UNIVERSITÄT BREMEN

FACULTY 3

ADVANCED METHODS OF AI

Bank Statement Classification

Proposal for an AI System for Bank Statement Classification

Submission date:	***23.02.2025***
Sören Töpper	4525913
Ann-Kathrin Mielebacher	2582519
Ashlesh Patil	6351941

Link to GitHub Repository: `https://github.com/soerentoepper/AM_of_AI_group_M`

Contents

1 Problem Description

One of the most common problems faced by business consultancies is the incomplete accounting records of their clients. The profit and loss statement (P&L) is a common method for structuring financial bookings. It provides a comprehensive overview of income, expenses, and the resulting profit or loss of a company over a specific period and includes categories such as revenue (or sales), cost of goods sold (or cost of sales), selling &administrative (SG&A) expenses, marketing and advertising, technology/research &development, interest expense, taxes and net income. A precise and efficient allocation of bank transactions to these categories is of enormous importance for this purpose (Tim Vipond 2022).
To automate and optimize this process, an AI system is introduced that addresses these challenges by automating the categorization process, thereby improving the accuracy and efficiency of financial reporting for clients.

2 Challenges

There are many challenges in developing an AI system. These are some of the major challenges:

- The sensitive nature of financial data requires strict security precautions.
- The descriptions of the transactions are often ambiguous and vary greatly.
- Bank statements come in different formats and have different structures.
- The system must handle large amounts of data and be suitable for companies of different sizes.
- The system must be context sensitive to guarantee an optimized category fitting.
- The system has to be flexible to integrate new categories or business models.

3 Related Work

Several studies have addressed similar problems:
(Cucchiara et al. 2023) developed a transformer-based model specifically designed for bank transactions, which significantly outperformed other deep learning models. They highlighted the impact of AI in finance across four key areas: client data (enhancing intermediation and customer engagement), fintech documents (enabling intelligent analysis for financial workflows), finance data (facilitating trend prediction for trading, risk, and asset management) and transactional data (supporting the classification and generation of time-series product data).
Additionally, they addressed the challenges associated with working on bank transaction data, including limited access to real datasets due to data protection, the diverse formats of available

features (numerical, categorical and textual), and variations in data structure across different banks.

In a separate work (Orestav et al. 2024) compared various neural network architectures for categorization of bank transactions. They used a dual branch architecture to include both textual and numerical inputs. For the textual inputs character-level embedding was crucial to overcome misspellings in the transaction detail. Furthermore the use of class weights in the training process helped to deal with imbalanced classes in the dataset. The highest F1-score across all models was achieved by the LSTM and CNN architectures.

4 Technology Selection

Related work shows that neural networks perform well in the classification of bank transactions. Since the classification does not depend on temporal information, a CNN architecture is chosen. The textual inputs will be encoded at the character level.

5 Justification of Technology Choice

The choice of a CNN for classifying bank transactions is based on key advantages. CNNs effectively capture local patterns in text, making them well-suited for transaction descriptions, which often contain category-relevant keywords. Since transaction classification is not time-dependent, RNN-based models are unnecessary. CNNs treat each transaction as an independent input, which is a more fitting approach for this classification. Transaction descriptions vary widely, often including abbreviations and misspellings. Character-level embedding helps capture these variations. CNNs have the ability to continuously learn and adapt to new categories over time, making them more flexible compared to traditional classifiers.

6 System Description

The goal of the system is to classify all the incoming transactions correctly and store the information. The sum of all the classes is visualized in a P&L table. To pursue this goal, the systems needs these core cognitive abilities:

Perception The system must be able to process new transactions. Information about new transactions is transmitted via the bank statements in .csv format.

Attention The bank statements contain a lot of information that is not relevant for the classification. For example the current balance of the account results from the transaction but has no impact on the transaction. Therefore, the system must extract the information that is important for classification.

Action selection The actions that the system can perform are limited but it can select between some actions. If the output of the CNN does not clearly indicate in which category the transaction belongs, it can decide to use further methods to classify the transaction.

Memory The system needs to store information in order to classify transactions. It requires both short-term and long-term memory. The CNN model is part of the procedural (long-term) memory.
Before being stored in long-term memory, the new transactions are first kept in working memory. After that, they are added to the already classified transactions in episodic memory.

Learning In the beginning the system learns by associating input patterns with corresponding labels, making this a form of associative learning. During operation, the system continuously learns and improves its classification abilities. This process is known as perceptual learning, as the system enhances its ability to recognize patterns over time and to add new categories to the P&L statement.

Reasoning For each new transaction, the CNN outputs probabilities for each category. Based on these probabilities, the system must infer which category the transaction belongs to. In addition to the probabilities, logical reasoning can assist the classifier. For example, a positive amount in the new transaction should not be classified as an expense. Therefore, the system usesuses a combination of probabilistic and logical inferences to make a final classification.

Meta Reasoning The system is able to monitor its own decision-making process. When classifying transactions the system does not purely rely on the outputs of the CNN but also evaluates its current performance to ensure accuracy and adapt its strategy if needed.

Prospection The system is able to anticipate future data trends and adjust its classification strategy proactively. While the CNN processes current transactions, prospection uses historical data and emerging signals to predict shifts in transaction patterns. For example, if a rising trend in a specific category is detected, the system can adjust thresholds or trigger extra verification steps.

7 Implementation in the Prototype

The prototype uses a synthetic dataset, as no suitable real datasets are available. It was designed to resemble a realistic bank statement. The categories in the dataset are balanced to ensure the classifier learns them equally.
Only the features from the dataset that are relevant for the classification are passed to the CNN. In the prototype he classification is only based on the outputs of the CNN. After classifying, the amount of the transaction is added to the corresponding category and then stored in the

database.
The results of the classification are visualized in a table. For comparism a second table with the actual categories of the data is added. A new transaction can be added to the P&L statement with a button.

8 System Limitations

Data Quality and Real World Diversity

- Synthetic vs. Real Data: Our prototype relies on synthesized bank statements, which does not capture the full or real variability and noise of real banking data. In an actual production environment the diverse transaction descriptions and new merchant types and irregular edge cases for example partially corrupted statements, unusual currencies shall degrade accuracy.

Integration with Existing Systems

- Downstream Dependencies: Once transactions are categorized the finance and accounting teams may rely on these labels for subsequent processes like tax calculations. Any Inconsistencies or incorrect classifications can propagate errors downstream.

Maintenance and Extensibility

- Adapting to New Categories: If a company introduces new transactions to the business or categories for example cryptographic transaction fees, the current CNN would require retraining or at least finetuning.

Interpretability and Explainability

- BlackBox CharacterLevel CNN: The implementation uses a CNN with characterlevel embeddings. While effective at capturing local text patterns, it can be required to business users and auditors who expect information for a classification. Posthoc explainers might be necessary in a regulated finance environment.

Broader Cognitive Abilities

- Reasoning and Context: Raw text classification might not just be enough as real financial data might require contextaware reasoning like linking transaction patterns over months or analyzing additional textual references. A purely CNN based approach cannot provide the broader cognitive architecture features such as metareasoning and prospective decisionmaking.

- Uncertainty Handling: The current classifier provides a “best guess” label with no explicit representation of uncertainty. For highstakes financial decisions, having probabilistic confidence scores or uncertainty estimates can be very crtical.

References

Cucchiara, Rita et al. (2023). “Deep Learning and Large Scale Models for Bank Transactions”. In: *Ital-IA 2023 Thematic Workshops*, pp. 512–516. URL: https://ceur-ws.org/Vol-3486/132.pdf.

Orestav, Filip and Kolumbus Lindh (2024). “Deep Learning Techniques for Bank Transaction Categorization: A collaborative study with Kreditz”. PhD thesis. Stockholm: KTH Royal Institute of Technology.

Tim Vipond (2022). *Profit and Loss Statement (P&L): A summary of income and expenditures for a business*. URL: https://corporatefinanceinstitute.com/resources/accounting/profit-and-loss-statement-pl/.