Improving the Click Prediction for Online Advertisement with the Integration of Recommended System using Neural Network Architecture

MSc Research Project
Data Analytics

Jayalakshmi Jayachandran
Student ID: x20213557

School of Computing
National College of Ireland

Supervisor: Abubakr Siddig

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Jayalakshmi Jayachandran

Student ID: X20213557

Programme: Msc in Data Analytics **Year:** 2021-2022

Module: Research Project

Supervisor: Abubakr Siddig

Submission Due Date: 15-Aug-22

Project Title: Improving the Click Prediction for online advertisement with Integration of recommended system using Neural Network Architecture

Word Count: 7717 **Page Count** 26

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ……………………………………………………………………………………………………

Date: ……………………………………………………………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Improving the Click Prediction for Online Advertisement with the Integration of Recommended System using Neural Network Architecture

Jayalakshmi Jayachandran
X20213557

Abstract

Nowadays shopping for products through online platforms is highly inclining due to better internet and technologies across the globe. Online advertisements play an important role in such business strategies. It provides the access to customers to view a large variety of the products of their requirements. Customers make purchases based on the rating of products. Every day, millions of customers are posting their reviews online about the products. In this business strategy clicking on online advertisements is an important factor because the more clicks lead the more probability of making a purchase. A noble idea of online advertising based on the recommendation system can be more effective in such scenarios but still, it is a tedious task. Prediction of clicks on the online advertisements shown based on the recommendation system will help in understanding the requirements of customers and companies both. In order to solve such a task, a product-based recommendation system is implemented to identify the top products based on ratings, and these products are randomly combined with the click dataset in order to predict clicks on the shown advertisements. After data combination, data pre-processing, and feature extraction, three deep learning-based models are implemented. These models are trained over training data and then tested on test data for the evaluation based on the metrics such as Accuracy, PRF score, and MCC score. This evaluation results in the LSTM model being the best optimal model that can address such tedious tasks and can help in new online business strategies.

1 Introduction

The development of the internet has unveiled multi-faceted benefits in the domain of online marketing and business venture. With the advent of 5G and rapid digitalization, a myriad of scaled-up business options is available to buyers and sellers. One can easily buy and sell products online to pitch their business to the next level. The seamless connection and smooth communication with the buyers offered by the internet have helped brands reinvent their operational techniques. One of the boons of technological advancement of the digital age is the process of online advertisements. By employing the etiquette of digital marketing and online branding, businesses and organizations have the potential of spanning across the globe to reach a new customer base. This method has also been quite helpful in dodging various usual advertisement challenges that hamper the growth of the business. The use of the internet has also widely helped businesses and brands to opt for more innovative tech-rich solutions when compared to traditional branding and promotion processes. So, the growth of E-commerce has skyrocketed to the maximum over the past few years.

The prudent usage of E-commerce and online advertisement has effectively lowered the share of capital involved in traditional marketing. Switching to the online mode of promotion and marketing has helped business owners' lower production costs, enhance customer relationships, boost sales pitches, minimize errors, and swiftly handle product transactions. Due to multi-faceted benefits, the majority of brands and organizations have switched over to the mode of online advertisements. Usually, the traditional website-based method of online advertisement involves the identification and recognition of keywords from the available sources on the internet to match the customer's search. This method was observed to experience certain challenges as the supply and demand of E-commerce grew. Promoting the products in a perfect way to research ideal customers began to become a critical issue. That is where the advertisement clicks prediction technique came into the picture.

Predicting the clicks on an advertisement is nothing but a brief prediction analysis model that takes into account the history of the customer base and anticipated the customer's click probability concerning the particular online advertisement. Following this predictive model of online advertisement, efficacy turns out to be effective in determining the reach and visibility of the business. In order to correctively examine the performance of an advertisement, the click through rate (CTR) is highly used metrics. Majority of the companies have employed the Click through rate prediction to determine their products' reach in the contemporary market. One of the major takes with calculation of CTR for an advertisement is that it is not capable of determining the click-to-purchase conversion ratio. The majority of the people who use the internet and social media, intend to just scroll through and do window shopping, instead of actually buying the product. Only a few are the actual potential customers among the massive internet crowd. This criterion paves way for further prediction challenges to arise in the considered framework.

To segregate and consolidate the vast customer data, the use of a recommender system is strongly advised. Building the proposed Deep learning-based advertisement model along with the strategy of the recommender system will actually increase the efficacy of the system, and also settles the data protection concerns in an ideal way. The utilization of the recommender system also serves as an effective alternative to settle the issues of data overload, which can limit the data to increase the engagement and interest of the customer to end up buying the product. Apart from the usual method of suggesting the targeted products to the potential customers by identifying them through the internet history and other data, this recommender system also takes into consideration the behavior, purchase data and preference of the customer to notify related products up in their feed. Since the recommender model takes all the user history and personal preferences into consideration, it also works top-notch in suggesting new related products for the customer according to their stipulated requirements.

The use of a recommender system along with the basic prediction of clicks on advertisement will greatly benefit the online advertisement process in targeting potential clients. It is very much effective in shortlisting the ideal customer from the pool of people data available on the web.

Considering the method of recommender models, there are three major directives called collaborative filter, content-based model, and hybrid technique that are prevalently used in contemporary research ideologies. The content-based model works by collecting the customer data, web search, shopping history and personal preferences to understand their narrative, while the collaborative filter approach tries to segregate the right product to the right customer based on the filtering out technique. The blend of both these approaches can be found in the hybrid recommender working model, where the products are filtered according to the taste of the customer. It brilliantly works in curating the perfect product recommendation system that is used for online marketing by many businesses and brands. In this research popularity-based recommendation system has been utilized.

Since the recommendation system is responsible for displaying a relevant product to the customer via advertisement on the customer's screen should significantly improve the chances of click on advertisement and thus, Click Through Rate (CTR). However, in this research, a brief analysis will be carried out to determine the influence of the recommender system model on the click of an advertisement. To accentuate the proposed research further, the advanced deep learning techniques of neural network architecture are employed here. Since the Click prediction on advertisement comes under the classification problem, the classification deep learning algorithm will be utilized for this research. The models such as Convolutional Neural network (CNN), Recurrent Neural Network (RNN) and Long Short-term Memory (LSTM) will be utilized for this research. To identify the most efficient model for predicting the click on advertisement the metrics such as Accuracy, Precision, Recall and MCC Score along with F1-Score will be calculated.

1.1 Research Question

Q.1. How does the popularity-based recommendation method using deep learning algorithms help to improve the click on Advertisement for a specific product?

Q.2. Which deep neural network architecture can correctly predict the click on advertisement and to what extent maximum accuracy can be achieved?

1.2 Research Objectives

Obj1: A critical Review of Click prediction on Advertisement data and discuss the implemented technologies.
Obj2: Implement and evaluate the results of CNN, RNN and LSTM algorithms.
Obj3: Utilization of recommendation system to improve the CTR of product advertisement.

2 Related Work

This section of the literature review presents all the contemporary research works that are related to the use of recommender systems and click-through rate (CTR) prediction models using machine learning and deep learning algorithms. With the help of the existing works, the claims and effectiveness of the recommender system in increasing the CTR prediction are justified.

2.1 Using machine learning for CTR prediction

In the recent study carried out by (Choi & Lim, 2020) the authors present the importance of utilizing machine learning techniques to carry out the method of targeting advertisements to a specific audience base. This research study also portrays the use of machine learning techniques with the blend of AI computational techniques to scale up the prediction process. Two different types of categories are taken into consideration in this work, which are termed content-based and user-based. A total of twenty-three machine learning strategies are taken into consideration to analyses the accuracy and effectiveness of the machine learning algorithms. Also, by predicting the shadow areas of click detection, this study tests the viability of machine learning algorithms. Another contemporary study proposed by (Shah et al., 2020) put forth the claim that the challenges of online advertising have been simplified with the help of machine learning and AI models. The added insights of statistical observations also help a great deal in formulating the methodology in the most perfect manner. The author also states that the digital realm is highly going to be influenced by the benefits of machine learning algorithms. From this research, the author tries to prove the efficacy of ML techniques. Machine learning architecture can also be used to identify and classify advertisements based on the preference of the users.

The same has been put out first-hand in the work done by (Yoganarasimhan, 2020). Here the author represents how effectively customer search personalization can be done with the help of machine learning algorithms. The query-based search technique is the one employed here. The customer's web search and profile history are the collected inputs to facilitate this research process. The three main modules of this research are feature extraction, LambdaMART and feature selection. The personalization is done for a hefty amount of data from Amazon EC2. There is also a significant 3.5% increase in the CTR ratio. Identifying the right customer and displaying them with the right products can be done effectively by using the keyword identification and analysis process.

One such brief research study performed by (Shi & Li, 2016) does the same. Search engine optimization (SEO) is taken as an underlying term for keyword identification in this work. The data from google ads are taken for analysis and its working is analyzed in a deep manner to comprehend the suggestion patterns. Along with the past data, the cost per click (CTC) and CTR are the main rudimentary factor concerning this research study. It is also found that the use of machine learning techniques to identify the keyword and predict the CTR from google AdSense worked perfectly.

The sponsored search is another different routine that takes the input of the users and fetch out the sponsored related products to the customers. This niche is dealt clearly in the work done by (Edizel et al., 2017). The majority of the commercial search engines employ this feature by partnering with brands and businesses to boost the visibility of the product. It involves a lot of features and attributes to process which is quite a hassle. That is why the machine learning algorithms are employed in the study to ease the data collection and processing technique for the better prediction of accurate click-through rate. Along with this step, a deep convolutional neural network is also used in this research to predict the CTR rate of query pair. The outcomes of this research are also quite impressive. Another similar work is presented by (X. Wang et al., 2019), where the author makes use of the traditional machine learning algorithms like random forest and decision trees to supervise the CTR prediction in online advertisements. The recommender method of factorization machine (FM) is also involved in this research work. To sort out the recommendations the new model of Gradient boosting decision trees (GBDT) models is employed here. The combined techniques are found to be far more effective than the regular models used for CTR.

2.2 Using deep learning algorithms for CTR prediction

While it is already an established fact that deep learning algorithms like neural networks perform excellently in the prediction and detection modules, the recent research work by (Xu et al., 2021) reiterates the same. Here the author employs the benefits of deep neural networks in recognizing the CTR of a series of online advertisements. The sampling, training and other feature-related problems are easily rectified with the help of the use of neural networks. The author takes Baidu's industrial etiquettes as an example in this research. From the outcomes of this study, it is clear that the use of neural network algorithms boosts the effectiveness and revenue increase by more than 1%. The multi-billion-dollar online advertisement industry is here forever to stay. In this study performed by (Effendi & Abbas, 2016), the CTR prediction is elevated with the usage of linear regression algorithms that effectively solve the learning problem in an enhanced manner. The motive of this research is to increase the CTR and the total revenue of the business taken into consideration. The basic steps of feature extraction, prediction and collection are done dynamically to show the stipulated results. The results of this study are found to be close to the range of excellence expected. Native advertisements are another popular genre of online advertisements.

A brief analysis concerning such a topic is dealt in a research work done by (Parsana et al., 2018). According to the author, processing the CTR for the native ads is quite difficult as there is no direct intent for the query. A massive scale embedding scheme along with the recurrent neural network algorithms are used in this paper to predict the preference of the customers. The pre-trained embedded vectors collect and bring in the user's past history and the CTR prediction is done in such a manner. From the outcomes, it is proven that the use of neural networks works better in producing the much-needed accuracy than other baseline methods. A very new approach of predicting the CTR on-point is presented by (Fei et al., 2021) in their research work. They take the Baidu search advertisements, which is currently known as Paddle box as a classic example to proceed with their analysis. As already well known, deep neural networks are great in the prediction and detection process. The same is utilized by the author in his research. Additionally, the authors also employ an innovative model called Gating enhanced multi task neural networks (GemNN), which greatly reduces the customer profile to pick out the right add display. The combined and blended approaches have always resulted in better outcomes.

One such research that uses fuzzy logic along with the deep neural network is presented by (Jiang et al., 2018). The method of Gaussian Bernoulli and restricted Boltzmann machine techniques are employed to fetch the data from the developed approach. The addition of a fuzzy logistic approach is also found to be an effective add-on to predict the CTR of the taken advertisements. The outcomes of this research are also found to be accurate and effective. The capability and the robustness have seemed to be improved than usual. From all the previous contemporary works it has been established that neural networks are the perfect pick for prediction methodologies. The same is been presented by (Chen et al., 2016) in their research study. Here the author employs the techniques of recurrent neural network to predict the estimated ad-clicks. The ring structure architecture of the neural network pattern is based on the insight of the Long short-term memory (LSTM) module. The stipulated results of this analysis states clearly that the recurrent neural network patterned with the LSTM seems to perform better than the regularly devised modules. The feature extraction is a very important step that determining the effectiveness of the CTR prediction process. A research that talks more about the light feature extraction in the deep neural network is presented by (P C D et al., 2021). To rule out the ineffectiveness of random forest algorithms, a graph-based convolution network is employed here. The model interference, latency and such other complications are effectively reduced with the help of neural network algorithms. The accuracy and effectiveness of the devised model is estimated to be very effective and on point.

2.3 AI-based Recommender system techniques

Now that the claims of deep learning algorithms are briefly started in the prediction of CTR, we switch over to the methodologies of the recommender system for further analysis. A very relatable research work that is close to the proposed study is done and presented by (Zhang et al., 2019). The author presents a brief survey and a comprehensive perspective of the deep learning-based recommender system that is used to predict the click-through rate of online ads. The persuasive influence of deep learning algorithms is clearly explained in this research study. It also serves as effective evidence of the spot-on work of deep learning and recommender systems in the prediction process. To prudently use the recommender system in the proposed research one needs to comprehend its fullest potential and applications. This comprehensive research work presented by (Lu et al., 2015) projects all the dimensions of recommender system practices and their applications. All the eight models of the systems are explained briefly in this study. From this work, it is identified that the deep learning and AI-based recommender models are apt for e-commerce and online promotions.

Another significant research analysis that involved the use of a recommender model with a deep learning algorithm to predict the CTR is drafted by (Guo et al., 2019). The key highlight of this research is that it is done in a live environment with all the parameters set. A very comprehensive PAL network is used along with the taken recommender system to predict the live CTR traffic of a proposed network in the research. From the drafted outcomes it is evident that the PAL model outperforms the usual strategy by more than 3% in total. It is also said that the click-to-purchase conversion rate is higher when estimated using this approach. The rise of AI-based automated recommender systems is also quite popular in the current era of CTR prediction. The research done by (T.-H. Wang et al., 2020) is a complete list of that. The author portrays the adaptability issues of a realistic recommender system to switch over to the automated ones for this analysis. He utilizes automated machine learning and the TensorFlow ecosystem to bring out the efficacy of the developed approach. The prediction process of CTR is facilitated in a very smooth, flexible and hassle-free way with the use of AutoRec. From the outcomes of this research, it is proven that the accuracy and reliability of this model are better than the realistic recommenders.

A very different approach to the same domain is presented by (Sheng et al., 2021). The ultimate goal of the author is to devise a one approach fits-all model that would be of great use to the next generation of CTR prediction researchers using the recommendation system. To achieve this milestone, the author uses a star-topology multi-domain model by leveraging all the collected information from the dataset. Since the design is based on a shared network, the access is unified. There has been significant growth in the sales and revenue per mille as per the stipulated outcomes. The works of (Hong et al., 2021) states that the use of recommender system will actually boost the user interaction with the business to a whole new level. To prove this claim, the author uses a novel neural network algorithms to predict the estimated CTR. There are two main parts to this research, which is the present interaction and anti-layer perceptron approach. Both the public and private business-related datasets are procured for this study and the outcome establishes the accuracy, relevance and effectiveness of AI based recommender systems.

2.4 Indentified Gaps:

After reviewing the certain set of research papers, it has been identified that most of the research paper utilizes the Avazu advertisement click data or Criteo Display Advertising dataset for click prediction. However, the relation with recommendation system in order to increase the CTR of advertisement has not been explored yet and still an open area of research. Also, in the previous research machine learning algorithms are highly implemented which has generated satisfactory results but in order to generate the best outcomes deep learning algorithms are preferred for large advertisement dataset. To overcome the shortcoming of the prior work, in this work the advertisement data is merged with the recommendation system for improving the CTR of advertisement. Also, the deep learning algorithms are explored in this research where we will be performing a comparative analysis among the obtained results from each model.

3 Research Methodology

Due to the covid-19 pandemic, business methods have been escalated to online shopping terminologies. Customers make online purchases of the products from various apps and websites and the online advertisements of products on various platforms play an important role in the escalation of online shopping. A more perfect and relevant advertisement leads to more shopping for the products which can be achieved by combining both a recommendation system and advertisement click systems. In this task first, what product should be shown in advertisements is extracted using product-based system and then this data is combined with the clicked advertisement data so that a prediction can be made on the click which indicates whether a consumer will click on the advertisement of the product or not which can help various companies in accessing the consumer requirements. In this research task, the main aim is to pick out the best deep learning-based model which can predict the click on the advertisement shown based on the product-based recommendation system. Data collecting, data cleaning, data preprocessing, data exploration, visualization, feature extraction, model setup, model training, and evaluation procedures are all covered in the uniform flow diagram offered to accomplish this task. Each stage is addressed in great depth in further sections.

Figure 1. Flow Diagram for Predicting the Click on Advertisement based on Recommendation System

3.1 Dataset Description

The advertisement click data has been accumulated from the Kaggle website i.e., the Avazu click dataset (Click-Through Rate Prediction | Kaggle, 2022). This data contains four million rows and twenty-five columns. These columns contain information such as time stamp, banner position, site category, app category, click, etc. This data has both numerical and categorical values. Since this research project ignites the concept of clicking on online advertisements based on a recommendation system therefore the data for the recommendation system is also collected from the Kaggle which is an Amazon product review dataset (Recommender System Using Amazon Reviews, 2022). This data contains seven million rows and four columns. These columns are accountable for product id, user id, timestamp, and ratings. All values in this data are in numerical format. The size of click datasets is about 6 GB and the Size of Amazon review dataset is 300 MB.

3.2 Data Pre-Processing

After accessing the data from the authenticate source next step performed is the pre-processing of the data. Since both datasets contain a large number of rows therefore random samples of 100 thousand rows are taken from both datasets. In this step processes like data cleaning, removal of null values, changing of data types of columns, and data normalization are executed. Click data contains categorical columns, which are converted into the required numerical format by deploying a label encoder. This dataset is unbalanced so a balanced dataset is obtained by applying smote oversampling technique. After pre-processing the review dataset, a product-based recommendation method is implemented which output the top 10 product. These products are concatenated randomly with click data while preserving the balance of distribution. A detailed overview of important features is expressed in succeeding sections.

3.3 Product Based Recommendation system:

Recommendation system can be integrated with the click prediction system to result out the most relevant products which can be shown to the users so that more clicks can be generated. A product-based recommendation system generates product suggestions by attempting to anticipate and present goods that the customer is likely to purchase. Even if something isn't quite right, if it allows you to see what you want to see, it has served its purpose. Recommender systems are used in movies, music, news, books, research articles, search queries, social tagging, and other situations. Its popularity has grown in recent years. The majority of e-commerce sites today, including eBay, Amazon, and Alibaba, use their own recommendation algorithms to better match products that buyers are likely to like. These algorithms are primarily used in the digital world.

3.4 Exploratory Data Analysis (EDA)

Data analysis and visualization is an important step in building deep learning or machine learning-based model because it gives better insights into the data which helps in tuning and optimization of the algorithms. In the click dataset, it has been observed that the class of click vs non-click is imbalanced as shown in Figure 2 and after applying smote oversampling technique that balanced dataset is shown in Figure 3.

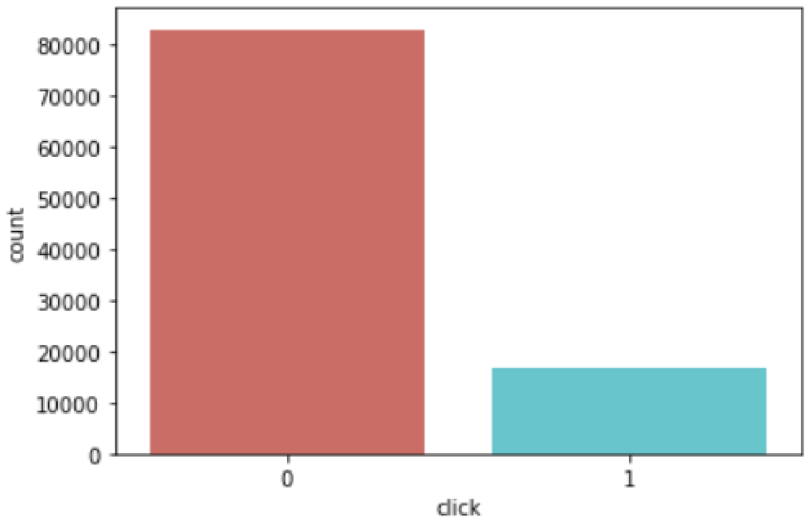

Figure 2. Imbalanced label for Advertisement Data

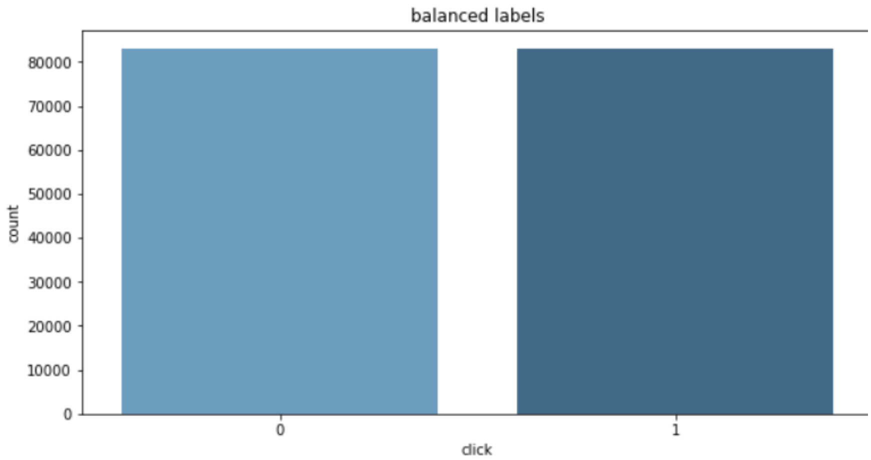

Figure 3. Balanced label in Advertisement Data

In further analysis trends of clicks based on hour, class-wise clicks based on hour, and ctr based on the hour are analysed which are shown in Figures 4 and Figure 5.

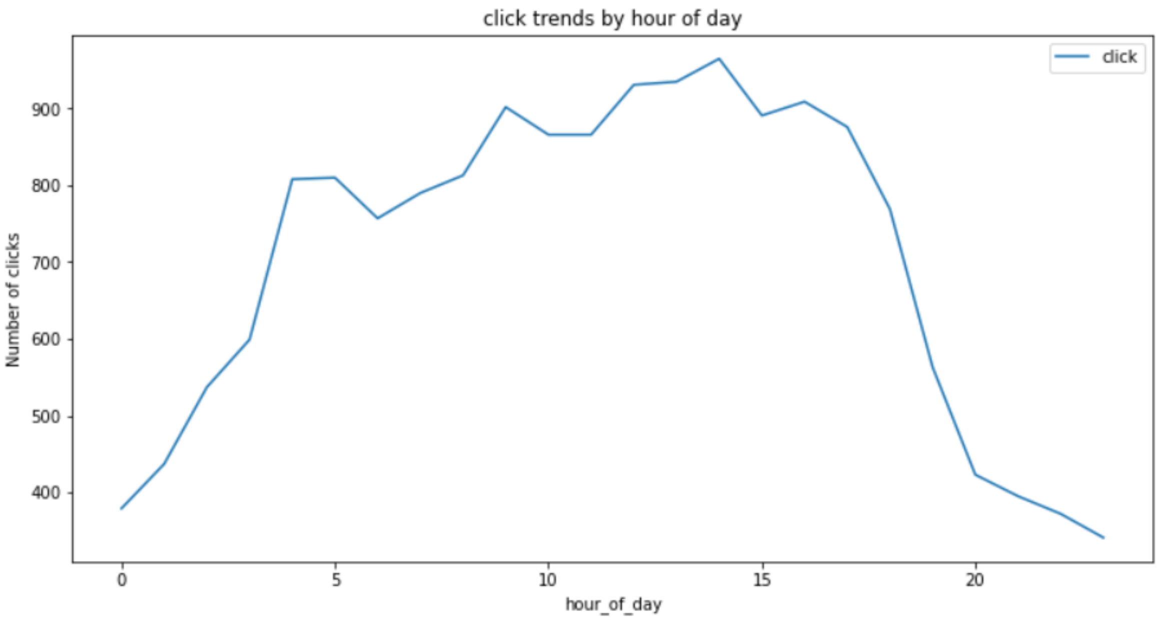

Figure 4. Trends of Clicks based on Hour of the Day

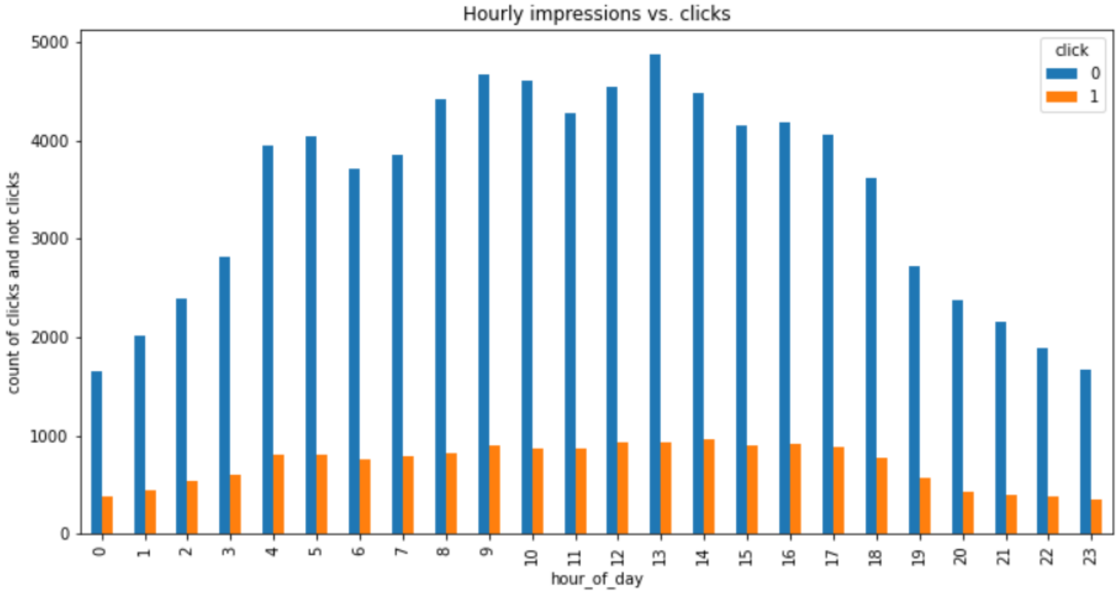

Figure 5. Bar Plot representing the Count of Impressions and Clicks based on the Hour of the Day

From Figure 4, it is observed that during 10-15 hours of the day the user click more on the advertisement. Figure 5 depicts that throughout the day the number of non-clicks is more than the clicks but from 4th hour to the 17th hour of the day the number of clicks also increases. Further, an analysis is carried out based on the banner positions and top 10 sites id as shown in Figures 6 and Figure 7.

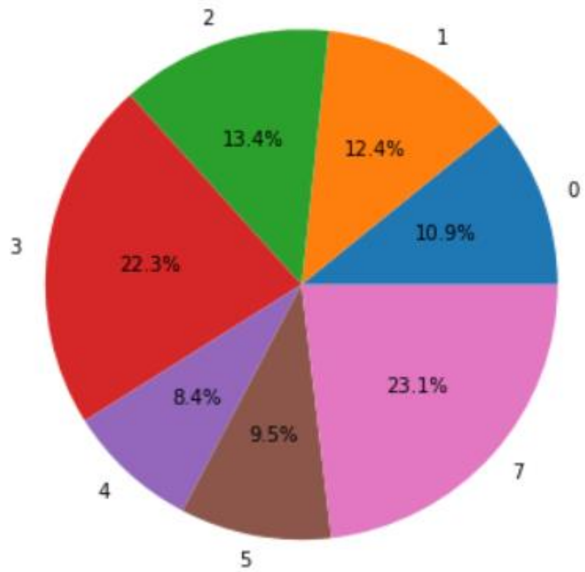

Figure 6 Click Through rate based on the Banner Position

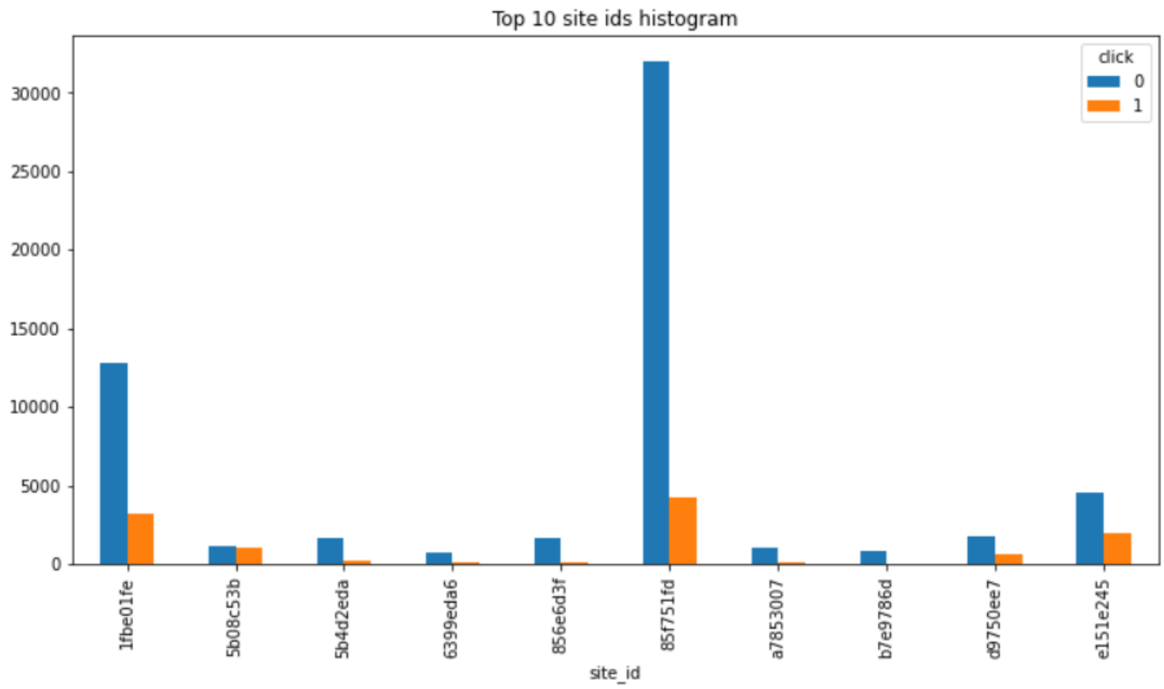

Figure 7 Distribution of Clicks based on Top 10 Sites

Here, Figure 6 represents the click-through rate of all banner positions. Numbers represent the different positions of the banner. Banner positions 3 and 7 have more CTR than other banner positions. From Figure 7, it can be deduced that out of 10 there are 4 site ids where no clicks are observed while one of the site id has almost the same ratio of clicks and non-clicks. For the amazon review dataset, the distribution of the rating is plotted to get insights into the review distribution as shown in Figure 8.

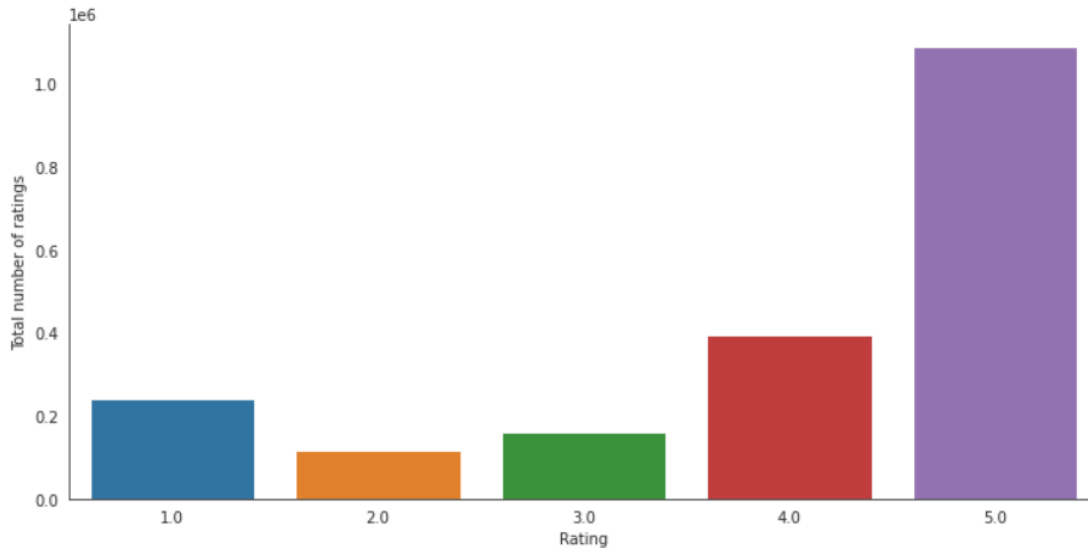

Figure 8. Distribution of ratings

An analysis is executed to identify the click-through rate based on the Top 10 product id. This analysis results in the deduction that the distribution of CTR is uniform among all the products which indicate the balanced distribution of product id among the click data which is shown in Figure 9.

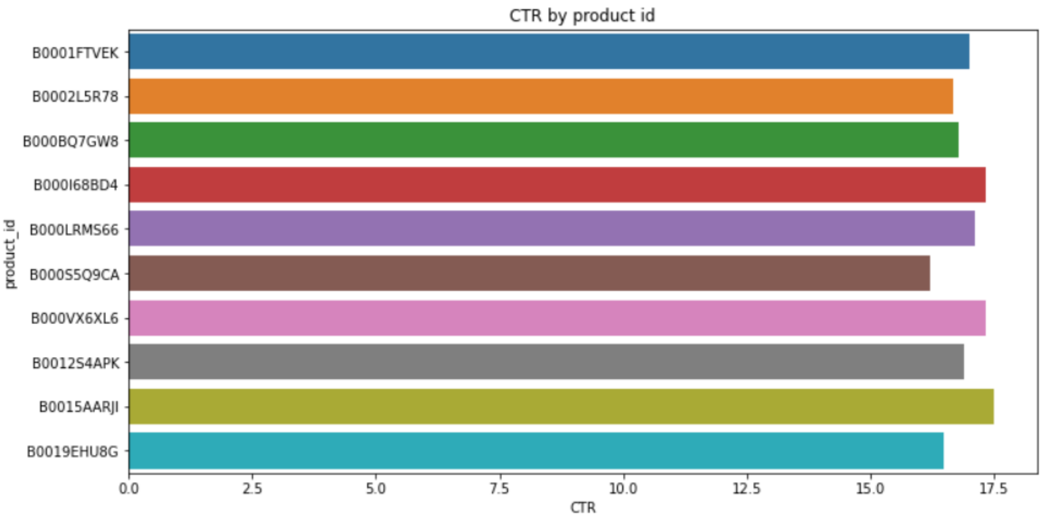

Figure 9. CTR based on product id

After the implementation of the product-based recommendation system, the top 10 products are extracted, where the count of rate for each associated product is shown in Figure 10.

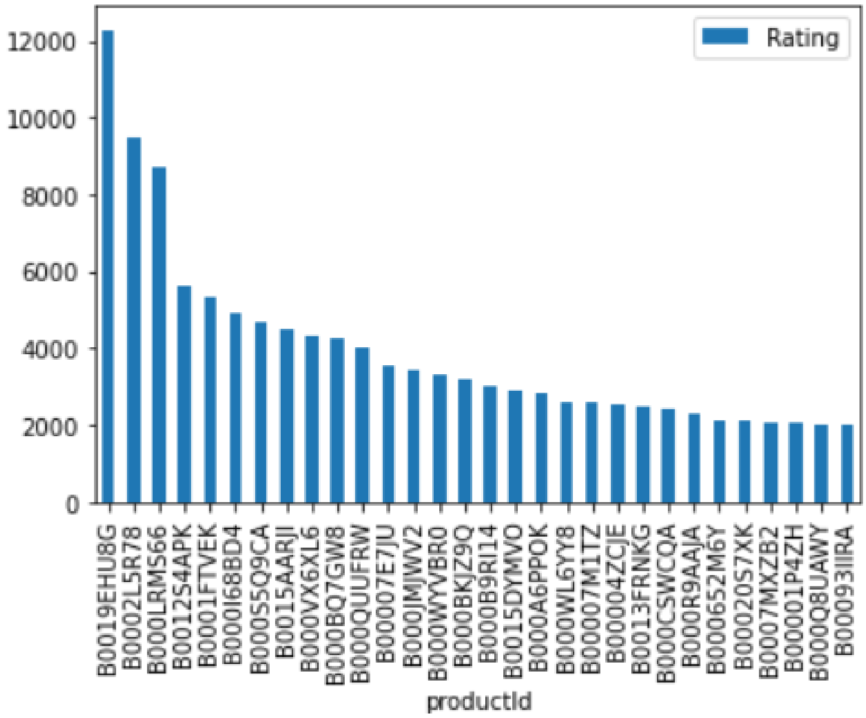

Figure 10. Top-10 product id based on recommendation system

3.5 Feature Extraction

Feature extraction and feature selection are important steps for this research as it helps to reduce the curse of dimensionality of data and helps to reduce the training time. For this task, hours are extracted from the date column. Unnecessary columns such as date, and id are dropped first and new columns are added to the existing data which contains cumulative values of other columns. Data is scaled using a standard scalar for the standardization of data.

This step results in a reduction in the size of the feature vector from 26 to 19. It has been observed that various attributes hold low co-relation with the target variable. Attributes such as various ids, and connection types can be dropped because these are noted at the instant of generating data and may have changed while working on this task. Click feature is allocated as the target variable while the remaining variables are assigned as feature variables.

3.6 Model Training and Testing

In this research task, three deep learning-based algorithms are executed for the purpose of click prediction based on a recommendation system using the advertisement and product review dataset. After pre-processing and feature engineering the dataset is divided into a training dataset and a test dataset in the ratio of 80 to 20. For the training of models, this training dataset is used while for testing and comparison of each executed model test dataset is used. The deep learning models implemented are Convolution Neural Network (CNN), Recurrent Neural Network (RNN), and Long Short-Term Memory Network (LSTM). The Keras tuner is executed while training the model to tune the hyperparameter of the model so that the most optimized model can be trained for better results.

3.7 Model Evaluation

The aim of this research task is to designate the optimal deep learning algorithm to predict whether a customer will click on the online advertisement shown to him. This online advertisement is based on a product-based recommendation system which makes it a binary classification task. Therefore, metrics based on classification tasks such as accuracy, precision, recall, f1-score, and MCC score are executed for the evaluation of each executed model. These four metrics are calculated for each executed model on test data after the training on the training dataset. The implemented model which scores the highest value in metrics is considered the best model. Validation loss is also inspected in such scenarios for evaluation purposes.

4 Design Specification

4.1 LSTM

LSTM stands for the Long short-term memory algorithms. RNNs (recurrent neural networks) are difficult to build using LSTMs. When given a phrase as input, the LSTM model will only retain the terms that have a bearing on the result and disregard the remainder. If the algorithm produces that information travelled through the LSTM's storage cells, if it is valuable, it will be kept within one of those cells, and if it is not valuable, it will be ignored. RNNs have issues, but LSTMs solve them because they know what to remember and what to forget. If the algorithm determines that the information is valuable, the gate allows it to be stored in memory. If the information or word is discovered to be irrelevant to the results, it is ignored. In this way, LSTM can remember relevant information. LSTM is slower but more powerful than RNN and uses more processing power than RNN.

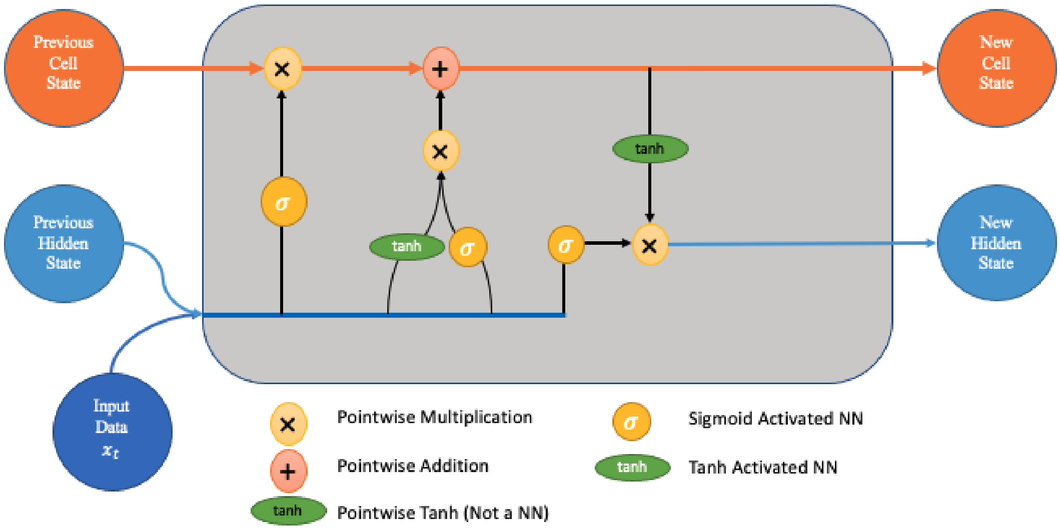

Figure 11. Architecture of the LSTM Model

4.2 CNN

A computerized convolutional neural network replaces the visual cortex. This is similar to how the human brain interprets and categorizes what it receives from the eye to distinguish it from other images. A CNN takes input in the form of an image, converts it into a matrix, biases each part of the matrix to distinguish it from other images, and outputs a classifier. Images are recognized and classified using CNN. A CNN emphasizes and remembers certain aspects of one image and tries to distinguish between the two by comparing those aspects to certain aspects of another image. Blank segments and segments with fewer features have less bias because they do not capture the original image as much as highlighted segments. Parts with more prominent features introduce more bias. This helps the CNN model classify the images. This CNN can also be implemented on text data by using different embedding layers provided by Keras.

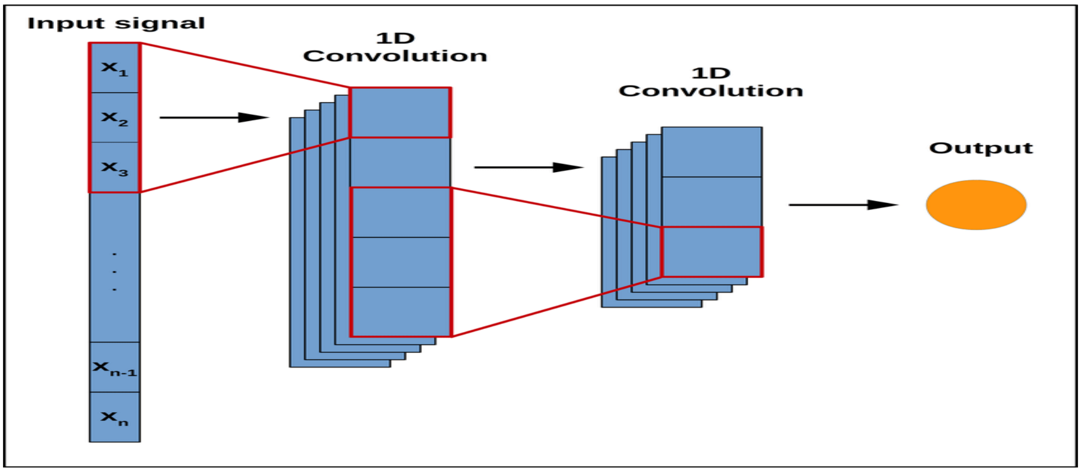

Figure 12. Architecture of CNN-1D

4.3 RNN

It is a type of neural network that excels at simulating time series data. An RNN layer iterates over a series of time steps using a for loop, encoding information from previous time steps while preserving the internal state. The network is initially phased. Based on this state and the current state, a new state is calculated, and the process is repeated until the network propagates forward. RNNs are trained by computing the output after each time step, comparing it to the actual output, and propagating the difference back into the network to update the weights.

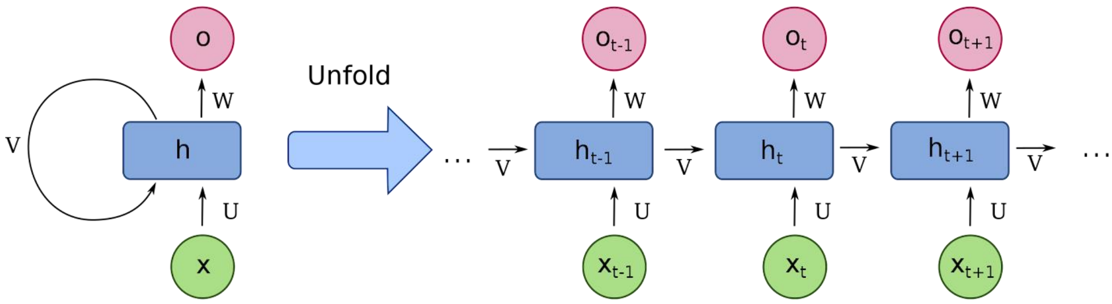

Figure 13. Architecture of the RNN Model

5 Implementation

In this research task, aim is to make prediction on clicks of online advertisements shown on various social platforms integrated with the product-based recommendation system. This task is achieved by executing three deep learning-based algorithms which are LSTM, CNN 1-D and RNN. These advanced deep learning models are capable of predicting the clicks on the product shown in the online advertisements by the help of recommendation system. Superlative model is selected based on the evaluation of model on the test data based on metrics such as accuracy, MCC score, Precision, recall, and F1-score. Each model is trained over 10 epochs first while having binary cross entropy as loss function and adam as optimizer. Keras tuner is implemented in order to tune the hyperparameter of the model in order to achieve the better outcomes. For data manipulation, pre-processing, model building and model evaluation several libraries are used which includes numpy, pandas, matplotlib, pyplot, sklearn, smote, TensorFlow, Keras etc. For training and testing purposes, the entire experiment is carried out in Jupyter notebook using Python as the programming language. The model had to be implemented according to the following requirements.

- Operating System: windows 10
- Random Access Memory (RAM): 16GB
- Hard disk: 15GB
- Languages: Python
- Cloud Platform: Anaconda, Jupyter
- Python libraries: numpy, Pandas, matplotlib, TensorFlow, numpy, seaborn and Keras

6 Evaluation

In this research task, the objective is to determine the foremost improved deep learning model capable of predicting the click on online advertisement, hence it is essential to assess each implemented model based on the different classification metrics. It is a classification task, more specifically a binary classification task so accuracy, recall, precision, f1-score, and MCC score on test data are calculated to assess various deep learning models. After the training of each model on training data, each model is assessed on test data and all metrics are intended. The model which achieves the highest value of MCC score, accuracy, f1-score, precision and recall will be selected as the finest model for heading prediction basis. Bar plots are plotted to visualize the comparison among the implemented models.

6.1 Evaluation based on accuracy

The LSTM model is trained on the training data for 10 epoch values. While training this model it is observed that for 5 value of epoch accuracy increases and for last the value of accuracy saturated at 83.16% therefore model is not forced to train to further epochs. A similar trend is observed for the validation accuracy. The training loss and validation loss start to decrease for the first five epochs and for the last five epochs it becomes constant. After training this model is tested on the test model and the test accuracy obtained is 83.13%.

The convolutional neural network (1-D CNN) is trained on the same training dataset for the same number of epochs i.e., 10. The observed training accuracy on the training data is 83% on the last epoch. The value of test accuracy obtained when this model is tested on the same test data is 82.3%. The training of the model is restricted to 10 epochs to prevent the model from overfitting because different training results in an increase in validation loss.

The recurrent neural network (RNN) was first trained on the same training data for 10 epochs. A number of epochs are restricted during training to prevent the model from overfitting because after five epochs the validation loss becomes constant and both training and testing accuracy are saturated. The training and testing accuracy obtained at the last epoch is observed to be 83.03% and 82% respectively.

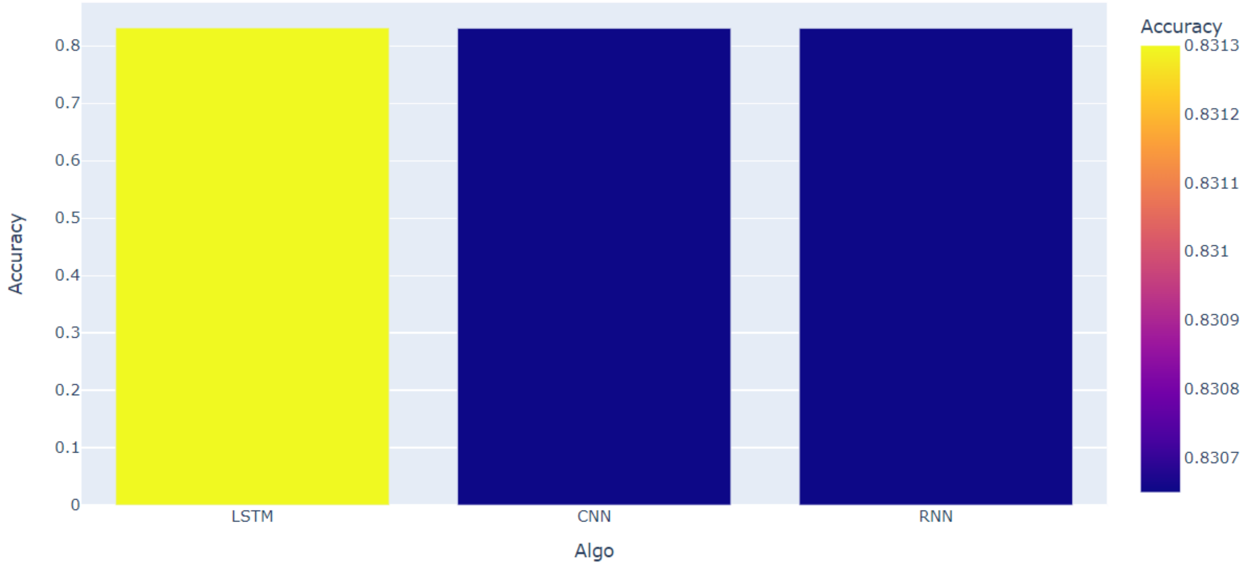

Figure 14. Accuracy comparison of implemented models

Bar graphs are plotted to compare all implemented models based on accuracy as shown in Figure 19. After a thorough analysis of plots, it has been observed that using LSTM architecture, a better accuracy score of 83.16% is achieved, followed by the CNN model with an accuracy score of 82.3%. The other model RNN achieves the results with an accuracy score of 82%.

6.2 Evaluation based on PRF score

PRF score stands for the precision, recall, and f1-score values. In the task of classification, one should not decide on the best performing model solely based on a single evaluation metric i.e., accuracy. PRF score helps in investigating the false positive and false negative predicted by the model which plays an important role in accessing the performance of the model in real-world scenarios. After the training of all implemented models on the training data, an evaluation of each model is performed on test data and the PRF score is calculated. The LSTM model achieved the PRF score of 0.7856, 0.8318, and 0.7670 respectively while the PRF score obtained by the CNN model is 0.7840, 0.8306, and 0.7590 respectively. The last model RNN achieved PRF scores of 0.7866, 0.8301, and 0.7602respectively. Comparison of PRF score is also visualized by plotting a bar graph shown in the figure. By analysing the bar plot, it can be concluded that the LSTM model achieved the highest PRF score followed by the CNN model. RNN model also achieved considerable values of PRF score.

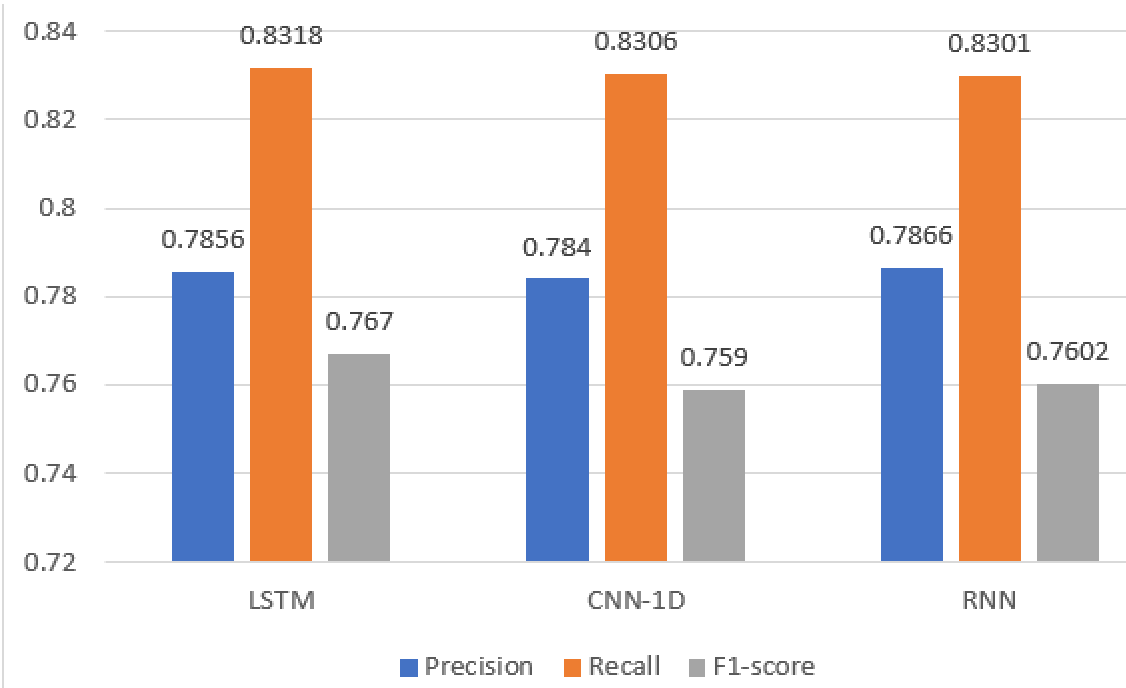

Figure 15 PRF score comparison of implemented models

6.3 Evaluation based on MCC score

MCC stands for Matthews Correlation Coefficient, which is used to validate the performance of classification model. The MCC values lies between the -1 to +1, where the score of +1 indicates the optimal model performance and -1 indicates the poor performance of model. In our case, model having the highest MCC score is considered as the most optimal model for classification. After calculating the MCC score for all the implemented models, it has been observed that LSTM generates the highest MCC score as compared to the other classes, which indicates highest agreement between the predicted classes and actual classes.

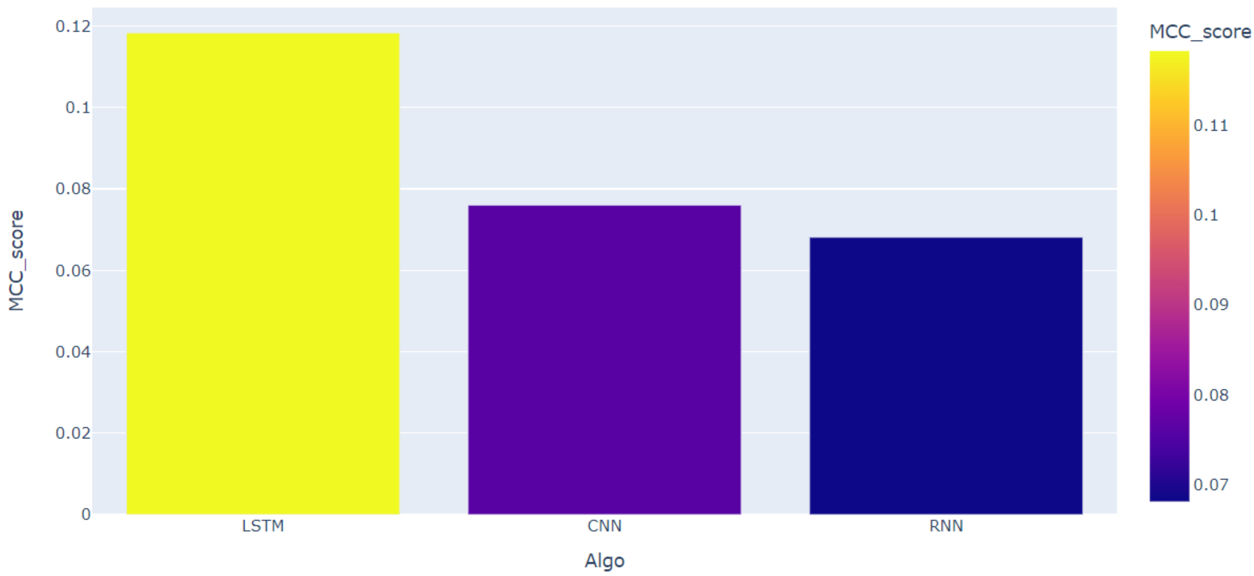

Figure 15: MCC score comparison of implemented models

6.4 Discussion

Prediction of the clicks on the online advertisement click dataset after the recommendation of advertisements by product-based recommendation is carried out in the set of experiments by introducing three different deep learning algorithms. The supreme model identified is the LSTM model, which can retain the last instances because such models have selective memory capabilities. This model surpasses the other implemented models in the terms of evaluation metrics which are accuracy, precision, recall, f1-score, and MCC score. However, CNN and RNN models also showed satisfying results which were close to the output of the LSTM model. Since the size of both datasets is large therefore, each model takes about 6 hours' time to train.

In order to inspect the miss-classified classes using the LSTM model, we have plotted the confusion matrix for every class over the test data, the image for which is shown in Figure 16. The diagonals in the confusion matrix represent the correct classification. Where non-diagonal elements are not classified correctly, from this confusion matrix, it can be deduced that the false negative and false positive output by the LSTM model is 3246 and 128 respectively.

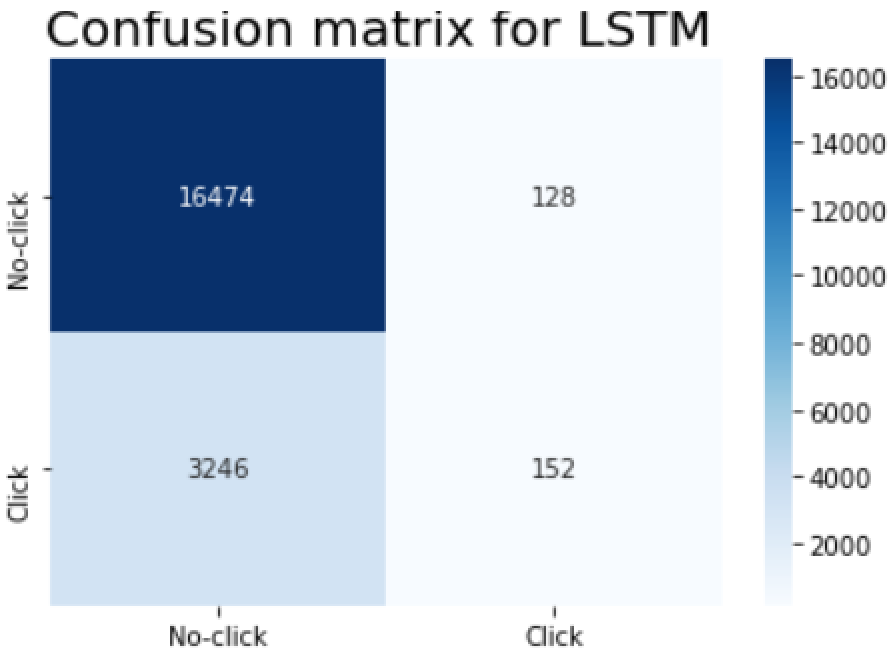

Figure 16. Confusion matrix of LSTM Model

In confusion matrix above, the number of false positive values over the test data are identified as 128. The false positive values here indicates the advertisements which are actually clicked but LSTM algorithm has predicted it as Not-Clicked. Similarly, the 3246 in the confusion matrix are false negative values over the test data. The false negative values in the confusion matrix indicates the advertisements which are actually Not-clicked but LSTM algorithm has predicted it as Clicked. In order to calculate the precision, recall and F1-Score all the True positive, true negative, false positive and false negative values has been calculated.

7 Conclusion and Future Work

Nowadays, due to advancements in various technologies and internet services, it has been observed that various business giants are shifting their advertisement strategies towards social platforms. Selling products online is customizing the business performance and this process is found to be cost-effective, easy, and satisfying to both sellers and customers in various aspects. Therefore, demand analyzes the advertisements, clicks, and what products should be shown to customers in order to fulfill their needs and relevance. Prediction of clicks on online advertisements and what advertisements should be shown is still a challenging task. In this research project work, the first product-based recommendation system is implemented to identify the list of the products shown in the advertisements so that number of clicks can be maximized, and thereafter a thorough analysis of the click dataset is made which helped in developing deep learning-based models to predict the clicks on advertisement shown. In the proposed methods and models the LSTM model is able to predict whether the person will click on the shown advertisement or not with an accuracy of 81.13%. The prime lead of implementing this model is the capacity of the model to retain its memory from previous predictions in further predictions through inbuilt memory and forget gates. In this prediction and analysis work, a 100k row dataset is used for both recommendations and click prediction purposes. This work helped in identifying the factors which affect the clicks on the advertisements and fulfills the objective of the combined recommendation and click prediction task. Due to the 100k row size, the achieved results are somewhat restricted to their optimum results because of limited computing capabilities. In future work, a more complex recommendation system can be integrated with the large size data of advertisement clicks which will surely result in higher accuracy and better results. With better computing resources, the time of training of such models will be reduced along with large computing capital.

References

Chen, Q.-H., Yu, S.-M., Guo, Z.-X., & Jia, Y.-B. (2016). Estimating Ads' Click through Rate with Recurrent Neural Network. *ITM Web of Conferences*, *7*, 04001. https://doi.org/10.1051/itmconf/20160704001

Choi, J.-A., & Lim, K. (2020). Identifying machine learning techniques for classification of target advertising. *ICT Express*, *6*(3), 175–180. https://doi.org/10.1016/j.icte.2020.04.012

Edizel, B., Mantrach, A., & Bai, X. (2017). Deep Character-Level Click-Through Rate Prediction for Sponsored Search. *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 305–314. https://doi.org/10.1145/3077136.3080811

Effendi, J., & Abbas, S. (2016). Click Through Rate Prediction for Contextual Advertisment Using Linear Regression. *International Journal of Computer Science and Information Security*, *14*.

Fei, H., Zhang, J., Zhou, X., Zhao, J., Qi, X., & Li, P. (2021). GemNN: Gating-enhanced Multi-task Neural Networks with Feature Interaction Learning for CTR Prediction. *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2166–2171. https://doi.org/10.1145/3404835.3463116

Guo, H., Yu, J., Liu, Q., Tang, R., & Zhang, Y. (2019). PAL: A position-bias aware learning framework for CTR prediction in live recommender systems. *Proceedings of the 13th ACM Conference on Recommender Systems*, 452–456. https://doi.org/10.1145/3298689.3347033

Hong, W., Xiong, Z., You, J., Wu, X., & Xia, M. (2021). CPIN: Comprehensive present-interest network for CTR prediction. *Expert Systems with Applications*, *168*, 114469. https://doi.org/10.1016/j.eswa.2020.114469

Jiang, Z., Gao, S., & Li, M. (2018). An improved advertising CTR prediction approach based on the fuzzy deep neural network. *PLOS ONE*, *13*(5), e0190831. https://doi.org/10.1371/journal.pone.0190831

Kaggle.com. 2022. Click-Through Rate Prediction | Kaggle. [online] Available at: <https://www.kaggle.com/competitions/avazu-ctr-prediction/data>.

Kaggle.com. 2022. Recommender System Using Amazon Reviews. [online] Available at: <https://www.kaggle.com/code/saurav9786/recommender-system-using-amazon-reviews/data>.

Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: A survey. *Decision Support Systems*, *74*, 12–32. https://doi.org/10.1016/j.dss.2015.03.008

P C D, K., V E, S., Hatamleh, W. A., Haouam, K. D., Venkatesh, B., & Sweidan, D. (2021). Advanced lightweight feature interaction in deep neural networks for improving the prediction in click through rate. *Annals of Operations Research*. https://doi.org/10.1007/s10479-021-04384-7

Parsana, M., Poola, K., Wang, Y., & Wang, Z. (2018). *Improving Native Ads CTR Prediction by Large Scale Event Embedding and Recurrent Networks*.

Shah, N., Engineer, S., Bhagat, N., Chauhan, H., & Shah, M. (2020). Research Trends on the Usage of Machine Learning and Artificial Intelligence in Advertising. *Augmented Human Research*, *5*(1), 19. https://doi.org/10.1007/s41133-020-00038-8

Sheng, X.-R., Zhao, L., Zhou, G., Ding, X., Dai, B., Luo, Q., Yang, S., Lv, J., Zhang, C., Deng, H., & Zhu, X. (2021). One Model to Serve All: Star Topology Adaptive Recommender for Multi-Domain CTR Prediction. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 4104–4113). Association for Computing Machinery. https://doi.org/10.1145/3459637.3481941

Shi, L., & Li, B. (2016). *Predict the Click-Through Rate and Average Cost Per Click for Keywords Using Machine Learning Methodologies*. 7.

Wang, T.-H., Hu, X., Jin, H., Song, Q., Han, X., & Liu, Z. (2020). AutoRec: An Automated Recommender System. In *Fourteenth ACM Conference on Recommender Systems* (pp. 582–584). Association for Computing Machinery. https://doi.org/10.1145/3383313.3411529

Wang, X., Hu, G., Lin, H., & Sun, J. (2019). A Novel Ensemble Approach for Click-Through Rate Prediction Based on Factorization Machines and Gradient Boosting Decision Trees. In J. Shao, M. L. Yiu, M. Toyoda, D. Zhang, W. Wang, & B. Cui (Eds.), *Web and Big Data* (pp. 152–162). Springer International Publishing. https://doi.org/10.1007/978-3-030-26075-0_12

Xu, Z., Li, D., Zhao, W., Shen, X., Huang, T., Li, X., & Li, P. (2021). Agile and Accurate CTR Prediction Model Training for Massive-Scale Online Advertising Systems. *Proceedings of the 2021 International Conference on Management of Data*, 2404–2409. https://doi.org/10.1145/3448016.3457236

Yoganarasimhan, H. (2020). Search Personalization Using Machine Learning. *Management Science*, *66*(3), 1045–1070. https://doi.org/10.1287/mnsc.2018.3255

Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep Learning Based Recommender System: A Survey and New Perspectives. *ACM Computing Surveys*, *52*(1), 5:1-5:38. https://doi.org/10.1145/3285029