2022AC05272
3. G Krishna Sameera – 2022AC05407

Part A: Literature Exploration and Comparison

Comparison Table

Group 46 Satya Prakash Pandit - 2022AC05040 Pushkar Kumar Verma – 2022AC05272 G Krishna Sameera – 2022AC05407			
Domain: Credit Risk (Banking)			
	PAPER 1 – 10171259	**PAPER 2- 9242905**	**PAPER 3 – 10229794**
Title of the paper	Proposal of a model for credit risk prediction based on deep learning methods and SMOTE techniques for imbalanced dataset	Transparent Decision Support System for Credit Risk Evaluation: An automated credit approval system	Research on Bank Credit Risk Assessment Based on BP Neural Network
Authors	Adaleta Gicić - Engineering Faculty, University of Sarajevo, Bosnia and Herzegovina Dženana Đonko - Engineering Faculty, University of Sarajevo, Bosnia and Herzegovina	Abhinaba Dattachaudhuri - NIT, Silchar, India Saroj Biswas - NIT, Silchar, India Sunita Sarkar - Assam University, Silchar, India Arpita Nath Boruah - NIT, Silchar, India	Qinghua Li Shanghai Donghai Vocational&Technical College, Shanghai, China
Year of publication	2023	2020	2023
Architecture of Deep Learning (including the number of layers, types of	The architecture of the deep learning models discussed in the paper involves Stacked LSTM and Stacked BiLSTM networks: **Number of Layers:** Both Stacked LSTM and Stacked BiLSTM architectures consist of three hidden layers.	The architecture utilized in the TDSSCRE model is specified as: **Neural Network Structure:** It employs a neural network with a single hidden layer. **Number of Layers**: One hidden layer and an output layer.	The architecture described in the provided research paper outlines a multi-level deep neural network for credit risk assessment. Here's a breakdown of the architecture based on the information given: **BP Neural Network (BPNN):** **Layers:** It consists of an input layer, hidden layer(s), and an output layer.

layers, activation functions, and any unique features)	**Types of Layers:** The models include LSTM (Long Short-Term Memory) layers, BiLSTM (Bidirectional LSTM) layers, and Dense layers for classification. **Activation Functions:** ReLU (Rectified Linear Unit) activation function is used in the Dense layers for processing weighted inputs and producing the output. Sigmoid activation functions are employed within the LSTM units for gating mechanisms. **Unique Features:** The models incorporate bidirectional temporal dependencies from historical data via BiLSTM layers. Stacked architectures feed the output of one hidden layer as input to the subsequent layer, enhancing the network's ability to capture sequential information and context. Additionally, the use of SMOTE technique for handling imbalanced data is a unique feature incorporated into the model architecture to address class imbalance in the dataset.	**Types of Layers**: The hidden layer utilizes neurons with activation functions, while the output layer is likely structured for classification. **Activation Functions**: The specific activation functions used in this context are not explicitly mentioned in the provided details. However, common activation functions for neural networks include ReLU (Rectified Linear Unit), Sigmoid, or Tanh for hidden layers, while the output layer might use a different activation depending on the classification task (e.g., softmax for multi-class classification). **Unique Features:** The uniqueness lies in the method of rule extraction from the neural network. The system focuses on generating concise and explainable rules from the neural network, which can justify the decisions made in credit risk assessment.	**Types of Layers:** Input layer, hidden layers, and output layer. **Activation Functions:** Although not explicitly mentioned in the provided excerpt, common activation functions used in BPNNs include Sigmoid, Tanh, ReLU, etc. The specific activation functions employed in this BPNN are not explicitly stated in the text. **Unique Features:** The network employs a multi-level structure, divided into **H-Net and L-Net:** **H-Net (High-Level Network)**:Utilizes secondary features of the loan applicant. Consists of multiple hidden layers.Specific numbers mentioned: 3 hidden layers with neurons - 32, 32, 16. **L-Net (Low-Level Network)**:Utilizes tertiary features of the loan applicant.Also consists of multiple hidden layers.Specific numbers mentioned: 6 hidden layers with neurons - 64, 48, 32, 32, 16, 8. **Training Process:** The training involves forward propagation and backpropagation of errors to adjust the network weights iteratively. **Feature Normalization:** Normalizes input data to a range of [0-1] before feeding it into the neural network. This preprocessing step ensures uniformity in the input data, improving the efficiency of the deep neural network. **Experimentation and Verification:** The architecture is experimentally verified using real credit data from a commercial bank over the past five years. The comparison includes the accuracy and sensitivity scores of the proposed multi-level DNN against other methods like random forest and decision tree algorithms. The provided paper gives an overview of the structure and working principle of the BP Neural Network used for credit risk assessment. However, specific details about the activation functions, loss functions, or additional techniques (besides the general BPNN structure)is not explicitly maintained in the excerpt provided.
How is the network helping the overall task? eg: feature engg or classification or regression or all	Here's how the network contributes to the task: **Feature Extraction:** The LSTM-based architectures, especially Stacked LSTM and BiLSTM, are effective in learning and extracting temporal patterns and dependencies from sequential credit data. They automatically capture and process relevant features from the input sequences, such as customer demographics, application forms, and transactional histories, without the need for explicit feature engineering. **Classification**: The primary focus of these models is on credit scoring, a classification problem that determines whether an applicant is likely to belong to a "good" or "bad" class. The LSTM and BiLSTM networks, along with Dense layers, are utilized to classify credit applicants	The network in the Transparent Decision Support System for Credit Risk Evaluation (TDSSCRE) primarily aids in classification for credit risk assessment. **Rule Extraction:** The neural network is used as a basis for rule extraction. Instead of directly focusing on predictive accuracy, the network is trained and then pruned to extract concise and understandable rules. These rules aim to explain and justify the decisions made in the credit risk assessment process. **Decision Support:** By deriving rules from the network, the system can offer transparent explanations for why credit applications are accepted or rejected. This transparency assists financial organizations in justifying their decisions, ultimately improving customer relations and trust.	The network described in the research paper serves the overall task of credit risk assessment. Here's how it aids in this task: **Feature Engineering:** • The network employs a multi-level structure, considering both secondary and tertiary features of loan applicants. • Features are preprocessed and normalized before being fed into the network, enhancing the network's ability to understand and utilize the input data effectively. • Analytic Hierarchy Process (AHP) is used to identify features with significant impact factors, which are then input into the deep neural network. This process aids in feature selection and engineering, allowing the model to focus on essential aspects for credit risk assessment. **Classification:**

	based on their historical data, attributing a risk label (default or non-default) to each applicant. **Handling Imbalanced Data:** The inclusion of the SMOTE technique within the model architecture addresses the issue of imbalanced datasets commonly encountered in credit scoring. By oversampling the minority class (e.g., "bad" credit applicants), SMOTE helps to balance the dataset, enabling the model to better learn from both classes and improve classification performance. In summary, the network's LSTM-based architectures aid in feature extraction from sequential credit data and contribute significantly to the classification task of credit risk assessment, especially when dealing with imbalanced datasets.	**Classification:** The extracted rules form the basis of a decision support system, which classifies credit applications into categories like 'approved' or 'rejected'. This classification is based on the rules generated from the neural network, enabling an explainable decision-making process. While the primary focus is on classification and decision support, the network indirectly contributes to feature engineering by automatically learning and deriving meaningful patterns and representations from the input data. However, the main emphasis of this system lies in the extraction and utilization of comprehensible rules for credit risk assessment.	• The primary goal of this network is credit risk assessment, which typically involves a classification task. It aims to predict whether a borrower is a high or low credit risk. • The H-Net and L-Net structures categorize applicants based on their risk scores. • The final output from the network assists in making decisions about granting or denying credit based on the assessed risk level. **Regression (Indirectly):** While the primary task is classification (high or low risk), the network might involve a regression-like approach within its layers. It doesn't directly predict a continuous output (like an exact default probability), but the multi-layer structure and processing might involve regression-like operations to map input features to risk scores. In summary, the network is designed to process and analyze input features to assess the credit risk of loan applicants. It performs feature engineering, classification, and indirectly utilizes regression-like operations within its structure to achieve this goal. The model aims to provide accurate risk scores for effective decision-making in credit management.
Training procedures (e.g, training strategy, including optimization algorithms, learning rates, batch sizes, and regularization techniques)	The training procedures used in the paper's deep learning models involve the following key aspects: **Optimization Algorithms:** The models utilize default optimization algorithms provided by Keras, likely variants of stochastic gradient descent (SGD) or adaptive methods like Adam or RMSprop. **Learning Rates:** The learning rates and other hyperparameters are set to default values for the optimization algorithms employed in the models. No specific values or tuning details are provided in the information shared. **Batch Sizes:** The batch size for training is set to 32, determining the number of input-output pairs passed through the network before updating the internal model parameters. **Regularization Techniques:** The models incorporate dropout regularization with a rate of 0.2, a technique aimed at preventing overfitting by randomly dropping a proportion of neurons during training. Overall, the training strategy in the paper appears to follow default settings for optimization algorithms and hyperparameters, including batch sizes and regularization techniques, with limited details on specific values or tuning procedures for these parameters.	The training procedures used in the Transparent Decision Support System for Credit Risk Evaluation (TDSSCRE) involve several steps and strategies: **1.Neural Network Training:** **Architecture Selection:** The system employs a backpropagation neural network with a single hidden layer and h hidden neurons. The optimal number of hidden neurons is determined based on mean square error, varying from l+1 to 2*l hidden neurons. **Training Algorithm:** Backpropagation is used for training the neural network, a common algorithm for updating network weights to minimize error. **2.Network Pruning:** The neural network undergoes pruning to identify insignificant input neurons. It iteratively removes inputs causing the least increase in network error, eventually eliminating these insignificant inputs. Sequential floating search is applied for network pruning. **3 Data Range Computation:** Data range for each remaining neuron is computed using both misclassified and properly classified patterns. It involves finding misclassified rates and properly classified rates for each input neuron to construct the data ranges. **4.Rule Generation:** Based on the computed data ranges, mandatory attribute ranges are extracted to construct classification rules. Rules are generated in the form of IF-THEN propositional clauses **5. Rule Pruning and Updation:** The generated rules undergo a pruning phase to remove insignificant ones, enhancing the overall accuracy.	The provided excerpt doesn't delve into exhaustive details about the training procedures of the neural network. However, it does offer some insights into the training process and optimization methods: **1. Optimization Algorithms:** **Backpropagation:** The paper mentions the backpropagation process used for training the neural network. This involves forward propagation of signals and backward propagation of errors to adjust the network weights iteratively. **Momentum Gradient Descent:** Specifically, the paper references the application of Bayesian regularization and momentum gradient descent methods to improve accuracy. These are optimization techniques that modify the standard gradient descent algorithm by incorporating momentum and regularization terms, respectively. **2. Learning Rates:** Unfortunately, the exact learning rates used in the training process are not explicitly mentioned in the provided text. Learning rates control the step size during weight updates in the training process. **3.Batch Sizes:**The paper doesn’t explicitly state the batch sizes used for training. Batch sizes determine the number of samples propagated through the network before updating the weights. **4.Regularization Techniques:** Bayesian Regularization: The paper mentions the use of Bayesian regularization as one of the techniques employed to enhance the accuracy of the model. Bayesian regularization adds penalty terms to the loss function to prevent overfitting by penalizing large weights.

		Additionally, rules are updated to overcome inherent overlaps. Attribute values within the overlap range are shifted to the class with the highest probability, reducing ambiguity. Regarding specific hyperparameters: Optimization Algorithm: The specifics of the optimization algorithm (like variations of gradient descent) aren't explicitly mentioned. Learning Rate: The learning rate utilized during backpropagation isn't specified. Batch Sizes: Batch sizes used in training are not explicitly defined. Regularization Techniques: No specific regularization techniques like dropout, L1/L2 regularization are mentioned in this description. Overall, while the general training strategy is outlined, specific hyperparameter details and optimization techniques seem to be missing or not specified in the provided text.	Overall, while the paper touches upon the optimization algorithms and regularization techniques used in training (such as backpropagation, Bayesian regularization, and momentum gradient descent), it doesn't provide specific details about learning rates or batch sizes employed during the training process. These elements are crucial in fine-tuning the training process and optimizing neural network performance.
Evaluation / Performance metric used	The paper utilizes various standard performance metrics to evaluate the effectiveness of the proposed deep learning models for credit risk assessment. These metrics are commonly employed in classification tasks and include: **Accuracy**: The proportion of correctly classified instances (both "good" and "bad" credit applicants) over the total number of instances. **Precision**: The ratio of correctly predicted "bad" credit applicants to the total predicted "bad" credit applicants. It assesses the accuracy of positive predictions. **Recall** (Sensitivity): The ratio of correctly predicted "bad" credit applicants to the actual number of "bad" credit applicants. It measures the model's ability to identify all relevant instances. **F1 Score**: The harmonic mean of precision and recall. It provides a balanced measure between precision and recall. **Additionally**, the paper mentions other evaluation metrics tailored for imbalanced datasets: **H-measure:** A metric that balances precision and recall to handle class imbalance effectively. **Brier Score:** Measures the accuracy of probabilistic predictions. **G-mean**: Geometric mean of sensitivity (recall) and specificity, providing a single metric to evaluate a model's performance across both classes in an imbalanced dataset.	The evaluation and performance metrics used in the Transparent Decision Support System for Credit Risk Evaluation (TDSSCRE) are: **10-Fold Cross-Validation (CV) Accuracy**:The system employs 10-fold cross-validation to assess the accuracy of the model. It divides the dataset into ten parts, trains the model on nine parts, and evaluates it on the remaining one. This process is repeated ten times, and the average accuracy is computed. **Average Number of Antecedents per Rule**:This metric measures the average number of conditions (antecedents) per rule generated by the system. It reflects the complexity or simplicity of the rules created. **Recall**:The recall metric measures the proportion of actual positive instances that are correctly identified by the system. It's calculated as the ratio of true positives to the sum of true positives and false negatives. **False Positive (FP) Rate:**The FP rate is the ratio of false positive predictions to the sum of false positives and true negatives. It measures the proportion of negative instances that were incorrectly classified as positive. These metrics are computed for each dataset used in the experiments (Australian Credit, German Credit, Bank Marketing). The performance of the TDSSCRE model is compared against existing pedagogical rule extraction algorithms (RxREN and RxNCM) using these metrics to evaluate its effectiveness in credit risk evaluation	This document excerpt doesn't explicitly mention the specific evaluation metrics used to assess the performance of the credit risk assessment model. However, in the field of credit risk assessment and classification tasks in general, several common evaluation metrics are typically utilized. These might include: **Accuracy:** Measures the overall correctness of predictions, i.e., the ratio of correctly predicted instances to the total number of instances. **Precision and Recall**: Especially relevant in imbalanced datasets, precision measures the proportion of correctly identified positive cases out of all cases identified as positive, while recall measures the proportion of correctly identified positive cases out of all actual positive cases. **F1-Score:** Harmonic mean of precision and recall, often used as a balanced measure when dealing with imbalanced classes. **ROC Curve and AUC:** Receiver Operating Characteristic (ROC) curve visualizes the trade-off between true positive rate and false positive rate across different thresholds. Area Under the ROC Curve (AUC) summarizes the ROC curve's performance, providing an aggregate measure of model performance. **Confusion Matrix:** A table showing the true positives, true negatives, false positives, and false negatives, providing deeper insights into the model's performance. This dataset contains credit information with labeled instances for risk assessment and is widely used in academic research. Unfortunately, the specific metrics used in the research paper aren't explicitly mentioned in the provided excerpt. However, these standard evaluation metrics are commonly employed in similar classification tasks within the realm of credit risk assessment.
Name of Dataset	The dataset used in the research paper for credit risk assessment is the "German Credit" dataset, a well-known	The research paper mentioned the use of three publicly available datasets for credit-risk evaluation:	The provided excerpt from the research paper doesn’t specify the name or source of the dataset used for the credit risk assessment. Often, academic research papers might utilize proprietary or private datasets,

used. If a public dataset, provide the URL.	public dataset frequently employed in credit scoring research. This dataset is available from the UCI Machine Learning Repository, and it contains credit information about customers, specifically applicants' personal and financial attributes. It comprises categorical and numerical features used to predict creditworthiness, categorized as "good" or "bad" credit risks. Dataset: https://archive.ics.uci.edu/dataset/144/statlog+german+credit+data	Australian Credit Dataset https://archive.ics.uci.edu/dataset/143/statlog+australian+credit+approval German Credit Dataset https://archive.ics.uci.edu/dataset/144/statlog+german+credit+data Bank Marketing Dataset https://archive.ics.uci.edu/dataset/222/bank+marketing These datasets are commonly used in the field of credit risk analysis and are available in the UCI Machine Learning Repository.	especially in domains like banking or finance, where sensitive information is involved. Unfortunately, without explicit details within the text, it's challenging to determine the exact dataset used for this specific study. The researchers might have used an internal or proprietary dataset from a commercial bank that couldn't be publicly accessed. However, in the domain of credit risk assessment, some public datasets are commonly used for research purposes, such as the German Credit dataset from the UCI Machine Learning Repository (https://archive.ics.uci.edu/ml/datasets/Statlog+%28German+Credit+Data%29).

Conclusion: In summary, these papers demonstrate how deep learning, through various architectures like LSTM, BiLSTM, and BP Neural Networks, significantly contributes to credit risk assessment. They leverage sophisticated models to automate predictions, handle imbalanced data, and extract meaningful rules for transparent decision-making. While some emphasize classification accuracy, others prioritize explainability, showcasing the diverse applications and potential of deep learning in effectively managing credit risk within the banking domain.

Looking across these papers, it's evident that deep learning models, especially variations like LSTM, BiLSTM, and BP neural networks, are prominently featured in credit risk assessment. Each paper brings a unique perspective to the table:

Paper 1:

1. Utilizes Stacked LSTM and BiLSTM networks, incorporating SMOTE for imbalanced data.
2. Focuses on feature extraction and classification, effectively handling imbalanced datasets.
3. Leverages deep learning to automate credit risk prediction.

Paper 2:

1. Proposes a Transparent Decision Support System using a neural network.
2. Emphasizes rule extraction for transparent credit risk assessment.
3. Prioritizes explainability and decision support.

Paper 3:

1. Employs a BP Neural Network with H-Net and L-Net structures.
2. Emphasizes multi-level feature analysis, classification, and regression.
3. Incorporates normalization and real credit data for validation.