Journal entries with deep learning model

Župan, Mario; Letinić, Svjetlana; Budimir, Verica

International Journal of Advanced Computational Engineering and Networking, ISSN(p): 2320-2106, ISSN(e): 2321-2063
Volume-6, Issue-10, Oct.-2018, http://iraj.in

JOURNAL ENTRIES WITH DEEP LEARNING MODEL

[1]MARIO ZUPAN, [2]SVJETLANA LETINIC, [3]VERICA BUDIMIR

[1,2,3] Polytechnic in Pozega, Pozega, Croatia
E-mail:[1] mzupan@vup.hr

Abstract - Deep learning is the most recent approach to achieve Artificial intelligence. Especially neural networks are used for solving many human problems - from repetitive operations to intelligent recognizing in image, sound and text processing. They are used in medicine, car industry, game industry and robotics. Business companies also try to find the way of exploitation of the latest technology despite the fact that it is the long way to the point where machines will be capable to replace the human intelligence. Authors of this paper explore possibilities of semi-supervised learning application in accounting. One of the latest deep learning algorithm is successfully used to reconstruct the journal entry key columns. The model was trained and tested on a real-world dataset so it could become base for developing the wide pallet of accounting and audit applications - as anomaly detection module of Enterprise Resource Planning (ERP) software or as a standalone application.

Index Terms - General ledger, journal entry, bookkeeping, accounting, deep learning, variational autoencoder, anomaly detection, accounting control system.

I. INTRODUCTION

Croatian small and medium enterprises have a wide pallet of quality developed Enterprise Resource Systems (ERP). As in every information system human work and interaction between the modules could cause the errors. An errors in accounting books happen on a daily basis especially because Croatian and European legislative are experiencing continuous changes. As a small and medium enterprises (SME's) do not have audit obligations regulated by the law, tax inspections are the only mechanism of their accounting and tax books controlling and their job is mainly manual. Generally, finding errors, purposely made or not, consumes the large part of time of bookkeepers or tax inspectors and afterwards correction of the errors is not the easy part of the job, especially because of the nature of accounting modules architecture and functioning. Namely, most of today's ERP systems have a specialized documents (forms) for every possible business event. Every document is connected with one or more journal entry schemes created by the senior accountants. Junior accountants or non-accountant employees do not have to be familiar with the journal schemes because they communicate only through documents (forms). Modern accounting modules inside ERP systems are generally functioning on described principle. Consequently, one error in only one scheme could cause an incorrect accounting entry for a whole set of connected documents. Most of existing controls integrated in accounting modules of modern ERP systems are created in compliance with the bookkeeping rules. As the statistics came into audit or customer relationship modules implementation of statistics together with deep learning algorithms, as the latest approach to machine learning, could make a business life of the people involved in accounting and tax processes a lot easier.

II. RELATED WORKS

When the idea for this research study began to form, the authors have found the great paper with the same challenge and methodology. The paper with the title Detection of Anomalies in Large Scale Accounting Data using Deep Autoencoder Networks [1] processed two dataset extracted from SAP ERP.First dataset represents accounting document header (e.g. document id, type, time currency) and the second contains journal entry details (e.g. general ledger account, debit, credit, amount). As the majority of the attributes correspond to categorical variables authors preprocessed the journal entry attributes to obtain a one-hot encoded representation of each attribute. They got 401 encoded dimensions for dataset A and the 576 encoded dimensions for dataset B. Each journal entry was labeled as either synthetic global anomaly, synthetic local anomaly or non-synthetic regular entry. Described datasets became inputs in five distinct autoencoder architectures. Authors use classical autoencoder in opposite to variational autoencoder which will be used in their future works. In dataset A the best architecture leaves only 0.15%, and 0.62% of entries in a dataset B, journal entries that exceeds a predefined reconstruction error threshold. Accordingly, accuracy of the model was 99.87% for dataset A and 99.43% for dataset B. These entries are leaving for a further manual inspection because they are not, potentially, created by
a regular business activity.

III. METHODOLOGY

Keras, the most popular deep learning library made a big contribution to the popularization of artificial intelligence(AI). This high-level neural networks Application programming interface (API) is capable of running Google's TensorFlow, Microsoft

Cognitive Toolkit (CNTK) or Theano deep learning library. Nowadays neural networks are mainly used for supervised learning - network was trained on labeled dataset. Supervised model could be well optimized but unuseful at the same time with new data if it is trained and tested on the same dataset - generalization of the model is sacrificed. But the neural networks can also be used for semi-supervised learning, through the autoencoders and with unlabeled datasets. An autoencoders predict the input given that same input. They can be trained and tested on the same data if the generalization is not the condition, like in this kind of dataset processed by this research study. Due to the Keras friendliness and to the TensorFlow power the neural network model was built as the result of this research. More precisely, the model is atuoencoder but not classical but rather variational autoencoder (VAE) simultaneously discovered by Kingma and Welling in December 2013 6 and Rezende, Mohamed, and Wierstra in January 2014 [2]. In opposite to classical autoencoder VAE is capable for sampling from a latent space to create entirely new output. That is possible because it turns the input into the parameters of a statistical distribution: a mean and a variance instead of compressing its input into a fixed code in the latent space how the classical autoencoder works. VAEs can be used to develop latent spaces of sound, music, or even text, but in practice, the most interesting results have been obtained with pictures. Accordingly, the VAE is potentially capable to generate a new journal entries from the latent space as well as reconstruct the existing ones.

A. Journal entries

Journal entries are generated by Synesis, one of the most popular Croatian enterprise resource planning(ERP) systems pretty widespread in small-sized enterprises as well as in accounting service enterprises. In Synesis, every bookkeeping document (form) is connected with the specific journal entry scheme. Scheme suggests to the bookkeeper a few general ledger accounts specific for connected document. For instance, Value-added tax (VAT) input invoices evidence scheme will suggest VAT receivables and trade payables accounts and then the bookkeeper will finished the entry to get a balance. The rest of the entry is sometimes connected with some account, or maybe with liabilities for not invoiced but received goods and services account if it is the word about the invoice for a merchandise goods in stores. Journal entries exported document from Synesis contains the following fields: date, priority, id of the journal entry (DIN), document name, document id, departure, account, analytic identification number of partner, description of account, description of the entry, debit amount and credit amount. Preprocessing part of the algorithm developed for this research study selected 4 fields: document id, general ledger account, debit and credit indicator without amount. Then the selected fields are one-hot encoded which generates 57 binary columns ready for the VAE model as an input. It is impotant to mention that date, amount and journal entry id attributes could become potential inputs for the future model development while this study has been tried to evaluate learning capabilities of bookkeeping rules depends on specific business activity stored and presented with specific ERP document. To clarify bookkeeping rule dependence on specific document with an example: revenue account most of the time records the credit side with the document xinvoicesb except at the end of the fiscal year when the bookkeeper needs to close the revenue accounts with the aim of annual profit and loss calculations and financial statements creation. Revenue and expenses accounts have been closing with acc7close document. To summarize, bookkeeping rules like the mentioned ones need to be learned by VAE model that accepts journal entries grouped only by document id to avoid partially redundant grouping by date and journal entry id at the same time.

B. Variational autoencoder modeling

The VAE model was trained and tested on 3.731 rows x 57 columns dataset - journal entries for 4 fiscal years, from 2014 to 2017. The dataset was divided to train and test part in 1:9 ratio, that is 374 x 57 train shape and 3.359 x 57 test shape. By respecting the deep learning rules, 10% of the dataset has never been seen by the model so that the generalization of the model was ensured although the generalization is not relevant in this specific reconstruction problem where journal entries cannot be reconstruct partially successful. Encoder and decoder was built by using dense layers with relu activation, except the last one when sigmoid activation was used. The model was built by using binary crossentropy loss function and was compiled with rmsprop optimization scheme. The model has asymmetric architecture. It was a key fact for a successful development. Great explanation of asymmetric architecture was given by Xi Chen and et.al. [3].

C. Hardware

Training and testing algorithms was performed in a virtual environment with Linux Ubuntu operating system installed on Asus Strix GL553V laptop with 16GB DDR4 2400MHz of RAM, NVIDIA GTX1050Ti 4GB GDDR5 GPU and Intel Core i7-7700HQ 2.8GHz CPU. Algorithm was made with Keras deep learning API developed by Francois Chollet with the Tensorflow deep learning API backend developed by Google Inc.

A training was performed in 1.000 epochs in 7.21 minutes by using only CPU power and 8 GB of RAM assigned to the virtual machine. For larger datasets Keras API can use GPU.

IV. RESULTS

The model was trained and tested on 3.731 rows x 57 columns dataset - journal entries for 4 fiscal years, from 2014 to 2017. The dataset was divided to the train part and to the test part in 1:9 ratio - that is 374 x 57 train shape and 3.359 x 57 test shape. The model has incorrectly reconstructed 4 of 347 journal entries of the test dataset, so the precision of the model is 4/374*100=99.9893%. Error analysis has shown that 4 entries with positive reconstruction error have the same common characteristic: they are single in a whole dataset i.e. there is only one record per error of these specific business activity. Namely, the whole dataset have 183 unique rows but they are repeated through the years except 4 marked by the model:
Reconstruction error #875 is the entry from 2014 on stockbalance document. That journal entry has been recorded finished goods in warehouse and the cost of goods sold as the counterpart entry. As it is written, this is the only production entry in the whole dataset (test and train part). It is interesting to notice that only one side of this double entry is not properly reconstructed. Reconstruction error #1421 is the entry from 2015 on iteminput document. That journal entry has been recorded production account and the semi-finished goods in warehouse.

What is written for the previous error is also valid for this error: that is the only activity in the whole dataset and the counterpart entry is reconstructed properly. Reconstruction error #1837 happened in 2015 and is also unique in the whole dataset - sale of the building land, while the purchase on that account was recorded for a few times.

Reconstruction error #1889, the last single and unique event recorded the transfer of half-finished products back to the production account. The research has shown that the model cannot reconstruct the journal entries if it sees entry for the first time in a test part of the dataset. The result of a research study is transparent and the robustness of the model has come to the test due to the fact that the input was generated by the micro-sized production enterprise which had a few business events for the first time and the only once in a four-year period.

Further research should find out does dataset contains more than 4 unique and single journal entries or the model has recognized them all. The whole test part of the dataset is shown in Figure 1. By using t-SNE [4] for visualizing high-dimensional data, 374 journal entries are visualized in 2 dimensional space. Figureshows that the entries are grouped into clusters. The clusters are more or less separated one from another and the errors are placed in one of the given clusters. Further research should find out the common characteristics of the given clusters and their differences in relationship to the others.

CONCLUSION AND FUTURE WORK

The model has incorrectly reconstructed 4 of 347 journal entries of the test dataset so the precision of the model is 99.98%. The analysis has shown that 4 entries with positive reconstruction error have the same common characteristic. They are errors in general ledger from the model point of view because the model was trained on a dataset which did not contains that kind of the record. Due to the chosen semi-supervised deep learning algorithm, the model learned double entry bookkeeping on unlabeled

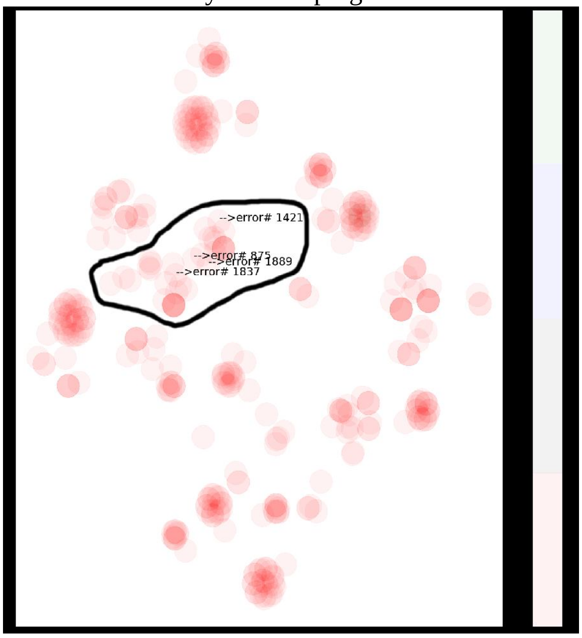

Fig. 1. Visualization of the test dataset

dataset and it is capable to generate another journal entries record or the whole fiscal year correctly. It is capable even to generate fictive fiscal year. That facts open a few interesting and useful applications of the model. One of them is highly accurate journal entry control system neural network. The development of the model has been time-consuming and it demanded careful research of the number of layers and neurons, optimal batch size, optimal loss function coding and suitable optimization scheme. However, experience collected with this research is unquestionable invaluable because deep learning researchers build intuition over time as to what works and what does not. Actually, researchers still do not know the possibilities of deep learning algorithms.
Future research should find out does dataset contains more than 4 unique and single journal entries or the model recognized the only ones. Accuracy of the model needs to be evaluate on synthetically made errors in a test part of the dataset. The authors will also research the idea of combining this model results and separately developed model which will be trained on date, amount, journal entry id and maybe some other inputs. Further, common characteristics of the journal entry clusters given by tSNE methodology and their relationship was not researched.

ACKNOWLEDGMENT

The authors would like to thank: Pascal d.o.o. – Croatian accounting service enterprise for a generous donation of a journal entries dataset; unselfish researches and developers from stackoverflow.com and keras.io on slack.com for sharing their experiences from the outside of the books.

REFERENCES

[1] Schreyer, M., Sattarov, T., Borth, D., Dengel, A., and Reimer, B. (2017). Detection of Anomalies in Large Scale Accounting Data using Deep Autoencoder Networks. arXiv preprint arXiv:1709.05254.
[2] Chollet, F. (2017). Deep learning with python. Manning Publications Co..
[3] Chen, X., Kingma, D. P., Salimans, T., Duan, Y., Dhariwal, P., Schulman, J., ... and Abbeel, P. (2016). Variational lossy autoencoder. arXiv preprint arXiv:1611.02731.
[4] Maaten, L. V. D., and Hinton, G. (2008). Visualizing data using t-SNE. Journal of machine learning research, 9(Nov), 2579-2605.

★★★