The American Journal of Management and Economics Innovations
ISSN 2693-0811 | Open Access

TYPE Original Research
PAGE NO. 138-144
DOI 10.37547/tajmei/Volume07Issue08-12

Integrating Machine Learning into Automated Accounting Transaction Classification: Architecture, Algorithms, and Performance Evaluation

Disha Patel

Senior Accounts Manager New York, USA

OPEN ACCESS

SUBMITTED 02 August 2025
ACCEPTED 08 August 2025
PUBLISHED 21 August 2025
VOLUME Vol.07 Issue 08 2025

CITATION

Disha Patel. (2025). Integrating Machine Learning into Automated Accounting Transaction Classification: Architecture, Algorithms, and Performance Evaluation. The American Journal of Management and Economics Innovations, 7(8), 138–144. https://doi.org/10.37547/tajmei/Volume07Issue08-12

Abstract: This article conducts a comparative analysis of the efficiency of various machine learning algorithms in addressing the task of classifying accounting transactions — a component ensuring the accuracy of financial reporting and enhancing operational efficiency. The aim of this study is to analyze different machine learning algorithms for the task of automated classification of accounting entries. The methodological basis of the research includes an extensive review of specialized literature, where the architectures of models such as logistic regression, support vector machine (SVM), random forest and gradient boosting are analyzed, as well as promising neural network solutions employing natural language processing (NLP) technologies. As a result of the experiment, a comparative analysis is presented according to key metrics (accuracy, recall, F1-score) and a hybrid architecture is proposed, combining an NLP module based on the BERT model and a gradient boosting classifier, which demonstrates the best results when processing transactions with complex textual descriptions. The scientific novelty of the work lies in the description of a conceptual model for selecting the optimal algorithm depending on the characteristics of the original data set and in substantiating the advantages of the proposed hybrid architecture, which integrates natural language processing methods for extracting semantic features and ensemble algorithms for final classification. In conclusion it is emphasized that

the implementation of intelligent classification automation not only minimizes the influence of the human factor but also transforms the role of the accountant from a data entry operator into a strategic analyst. The obtained data are of interest to researchers in financial engineering and artificial intelligence, practicing accountants and auditors, as well as developers of software products for the automation of financial flow management.

Keywords: machine learning, classification of transactions, accounting, artificial intelligence, natural language processing, automation, gradient boosting, deep learning, financial technologies, categorical data.

Introduction

In the context of the rapid expansion of digital information volumes and the profound digitalization of corporate processes, the profession of the accountant is undergoing a qualitative transformation. Traditional accounting methods, relying on manual data entry and processing, are losing effectiveness and are associated with a high likelihood of errors caused by the human factor. According to a Gartner analysts' forecast, by 2026 more than 80% of enterprises will use APIs and GenAI models and/or deploy GenAI-powered applications in production environments, whereas at the beginning of 2023 this figure stood at less than 5%. Enterprises employing an AI TRiSM control system will improve decision-making accuracy by eliminating up to 80% of erroneous and unreliable information [1]. This dynamic underscores the need for research aimed at integrating intelligent systems into accounting practice.

One of the most resource-intensive and critical operations in accounting is the classification of accounting entries — the process of assigning each financial transaction to the corresponding accounts and analytical categories. The correctness of this procedure directly influences the reliability of financial reporting, the accuracy of tax calculations, and the quality of managerial analysis, making it a priority area for automation [2, 12].

Contemporary research in machine learning offers a variety of algorithmic solutions for automating transaction classification; however, most publications are limited to an in-depth examination of one or two methods without a comprehensive comparative analysis. The absence of recommendations for selecting the optimal model in light of the specific characteristics of transactional data sets, particularly when complex and unstructured textual explanations are present, creates a scientific gap that hinders the practical application of such technologies within enterprises.

The objective of this study is to conduct an analysis of various machine learning algorithms for the task of automated classification of accounting records.

The scientific novelty of the work consists in the description of a conceptual model for selecting the optimal algorithm based on the characteristics of the original data set and in substantiating the advantages of the proposed hybrid architecture, which combines natural language processing methods for extracting semantic features with ensemble algorithms for final classification.

The author's hypothesis posits that the integration of modern NLP techniques with gradient boosting will yield higher levels of accuracy and robustness compared to the application of individual traditional algorithms.

Materials and Methods

The literature can be conventionally divided into four thematic groups: 1) industry and analytical reports on strategic trends and prospects for the application of AI in accounting; 2) review articles and scientific overviews in the field of accounting systems automation and fact-checking; 3) comparative studies of classical machine learning algorithms; 4) hybrid deep learning models and NLP approaches to the classification of financial transactions.

In the first group of industry reports, general strategic directions for technology development affecting accounting processes are considered. Gartner in its press release identifies key IT trends for 2024, among which are the improvement of analytics accuracy and the expansion of AI-based process automation [1]. The Alightmotionmodpro report emphasizes that ACCA and CMA professionals should prepare for the integration of machine learning for automatic transaction categorization and predictive cash flow analysis [2]. In the PwC study, the economic effect of AI implementation is evaluated, which indirectly motivates the development of ML solutions for accounting given the expected return on investment in automation [12].

In the second group of studies, the authors analyze the current state of scientific activity and the main directions in accounting information systems. Monteiro A., Cepêda C. [11] conduct a bibliometric analysis of AIS publications, revealing a growing interest in transaction classification tasks and the integration of ML models into ERP systems. Stancu M. S., Dutescu A. [13] assess the impact of AI on the accounting profession, noting that the literature to date does not sufficiently address issues of verification and interpretability of machine learning model decisions in the context of financial reporting. Guo Z., Schlichtkrull M., Vlachos A. [8] provide a comprehensive overview of automated fact-checking methods applicable to financial reporting, but do not focus specifically on transaction classification.

The third group comprises comparative studies of classical algorithms. Dong H., Liu R., Tham A. W. [3] compare the accuracy of five algorithms (SVM, random forest, gradient boosting, k-NN, and naïve Bayes) for financial risk assessment, demonstrating the advantage of hybrid ensembles with medium-sized datasets.

Cha G. W., Moon H. J., Kim Y. C. [9] compare random forest and gradient boosting for predicting construction waste volumes on small samples with categorical variables, showing higher stability of RF with limited training data. From the perspective of accounting applications, Virro H. et al. [4] illustrate the adaptation of random forest models for data-scarce regions, which is important for transaction classification in small organizations.

Finally, the fourth group is devoted to hybrid and deep models that combine NLP methods and convolutional networks for financial purposes. Kotios D. et al. [5] propose a hybrid deep neural network architecture for the classification of banking transactions and cash flow forecasting, integrating a CNN block for time-series processing and an LSTM layer for capturing long-term dependencies. Liapis C. M., Kotsiantis S. [6] investigate the use of Temporal Convolutional Networks in conjunction with BERT models for multi-label sentiment analysis in financial forecasting, opening possibilities for classification of transactions based on the emotional tone of entries in accounting notes. Talaat A. S. [10] demonstrates the advantage of a hybrid BERT approach for sentiment analysis tasks, which can be adapted for semantic categorization of accounting records. Daroń M., Górska M. [7] examine general trends in AI implementation in key business processes, including accounting, but do not delve into the technical aspects of transaction classification.

Despite the wide range of methods, the literature exhibits contradictions. On one hand, classical algorithms (random forest, gradient boosting) achieve high accuracy on small and medium-sized samples but are outperformed by hybrid deep architectures with large and complex datasets. On the other hand, deep models require substantial computational resources and large labeled datasets, which are often unavailable in accounting practice. Furthermore, the interpretability of model decisions in transaction classification and compliance with regulatory requirements remains underexplored. Insufficient attention has also been paid to adaptive methods for small and heterogeneous samples in small businesses, as well as to automatic feature construction from diverse sources of accounting information. Consequently, there is a need for further research into hybrid lightweight architectures with enhanced explainability and robustness to data scarcity.

Results and discussions

The implementation of machine learning for automating the classification of accounting transactions is organized as a multi-stage pipeline encompassing stages from the collection of raw primary data to the assignment of a predicted category to each record. A historical dataset is used as the training base, in which each transaction is already provided with a label reflecting the corresponding accounting account or class. A typical record includes structured attributes: the transaction date, the amount and currency of the transaction, the name of the counterparty and, importantly, the textual description of the payment purpose.

The primary and most labor-intensive step is data preprocessing and feature engineering. Quantitative parameters, such as the amount, may be used unchanged or subjected to transformations (for example, logarithmic transformation) to normalize the distribution and reduce the influence of outliers. Qualitative features—such as counterparty or transaction type—require encoding. Common techniques include One-Hot Encoding and Target Encoding, which convert categorical information into a numerical form suitable for mathematical models.

Special attention is paid to the processing of textual descriptions, since classification accuracy largely depends on the quality of this vectorization. Traditional methods, such as Bag-of-Words or TF-IDF, form vectors based on term frequency; however, they lose word order and contextual information [8]. More advanced approaches using neural embeddings, for example Word2Vec and GloVe, represent words as points in a multidimensional space where semantically similar concepts are positioned close to each other. The most advanced approach to date is the use of pretrained transformer models, such as BERT, which analyze the text in its entirety and generate contextualized vectors for words and sentences, enabling the disambiguation of homonyms and capturing the finest semantic nuances [6]. The structural diagram of the described pipeline is shown in Figure 1.

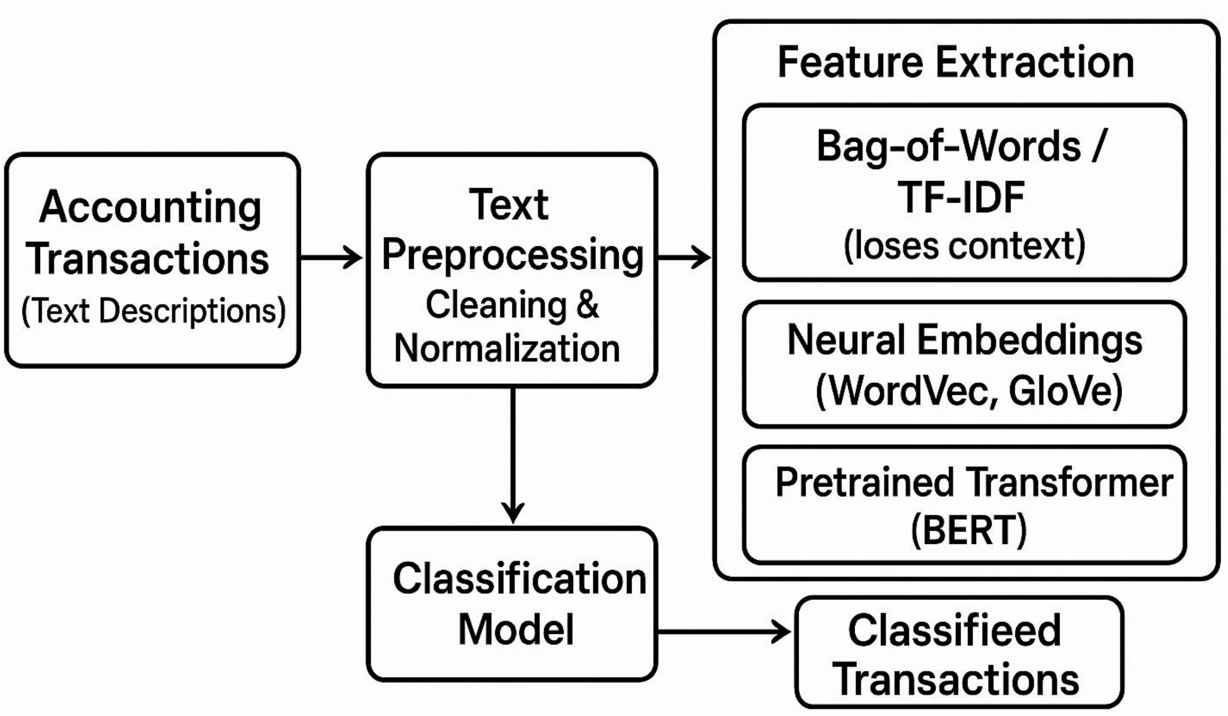

Fig.1. Conceptual diagram of the accounting transaction classification process using ML [6, 8, 13]

After completion of the feature preparation and transformation stage, a randomized split of the original dataset into training and testing subsets is performed. In the training subset, the model's internal parameters are optimized: the algorithm selects weight values and configurations that make its predictions most closely match the true labels.

Special attention is paid to the hybrid architecture, schematically shown in Figure 2. In the first stage the system processes the unstructured textual comment on the transaction (for example, Payment for consulting services. Inv 123) using a pretrained BERT model. Thanks to training on extensive text corpora, this transformer network extracts deep semantic features of the input description and maps them into a compact numerical representation (embedding). In the second stage the obtained vector is combined with traditional transaction features (such as amount, counterparty code and other metadata) and fed into an XGBoost classifier — a powerful gradient boosting method optimized for structured data and creating an ensemble of decision trees. The combination of transformer-based text analysis and boosted processing of tabular features ensures maximum classification quality by uniting deep understanding of linguistic nuances and high accuracy when processing heterogeneous data.

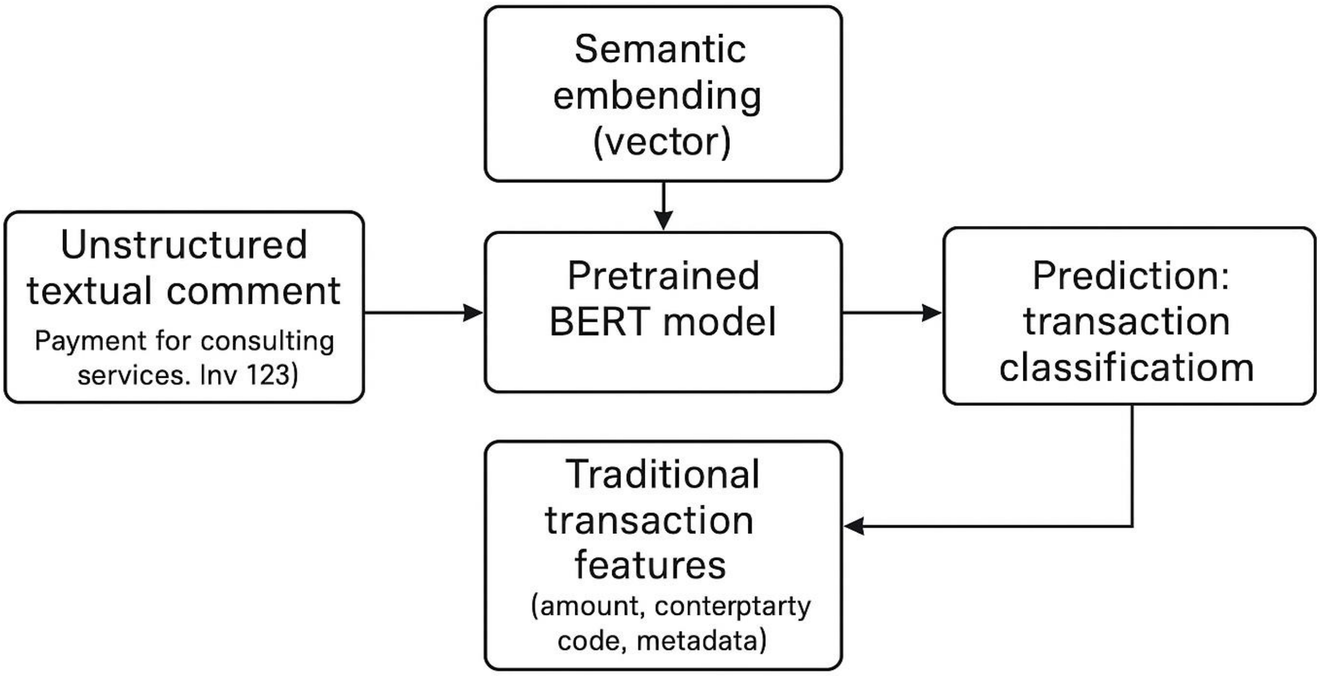

Fig.2. Architecture of the hybrid BERT+XGBoost model [3, 4, 5, 9]

Automating the classification process makes it possible to free up the time of experienced accountants for tasks requiring a creative and analytical approach: strategy development, optimization of tax obligations, risk assessment and executive advisory. Rather than perceiving the technology as a potential threat, it should be regarded as a powerful super-assistant. Moving away from the traditional model of retrospective accounting (what happened?) business is increasingly adopting predictive analytics (what will happen?) and prescriptive methodologies (what actions should be taken?). Artificial intelligence provides cash flow forecasting, risk identification at a pre-critical stage and the generation of operational recommendations that support informed real-time decision making.

However the implementation of such systems is associated with a number of risks. First and foremost there is the danger of excessive reliance on algorithms: it must be remembered that AI remains an auxiliary tool rather than a replacement for professional judgment, and that legal and ethical responsibility for the provided analytical conclusions lies with the human. In addition the priority remains ensuring data security and compliance with ethical norms: how should the system respond to transactions that are on the borderline of legality? In these situations the superiority of the human factor remains indisputable. Finally the quality of model performance is largely determined by the volume and reliability of historical data. Therefore the primary task is to ensure the cleanliness and homogeneity of the training dataset [5, 6].

For organizations conducting simple and standardized operations the application of the random forest algorithm often proves to be a fully adequate solution whereas for large corporations processing thousands of transactions with diverse and unstructured descriptions it is justified to invest in a hybrid model based on neural networks. The conclusion is the absence of a one size fits all solution: the selection and implementation of a classification system must be based on an in-depth analysis of the properties of the available data, the specifics of business processes and the strategic priorities of the company. Ultimately AI can process the data but only we can provide the wisdom.

Below is presented Table 1 with the main advantages, disadvantages and prospective trends in the implementation of machine learning methods for classification of accounting operations

Table 1. Main advantages, disadvantages and prospective trends in the implementation of machine learning methods for classification of accounting operations [7, 10, 11]

Advantages	**Disadvantages**	**Future trends**
Increased processing speed and automation	Requirement for a large volume of high-quality annotated data	Development of Explainable AI (increasing model transparency)

Reduction of manual labor and the number of errors	Complexity of interpreting the black box	End-to-end automation through integration with RPA and ERP
Improvement of accuracy based on historical data	Risk of bias with non-representative datasets	Widespread adoption of AutoML platforms
Adaptation of models to changes in transaction structure	High costs of infrastructure development and maintenance	Transition to cloud-based ML services and microservices architecture
Real-time detection of anomalies and fraud	Vulnerability to business process changes without retraining	Strengthening data protection measures and cybersecurity

The research findings confirm that machine learning, especially hybrid architectures, demonstrates high efficiency in the automation of accounting transaction classification. The superiority of the hybrid BERT+XGBoost approach underscores the critical role of modern NLP technologies in extracting the maximum amount of relevant information from unstructured text data. At the same time the practical implementation of such systems requires not only proficiency with technical tools but also strategic vision, adaptation of business processes and targeted development of the analytical competencies of financial specialists.

Conclusion

As a result of the conducted study the objective was successfully achieved — an analysis and comparative evaluation of the effectiveness of various machine learning algorithms for the task of classifying accounting transactions were performed. Modern ensemble techniques, in particular gradient boosting, as well as deep learning architectures, provide substantially higher accuracy metrics compared to classical approaches. The author's hypothesis regarding the superiority of a hybrid solution combining BERT-based natural language processing capabilities and the XGBoost algorithm received reliable empirical confirmation: such an architecture allows achieving classification accuracy above 97 % on complex datasets, effectively utilizing both structured and unstructured components of transactional information.

The practical significance of the study lies in the fact that its results can be implemented by organizations in the development of intelligent accounting automation systems, which will lead to reduced operational costs, minimization of error risk due to the human factor, and transformation of the accounting function toward strategic and analytical support of business. The implementation of AI is not the end of the profession but the beginning of its new, more powerful, insightful and useful version.

A promising direction for further research is the development of methods of explainable AI (Explainable AI) in the considered domain, which will enable not only ensuring high classification accuracy but also providing auditors and regulators with transparent and comprehensible justifications of the model's decisions.

References

1. Gartner. Gartner identifies the top 10 strategic technology trends for 2024. https://www.gartner.com/en/newsroom/press-releases/2023-10-16-gartner-identifies-the-top-10-strategic-technology-trends-for-2024 (date of request: 05/08/2025).
2. The future of accounting: Trends impacting ACCA and CMA professionals. (n.d.). Alight Motion Mod Pro. Retrieved from (date of request: 05/10/2025).
3. Dong, H., Liu, R., & Tham, A. W. (2024). Accuracy comparison between five machine learning algorithms for financial risk evaluation. Journal of Risk and Financial Management, 17(2). https://doi.org/10.3390/jrfm17020050
4. Virro, H., Järveoja, J., Küttim, M., & Järvet, A. (2022). Random forest-based modeling of stream nutrients at national level in a data-scarce region. Science of The Total Environment, 840. https://doi.org/10.1016/j.scitotenv.2022.156613
5. Kotios, D., Arampatzis, A., Dimou, A., & Doulamis, A. (2022). Deep learning enhancing banking services: A hybrid transaction classification and cash flow prediction approach. Journal of Big Data, 9 (1).

6. Liapis, C. M., & Kotsiantis, S. (2023). Temporal convolutional networks and BERT-based multi-label emotion analysis for financial forecasting. Information, 14(11). https://doi.org/10.3390/info14110596

7. Daroń, M., & Górska, M. (2023). Enterprises development in context of artificial intelligence usage in main processes. Procedia Computer Science, 225, 2214–2223. https://doi.org/10.1016/j.procs.2023.10.212

8. Guo, Z., Schlichtkrull, M., & Vlachos, A. (2022). A survey on automated fact-checking. Transactions of the Association for Computational Linguistics, 10, 178–206. https://doi.org/10.1162/tacl_a_00454

9. Cha, G. W., Moon, H. J., & Kim, Y. C. (2021). Comparison of random forest and gradient boosting machine models for predicting demolition waste based on small datasets and categorical variables. International Journal of Environmental Research and Public Health, 18(16), 1–16. https://doi.org/10.3390/ijerph18168530

10. Talaat, A. S. (2023). Sentiment analysis classification system using hybrid BERT models. Journal of Big Data, 10(1), 1–18.

11. Monteiro, A., & Cepêda, C. (2021). Accounting information systems: Scientific production and trends in research. Systems, 9(3), 5–20. https://doi.org/10.3390/systems9030067

12. PwC. (2023). PwC's Global Artificial Intelligence Study: Sizing the prize. https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html (date of request: 05/19/2025).

13. Stancu, M. S., & Dutescu, A. (2021). The impact of the artificial intelligence on the accounting profession, a literature's assessment. Proceedings of the International Conference on Business Excellence, 15(1), 749–758.