

Deep Learning and Large Scale Models for Bank Transactions

Fabrizio Garuti^{1,3}, Simone Luetto², Rita Cucchiara^{3,4} and Enver Sangineto³

¹Prometeia Associazione, Bologna, Italy

²Prometeia SpA, Bologna, Italy

³AlmageLab, UNIMORE, Modena, Italy

⁴IIT-CNR, Italy

Abstract

The success of Artificial Intelligence (AI) in different research and application areas has increased the interest in adopting Deep Learning techniques also in the financial field. Particularly interesting is the case of financial transactional data, which represent one of the most valuable sources of information for banks and other financial institutes. However, the heterogeneity of the data, composed of both numerical and categorical attributes, makes the use of standard Deep Learning methods difficult. In this paper, we present UniTTAB, a Transformer network for transactional time series, which can uniformly represent heterogeneous time-dependent data, and which is trained on a very large scale of real transactional data. As far as we know, the dataset we used for training is the largest real bank transactions dataset used for Deep Learning methods in this field, being all the other common datasets either much smaller or synthetically generated. The use of this very large real training dataset, makes our UniTTAB the first foundation model for transactional data.

Keywords

Transactional data, Deep Learning for finance, fraud detection, financial predictions

1. Introduction. Why large-scale models on transactional data?

The adoption of Artificial Intelligence (AI) in Finance and Banking has long been talked about. In 2017, J.P. Morgan presented the own disruptive AI-based software for financial document processing called COIN (CONtract INtelligence [1]), and few years later, OECD opened the AI Observatory on Fintech [2] focusing on both the opportunities, e.g., in asset management, credit intermediation and finance data analysis, and on the related risks, e.g., lack of explainability, learning bias, etc. Also Europe and Italy have gone in this direction, so that in the National Strategic Program on Artificial Intelligence launched in November 2021, one of the eleven Italian priorities is indeed “AI for banking, finance and insurance”. This is also a topic of interest for the new large National research project on foundational AI “FAIR” [3] that has just started in 2023, funded by Next Generation Europe funds.

AI is affecting Finance in several fields, and a short taxonomy in four main impact areas from the data point of view focuses on:

- Client data: Client intermediation and Customer engagement;

- Fintech Documents: intelligent document analysis for whichever finance and bank workflow;
- Finance data: analyzing and predicting financial trends for trading, risk and asset management;
- Transactional data: analysis classification and generation of product data described as time series.

The first two categories, AI on Client Data and AI for Intelligent Document Processing, represent the two most mature areas. Regarding the former, recent deep learning models for visual and text data interpretation and generation offer new solutions for client interfaces with digital platforms, for client biometric identifications, chatbots for customer interactions and other AI-based apps for client disintermediating. On the other hand, the recent advances in Natural Language Processing (e.g., BERT [4], GPT [5] and BART, Chat-GPT) can directly be used on finance document repositories for classification, search and retrieval.

In the third category, based on finance time-series, machine learning has been often applied with success, e.g. for macroeconomic analysis [6]. Similarly, deep learning has been applied to stock exchange prediction: thanks to the large available stock exchange data, neural network jointly with more traditional statistical methods are largely effective.

Conversely, the adoption of deep learning for transactional bank data is still under-explored. So far, these data have been usually processed with symbolic AI (e.g., rule-based expert systems) and/or with traditional machine

Ital-IA 2023: 3rd National Conference on Artificial Intelligence, organized by CINI, May 29–31, 2023, Pisa, Italy


*Corresponding author.

✉ fabrizio.garuti@prometeia.com (F. Garuti);

simone.luetto@prometeia.com (S. Luetto)

© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License

Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

learning approaches (e.g., SVMs or gradient boosted decision trees [7]). The multimodal nature of transactional data and the lack of large public annotated dataset (due to privacy and commercial reasons) make these data extremely difficult to be handled by deep neural networks. However, transactional data represent the largest source of information for banks: transactions categorization, Client profiling, Fraud detection, Dynamic prediction (e.g., churn prevention) to mention a few.

The UniTTab Italian project [8], a research collaboration of **Prometeia Associazione** and the University of Modena and Reggio Emilia, is one of the pioneering approaches exploring the use of attentive deep learning for transactional bank time series. Specifically, the project achieved some preliminary important results in the creation of Foundation models [9] for fintech. UniTTAB is based on a new self-supervised Transformer architecture [10], trained on tens of millions of real transactional data for different financial tasks, including the generation of synthetic data, useful also for secure and anonymized processes.

2. Related works

Padhi et al. [11] recently proposed one of the first deep learning architectures for transactional data. Specifically, the authors present two different architectures: TabBERT which is used for classification tasks, and TabGPT used for forecasting / generation tasks. As a solution to the data heterogeneity problem, the authors quantize continuous attributes so that each field is defined on its own finite vocabulary. Then they define a data sample as a sequence of transactions. The main difference with NLP is that they have a sequence of structured data consisting each of fields defined on a dedicated vocabulary.

Another recent work is TabAConvBERT, proposed by Shankaranarayana & Runje (2021) [12]. They present an architecture which can deal with both categorical inputs (by using an embedding neural network) and numerical inputs (by using a shallow neural network). They also propose a special timestamp embedding block, where they break the original timestamp into multiple components, such as year, month, day and hour. The obtained time embedding is then added with input features' embedding and positional encoding.

The architecture presented by X. Huang et. al. [13] is able to handle both categorical and numerical features, providing a solution to data heterogeneity. However, the main drawback is that this method cannot deal with the temporal component of the data, and therefore is unable to solve task involving transaction sequences.

Ours proposal differs from the aforementioned works in different aspects. On the one hand, we deal with all the variability dimensions of the problem: numerical,

categorical and temporal. On the other hand, thanks to the collaboration with a private financial institution, we scale the size of the pre-training dataset to 48 million transactions. In fact, as far as we know, our Real Bank Account Transaction Dataset (in short RBAT dataset) is larger than all the other *real* transactional datasets used for training Deep Learning methods.

3. The challenges of transactional data and Deep Learning

Dealing with transactional data is more complex than working with music, images or text because of the heterogeneity of the input. It is also more difficult than working with multimodal data; in addition the public transactional datasets are often too small and limited in diversity.

Indeed, the impactful results of UniTTab (Unified Transformer model for Tabular data) were driven by the availability of large transactional datasets as well as the availability of NVIDIA GPUs, which facilitated the definition of new neural architectural models, specifically designed for banking transactional data.

The main challenges addressed by UniTTab depend on the aspects of these data, briefly outlined in the following:

- Tabular data, usually collected from different sources (e.g., separate databases) thus they require an intensive pre-processing for data cleaning and interoperability.
- Time dependence. Transactional data represent a special case of time-series with non-regular frequency: bank customers carry out a variable number of transactions per year, ranging from very few transactions up to several thousand transactions per year.
- Heterogeneous data. Transactions have not homogeneous fields: some of them are numerical (e.g. the amount), some categorical (e.g. the type of transaction), some textual (e.g. the bank transfer description) or with a specific structure (e.g. the date).
- Multiple-structure data. The transactions of a client account have a different field-structure according with the type of transaction (e.g. a POS, a credit card, an ATM or a bank transfer).
- Correlated data. The transaction fields are often correlated to each other in the same time series (e.g. in periodical payments) and among different time series: each client can own different bank products, different accounts, and some accounts have different owners (join accounts). Finally, some transactions are correlated with external

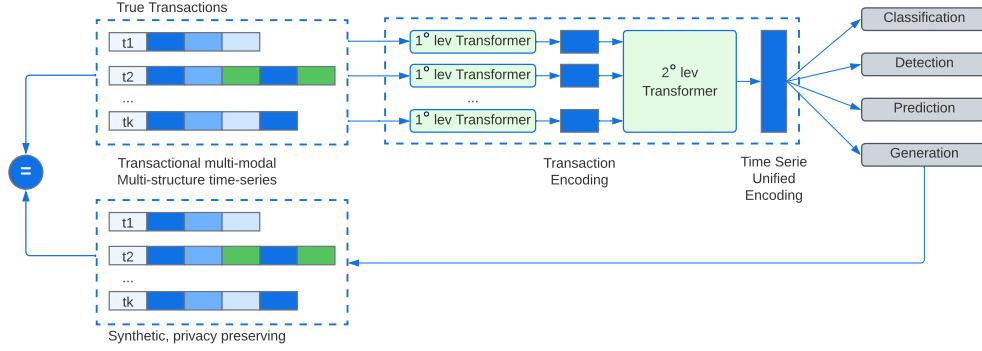


Figure 1: A schematic illustration of the UniTTAB architecture for financial data.

but unknown conditions (e.g. holiday times or the lockdown in the pandemic period).

4. The architecture

Given the previously discussed challenges of data variability, quantity and heterogeneity, deep learning for transactional data has been largely underexplored, with only a few public experiments and a small number of private institutions and banking research centers. One interesting recent position paper (J.P Morgan 2021 [14]) concerns synthetic data generation, even if it actually defines the problems but does not offer any solutions.

In non-financial AI, state-of-the-art models are usually based on Transformer architectures, usually trained using self-supervised learning (e.g., using “masked word” prediction tasks or through generative or “contrastive learning”). Initially defined as language models in the field of NLP, they are now common also in other AI areas such as Computer Vision, making Transformer networks the basic paradigm for contemporary AI. Specifically, “Foundational Models” are typically large Transformers pre-trained on huge datasets (e.g., the Wikipedia documents, or billions of Web-collected images). Their goal is to create compact, intermediate representations of the input in a latent space, useful for different tasks, such as classification, generation, recognition, image segmentation, anomaly detection, etc. Examples are BERT [4], GPT2/3/3.5 [5], CLIP [15], etc. They can be used in a simple way, see the worldwide success of Chat-GPT3, and “fine-tuned” to be adapted to specific tasks. On the other hand, using Transformers to create Foundation models for new types of data -such as the transactional data- is more complex, both because training and testing requires days of GPU computation, jointly with very large datasets, and because the Transformer architecture should be re-defined for the specific domain.

In the UniTTAB project we explore this trend by re-designing attentive and generative models, i.e. Transformers. This allowed us to deal with the heterogeneous nature of transactional data: tabular time series with multiple-structure and multimodal fields. The designed architecture is resumed in Figure 1.

Borrowing the techniques used in text analysis in BERT or GPT models, we used input time series with variable length. We varied the sequence length from 50 to 150 transactions, where each transaction is composed of a fixed number of 10 fields. As a result, each time series can vary in length from 500 to 1500 items, a challenging length to be managed even for text sentences.

Given the data structure we use the hierarchical architecture shown in Figure 1. First, we endow a field-level transformer, which encode individual transactions into embeddings. Then these embeddings are fed into the second-level transformer, that processes the time-series to encode them as a single element in the latent space. This is the foundation latent space where the representation can be potentially exploited for many tasks, such as Classification (e.g., to classify the client behavior), Detection (e.g., to detect anomalies, frauds, etc.), Prediction (e.g., to predict product churn in next few months). As shown in Figure 1, this model can also be used for a generative task. For instance generating time series data has the advantage of preserving the content but also the privacy of the client.

5. Experiments on available datasets

Details of the architecture, at least for some tasks of detection and prediction, are described in UniTTaB proposed by Simone Luetto et al. [8]. The effectiveness of the model has been tested over various datasets, used as benchmarks for different tasks, according with the provided manual annotation.

Table 1

A quantitative comparison between UniTTAB and the state-of-the-art Deep Learning methods for time series data on three tasks: (1) a Fraud detection task on the (synthetic) Transactions Dataset, (2) a Loan default prediction task on the (real) PKDD'99 Financial Dataset, (3) a Churn prediction task on our (real) RBAT Dataset.

Model	Transactions Dataset	PKDD'99 Financial Dataset		RBAT Dataset	
	F1 score	F1 score	Accuracy	F1 score	Accuracy
TabBERT [11]	0.860	0.620	91.6	0.526	86.5
LUNA [16]	0.862	-	-	-	-
TabAConvBERT [12]	0.896	-	-	-	-
UniTTab (ours)	0.915	0.673	92.3	0.604	90.8

- A Fraud Detection task has been tested on the Transactions Dataset [17] proposed in 2021 with synthetic credit card transactions, composed of multimodal data (some fields are categorical, some are numerical). Trained on about 2 Million samples, tests have been provided on a sets of about 450K sequence of transactions. As reported in Table 1 UniTTab strongly outperform any competitor, with an accuracy 93.5 and an F1 measure of 0.915.
- A Loan Default Prediction task is evaluated on the PKDD'99 Financial Dataset [18]: it is a relatively small dataset with “only” 45K clients, each performing 200 transactions in average. Although the dataset is very unbalanced, loan Default is correctly predicted with a F1 measure of 0.673 and an accuracy of 92.3 (Table 1). Also in this case results are the state-of-the art.
- A Churn Rate Prediction task is finally evaluated on the RBAT Dataset. A subset of approximately 100K bank accounts with about 50 Million of transactions have been adopted for training the complete architecture of Figure 1. As reported in Table 1, prediction is very precise – absolutely better than other competitors - with an accuracy of 90.8 and an F1 measure of 0.604.

A brief comparison of the results of our UniTTab model against the competitors on the previously mentioned tasks is provided in Table 1.

6. Conclusions

The project carried out by **Prometeia Associazione** and UNIMORE, presented in this paper, is a first step towards the creation of foundation models for transactional time series data. The empirical results show that our model drastically outperforms both deep learning and standard machine learning based predictive models on different benchmarks. We believe that our work and our results

can stimulate this research field and the adoption of self-supervised deep learning in banking data.

References

- [1] AI Case Study JPMorgan. JPMorgan reduced lawyers hours by 360,000 annually by automating loan agreement analysis with machine learning software coin, 2018. URL: <https://www.bestpractice.ai/studies/jpmorgan-reduced-lawyers-hours-by-360-000-annually-by-automating-loan-agreement-analysis-with-machine-learning-software-coin>.
- [2] AIFinance OECD, Artificial Intelligence, Machine Learning and Big Data in Finance, 2021. URL: <https://www.oecd.org/finance/financial-markets/Artificial-intelligence-machine-learning-big-data-in-finance.pdf>.
- [3] FAIR, Future AI Research is the Italian three-year project under the PNRR “Partneriati Estesi” program, coordinated by Italian National Council, aggregating several Italian universities and industries, 2023. URL: <https://future-ai-research.it/>.
- [4] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, in: NAACL, 2019.
- [5] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, D. Amodei, Language Models are Few-Shot Learners, arXiv:2005.14165 (2020).
- [6] E. Casabianca, M. Catalano, L. Forni, E. Giarda, S. Passeri, A machine learning approach to rank the determinants of banking crises over time and across countries, Journal of International Money and Finance (2022).
- [7] T. Chen, C. Guestrin, XGBoost: A Scalable Tree

- Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.
- [8] S. Luetto, F. Garuti, E. Sangineto, L. Forni, R. Cucchiara, One Transformer for All Time Series: Representing and Training with Time-Dependent Heterogeneous Tabular Data, 2023. [arXiv:2302.06375](https://arxiv.org/abs/2302.06375).
 - [9] Large scale models based on unsupervised learning have been defined by Stanford HAI "Foundational Models", a term now in common use in science, 2021. URL: <https://fsi.stanford.edu/publication/opportunities-and-risks-foundation-models>.
 - [10] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is All you Need, in: NeurIPS, 2017.
 - [11] I. Padhi, Y. Schiff, I. Melnyk, M. Rigotti, Y. Mroueh, P. L. Dognin, J. Ross, R. Nair, E. Altman, Tabular Transformers for Modeling Multivariate Time Series, in: IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, 2021.
 - [12] S. M. Shankaranarayana, D. Runje, Attention Augmented Convolutional Transformer for Tabular Time-series, in: 2021 International Conference on Data Mining, ICDM 2021 - Workshops, 2021.
 - [13] X. Huang, A. Khetan, M. Cvitkovic, Z. S. Karnin, TabTransformer: Tabular Data Modeling Using Contextual Embeddings, [arXiv:2012.06678](https://arxiv.org/abs/2012.06678) (2020).
 - [14] S. Assefa, D. Dervovic, M. Mahfouz, T. Balch, P. Reddy, M. Veloso, Generating synthetic data in finance: opportunities, challenges and pitfalls, Workshop on AI in Finance Neurips (2021).
 - [15] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, I. Sutskever, Learning Transferable Visual Models From Natural Language Supervision, 2021. [arXiv:2103.00020](https://arxiv.org/abs/2103.00020).
 - [16] H. Han, J. Xu, M. Zhou, Y. Shao, S. Han, D. Zhang, LUNA: Language Understanding with Number Augmentations on Transformers via Number Plugins and Pre-training, [arXiv:2212.02691](https://arxiv.org/abs/2212.02691) (2022).
 - [17] E. R. Altman, Synthesizing Credit Card Transactions, 2019. [arXiv:1910.03033](https://arxiv.org/abs/1910.03033).
 - [18] P. Berka, Workshop notes on Discovery Challenge PKDD'99, 1999. URL: <https://sorry.vse.cz/~berka/challenge/pkdd1999/berka.htm>.