Prediction of Next Events in Business Processes: A Deep Learning Approach

Tahani Hussein Abu Musa[1[00000-0003-0661-4962]] and Abdelaziz Bouras [2[0000-0001-5765-1259]]

[1] Department of Computer Science and Engineering, College of Engineering (Qatar University)
Doha, Qatar
ta090001@student.qu.edu.qa.
[2] Department of Computer Science and Engineering, College of Engineering (Qatar University)
Doha, Qatar
abdelaziz.bouras@qu.edu.qa

Abstract. Business Process Mining is considered one of the merging fields that focusses on analyzing Business Process Models (BPM), by extracting knowledge from event logs generated by various information systems, for the sake of auditing, monitoring, and analysis of business activities for future improvement and optimization throughout the entire lifecycle of such processes, from creation to conclusion. In this work, Long Short-Term Memory (LSTM) Neural Network was utilized for the prediction of the execution of cases, through training and testing the model on event traces extracted from event logs related to a given business process model. From the initial results we obtained, our model was able to predict the next activity in the sequence with high accuracy. The approach consisted of three phases: preprocessing the logs, classification, and categorization and all the activities related to implementing the LSTM model, including network design, training, and model selection. The predictive analysis achieved in this work can be extended to include anomaly detection capabilities, to detect any anomalous events or activities captured in the event logs.

Keywords: Business Process Mining, Event log, LSTM, Business Process.

1 Introduction

Business process mining or process mining is an emerging research area that gained increasing attention in both academic and industry. It aims at analyzing Business Process Models (BPM), by extracting knowledge from event logs generated by various information systems, for the sake of auditing, monitoring and analysis of business activities for future improvement and optimization throughout the entire lifecycle of such processes, from creation to conclusion [1]. It is also useful when detecting potential exceptions or anomalies in the logged workflow, and hence help identify the source of such exceptions and modify them accordingly.

Typically, a business process may be defined as a group of linked activities produced for a particular purpose [2]. A business process has a number of attributes, has a particular goal, needs a certain input, produces a certain output, has a series of events or activities carried out in a predefined order, might need some resources and involve a number of owners [3]. There exist various representations and notations of business processes, including Unified Modeling Language (UML), Business Process Model Notation (BPMN), Event-driven process chain (EPC) and Petri Net Markup Language (PNML) [2].

Due to the availability of event logs produced from different information systems, typical process mining techniques are utilized in process discovery, conformance checking, and model enhancement. Hence, providing valuable business perceptions and insights [4].

In this context, event logs can be considered important means for extracting knowledge about the sequence of activities, making it possible to discover, monitor and improve such processes in different application domains of interest.

However, those typical systems do not have the ability of making predictions about the next activity in business processes.

Recently, the growth and spread of Machine Learning (ML) algorithms in the field of process mining, and predictive analysis has become a motivation of using them in performing accurate predictions of future events and activities in business processes, by analyzing logged historical activities and events, extracted from event logs. Such methods include parametric regression [7], Naive Bayes classifier [5], and predictive clustering tree inducer [6].

Besides, and due to the large volumes of data presented by the extracted event logs, deep learning approaches can outperform those typical process mining approaches, bringing more accurate predictions about future events in each event log trace, through analyzing data from previous events. Deep learning approaches including deep neural networks have been widely used in predictive analysis and process mining [8].

In this paper, a Long Short-Term Memory (LSTM) Deep learning model is used to predict the next activity of a given business process. LSTM is one type of Recurrent Neural Networks (RNNs) that based on the literature, has been used extensively in time series analysis and sequential problems [9], [10], [11]. The model presented in this work has been trained on event logs that represent procurement scenario, allowing for predicting the next activity on a given trace of events that follow another activity or a series of activities as an input. To validate our approach, we obtained our preliminary results based on a dataset comprising of 300 traces. The test performed on the trained model showed that it was able to predict the next activity in our business process model presented.

The structure of this research is organized as follows: Section 2 survey some related work in literature, Section 3 describes our proposed approach, Section 4 demonstrates results and analysis, and Section 5 presents a discussion of our research findings and future work to be done in this area.

2 Literature Review

2.1 Process Mining

Process mining may be defined as the area that resides between data mining and Business Process Modelling (BPM) and analysis. The main goal of process mining is analyzing Business Processes through the extraction of knowledge from event logs generated by different information systems, to support the activities related to auditing, monitoring and analysis of business activities for future improvement and optimization throughout the entire lifecycle of such processes, from creation to conclusion. Several process mining algorithms exist in literature to support the activities of conformance checking, performance analysis, decision mining, organizational mining, predictions, and recommendations [12].

In process mining, data related to events and activities of a given business process are recorded in event logs, which may be defined as a "hierarchically structured file with data on the executions of business processes" [13]. Event logs are considered the input to process mining, and typically contain traces of several executions of the same business process. A ***case*** or ***trace*** is a collection of related events within a single instance of the business process. An ***event*** or ***activity*** is a single atomic part of the trace, that has several attributes describing it, beside other information related to updating the state of such events, for example (Notify Delivery of the parts). Other attributes of events may include resources allocated and timestamp [14]. The following figure illustrates the relation between and event log, a case and event.

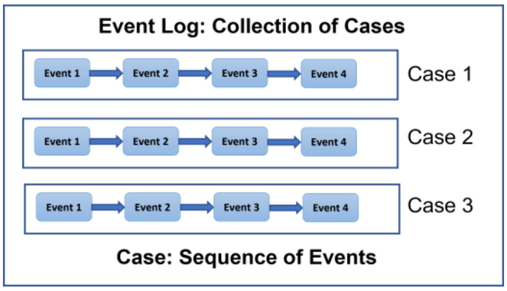

Fig. 1. Relation between Log, case, and event

According to [13],the three main tasks of process mining are: Process Discovery, Process Conformance and Process Improvement. ***Process discovery*** involves the use of event log as an input and the output will be the corresponding business process model without any previous knowledge of it [13]. ***Process Conformance Checking*** involves the comparison between a business process model and the event log generated via executing the same process model [13]. Such comparison helps evaluates whether the log information and the model are equivalent. ***Process Improvement*** involves applying any enhancements on the current model using the current information from the log.

Although event logs can be in either structured or unstructured, but nowadays, the eXtensible Event Stream (XES) format is the standard format for event log specification [15], considered as the 'de facto' standard for storing event logs by the IEEE task force, and it was preceded by the format Mining eXtensible Markup Language (MXML) , enabling the exchange of event data between diverse systems and employing process mining techniques[15].

2.2 LSTM Neural Networks

Recurrent Neural Networks (RNNs) and its variant the Long-Term-Short Memory (LSTM) have been used extensively in time series analysis and sequential problems.

RNNs have two types of input, present and recent past. RNN use both types of input to verify how they behave in case of new data. In other words, at any given time, the output of a RNN at time step $_{t-1}$ affects its output at time step $_{t}$. LSTMs are considered more efficient at obtaining long-term temporal dependencies [16]. The information contained in LSTMs are beyond the normal flow of the recurrent network in a gated cell. Such Information can be stored, written, or be read from a cell, like data in a computer's memory. The cell is able to make decisions about what should be stored and when it should be allowed to read, write, and delete, through gates that open and close. Such gates are implemented with the multiplication of elements by sigmoid, resulting in having all in the range of 0-1 [16].

In the area of process mining, there exists several research works been conducted related to utilizing LSTM in predicting next event in business processes. In [17], the authors proposed the question of how to use Deep Learning techniques to train accurate models of Business Processes behavior from event logs. The proposed approach trained a neural network with LSTM architecture, in order to predict the sequence of next events, their timestamp and associated resources. An experimental evaluation on real life event logs has been performed and showed that the proposed approach outperforms previously proposed LSTM architectures targeted at this problem [17].

In [18], authors extended the body of research in this area by testing four different variants of Graph Neural Networks (GNN) and a fully connected Multi-layer Perceptron (MLP) with dropout for the tasks of predicting the nature and timestamp of the next process activity.

On the other hand, [19] proposed a model that is composed of Convolutional Neural Network (CNN) to forecast the next business activity. The use of CNN in process mining was justified by guaranteeing high accuracy in predictions of next activity, and according to the authors, CNN has achieved faster training and inference than LSTM, even when the processes tend to have longer traces.

3 Proposed Approach

3.1 Example Business Process

For demonstration purpose in this paper, we focus on a generic business process for a procurement scenario illustrated in **Fig 1,** (The real work is also performed on similar scenarios provided by our project industrial stakeholders).

3.2 Simulate Business Process and Generate Event Logs

For this purpose, we used ProM 6.10[1], that is a well-known Process Mining tool developed by Developed by Process Mining Group, Math and CS department, Eindhoven University of Technology, and has several capabilities related to simulation of business processes, as well as log generation capabilities. Table **1**. provides a snapshot of a generated event log resulting from the simulation.

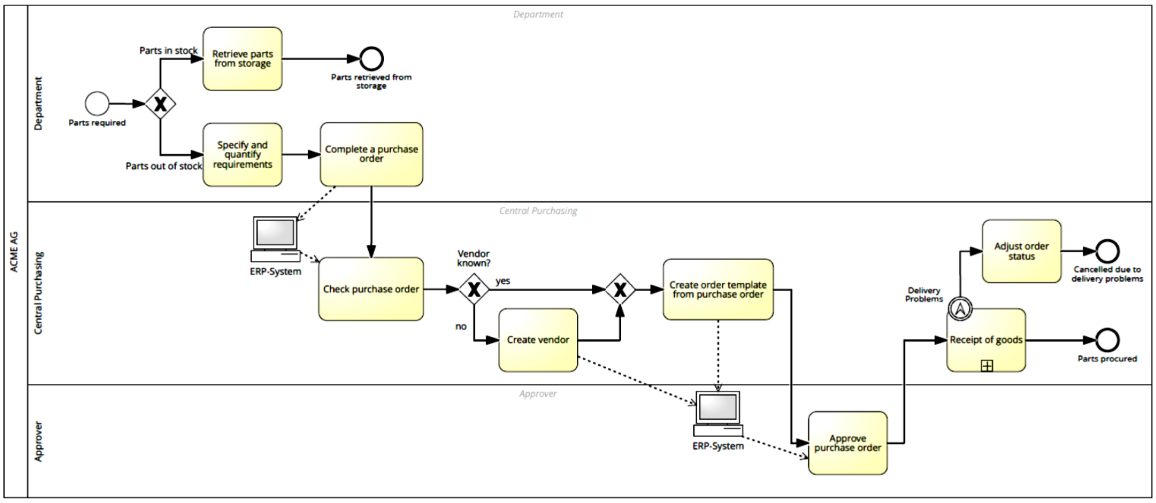

Fig. 1. Procurement Business Process

3.3 Architecture.

Fig **2**. Illustrates the architecture of our approach to predict the next activity in the procurement business process using LSTM , adopted from [20]. The approach is composed of three phases: 1) Preprocessing of Event log, 2) Classification and 3) Prediction.

Phase 1: Event Log Preprocessing:

This phase is divided into three steps:

Data Extraction:

In which we decide on which attributes we will select from the event log needed for the prediction, in our case the “Activity” attribute.

Trace Identification:

[1] http://www.promtools.org/doku.php?id=prom610

In this step we identify the traces representing different cases in the event log, each case is represented by the corresponding trace. Those traces are extracted and appended to a text file, preserving their order in the original log.

Segmentations:

After identifying the traces in the previous step, we *divide* each trace into a set of activities/events. Each event is then converted to a sequence of integers, comprising a two sequence lists of “integers”, the fist contains is the list of activities (X) and the

Table 1. Snapshot of a generated Event Log

Process ID	Activity	Timestamp
1	Parts Required	2010-12-30T11:02:00.000+01:00
1	Retrieve Parts from Storage	2010-12-31T10:06:00.000+01:00
1	Parts Retrieved from Storage	2011-01-05T15:12:00.000+01:00
2	Parts Required	2021-01-26T21:06:18.849+03:00
2	Specify Quantify Requirements	2021-01-26T21:12:18.849+03:00
2	Specify Quantify Requirements	2021-01-26T21:13:18.849+03:00
2	Complete PO	2021-01-27T21:23:03.707+03:00
2	Check PO	2021-01-27T21:25:03.707+03:00
2	….	…..
3	Parts Required	2021-01-30T22:18:32.555+03:00
3	Specify Quantify Requirements	2021-01-30T22:19:32.555+03:00
3	Complete PO	2021-01-31T22:28:32.555+03:00
3	Check PO	2021-02-01T22:14:41.262+03:00
3	Specify Quantify Requirements	2021-02-02T22:27:41.262+03:00
3	…….	……
4	Retrieve Parts from Storage	2021-02-22T23:03:01.460+03:00
4	Parts Retrieved from Storage	2021-02-22T23:04:01.460+03:00
4	…….	……
5	Parts Required	2021-03-26T23:00:32.002+03:00
5	Retrieve Parts from Storage	2021-03-26T23:00:38.002+03:00
5	…….	……

second of output activities (Y). Then, the sequence of input activities (X) is converted into a 2-dimensional matrix composed of both number of sequences and the maximum length of sequences.

Phase 2: Classification

In this phase we convert the sequence of integers of output activities (Y) obtained in the previous step into one hot encoding representation, identifying that the number of classes will be equal to the size of the vocabulary.

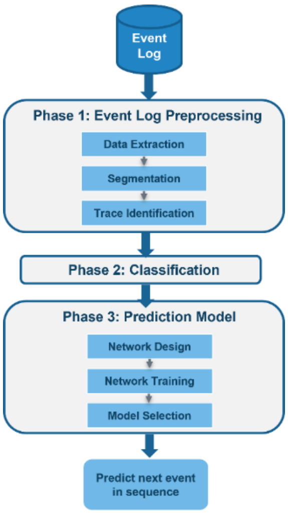

Fig. 2. Proposed Architecture

Phase 3: Prediction

This step is furtherly divided into three sub-steps:

Network Design:

In which we defined the design of the LSTM network in terms of input, hidden and output layers, along with specification of needed parameters.

Network Training:

In this step we train the LSTM network using dataset composed of the sequence list of integers represented by the activities in matrix (X), and the one hot representation of matrix (Y).

Model Selection:

Proceeding the training phase, in this step the model of LSTM with best results will be identified as the final model. Such 'best results' including high accuracy would be the best one to make accurate predictions. Otherwise, another iteration on training and adjustments of the parameters will be needed.

Prediction:

This is the output generated by the LSTM model, which is the prediction of the next activity in the in the business process stated earlier, from a single input activity or a series of activities.

3.4 Implementation.

Table **2** lists the parameters configured for the LSTM network. For training the network, we used an event log of with 300 traces and 60 different activities in the log. During the training phase, the number of sequences that were identified was 4321. To elaborate more on the ways LSTM network works, the network accepts as input a single activity or a sequence of activities, and based on a value specified by the user, a number of predicted activities will be performed. Table **3** lists the names of the activities captured in the business process, along with their acronyms for simplifying the representation of the activity name instead of the full name. For example, SAQR represents the activity "Specify and Quantify Requirements", and CPO is the acronym for "Complete Purchase Order". As part of the implementation, Keras [21], which is a python library that enables us to build deep learning models.

Table 2. LSTM Model Parameters with their configured values

Parameter	**Value**
Epochs	200
Optimizer	Adam
Loss	Categorial Cross entropy
LSTM Units	100
Batch size	20

Table 3. Activity Names with their acronyms

Activity Name	**Acronym**
Part Required	PR
Retrieve Parts from Storage	RPFS
Specify and Quantify Requirements	SAQR
Complete Purchase Order	CAPO
Check Purchase Order	CPO
Create Order Template from PO	COTFPO
Approve Purchase Order	APO
Receipt of Goods	ROG
Adjust Order Status	AOS

4 Results

Table **4** exhibits some of the obtained results from our proposed model. The first column in the table shows the input activity, which is the one that was passed to the LSTM model as an input to the prediction phase. The target activity is the anticipated activity (or activities), i.e., those activities with highest probabilities of the targeted prediction, according to the weight of each activity. Results showed that the model proposed was able to predict the next target activity with high precision, with 85% accuracy. Fig **3** depicts the training versus validation loss for the model, and **fig 4** illustrates the model accuracy.

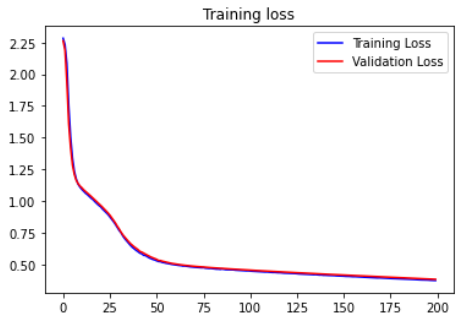

Fig. 3. Training vs. Validation Loss

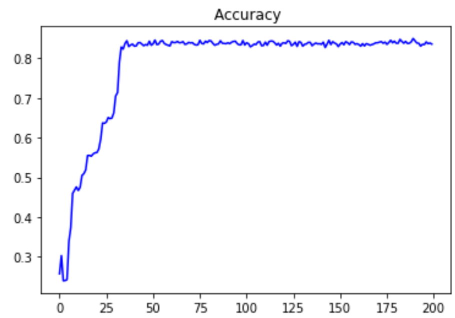

Fig. 3. Model Accuracy

Table 4. Sample results of Prediction

No.	Input Activity	Target Activity	Output Activity 1	Output Activity 2
1	**PR**	**SAQR|RPFS**	**SAQR**	**RPFS**
2	**CAPO**	**CPO|CV|**	**CPO**	
3	**COTFPO**	**APO|ROG|AOS**	**APO**	**ROG**
4	**CAPO**	**CPO|CV|COTPFO**	**CPO**	**CV**
5	**PR**	**RPFS**	**RPFS**	

5 Conclusion

Deep learning methods including LSTM can be used in predicting the next activity in a given business process model, with 85% accuracy. In this work, we utilized LSTM neural network model for such a task, by training it on traces of events extracted from event log of a simulated procurement business process. Such an approach can be used for process improvement and auditing business process models. As future work, we may test our proposed model on real world event data, and we may extend it to involve anomaly detection component, to detect anomalous events in the sequence of traces, namely, contextual anomalies using LSTM neural networks.

6 Acknowledgement:

This research is supported by Qatar University.

References

1. Saylam R., Sahingoz. O. K., "Process mining in business process management: Concepts and challenges," International Conference on Electronics, Computer and Computation (ICECCO), Ankara, 2013, pp. 131-134, doi: 10.1109/ICECCO.2013.6718246 (2013).
2. Sarno R., Wibowo W. A., Kartini, Haryadita F., Effendi Y., and Sungkono K., “Determining Model Using Non-Linear Heuristics Miner and Control-Flow Pattern,” TELKOMNIKA Telecommunication Computing Electronics and Control Journal, vol. 14 (1), (2016).
3. Sarno R., Pamungkas E. W., Sunaryono D., and Sarwosri, “Business process composition based on meta models,” International Seminar on Intelligent Technology and Its Applications (ISITIA), (2015).
4. Aalst W. M. P. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Heidelberg: Springer. (2011)
5. Ceci, M.; Spagnoletta, M.; Lanotte, P.F.; Malerba, D. Distributed learning of process models for next activity prediction. In Proceedings of the 22nd International Database Engineering & Applications Symposium (IDEAS 2018), Villa San Giovanni, Italy, 18–20 June (2018; Desai, V.B.C., Fresca, S., Zumpano, E., Masciari, E., Caroprese, L., Eds.; ACM: New York, NY, USA; pp. 278–282. (2018)
6. Pravilovic, S.; Appice, A.; Malerba, D. Process mining to forecast the future of running cases. In New Frontiers in Mining Complex Patterns—Second International Workshop; NFMCP, Revised Selected Papers, Ser. LNCS; Appice, A., Ceci, M., Loglisci, C., Manco, G., Masciari, E., Ras, Z.W., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; Volume 8399, pp. 67–81. (2013)
7. Aalst, W.M.P.; Schonenberg, M.H.; Song, M. Time prediction based on process mining. Inf. Syst. 36, 450–475. (2011)
8. Evermann, J.; Rehse, J.-R.; Fettke, P. Predicting process behavior using deep learning. Decision Support Syst, 100, 129–140. (2017).
9. Li Y., Cao H., “Prediction for tourism flow based on LSTM neural network,” Procedia Computer Science, vol. 129, pp. 277–283, (2018).
10. Cortez B., Carrera B. Kim ., Y.-J., and J.-Y. Jung, “An architecture for emergency event prediction using LSTM recurrent neural networks,” Expert Systems with Applications, (2017).
11. Liu F., Chen Z., and Wang J., “Video image target monitoring based on RNN-LSTM,” Multimedia Tools and Applications, pp. 1–18, (2018).

12. Aalst, Wil. Process discovery from event data: Relating models and logs through abstractions. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 8. e1244. 10.1002/widm.1244. (2018).
13. W. M. Van der Aalst, Process mining: data science in action. Springer, (2016).
14. Baier T., Mendling J., and Weske M., “Bridging abstraction layers in process mining,” Information Systems, vol. 46, pp. 123–139, (2014)
15. IEEE Task Force on Process Mining. IEEE 1849-2016 XES Standard Definition. 2016. Available online: https://www.xes-standard. org/ (accessed on 9 February 2023).
16. Greff K., Srivastava R. K., Koutník J., Steunebrink B. R., and Schmidhuber J., “LSTM: A search space odyssey,” IEEE transactions on neural networks and learning systems, vol. 28, no. 10, pp. 2222–2232, (2017).
17. Camargo, Manuel & Dumas, Marlon & González-Rojas, Oscar. Learning Accurate LSTM Models of Business Processes. 10.1007/978-3-030-26619-6_19. (2019).
18. Venugopal I., T¨ollich J., Fairbank M., Scherp A., “A Comparison of Deep-Learning Methods for Analysing and Predicting Business Processes” in Proceedings of International Joint Conference on Neural Networks, IJCNN, IEEE Press (2021).
19. Obodoekwe, E.; Fang, X.; Lu, K. Convolutional Neural Networks in Process Mining and Data Analytics for Prediction Accuracy. Electronics 2022, 11, 2128. https://doi.org/10.3390/electronics11142128 (2022).
20. Tello-Leal E., Roa J., Rubiolo M. and Ramirez-Alcocer U. M., "Predicting Activities in Business Processes with LSTM Recurrent Neural Networks," 2018 ITU Kaleidoscope: Machine Learning for a 5G Future (ITU K), Santa Fe, Argentina, pp. 1-7, doi: 10.23919/ITU-WT.2018.8598069. (2018)
21. Chollet, F. Keras, “Keras: The Python Deep Learning library,” https://github.com/fchollet/keras, (2018).