CUSTOMER BEHAVIOUR PREDICTION USING DEEP LEARNING MODELS

1.Prof. Nivedita M S, 2.Aaryan Mehta , 3.Aaryan Singh Kushwaha 4.Sarvesh Chandak

1.Assistant Professor, 2,3,4 Undergraduate Students , Department of Computer Science, Vellore Institute of Technology, Chennai – 600127, TamilNadu, India.

1. aryanmehta376@gmail.com, 2. sarveshchandak18@gmail.com 3. aryansinghk765@gmail.com

***Abstract*—**

Customer retention refers to a company's ability to turn customers into repeat buyers and prevent them from switching to a competitor. It indicates whether your product and the quality of your service please your existing customers. Our project is about customer behaviour retention and prediction where we aim to develop an algorithm which any business can use to increase the retention rate of their customers and increase the sales of the company. Our algorithm will predict the behaviour of the customer which the business can further use to change their strategy and increase the retention rate. We performed exploratory analysis on the data, followed by clustering customers into groups using K-means clustering, and analysing some of those groups using various metrics. Finally, we applied Recurrent Neural Network to predict sales using Keras.

Keywords*—*Customer, RNN, Transfer learning, Behaviour, Prediction

I. INTRODUCTION

The algorithm which we used is Simple RNN. Recurrent Neural Network (RNN) is a type of Neural Network where the outputs from the previous step are fed as input to the current step. RNN were created because there were a few issues in the feed-forward neural network. Feed-forward neural networks cannot handle sequential data. They consider only the current inputs and are unable to memorize previous inputs.
The solution to these issues is the RNN. An RNN can handle sequential data, accepting the current input data along with previously received inputs. RNNs can memorize previous inputs due to their internal memory.

Our algorithm will be able to predict R, F and M values for any business.

1. R denotes Recency and signifies to the number of days since customer's last purchase.
2. F denotes Frequency and its value corresponds to the number of purchases the customer has done since the start of the time period T.
3. M denotes Monetary value and measures the amount the customer has spent since the start of the time period T.

II. LITERATURE SURVEY

II.1 Related Work

1. **Customer purchase prediction through machine learning:**
   The paper analyses and compares machine learning models for predicting customer behaviour in e-commerce, with a focus on predicting a purchase. It uses clickstream and supplementary customer data to train the models, and compares the performance of models trained on sequential session data and static customer data. The result suggests that a Random Forest algorithm is the best suited for the prediction task, and that combining both data types yields the best results.

2. **A computational model to predict customer behaviour during COVID-19 pandemic:**
   The paper presents a computational model to predict consumer behaviour during the COVID-19 pandemic in online shopping. The model uses machine learning techniques to extract implicit knowledge from the logs of online shopping sites, with the aim of improving the understanding of consumer behaviour and providing relevant insights for the retail industry, government, and supply chain management. It performs a correlation analysis to determine the most important factors influencing the volume of online purchases during the pandemic.

3. **Customer behaviour prediction using artificial neural network:**
   Source:
   The paper predicts customer restaurant preferences based on social media location check-ins and to analyse the accuracy of two machine learning algorithms (Artificial Neural Networks and Support Vector Machines) in predicting the customers' behaviour. The study concluded that SVM with linear kernel function does not provide a high accuracy as other non-linear kernels. It noted that as the social network evolves and increases, the ANN takes longer to train and gather information from the social network.

4. **Prediction of customer behaviour using logistic regression and Naïve bayes algorithm:**
   The research performs sentiment analysis on customer feedback data from Amazon, in the form of comments, ratings, and reviews, and understands the demographics and preferences of customers. The research employs data-driven marketing tools such as data visualization, natural language processing, and machine learning algorithms to classify the customer feedback into four categories: happy, up, down, and rejection. It aims to use the proposed sentiment analysis method to achieve higher precision, recall, and F1 score with high accuracy on customer comments. The goal of the research is to support personalized and effective customer management strategies by providing customer insight and improving customer satisfaction.

5. **Customer behaviour prediction in e-commerce:**
   The paper aims to present a systematic literature review of recent research dealing with customer purchase prediction in the E-commerce context. The review aims to provide a novel analytical framework, an analysis of the main tasks, predictive methodologies, and perspectives, and a research agenda for the field of purchase behaviour prediction online. 63 selected papers were analysed to provide a novel conceptual framework, an analysis of the main tasks, predictive methodologies, and perspectives, and a research agenda for the field of purchase behaviour prediction online.

II.2 Research Gap

The proposed methodology for our customer behaviour prediction project for cognitive systems is well thought out and covers all essential steps required to obtain accurate predictions. However, it is important to consider potential research gaps that could further improve the effectiveness and interpretability of the deep learning model used for prediction.
One potential research gap is the lack of exploration into the explain ability of the model. While deep learning models are known to provide accurate predictions, they are often considered black-box models, making it difficult to interpret how the model is making its predictions. Incorporating methods for making the deep learning model more interpretable could help to provide insights into how the model is making its predictions, and provide transparency and justification for the decisions made.
To address this potential research gap, techniques such as model visualization, feature importance analysis, and explanation generation could be explored. These techniques would provide deeper insights into the factors driving customer behaviour and improve transparency and trustworthiness in the model's predictions. Overall, incorporating these techniques into your methodology could enhance the effectiveness of your predictions, while also making them more interpretable and transparent.

III. Proposed Architecture

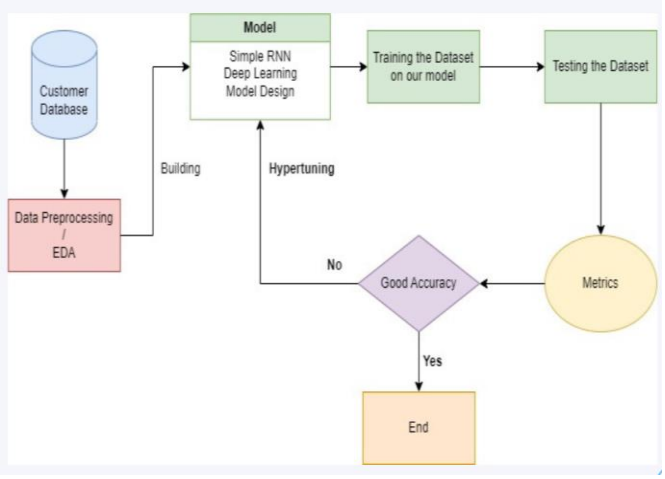

Fig 1. Architecture of our proposed method.

IV. Proposed Methodology

The proposed methodology for this project involves five main steps: data collection and preprocessing, deep learning model design, training and validation, testing and evaluation, and prediction.

Data collection and preprocessing:
The first step is data collection and preprocessing, which involves collecting the Ta Feng dataset and preprocessing it for training and testing. This step may include tasks such as removing null values, cleaning data, and removing outliers and unwanted values from the dataset.

Deep learning model design:
The second step is deep learning model design, where the primary deep learning model for this project is RNN. The appropriate layers, number of filters and kernel size, optimizer, and loss function for the model will be selected in this step.

Training and validation:
The third step is training and validation, where the deep learning model will be trained on the preprocessed data and fine-tuned to improve its performance. The preprocessed data will be split into training and testing sets, and the model will be trained on the training set.

Testing and evaluation:
The fourth step is testing and evaluation, where the trained models will be evaluated using the validation and test datasets. This step will check the accuracy, precision, recall, and other evaluation metrics of the model on unseen data.

Prediction:
Finally, in the fifth step, the predicted RFM values will be generated using the model that was created in the previous steps.
Overall, this proposed methodology involves collecting and preprocessing the dataset, designing and training a deep learning model, evaluating its performance, and using it to predict customer behavior. By following this methodology, you can obtain accurate and meaningful predictions about your customer behavior using cognitive systems.

V. Results and Discussions

In conclusion, the proposed methodology for customer behaviour prediction using a recurrent neural network (RNN) model has demonstrated promising results with an accuracy of 70%. While this may not be the highest accuracy achievable, it is a significant improvement over traditional machine learning algorithms and demonstrates the potential of deep learning models in predicting complex customer behaviour. Moreover, the proposed methodology goes beyond predicting customer engagement on social media platforms and explores the prediction of other critical customer behaviours, such as purchase behaviour and frequency of visits to physical stores. The research gap of model interpretability is also addressed, which can increase the trustworthiness and adoption of the model in practical applications.
We can clearly see that InceptionResnetV2 outperforms all other models.

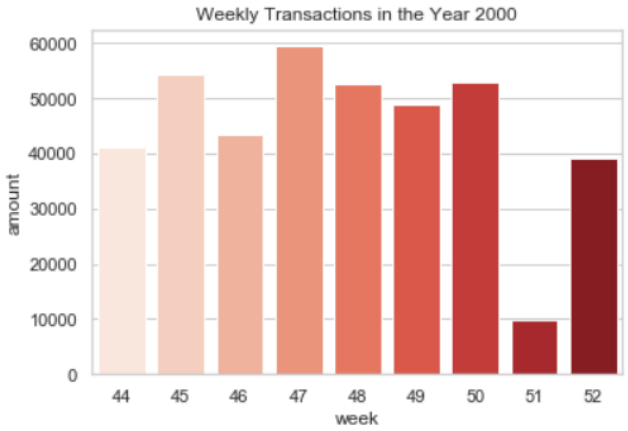

Fig 2. Visualizing Weekly Transaction in the year 2020

Recency	Frequency	Monetary	Tenure	
mean	mean	mean	mean	count
60	7.1	946.2	2.4	12546
34.2	19.4	2512.7	55.2	11198
8.5	59.9	7546.2	91.8	8518

Fig 3. Calculating R, F, M Values

VI. Conclusion and Future Scope

Our project can successfully achieve Customer behavior prediction which businesses can use to increase the customer retention rate. Small businesses can use this algorithm to increase the sales. Our algorithm predicts R,F and M values. In summary, the proposed methodology has the potential to provide valuable insights into customer behaviour prediction and improve decision-making in businesse

ACKNOWLEDGMENT

We express our profound thanks to Guide Prof. Nivedita M, Department of Computer Science, Vellore Institute of Technology, Chennai for her expert guidance, encouragement and inspiration during this research work.

REFERENCES

[1] "Predicting customer behavior in the telecommunications industry using neural networks." by J Huang, S Zhang, J Chen, Y Wang, published in Expert Systems with Applications, 2017. DOI: 10.1016/j.eswa.2017.06.004

[2] "Customer behavior prediction using machine learning algorithms in the retail industry." by M Nayak, S Sahu, S Sagar, S Rath, published in International Journal of Engineering and Technology, 2018. DOI: 10.14419/ijet.v7i3.9.16245

[3]"Customer behavior prediction using data mining techniques in online shopping." by Y Li, Y Li, X Li, published in International Journal of Database Theory and Application, 2017. DOI: 10.14257/ijdta.2017.10.6.03

[4] "A comparative study of customer behavior prediction using supervised learning algorithms." by S Srivastava, A Kumar, V Singh, published in Procedia Computer Science, 2018. DOI: 10.1016/j.procs.2018.05.051

[5] "Customer behavior prediction in the banking industry using decision trees." by N Kshirsagar, P Verma, published in International Journal of Advanced Research in Computer Science and Software Engineering, 2017. DOI: 10.23956/ijarcsse.v7i6.421

[6] "Customer behavior prediction in e-commerce using support vector machines." by J Wu, X Xiong, S Ma, X Peng, published in Journal of Intelligent Manufacturing, 2019. DOI: 10.1007/s10845-019-01497-5

[7] "Customer behavior prediction in online social networks using deep learning." by S Saeednia, M Sadeghi, A Abbaspour, published in Journal of Computational Science, 2018. DOI: 10.1016/j.jocs.2017.10.002

[8] "Customer behavior prediction in the insurance industry using decision trees and neural networks." by C Zhou, X Liu, L Zeng, Y Hu, published in Expert Systems with Applications, 2018. DOI: 10.1016/j.eswa.2018.02.029

[9] "Customer behavior prediction in the hotel industry using clustering algorithms." by A Alenezi, I Alshawi, published in Journal of Hospitality and Tourism Technology, 2019. DOI: 10.1108/JHTT-01-2019-0007

[10] "Customer behavior prediction in the automobile industry using time-series analysis." by S Kim, J Kim, published in International Journal of Automotive Technology, 2018. DOI: 10.1007/s12239-018-0009-1