International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5946
***Article Received**: 15 October 2022 **Revised**: 24 November 2022 **Accepted**: 22 December 2022*

Deep Learning for User Behaviour Prediction Using Streaming Analytics

[1]Mantri Gayatri, [2]Dr. P. Satheesh, [3]Dr. R. Rajeswara Rao
[1]Research Scholar, JNTU, Kakinada, Andhra Pradesh, India
[2]CSE Department ,MVGR College of Engineering, Vizianagaram, Andhra Pradesh, India
[3]Professor ,CSE Department,JNTUKUCE,Vizianagaram, Andhra Pradesh, India
gayatricse312@gmail.com, patchikolla@yahoo.com,**rrraao.cse@jntukucev.ac.in**

Abstract
Streams of web user interactions reflect behaviour of customers or users of a web application through which a company is being operated online. The interactions may be in the form of visits to web components and even purchases made by users in case of e-Commerce applications. Modelling user behaviour can help the organizations to ascertain patterns of user behaviours and improve their products and services to meet their needs besides making promotional schemes. There are many existing methods for modelling user behaviour. However, of late, deep learning models are found to be more accurate and useful. In this paper a deep learning based framework is proposed for predicting web user behaviour from streams of user interactions. The framework is based on the mechanisms that exploit Recurrent Neural Network (RNN), one of the deep learning approaches, to learn from low-level features of sequential and streaming data. The mechanisms are used to model user interactions and predict the user behaviour with respect to purchasing items in future. In presence of plenty of items, item embeddings is explored for better results. In addition to this, attention mechanisms are employed to achieve RNN model interoperability. The empirical study revealed that the proposed framework is useful besides helping to evaluate different variants of attention mechanisms and item embeddings.

Index Terms – User behaviour analysis, deep learning, Recurrent Neural Network, item embeddings, attention mechanisms

I. INTRODUCTION

Web user behaviour analysis provides knowledge about the users' behaviours and possible future moves in a web application. In the wake of e-Commerce web applications that drive business of organizations, it is essential to understand the thinking patterns of customers. In the presence of streams of web user interactions, it is challenging to model user behaviour accurately [1]. Enterprises that make use of web based applications to drive home their businesses are constantly striving to improve their web design and services based on the behaviour of customers. Modelling customer behaviour indirectly provides the feedback from customers. Such feedback forms business intelligence from time to time leading to enhancements in customer satisfaction and rich user experience.

Many researchers contributed towards modelling user behaviour using machine learning techniques. For instance, Association Rule Mining (ARM) is used in [1] for prediction of users accessing web pages. Word embeddings concept is employed in [2] to model customer behaviour. Multi-user environment is explored in [3] and [4] for modelling user behaviour. Web user events are related to temporal domain. It does mean that the data related to web users is essentially time series data. Deep learning models came into existence to characterise web user behaviour. Deep learning based approaches are studied in [7], [8] and [12]. However, from the literature, it is found that deep learning model with different embeddings and attention mechanisms leads to better performance. This idea is used in this paper for proposing a deep learning based framework. Our contributions in this paper are as follows.

- A deep learning based framework is proposed for predicting web user behaviour from streams of user interactions since predicting future consumer behaviour is fundamental to many use-cases in e-commerce. In trials using substantial data sets, we show the benefits of RNNs and how they are useful in predictions of user behaviour.
- The proposed framework is enhanced with different kinds of embeddings and attention mechanisms for better performance so that the model has the ability to focus on important elements of the input sequence and subsequently learn the relationships between them.
- A prototype application is built to evaluate the proposed framework and the underlying deep learning and other mechanisms.

The remainder of the paper is structured as follows. Section 2 reviews literature on various techniques used for web user behaviour analysis. Section 3 describes the proposed deep learning based framework. Section 4 provides experimental setup details. Section 5 presents experimental results. Section 6 concludes the paper and gives directions for future work.

II. RELATED WORK

This section reviews literature on various techniques used for web user behaviour analysis. Geetharamani*et al*. [1] proposed an association rule mining based solution. They proposed an Apriori based algorithm known as Apriori Prefix Tree for user behaviour analysis. They used measures like lift, support and confidence to rank the rules. They found that news and sports attracted more number of online users. Boratto*et al*. [2] proposed a methodology for modelling user behaviour. First, they segmented users and then proposed a model based on neural word embedding. The kind of embedding is known as Neural Class Embedding. In future, they intended to incorporate semantic approaches in the analysis. Yan *et al*. [3] proposed a generative process model known as Multi-Site Probabilistic Factorization (MPF) to model user behaviour related to cross-site and same site. In future they intended to use neural networks to improve performance of their method. Yang *et al*. [4] on the other hand employed MPF for application in video recommendations. Buettner [5] proposed a framework known as Personality Based Product Recommender (PBPR) user behaviour related to online social networks (OSNs). Product recommendations are made based on user personalities. They intended to apply their framework to other kinds of web sites.

Jalal and Mahmood [6] proposed a methodology with socio temporal features in order to understand the student behaviour with respect to e-Learning. In future, they intended to work on virtual invigilators to monitor student behaviour. Alharbi and Doncker [7] proposed a deep neural network based solution based on Convolutional Neural Network (CNN) to analyse the opinions of users over Twitter posts. They intended to improve their model using Recurrent Neural Network (RNN) in future. Almeida and Azkune [12] used RNN and LSTM to achieve user behaviour analysis. They could predict user's next action. In future, they indented to use CNN along with LSTM for better performance. Armando [8] proposed robust classification method based on different approaches such as auto-encoder and Deep Belief Networks (DBNs). They employed the model to e-Commerce platform and found its utility. They intended to improve the time efficiency of their methods. They compared different methods that include Decision Tree (DT), K-Nearest Neighbour (KNN) and Long Short-Term Memory (LSTM). They found that LSTM showed highest accuracy.

Atta-ur-Rahman *et al*. [11] proposed a neuro-fuzzy approach for supervised learning of user behaviour prediction. They used temporal logs of user's interaction with web site as dataset and found that their approach performed well. They wanted to use different weights for each activity in future for improving their method. From the literature, it is ascertained that there is need for deep learning based framework for analysing user interactions with web pages and find the web user behaviour with respect to different activities of the web site.

III. PROPOSED DEEP LEARNIING BASED FRAMEWORK

This section presents the framework proposed for web user behaviour prediction based on the streams of user interactions. The study applies primarily RNN without any attention mechanisms and embeddings. To predict the user behaviours while visiting the pages or any through any interaction with the e-commerce sites the hidden vectors of the input sequence are identified. In the next stage it is incorporated with embeddings for each interaction of user and the item using Word2vec, linear, non-linear, joint and predicting embeddings. For the better interpretation of prediction models along with embeddings, the attention mechanisms (linear and non linear attention weights) were applied with RNN. However RNN is not efficient in modelling temporal domain, LSTM is introduced to overcome this problem as well as for the better classification of streaming. This analysis is performed on two datasets namely movielens [24] and Santander [25] product recommendation using prototype application built using Python data science platform.

3.1 Problem Definition

Let u_n represents a web user while his/her past interactions are represented by $x^{(n)}= (x_1, x_2, ..., x_T)$. In a sequence of elements each sequence is denoted as x_i. A set of interactions in general is denoted as RD. the interactions of an e-Commerce application are considered. They include views to web pages (products), purchases and removal of items from portfolio. Predicting customer/user behaviour with respect to product purchase, denoted as $y^{(n)}{}_{T+1}$, in future is to be made provided the input sequence denoted as $x^{(n)}$.

3.2 Recurrent Neural Network

In the proposed prediction model RNN is used first without attention mechanisms and embedding. RNN is used to have the baseline prediction model to know the user behaviour with respect to purchases, page visits and removal of items from the portfolios in case of e-Commerce web sites.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5946
***Article Received**: 15 October 2022 **Revised**: 24 November 2022 **Accepted**: 22 December 2022*

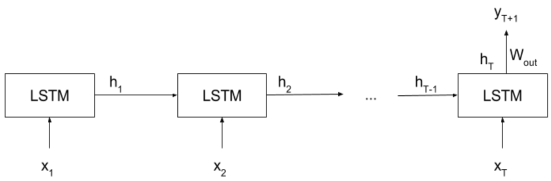

Figure 1: Recurrent Neural Network (RNN) model

As shown in Figure 1, the input sequence is denoted as $x=(x_1,x_2,...,x_T)$ while each element is denoted as x_t. Each element is nothing but an encoding vector. Every element is given to LSTM block thus resulting in a vector h_t which is hidden. It is shown in Eq. 1.

$$h_t = LSTM(x_t, h_{t-1}) h_t = LSTM(x_t, h_{t-1}) \quad (1)$$

The vector's hidden state is acquired using Eq. 2 and Eq. 3.

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \quad (2)$$

$$h_t = o_t .\tanh(C_t) h_t = o_t .\tanh(C_t) \quad (3)$$

After obtaining the hidden state vector h_T along with all elements in the sequence, the predicted probabilities are obtained as in Eq. 4.

$$yt + 1 = g(W_{out} h_T + b_{out}) yt + 1 = g(W_{out} h_T + b_{out}) \quad (4)$$

Each sample can have a valid class and there is need for dealing with multiple classes to solve the problem. A soft max function is required. Then the weights of the model are optimized as in Eq. 5.

$$L = \frac{1}{N}\sum_{n=1}^{N} yt + 1^{log}\ (yt+1)$$
$$L = \frac{1}{N}\sum_{n=1}^{N} yt + 1^{log}\ (yt+1) \quad (5)$$

In case of dealing with multiple classes and there each sample can have multiple valid classes, the optimization is made using Eq. 6.

$$L = \frac{1}{N}\sum_{n=1}^{N} yt + 1^{log}\ (yt+1) + (1 - yt + 1) log(1 - yt + 1)$$
$$L = \frac{1}{N}\sum_{n=1}^{N} yt + 1^{log}\ (yt+1) + (1 - yt + 1) log(1 - yt + 1) \quad (6)$$

After working with RNN without embeddings, it is combined with embeddings using Word2vec, linear, non-linear, joint and predicting embeddings.

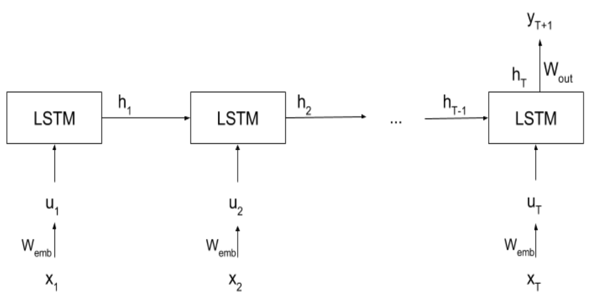

Figure 2: RNN with embeddings

As presented in Figure 2, it is based on RNN along with embeddings described in Section 3.3. Unlike in the Figure 1, it shows embeddings of different kinds for users and items. It also works for temporal domain as LSTM is used in the process.

3.3 RNN with Word2vec Embeddings

In Natural Language Processing, word embeddings are widely used. They are suitable for user behaviour prediction also for two reasons. First, there are number of online users for e-Commerce applications and they are to be used for analysis. Second, there are number of items involved in the e-Commerce applications and they are to be dealt with. Word2vec is widely used NLP tool for various real word applications where text corpus needs to be processed. When RNN is used with Wrod2vec it makes the prediction model more effective.

3.4 RNN with Embeddings (Jointly with Classification Model)

In this case embedding representations are made by using a classification model jointly. Here a linear function is used for learning embeddings. The function is as shown in Eq. 7.

$$u_t = W_{emb}{}^{x}{}_{t} \quad (7)$$

in the same fashion, non-linear embeddings representation is as in Eq. 8.

ut = tanh (Wembxt + bemb) (8)

The results of embeddings are given to LSTM along with the RNN baseline model used for the prediction.

3.5 RNN with Embeddings (Jointly Fine-tuned)

When embeddings are made with classification models from the beginning, it is costly and complex in nature. To overcome this problem, the jointly fine-tuned model makes use of pre-train embedding representations.

3.6 RNN with Embeddings (Separately Learned)

This model learns embeddings separately. It predicts the embedding of the product that is going to be purchased by the customer. Instead of finding probabilities for all items, it finds next item used by the customer. This embedding is computed as in Eq. 9.

^yT+1 = WouthT + bout (9)

The embedding size is M in this case and the for dimensionality of yT+1.

3.7 Attention Mechanisms

Different attention mechanisms are used to in the empirical study along with RNN. They are known as RNN along with linear attention weights with hidden states, RNN along with non-linear attention weights with hidden states, RNN with

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5946
***Article Received**: 15 October 2022 **Revised**: 24 November 2022 **Accepted**: 22 December 2022*

embeddings and attention with linear approach and RNN with embeddings along with attention weights with non-linear approach. These mechanisms are used for interpretation of prediction models.

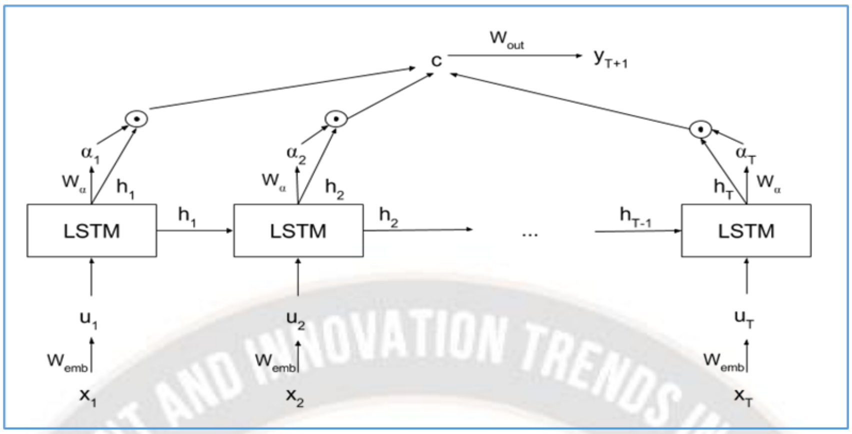

Figure 3: RNN (LSTM) with hidden states having linear and non-linear models

As presented in Figure 3, the LSTM is used along with hidden states and the linear and non-linear models. Attention to the hidden states is used for better interpretations. Different weights are used for model interpretations. The focus of the given model in hidden states is the main focus of the weights. A weight is denoted as α_t. First energy e_t is computed with both linear and non-linear approaches as shown Eq.10 and Eq. 11.

$$e_t = W_\alpha h_t + b_\alpha \quad (10)$$

$$e_t = \tanh(W_\alpha h_t + b_\alpha) \quad (11)$$

Attention to the embeddings is also considered in this paper. By using this approach, the focus is exclusively on various elements in the sequence.

IV. EXPERIMENTAL SETUP

The environment used for empirical study is the Anaconda, the Python data science platform. The datasets used for experiments are Movielens dataset and Santander Product Recommendation dataset. The former contains movie ratings while the latter contains views, purchases and removal of items from portfolios as part of user behaviour prediction. For the movielens dataset [24] Eq. 12 shows the evaluation measure.

R-Precision = # positives top R / R (12)

Similarly, for the Santander product recommendation dataset, the Eq. 13 is used for evaluation.

recall@k = # of TPs in the top k predictions / # of TPs (13)

Here the k value is considered to be 3. It does mean that minimum number of purchases as part of user behaver is 3. Tensor Flow is used for deep learning based study.

V. EXPERIMENTAL RESULTS

Experiments are made with the prototype application built using Python data science platform. The results are observed with the RNN along with different embeddings and attention methods. The deep learning based approaches are thus verified with two datasets namely movielens and Santander product recommendation.

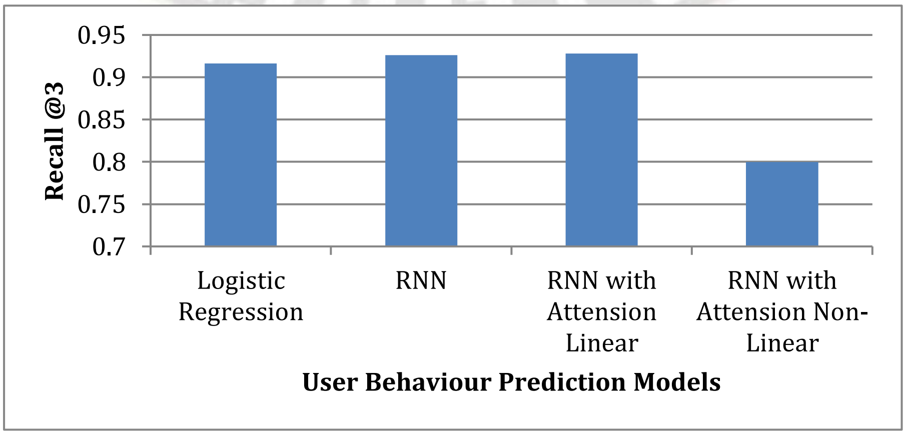

Figure 4: Results of Santander dataset

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5946
***Article Received**: 15 October 2022 **Revised**: 24 November 2022 **Accepted**: 22 December 2022*

As presented in Figure 4, the user behaviour prediction models are compared for recall@3. The prediction performance in terms of Recall@3 is observed for all the methods. The prediction performance is high with deep learning models when compared with the logistic regression model. However, RNN with linear attention model is better than baseline RNN. When RNN is used with non-linear attention model, it showed least performance.

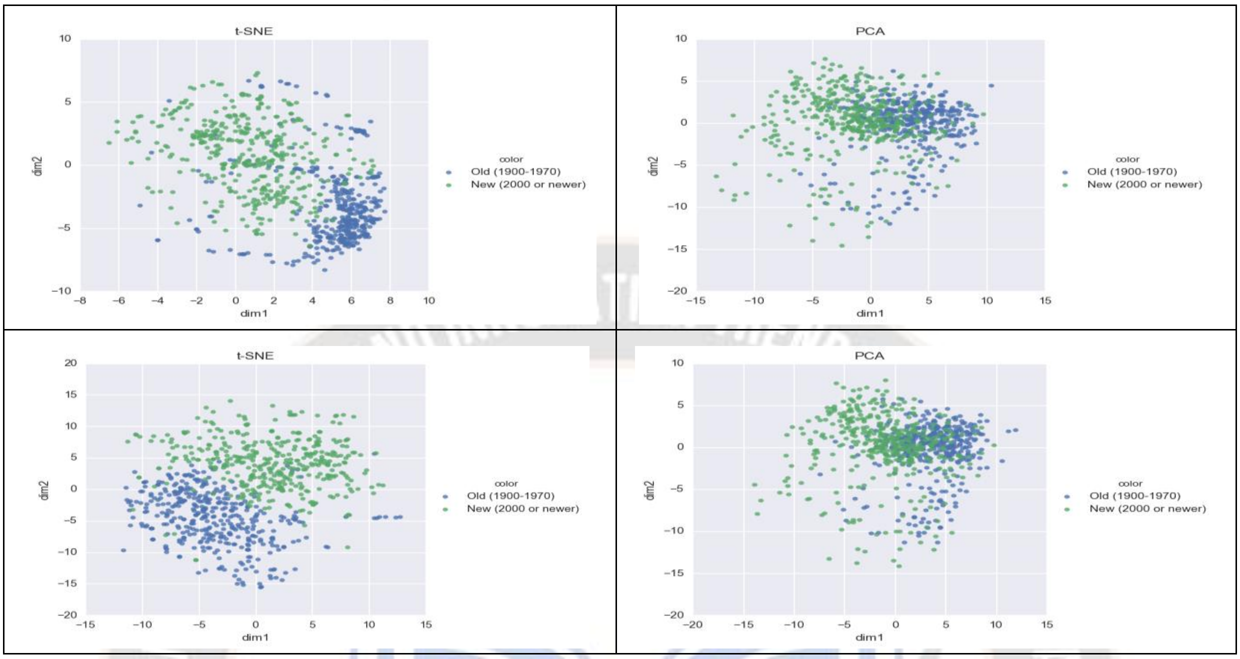

Figure 5: Representation of embeddings with movielens dataset

As presented in Figure 5, the t-SNE and PCA are used for dimensionality reduction. The 64 dimensions are decreased to two dimensions. The results of RNN with embeddings using word2vec are presented at top left. RNN with embeddings made using joint and linear approach is presented at top right. The results of RNN with embeddings made jointly with non-linear approach is provided at bottom left while the RNN with embedding with word2ec fine-tuned is presented at bottom right.

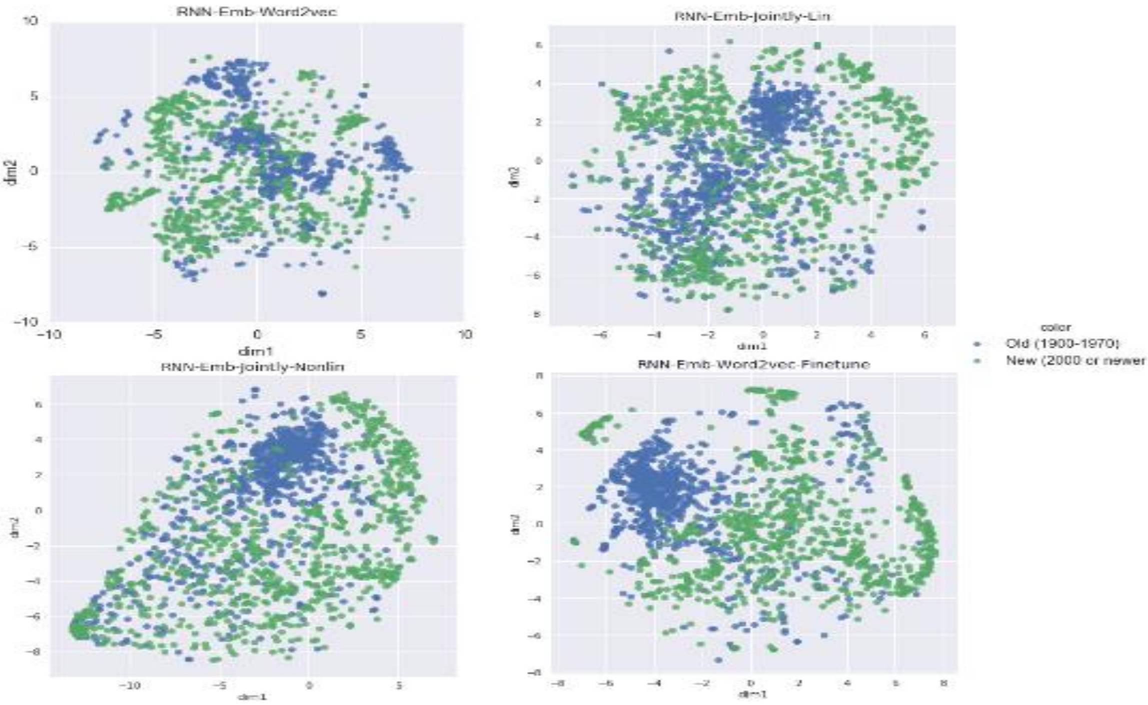

Figure 6: RNN with embedding representations with movielens dataset

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5946
***Article Received**: 15 October 2022 **Revised**: 24 November 2022 **Accepted**: 22 December 2022*

As presented in Figure 6, the results are 2 dimensions made by t-SNE. RNN embeddings with word2vec is shown at top left. RNN embeddings with joint optimization with linear approach is shown at top right. RNN embeddings with joint optimization with non-linear approach is shown at bottom left while the RNN embeddings with word2vec fine-tuned is shown at bottom right.

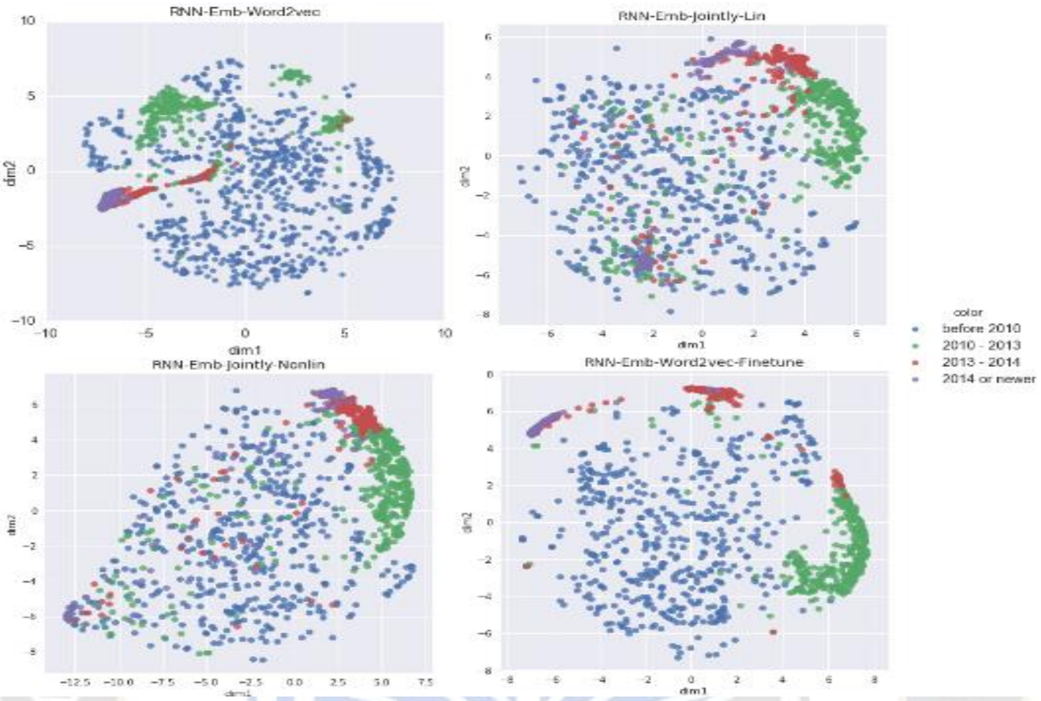

Figure 7: RNN with embedding representations with movielens dataset (year wise)

As presented in Figure 7, the results are 2 dimensions made by t-SNE. RNN embeddings with word2vec is shown at top left. RNN embeddings with joint optimization with linear approach is shown at top right. RNN embeddings with joint optimization with non-linear approach is shown at bottom left while the RNN embeddings with word2vec fine-tuned is shown at bottom right.

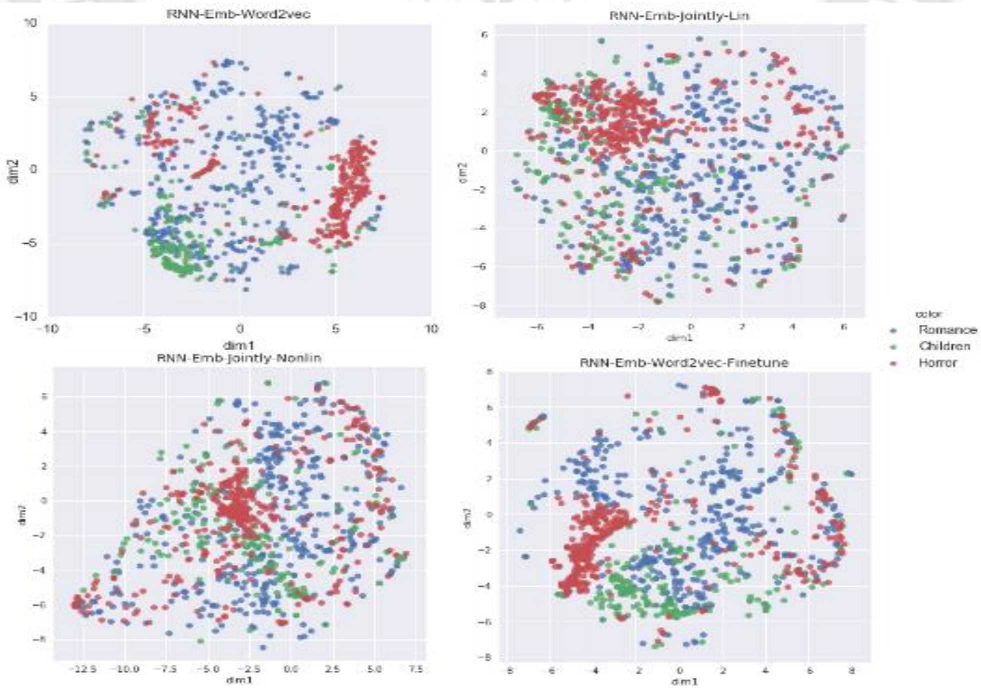

Figure 8: RNN with embedding representations with movielens dataset (genre wise)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5946
***Article Received**: 15 October 2022 **Revised**: 24 November 2022 **Accepted**: 22 December 2022*

As presented in Figure 8, the results are 2 dimensions made by t-SNE. RNN embeddings with word2vec is shown at top left. RNN embeddings with joint optimization with linear approach is shown at top right. RNN embeddings with joint optimization with non-linear approach is shown at bottom left while the RNN embeddings with word2vec fine-tuned is shown at bottom right.

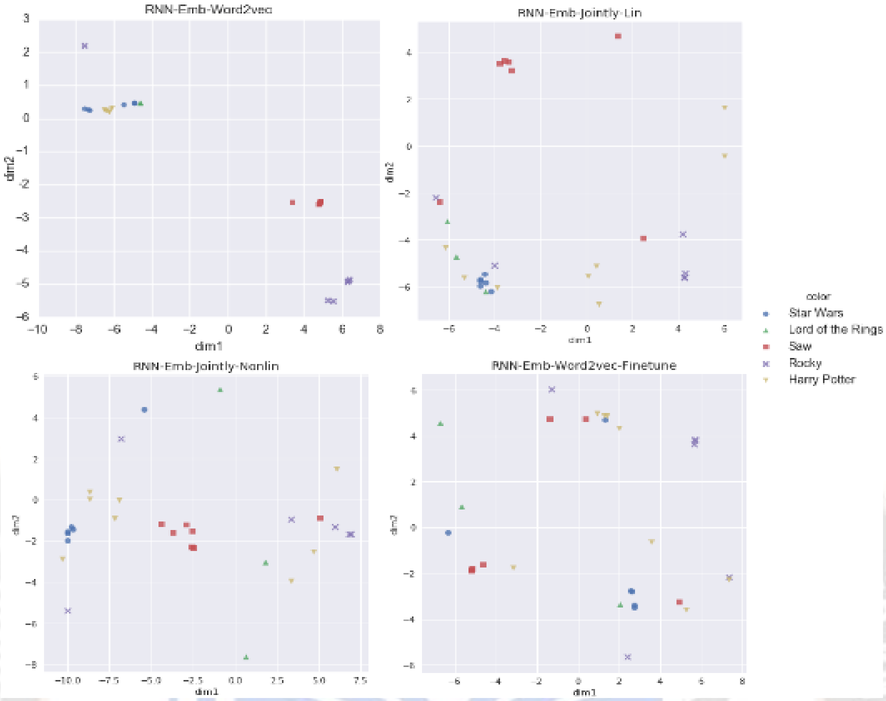

Figure 9: RNN with embedding representations with movielens dataset (sagas wise)

As presented in Figure 9, the results are 2 dimensions made by t-SNE. RNN embeddings with word2vec is shown at top left. RNN embeddings with joint optimization with linear approach is shown at top right. RNN embeddings with joint optimization with non-linear approach is shown at bottom left while the RNN embeddings with word2vec fine-tuned is shown at bottom right.

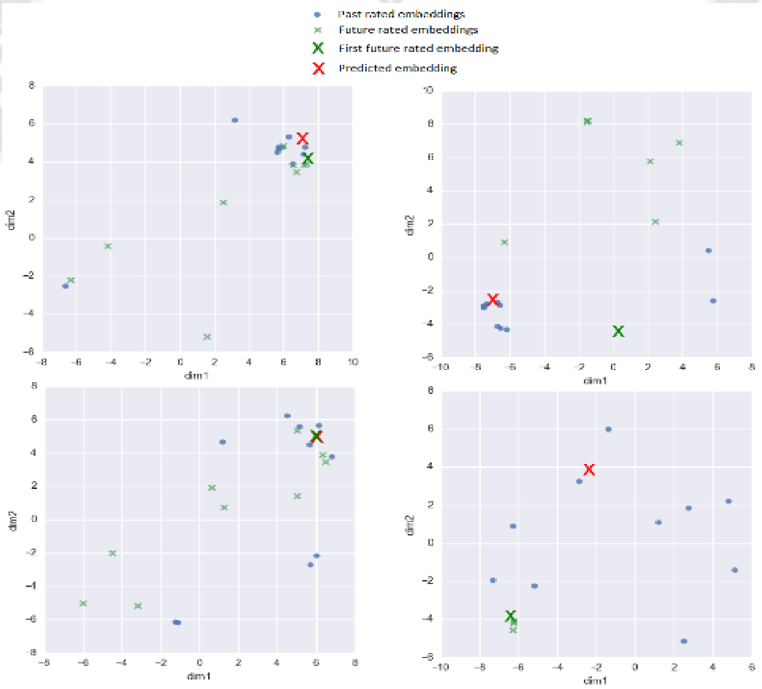

Figure 10: Predicted embeddings with movielens dataset

As presented in Figure 10, t-SNE is used to reduce dimensions from 64 to 2. The last 10 rated movies are shown in blue colour. The predicted embeddings are represented by big red cross while the green crosses reflect the embeddings pertaining to first 5 movies rated by uses. The big green cross indicates the first embedding reflecting the firstly ranked movie. The examples where at least one good prediction exists are shown left while the examples where incorrect predictions exist are shown at right side.

VI. CONCLUSION AND FUTURE WORK

In this paper we proposed a deep learning based framework with an underlying mechanism that exploit RNN to predict user behaviour from streams of user interactions (historical data). The predictions are to reflect the visiting and purchasing decisions of users in an e-Commerce case study. RNN is one of the deep learning methods that is found to be effective in predicting customer interactions with a web application. From the empirical study it is understood that RNN shows better performance over other models such as logistic regression. RNN is also evaluated with user and item embeddings and found to have performance improvement. Word2vec is used as part of empirical study and found that it represented the data very clearly. Afterwards, attention mechanisms are used to describe the results of the prediction models. Results also revealed that RNN is very useful for modelling sequential data. However, RNN is not efficient in modelling temporal domain. To overcome this problem, we used Long Short-Term Memory (LSTM). With LSTM the time related findings are made possible. In future, we intend to improve the deep learning framework proposed in this paper with joint optimization with linear and non-linear models for predicting user behaviour more efficiently.

References

[1]. GEETHARAMANI, R., REVATHY, P., & JACOB, S. G. (2015). *Prediction of users webpage access behaviour using association rule mining. Sadhana, 40(8), 2353–2365.*

[2]. Boratto, L., Carta, S., Fenu, G., & Saia, R. (2016). *Using neural word embeddings to model user behavior and detect user segments. Knowledge-Based Systems, 108, 5–14.*

[3]. Yan, H., Yang, C., Yu, D., Li, Y., Jin, D., & Chiu, D.-M. (2019). *Multi-site User BehaviorModelling and Its Application in Video Recommendation. IEEE Transactions on Knowledge and Data Engineering, 1–14.*

[4]. Yang, C., Yan, H., Yu, D., Li, Y., & Chiu, D. M. (2017). *Multi-site User BehaviorModelling and Its Application in Video Recommendation. Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '17.* P1-10.

[5]. Buettner, R. (2016). *Predicting user behavior in electronic markets based on personality-mining in large online social networks. Electronic Markets, 27(3), 247–265.*

[6]. Jalal, A., & Mahmood, M. (2019). *Students' behavior mining in e-learning environment using cognitive processes with information technologies. Education and Information Technologies.* P1-25.

[7]. MAlharbi, A. S., &Doncker, E. (2018). *Twitter Sentiment Analysis with a Deep Neural Network: An Enhanced Approach using User Behavioral Information. Cognitive Systems Research.* P1-36.

[8]. Armando. (2016). Predicting online user behaviour using deep learning algorithms, p1-21.

[9]. Román, P. E., Velásquez, J. D., Palade, V., & Jain, L. C. (2013). *New Trends in Web User Behaviour Analysis. Studies in Computational Intelligence, 1–10.*

[10]. Maghdid, H. S. (2018). *Web News Mining Using New Features: A Comparative Study. IEEE Access, 1–17.*

[11]. Atta-ur-Rahman, Dash, S., Luhach, A. K., Chilamkurti, N., Baek, S., & Nam, Y. (2019). *A Neuro-fuzzy approach for user behaviour classification and prediction. Journal of Cloud Computing, 8(1).* P1-15.

[12]. Almeida, A., &Azkune, G. (2018). *Predicting Human Behaviour with Recurrent Neural Networks. Applied Sciences, 8(2),* p1-13.

[13]. Sisodia, D. S., Khandal, V., &Singhal, R. (2016). *Fast prediction of web user browsing behaviours using most interesting patterns. Journal of Information Science, 44(1), 74–90.*

[14]. Yunsheng, G., Qianqian, Z., &Jie, K. (2018). Research on the prediction of user behavior based on neural network. Proceedings of the 3rd International Conference on Intelligent Information Processing - ICIIP '18. P1-6.

[15]. Ortigosa, A., Carro, R. M., &Quiroga, J. I. (2014). *Predicting user personality by mining social interactions in Facebook. Journal of Computer and System Sciences, 80(1), 57–71.*

[16]. Gokulapriya R And Ganesh Kumar R.. (2019). Research Aligned Analysis on Web Access Behavioral Pattern Mining for User Identification. *International Journal of Engineering and Advanced Technology*. 8 (6), p1-6.

[17]. Asiya N. Khan And Pallavi S. Pandhare. (2018). Web Usage Mining and User Behaviour Prediction. *International Journal on Future Revolution in Computer Science & Communication Engineering* . 4 (4), p30 – 33.

[18]. Lakshmi N And Krishnamurthy . (2019). Social Networking Data Research Using Frequent Pattern Mining and Machine Learning Data. *International Journal of Engineering and Advanced Technology*. 8 (6), p1-6.

[19]. Beatric , Ryan Fernandes, Leo. J. Peo, NikhilaKamatAndSergius Miranda. (2014). New Approaches to Web Personalization Using Web Mining

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 2s
DOI: https://doi.org/10.17762/ijritcc.v10i2s.5946
***Article Received**: 15 October 2022 **Revised**: 24 November 2022 **Accepted**: 22 December 2022*

Techniques. *International Journal of Computer Science and Information Technologies*. 5 (2), p 2195-2201.

[20]. Liu, Y., Wang, C., Zhang, M., & Ma, S. (2017). *User behaviormodeling for better Web search ranking. Frontiers of Computer Science, 11(6), 923–936.*

[21]. ZhabizGharibshah, Xingquan Zhu, Arthur Hainline and Michael Conway. (2020). Deep Learning for User Interest and Response Prediction in Online Display Advertising. SPRINGER. p1-15.

[22]. Alessandro Crivellari and Euro Beinat. (2020). LSTM-Based Deep Learning Model for Predicting Individual Mobility Traces of Short-Term Foreign Tourists. Sustainability. 12, p1-18.

[23]. Atta-ur-Rahman, Dash, S., Luhach, A. K., Chilamkurti, N., Baek, S., & Nam, Y. (2019). A Neuro-fuzzy approach for user behaviour classification and prediction. Journal of Cloud Computing, 8(1). p1-15.

[24]. https://www.kaggle.com/datasets/grouplens/movielens-20m-dataset

[25]. https://www.kaggle.com/c/santander-product-recommendation/data