Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
2326-9865

Integrated Weighted PageRank algorithm with Multi-Layer Perceptron for Predicting Web User Behaviour from Streams of User Interactions

Corresponding author
Mantri Gayatri[1],
P. Satheesh[2], and R. Rajeswara Rao[3]
[1]Research Scholar, JNTU, Kakinada, Andhra Pradesh, India
[2]Professor, Dept. of CSE, MVGR College of Engineering, Vizianagaram, Andhra Pradesh, India
[3]Professor, Dept. of CSE, JNTUK, University College of Engineering, Vizianagaram, Andhra Pradesh, India.
Email ID: gayatricse312@gmail.com[1], patchikolla@yahoo.com[2], rrraao.cse@jntukucev.ac.in[3]

Article Info
Page Number: 1331 – 1349
Publication Issue:
Vol. 71 No. 3s2 (2022)

Abstract
In this paper, a robust Multi-Layer Perceptron (MLP) classifier is designed to improve the efficacy of weighted PageRank algorithm that predicts the user behaviour on online e-commerce websites. Initially, the Weighted PageRank is designed to serve this purpose but with increasing users and e-commerce website, it is essential to design a fast response mechanism that assigns the weight for PageRank to possible estimate the user behaviour and tracking their behaviour over time. To achieve this, the study uses two different mechanisms i.e. user behaviour modelling and user interaction modelling to estimate their behaviour and both the models uses MLP to find the maximum extent of weights based on their interaction and behaviour. The simulation is conducted to check the efficacy of MLP-Weighted PageRank method with other existing machine and deep learning frameworks. The simulation carried out consists of following activities: Finding a web site, Building a web map, Finding the root set, Finding the base set and Check the clickstream. The results of simulation show a substantial improvement in extracting the relevant features by breaking down the problem of data dimensionality through the process of pre-training. Most of the conventional methods reported in the study failed to do so. The results show that the MLP-Weighted PageRank achieves higher rate of classification, F-measure, sensitivity, specificity and reduced mean absolute percentage error (MAPE) than other methods.

Keywords:
Multi-Layer Perceptron, Classification, User Behaviour, Weighted PageRank, Data Dimensionality

Article History
Article Received: *28 April 2022*
Revised: *15 May 2022*
Accepted: *20 June 2022*
Publication: *21 July 2022*

1. Introduction

The size of user behaviour on the web has recently increased quite rapidly, particularly human interaction and their web communication. This has greatly increased the data size. For example, Facebook held 0.1% of the content during its early creation in 2016 and it continues to increase to thousands of terabytes daily. However, any of these details were unused and discarded, though.

Web-based user behaviour is characterized as human interactions or activity and web-based relationships. The interaction of humans requires user messaging, e.g. submitting updates on social media to users, uploading and exchanging images, audio and video content. The

second aspect of consumer conduct is the development of relationships, which varies on different social media platforms. These interactions are complex. When overtime can evolve, a relationship can develop and then vanish later.

Web user conduct contributes to data intake that can be streamed in real time. Typically, this data is stored on a server. However, the data size is large and a preprocessing phase is required before machine learning or deep learning models are built. A data cleaning and data transformation process is carried out for this reason. This stage involves an extraction of features with which the data can be minimized in size and only valid data can be used. A model training and test phase is then taken and a model implementation and integration is eventually carried out.

However, for various prediction tasks, researchers tried to use the web-generated results. Predictions of web user support uses user satisfaction [1] [2], social media user behaviour [3], response times to questions [4] [5] and tweets [6], user churn [7], online advertising effectiveness prediction [8] are, for example, predictions of online customer support satisfaction.

Some researchers have focused on determining factors that influence the behavior of users on the Internet. However, an early stage of research is yet to consider the factors that influence consumer behaviour and in this the knowledge is extracted from the Web as well as the techniques for predicting behavior. Indeed, recent studies of user behavior prediction have shown that behavior is based upon multiple variables including contextual, social, temporal and spatial features [9]-[11].

For e-commerce sites, particularly for retargeting, it is important to predict consumer intention or behaviour for a particular product based on interactions within a website. Online sellers can grasp their habits and motivations more easily by keeping track of the customer search trends [12]. A rich collection of data is accessible on e-commerce sites and the users look for information on the products prior to purchasing them, thereby representing their concerns when buying. Users tend to show various search patterns that involve search frequency, time spent per object and returning visits [13].

Dimensionality reduction methods include feature selection, linear algebra methods, projection methods, and autoencoders are the methods that faces the uses related to data dimensionality. Feature extraction aids in the reduction of the quantity of redundant data in a data set by eliminating it. Finally, the reduction of the data allows for the construction of the model with less machine effort while also increasing the pace of the learning and generalisation processes in the machine learning process.

Using machine learning techniques [14], clickstream data may be used to measure search behaviour, which primarily concentrates on purchasing records. If the purchase shows the final preference of consumers in the same group, searching even for separate categories is an important element in measuring intention.

This paper addresses this intent in which the reliability of predicting web usage behaviour is critical. If a precise machine learning model is constructed and the model predicts the behaviour, in general, it takes more time to form such predictions, so it becomes worthless in practical application. Also, a collection of features must be derived during the time if a user raises a query to make a prediction using the machine learning method, where a predicting model on these features must be trained and a prediction must eventually be carried out. However, the duration to answer the query may take longer time since the prediction cannot be implemented in effect.

Modelling is a tool that allows users to understand and analyze the data generated on the website. Researchers tried using data to create web-based models of consumer behavior [15] [16]. The models are still incomplete and unfinished because of the difficulty of the task, but for particular smaller tasks, the models can be helpful. The use of graphical patterns can be

used for investigation and prediction of user interactions as well as to examine the communication dynamics [17] – [19]. Modeling user activity can also promote data discovery and help to recognize latent dynamics in human communication and website interaction.

In this paper, a Multi-Layer Perceptron (MLP) based Weighted PageRank classifier is used to predict the user behaviour on e-commerce websites. Customer Behavior Modeling is the creation of a mathematical construct using integrated page rank algorithm to represent the common behaviors observed among particular groups of customers in order to predict how similar customers will behave under similar circumstances. The interaction model is a design model that binds the MLP together in a way that supports the conceptual models of its target users. This method predicts and understands user behaviour in finding user interactions and behaviour on e-commerce websites. The former is used to assign the weights to the Weighted PageRank model by analysing the user interaction, clickstream analysis and user behaviour on the website. Secondly, the weighted PageRank and uses the link structures on other pages.

The main contribution of the paper is given below:

- The authors used MLP to obtain the weights in a faster and optimal way so as to improve the robustness of the classifiers in predicting the weights and to obtain higher summarisation quality of web documents or links.
- The MLP is combined additionally with a weighted PageRank algorithm and then the application of page rank to predict the user behaviour.
- The study uses two different models that include user interaction and user behaviour. The user interaction is then modelled with the clickstream analysis that traces the entire interarrival times and types of clicks.
- The study is finally analysed with other existing models on different datasets to test its efficacy in terms of classification accuracy and mean absolute percentage error (MAPE).

The outline of the paper is as follows: Section 2 provides the related works. Section 3 discusses the preliminaries. Section 4 provides the details of the proposed model. Section 5 evaluates the entire work. Section 6 concludes the entire work with possible directions for future scope.

2. Related works

An algorithm for online usage prediction based on an ant colony optimisation was introduced by Loyola, P. et al. [20]. The qualified ants are then reported on a new network graph, contrasted with individual sessions, which have been historically collected by Web log analysis. The key findings in this work concern an efficient forecast, which reaches about 80%, of the aggregated trends in actual usage. Second, this method enables the quantitative interpretation of the keywords that affect the navigation sessions to be achieved.

Vieira, A. [21] has proposed a comprehensive e-commerce classifier to anticipate user-based shopping intentions. We equate conventional methods of machine learning with the most recent approaches to deep learning. During the pre-training process, we demonstrated that Deep Belief and Stacked Denoising auto-encoding networks have significantly improved by extracting features from high dimensional data. They also seem to be better at addressing serious class imbalances.

Park, S. & Vasudev, V. [22] developed a software interface behaviour simulation application for web mining. Clustering helps forecast the navigation behaviour of web user by finding clusters of web user who exhibit common surfing habits. The usage of AMC enables a transitional probability estimation and the absorption at every time of successful sessions, leading to a greater personalization of the Internet and a better online publicity result. This study also provides a model-based performance assessment process and proposes a web mining scheme to enhance ad placement and optimize messaging on a website.

Dash, S., et al. [23] suggested a neurofuzzy classification and consumer behavior. A dataset is targeted, and contains device time logs containing three categories of data: local computer, network logs, and site usage logs. The 360-degree input of each individual is often used to complement the analysis. Different guidelines have been introduced to determine a user behavior policy, and may be useful in management decisions.

Setia, S., et al. [24] suggested a hybrid prediction model which would include usage mining and material mining to overcome both of those approaches' individual challenges. The suggested approach uses N-gram (continuous sequences of symbols or words or tokens in a document) with the click count for gathering contextual information. An evaluation of the hybrid method proposed was carried out using the AOL search logs, which showed a 26% improvement in estimation accuracy, and an overall increase of 10% in the hit ratio compared with other mining strategies.

Safara, F. [25] suggested a paradigm for forecasting the actions of consumers through machine learning. On the data set of an internet shopping website, five person classifiers and their groups are analyzed with bagging and boosting. The findings show that with Bagging, the highest predictor of customer behaviour was obtained using Decision Tree ensembles, with 95.3% precision. In order to assess the most significant features affecting online shopping volume during the disease of coronavirus, a correlation study will be conducted.

The reaction time and processing speed of such a learning pipeline can be problematic for many. In the first place, the function extraction step must be done in an effective manner since much data is produced by user behavior. If not, the data for constructing a model are analyzed in the event this phase takes a long time. As a consequence, the model is obsolete or even the model prediction is meaningless [26]. Secondly, model construction and testing may also be a key step. Like the feature extraction phase, if it takes a long time for the model training and testing, the predictions of this model can be reached too late for usefulness.

2.1. Problem Definition

Consider U as the entire user set accessing the e-commerce site: $\{u_1,u_2,u_3,\ldots,u_n\}$ by considering R as the total event set. Each click interaction is denoted as rt_iu_j that representing the total occurrence of the specific product displayed to a user u_j at a time interval t_i. In such cases, a real-valued vector encodes each event as $rt_iu_i \in R_d$. In the context of displaying the specific product, the users' tends to visits the page and a hierarchy of a category identity is created for the webpage with different levels of granularity that corresponds to different background information.

In this case, the categories of pre-defined webpage is defined as

$$C = \{c_1,c_2,\ldots,c_{|C|}\} \quad (1)$$

where

$|C|$ represents the total number of categories for a webpage.

If a user u_j visits a page at a time t_i, then the categories of that page is represented as an array $[c_1,c_2,c_3,\ldots]$. Further, the visited webpages by each users $u_j \in U$ is noted and it is represented as:

$$ru_j = \{rt_1u_j,rt_2u_j,\ldots,r_tmu_j\}. \quad (2)$$

On other hand, if a variety of webpages are visited by the users, the equation is modified as

$$ru_j \in R_m \times d \quad (3)$$

where

m represents the length of maximum sequences.

$R=\{ru_1,ru_2,\ldots,ru_n\}$ represents the history of webpages visited by the users, where $R \in R_n \times m \times d$, $d<|C|$

This study thus considers the user activity or behaviour from the past that are considered in the chronological order prior an arbitrary time period t_i. Such consideration is taken into account to achieve the prediction tasks as stated below:

- *User Interaction Prediction*: To predict the probability of interaction on an e-commerce website by a user at the time instant t_i based on the click response generated. This is considered as a formulation or objective function for the classification.
- *User behaviour Prediction*: To predict the product type the user purchases via a click. This is considered as a multi-objective function for the classification task.

3. Proposed Model

This section provides the problem definition, user interaction and behaviour prediction as a representation for the proposed system as illustrated in Figure 1a.

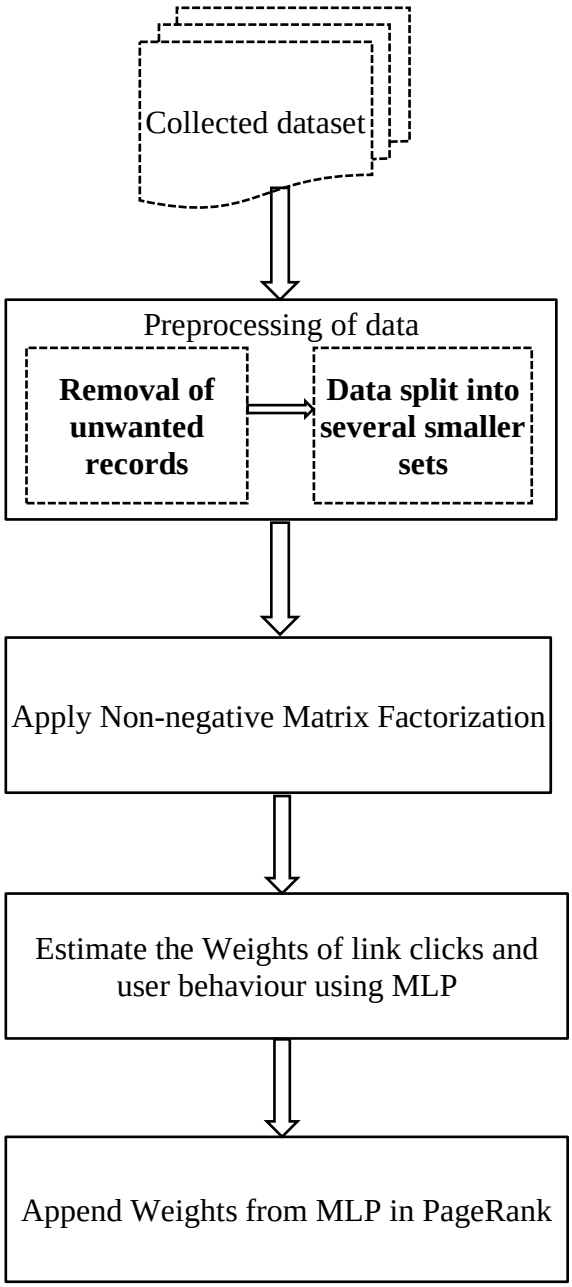


Figure 1a: Model of the Proposed Method

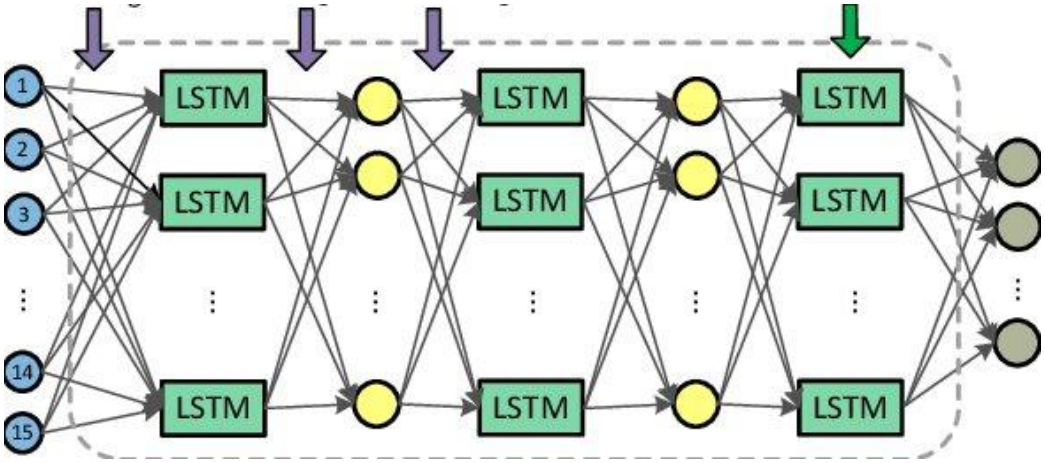


Figure 1b: LSTM Layers [31]

3.1. User Interaction Prediction

The behavior of each users is modeled as a multi-objective formulation for the task of binary classification. The study comprises the MLP model [27] with three LSTM layers [31] as in Figure 1b and a fully connected sigmoid activation layer for combining the outputs of previous hidden neurons to predict click behaviour by the users.

In such a scenario, the binary cross-entropy weighted loss function is defined that optimizing the likelihood of prediction. In the case of binary cross entropy, each of the predicted probabilities is compared to the actual class output, which can either be 0 or 1. The weights added calls for a trade-off between the precision and recall in both the classes that tends to reduce the negative effective in in classifying the instances with the class imbalance problem [28]:

$$L = -\frac{1}{N}\sum_{j=1}^{N}\left(wy_i \log\left(p\left(x_i\right)\right) + \left(1-y_i\right)\times\log\left(1-p\left(x_i\right)\right)\right) \quad (4)$$

where

N represents samples.

$y_i \in [0,1]$ represents target label, and

$p(x_i) \in [0,1]$ represents the predicted output.

$p(x_i)$ represents the maximum likelihood onto a click.

w represent the cost coefficient that determines the positive error in relationship with errors of misclassifying the negative ones.

Conventional research uses clickstream data to capture the events in web usage mining. The user history on webpages gets captured through the MLP model, which considers the user multiple sophisticated user behaviours in purchasing a product. The MLP further serves in identifying the user group with similar clickstream behaviour. This clickstream behaviour interaction is used to predict the future behaviour of users. Finally, the MLP model combines both the click interarrival times and click types are captured in the sequential order to predict well the user behaviour.

3.2. User Behaviour Prediction

The multi-class classification is established in order of classifying clicks in purchasing 10 different products by 10 different users. The number of buckets is uniformly specified in the dataset over the length of the chain. One data representation is provided by trimming the duration of the series chosen for all the longer samples. Then estimation is made via the MLP learning module. The MLP module with LSTM block is following the structure as in Figure 2 with a Softmax activation function as its last layer.

A LSTM unit is composed of a cell that includes an input gate, an output gate and a forget gate. Each cell remembers the values over arbitrary time intervals and the flow of information is regulated by the three gates into and out of the cell. Inputs are cell state from previous cell i.e., "c" superscript (t-1) and output of LSTM cell "a" super script (t-1) and input x super script (t). Outputs for LSTM cell is current cell state i.e., "c" superscript (t) and output of LSTM cell "a" super script (t).

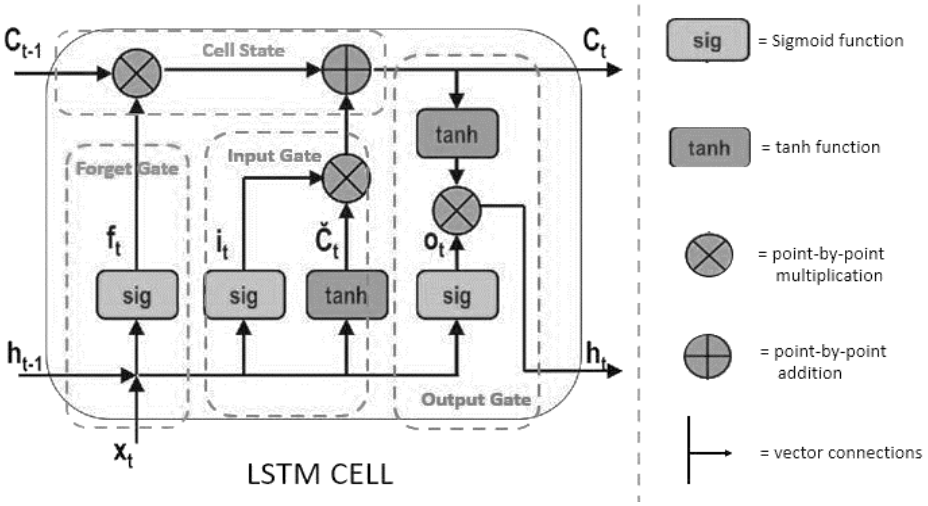


Figure 2: MLP with LSTM and Sigmoid Activation Function

The objective function in this case is equivalent to Eq.(4) when $w = 1$. In reality, $p(x_i)$ defines the MLP output after the softmax layer, which is a cross-entropy loss function.

3.3. Pre-processing

For any purchasing activity on the e-commerce website a time stamp (unique id) is registered such that the clicking instances of a user in the session is recorded. A click length is conveniently defined by finding the duration between the click times from the next click. Upon summing the total duration of the clicks for each individual item during a session, the model sets the item duration during that session. After sorting by timestamp, the study applies the itemDuration to a click data (inspection time of an item during a session). Other properties unique to a particular object are collected and added to each click data. By averaging all the features relevant to a click on a specific product, a purchasing power of a user is hence defined.

At times, the objects to be purchased are described in the form of a small texts. For handling textual data, the study uses word2vec [13] embedding and then average vector arithmetic of entire word definitions are utilised. Word2vec is a two-layer neural network that processes text by vectorizing words, as opposed to other neural networks. It takes as input a text corpus and produces as output a set of vectors: feature vectors that reflect the words in the corpus as represented by the feature vectors. Further, the entire datasets are balanced that may include both sales and non-sales events. The study mainly analyse the significance of the sample size and the efficiency of MLP, which address the data dimensionality because of its large size.

Data is presented in the JSON format and all the clicks and sessions needed to purchase a product are sorted. The click data for purchasing a product includes total number of items purchased and it is represented as B_s. The study extracts the object and session based features, where each item belongs to the entire item set.

3.4. Non-Negative Matrix Factorization

After the exclusion of unpurchased click behaviour, the study uses Non-Negative Matrix Factorization (NMF) [29] for handing the huge dimensional search space. Since NMF helps in reducing the dimensionality. NMF is an unsupervised class of learning algorithm [30] that factorises a constrained data matrix similar to learning vector quantization (LVQ) or principle components analysis (PCA). LVQ is an ANN that let's to choose training instances and learns how should those instances look like. The cons include poor generation of codebook generation procedure.

Assume V as a non-negative matrix after preprocessing results, the NMF learns well the variables of non-negative matrix W and H, in such case

$$V \cong W H \qquad (5)$$

A linear combination of the W columns, which weigh the H matrix patterns can be used to estimate each data vector V. Therefore, W can be assumed to be the basis for linear data approximation in V. As few vectors represent several data vectors, it is only possible to obtain a reasonable approximation by the base vectors to detect the structure that is latent.

NMF was effectively used for sparse data to solve high dimension issues. We used NMF in our case to compact data into a manageable subset. The main problem with NMF is that the factor matrix and stop parameters are not computed with an efficient approach to find the optimum number of features that can be chosen.

3.5. PageRank Algorithm

The PageRank algorithm is an important weighting tool for weighting a Web text. It has been developed for search engine testing and is focused on this development and servicing of the Google search engine. The more a website is popular, the more links it has to other webpages. The PageRank weighted algorithm is an extension of the PageRank traditional algorithm built on the same principle.

The PageRank algorithm focuses on the link relationship between webpages and uses the link structure existing on the webpages. In the graph, all webpages become nodes, and link relationships among webpages become edges. Every node has a unique PageRank value.

PageRank algorithm assigns values (higher rank) to common pages rather than dividing a page rank between the outlink pages. Each link page gets its popularity, i.e. its number of in-/out-links, according to the webpage popularity.

Consider a webpage u with an inlink, which is considered as the URL of another webpage that points to the link u. On other hand, the outlink points to the same link u that even refers to another webpage. Therefore the total inlinks is hence referred as $W^{in}(v,u)$ and total outlink is represented as $W^{out}(v,u)$.

$W^{in}(v,u)$ is regarded as the weight obtained via MLP for an inlink (v, u), which is estimated depending on the total inlinks for a webpage u and for the reference webpages v.

$$W^{in}(v,u) = \frac{I_u}{\sum_{p \in R(u)} I_p} \qquad (6)$$

where

I_p represent the total inlinks for a webpage p,

I_u represent the total inlinks for a webpage u

$R(v)$ represents the total reference pages for a webpage v.

Similarly, $W^{out}(v,u)$ is regarded as the weight obtained via MLP for an outlink (v, u), which is estimated depending on the total ouylinks for a webpage u and for the reference webpages v.

$$W^{out}(v,u) = \frac{O_u}{\sum_{p \in R(v)} O_p} \qquad (7)$$

where
I_p represent the total outlinks for a webpage p,
I_u represent the total outlinks for a webpage u
$R(v)$ represents the total reference pages for a webpage v.
The Weighted PageRank Formula is based on the validity of all pages defined by the number of links and outlinks:

$$PR(u) = (1-d) + d \sum_{v \in B(u)} PR(v) W^{out}(v,u) W^{in}(v,u) \qquad (8)$$

Here,
$PR(u)$ represents the Weighted PageRank for a webpage u.
d represents the damping factor.
The Weighted PageRank states that a surfer who clicks blindly on links finally stops clicking. The damping factor is the likelihood, at any point, that the user will proceed.

3.6. MLP

Networks that attempt to map input into output are known as Multilayer Perceptron (MLPs). In an MLP, the nodes are arranged into layers, each of which is connected to every other layer. Nonlinear activation functions are used for each node in the hidden layers. The neural network is honed using a backpropagation technique. Backpropagation activation functions is explained in the next section.
For example, MLPs use feedforward neural networks that are trained on incoming data. Feedforward neural networks are neural networks that transmit data from their input to their output. An MLP consists of numerous layers of neural networks, each of which is linked to the next. Each hidden layer node is activated by a nonlinear activation function. An algorithm called backpropagation is used to train this network, and features for activating and learning the process are provided below. Backpropagation is a technique that is used to calculate derivatives in a short amount of time. Backpropagation is used as a learning technique in artificial neural networks, and it is used to compute a gradient descent with respect to weights.
The linear regression model is given as below:

$$f(x) = w^T x + b \qquad (9)$$

where
x - input
w - weighted matrix
b – bias

Activation Function:

For the purpose of training, two distinct activation functions are being used in the research. Hyperbolic tangent is an activation function, and its evaluation runs from -1 to 1. It is detailed below.

$$y(v_i) = \tanh(v_i) \qquad (10)$$

Logistic function is utilized to validate the evaluation criteria between the range 0 and 1 and it is given as follows:

$$y(v_i) = \left(1 + e^{-v_i}\right)^{-1} \qquad (11)$$

where
y_i - i^{th} neuron output and
v_i - weighted input sum.

LSTM Learning:

Adjusting link weights in an MLP network can be done once the data for each neuron has been analysed. It does not matter how many errors there are in the output compared to the projected results in the study.

$$error_j(n) = d_j(n) - y_j(n) \quad (12)$$

where
d - Expected value and
y - Target value.

The model ensures modification for reducing the error probability onto output of MLP and it is given below:

$$e(n) = 0.5\sum_i error_j^2(n) \quad (13)$$

The changing weight upon the gradient descent application is hence modelled as below:

$$\Delta w_{ji}(n) = -\eta \frac{\partial e(n)}{\partial v_j(n)} y_i(n) \quad (14)$$

where
y - output of backward layer
η - Learning rate.

A derivative is defined for the overall output node in the study based on numerous induced local fields:

$$-\frac{\partial e(n)}{\partial v_j(n)} = error_j(n)\phi'(v_j(n)) \quad (15)$$

where
ϕ' - constant or activation function derivative.

In order to make the analysis easier, it is important to offer the following expression of a relevant derivative when there is a change in weights in the hidden layers:

$$-\frac{\partial e(n)}{\partial v_j(n)} = \phi'(v_j(n))\sum_k -\frac{\partial e(n)}{\partial v_k(n)} w_{kj}(n) \quad (16)$$

The change in node weights in the output layer affects this relative derivative. For this reason, we must first change the output layer weights in order to alter the hidden layer weights. The activation function is thus backpropagated in this approach.

The output layer node weight shift determines the value of this relative derivative. In order to update the hidden layer weights, the ML adjusts the output layer weights based on the activation function derivative.

The PageRank algorithm measures the weights equally for all node-related edges. However, if PageRank is used in many places, separate weights are used at the edges. Weighted-PageRank algorithm assigns various weights (estimated by MLP) based on the types of web links. Various relationships (i.e. blocks, duplicates, and dependencies) are also taken into consideration in addition to the user behaviour. In addition, MLP attains the weight of these pages and sends as input to the PageRank algorithm to obtain the maximum quality in its output, where the weights may vary based on the various types of relationship.

MLP first assigns each weight to the correlation report graph to implement the weighted-PageRank algorithm. First, MLP assigns all nodes a *PageRank* = 1/*Node*. Then each value of

the PageRank converged to a PageRank value of the corresponding sentence using the PageRank Algorithm. The higher the score, the higher the value of the phrase, and the more similar the phrase. The weighted score propagation is close to the original PageRank. If weighted PageRank are used for MLP, $PR(A)/C(A)$ is then multiplied with W, where W=[0,1].

3.7. Ranking Merger

This section uses MLP to determine the end score for each click by means of the scores of the sentences computed by a PageRank algorithm. A modified weighted PageRank algorithm score of a link clicked is specified as SWPR and every score is standardized for a value from 0 to 1 and then summed up. Therefore the score is estimated as in Eq.(17)

$$Score = \alpha \times SWPR + (1-\alpha) \times S_{pred} \quad (17)$$

S_{pred} represents the prediction score of MLP

SWPR represents the score of links clicked by a user and

α is the weight; $0 \leq \alpha \leq 1$.

The optimal score is obtained when α is 0.25 and the final score is estimated by the weighted sum of all these scores. Finally, the scores are sorted and the top results are extracted.

4. Results and Discussions

In this section, the proposed weighted PageRank algorithm is evaluated in terms user interaction and user behaviour models in comparison with machine learning classifiers namely Naive Bayes (NB), Random Forest (RF), Logistic Regression (LR) and Support Vector Machine (SVM). These methods are compared with similar set of training or testing data with same weighted PageRank algorithm. The model is tested with two different datasets that includes Post-View Click Dataset and Multi-Campaign Click Dataset.

4.1. Experimental Settings

Five comparative approaches are used including the proposed model. TensorFlow and CUDA are takes the advantage of GPU and trained with an Adam optimization in the form of gradient descent variant for the MLP model. The existing models are created using the Python scikit-learn library.

The simulation carried out consists of following activities:

1. Finding a web site:
2. Building a web map using JSpider software
3. Finding the root set: retrieving page set related to a given query using IR search engine embedded in the web site.
4. Finding the base set: A base set is created by expanding the root set with pages that directly point to or are pointed to by the pages in the root set.
5. Check the clickstream from the base set using Integrated WPR with algorithms that is applied to the base set.

The dataset is divided into training, testing and validation sets for training models with a ratio of 70:20:10 ratio. The evaluation uses one-hot encoding to transform sequential data into binary vector that eliminates the less commonly used links. The premature convergence is utilised to handle overfitting problems in MLP to stop the validation after 10 epochs. For MLP, the drop-out rate is set as 0.4 with L2 regularisation and a five-fold cross validation is conducted on all datasets. The entire set of parameters considered for simulation is given in Table.1.

Table.1. Parameters used by the MLP

Parameters	Value
Dropout ratio	[0: 0.3]
Training epochs for single hidden layer	[10: 100]
Training epochs for more hidden layers	[10: 150]
Total hidden layer	500 units.
Minimum hidden units in a layer	16
Annealing delay fraction	[0: 1]
Initial learning rate	[0.001: 0:25]
Momentum	[0: 0.95]
L2 weight cost	[0: 0.01]
Hidden unit activation function	logistic sigmoids
Noise level in input layer	[0: 0.2].

4.2. Performance Metrics

The result of the model is compared in terms of accuracy, precision, recall, F1-score and MAPE.

Accuracy is a metric that summarizes the performance of the proposed model as the number of correct predictions divided by the total number of predictions. Precision (positive predictive value) is the ratio of relevant click instances among the retrieved click instances, while recall (sensitivity) is the ratio of relevant click instances that were retrieved.

The accuracy for the classifier is defined as below

Accuracy = $(TP+TN) / (TP+TN+FP+FN)$ (16)

where,

TP is defined as the True Positive rate

TN is defined as the True Negative rate

FP is defined as the False Positive rate

FN is defined as the False Negative rate

The precision of the classifier is defined as below:

Precision = $TP / (TP+FP)$ (17)

The Recall of the classifier is defined as below:

Recall = TP/(TP+FP) (18)

The MAPE of the classifier is obtained as below:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \quad (19)$$

M = mean absolute percentage error

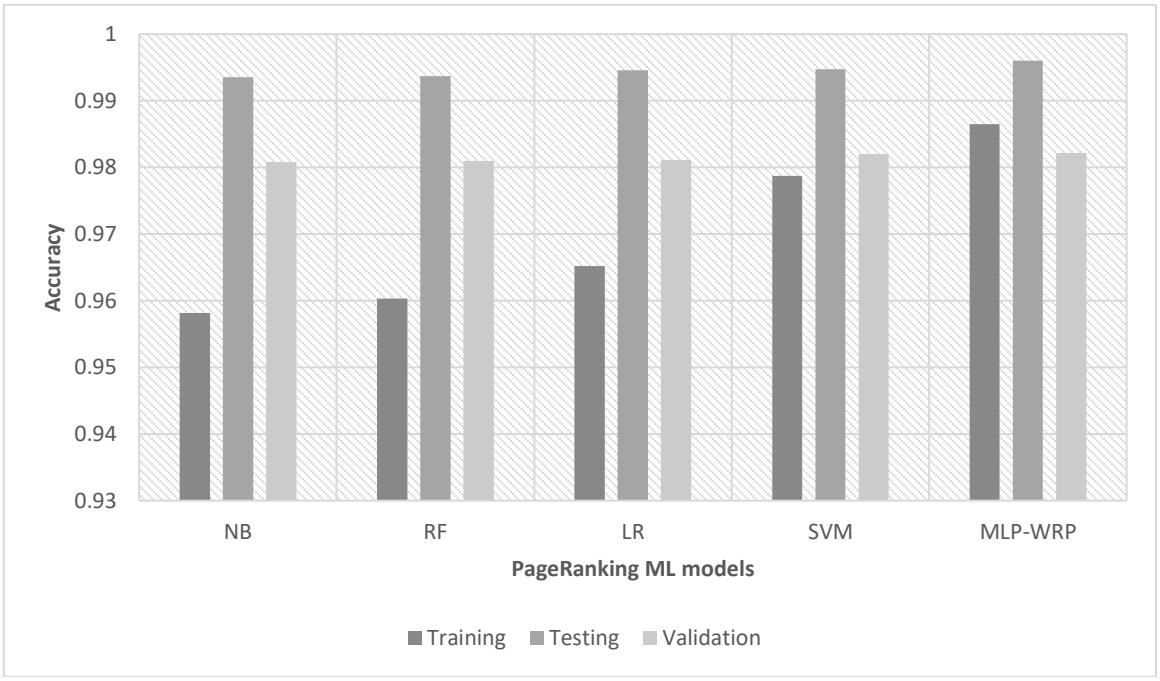
N = number of times the summation iteration happens

A_t = actual value

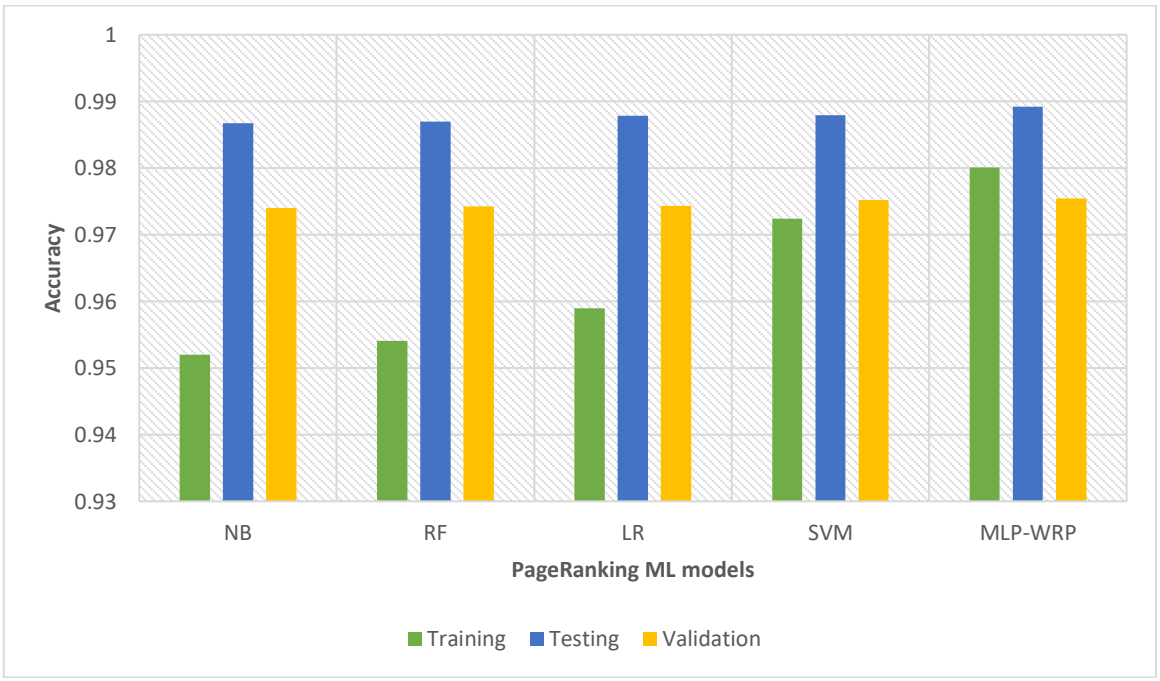
F_t = forecast value

4.3. Performance Metrics

The Figure 3 shows the results of classification accuracy, where the proposed method achieves higher classification accuracy than other methods in both the datasets. In the Post-View Click Dataset, the training accuracy is 98.65%, the testing accuracy is 99.60% and validation accuracy is 98.22%. Similarly, the Multi-Campaign Click Dataset, the training accuracy is 98.01%, the testing accuracy is 98.92% and validation accuracy is 97.54%. The utilisation of Weighted PageRank model with the application of MLP on finding the weights of inlinks and outlinks enables the proposed method to perform better than other methods.

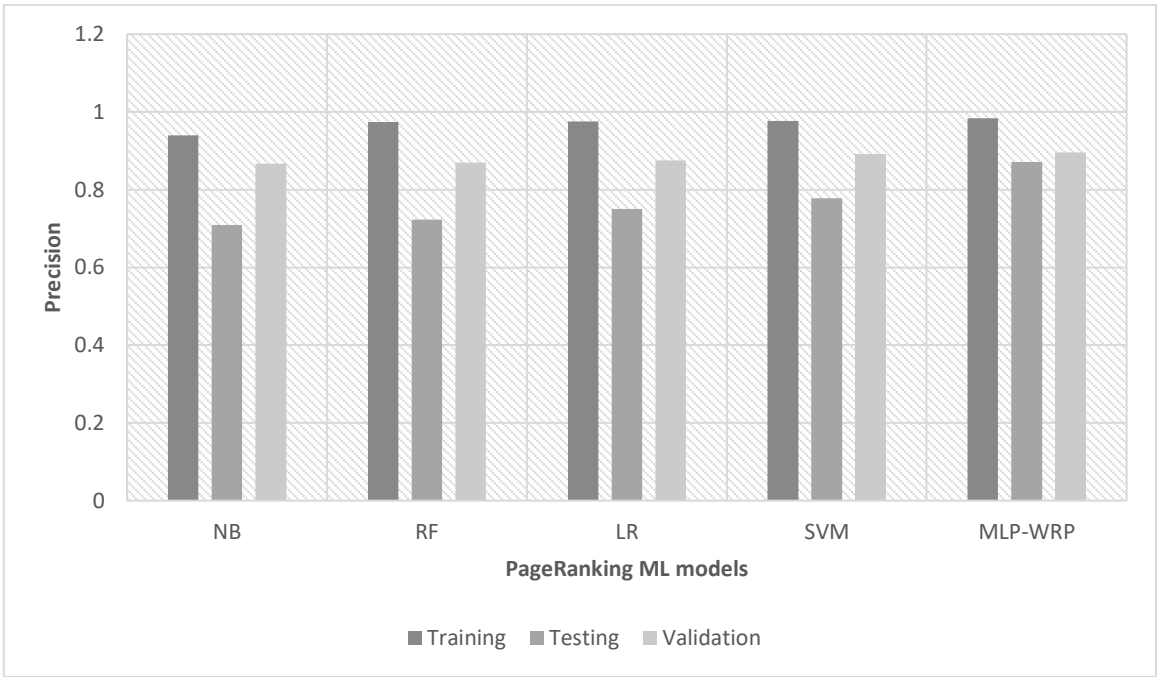


(a) Post-View Click Dataset

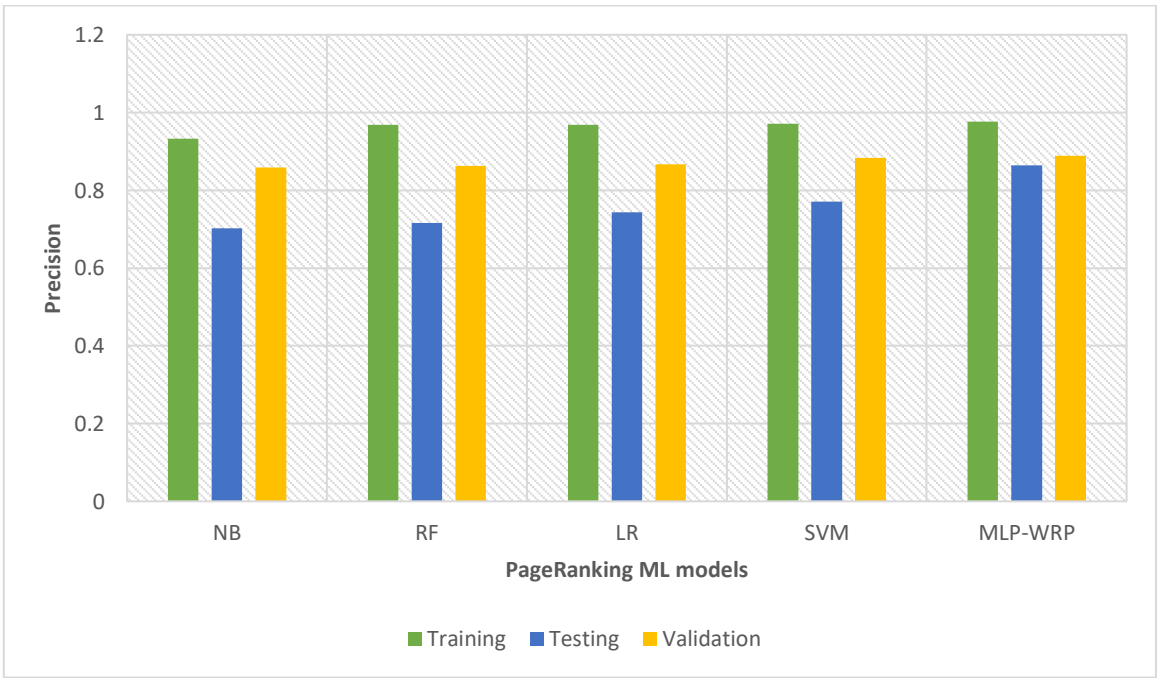


(b) Multi-Campaign Click Dataset
Figure 3: Classification Accuracy

The Figure 4 shows the results of precision, where the proposed method achieves higher precision rate than other methods in both the datasets. In the Post-View Click Dataset, the training precision is 98.37%, the testing precision is 87.16% and validation precision is 89.63%. Similarly, the Multi-Campaign Click Dataset, the training precision is 97.73%, the testing precision is 86.40% and validation precision is 88.83%. The utilisation of Weighted PageRank model with the application of MLP on finding the weights of relevant inlinks and outlinks in relation with the true inlinks and outlinks enables the proposed method to perform better than other methods.

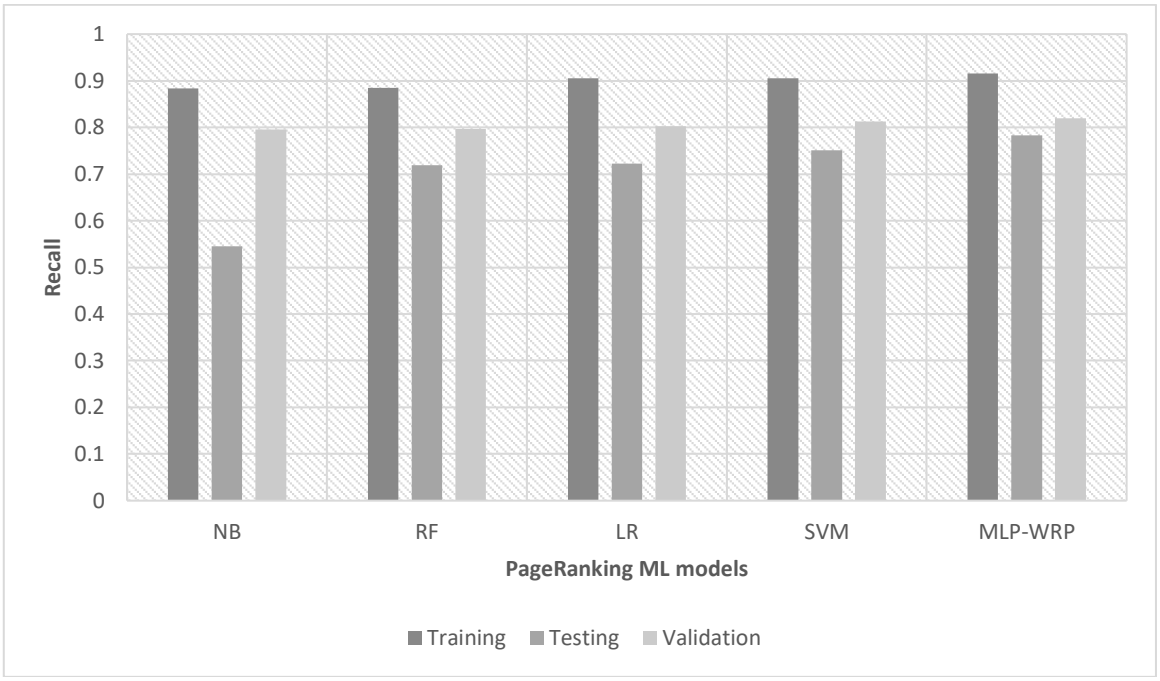


(a) Post-View Click Dataset

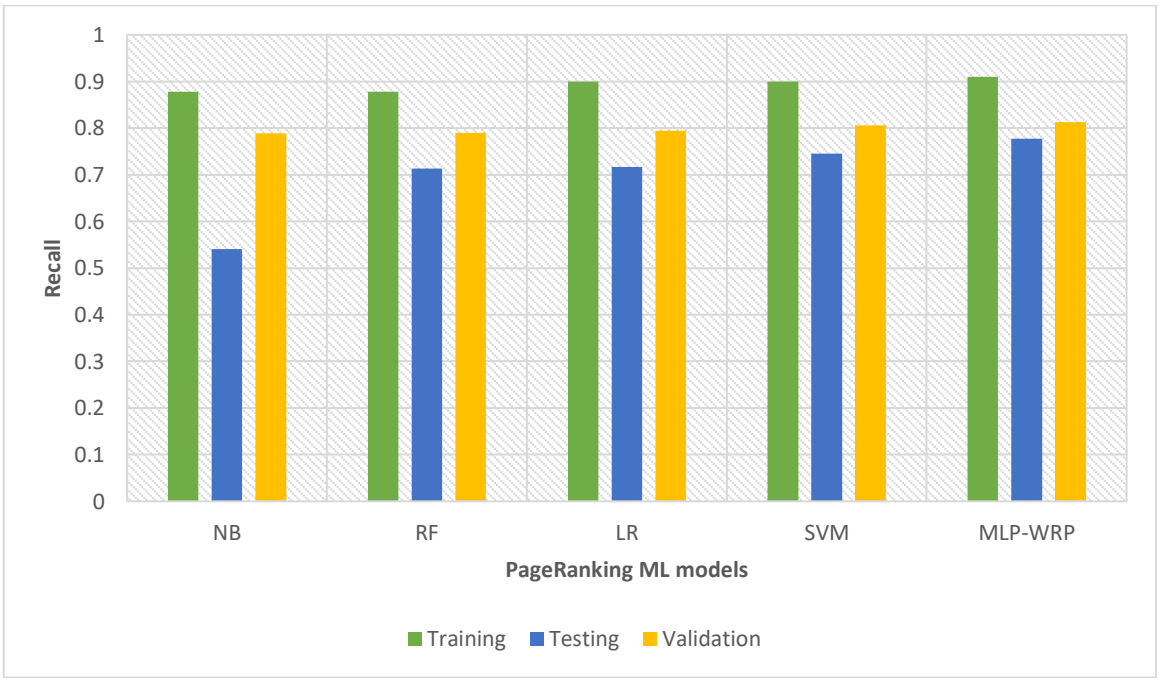


(b) Multi-Campaign Click Dataset
Figure 4: Precision

The Figure 5 shows the results of recall, where the proposed method achieves higher recall rate than other methods in both the datasets. In the Post-View Click Dataset, the training recall is 91.56%, the testing recall is 78.34% and validation recall is 81.97%. Similarly, the Multi-Campaign Click Dataset, the training recall is 90.96%, the testing recall is 77.67% and validation recall is 81.24%. The utilisation of Weighted PageRank model with the application of MLP on finding the weights of true inlinks and outlinks belonging to the positive class enables the proposed method to perform better than other methods.

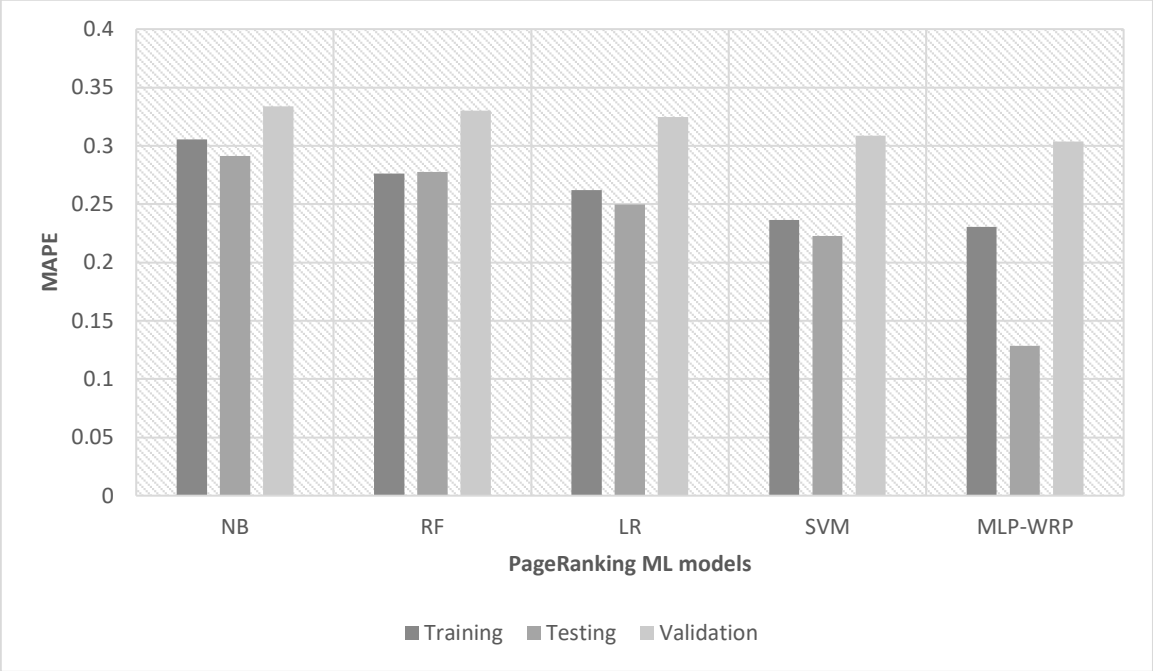


(a) Post-View Click Dataset



(b) Multi-Campaign Click Dataset
Figure 5: Recall

The Figure 6 shows the results of MAPE, where the proposed method achieves reduced classification accuracy than other methods in both the datasets. In the Post-View Click Dataset, the training MAPE is 23.04%, the testing MAPE is 12.82% and validation MAPE is 30.36%. Similarly, the Multi-Campaign Click Dataset, the training MAPE is 22.88%, the testing MAPE is 12.72% and validation MAPE is 30.09%. The utilisation of Weighted PageRank model with the application of MLP enables maximum prediction accuracy in finding the related links on the user clicks than other models.



(a) Post-View Click Dataset

(b) Multi-Campaign Click Dataset
Figure 6: MAPE

5. Conclusions

In this paper, the study applies MLP-Weighted PageRank classifier for modelling the user behaviour interaction in e-commerce website. The MLP-Weighted PageRank classifier efficacy of weighted PageRank algorithm in predicting the user behaviour than existing deep or machine learning classifiers. The MLP tracks and predicts the user behaviour, thereby the weights are generated in faster manner such that the Weighted PageRank using its graph based theory acquires its weights via MLP to find the relative weights of a webpage document and its link relationship between the e-commerce webpages and the link structures

of the webpage. The simulation conducted on robust environment with multiple users in e-commerce website to predict the user behaviour attained an improved response by MLP-Weighted PageRank classifier than conventional classifiers.

6. Future Work

Additional research may involve real-time data testing and real-time efficiency impacts. More work needs to be undertaken, however, to improve the time quality. The data is incredibly scarce in terms of scalability and the algorithms used fails to parallelise the tasks on multi-core machines. As long as a strong difference in prediction accuracy between the algorithms used with large data sizes is shown, the implications of using even greater training data will be fascinating to see. Moreover, multiple ID-based features or words may be considered as beneficial for initialising the unsupervised deep or machine learning classifier on large datasets.

References

[1] Karnstedt, M., Hennessy, T., Chan, J., & Hayes, C. (2010, August). Churn in social networks: A discussion boards case study. In *2010 IEEE Second International Conference on Social Computing* (pp. 233-240). IEEE.

[2] Glady, N., Baesens, B., & Croux, C. (2009). Modeling churn using customer lifetime value. *European Journal of Operational Research*, *197*(1), 402-411.

[3] Zhu, Y., Zhong, E., Pan, S. J., Wang, X., Zhou, M., & Yang, Q. (2013, October). Predicting user activity level in social networks. In *Proceedings of the 22nd ACM international conference on Information & Knowledge Management* (pp. 159-168).

[4] Yang, L., Bao, S., Lin, Q., Wu, X., Han, D., Su, Z., & Yu, Y. (2011, August). Analyzing and predicting not-answered questions in community-based question answering services. In *Twenty-Fifth AAAI Conference on Artificial Intelligence*.

[5] Dror, G., Maarek, Y., & Szpektor, I. (2013, September). Will my question be answered? predicting "question answerability" in community question-answering sites. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 499-514). Springer, Berlin, Heidelberg.

[6] Weerkamp, W., & De Rijke, M. (2012, August). Activity prediction: A twitter-based exploration. In *SIGIR workshop on time-aware information access* (pp. 1-4).

[7] Kousik, N., Natarajan, Y., Raja, R. A., Kallam, S., Patan, R., & Gandomi, A. H. (2021). Improved salient object detection using hybrid Convolution Recurrent Neural Network. *Expert Systems with Applications*, *166*, 114064.

[8] Kadhim, R. R., and M. Y. Kamil. "Evaluation of Machine Learning Models for Breast Cancer Diagnosis Via Histogram of Oriented Gradients Method and Histopathology Images". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 4, Apr. 2022, pp. 36-42, doi:10.17762/ijritcc.v10i4.5532.

[9] McDuff, D., El Kaliouby, R., Cohn, J. F., & Picard, R. W. (2014). Predicting ad liking and purchase intent: Large-scale analysis of facial responses to ads. *IEEE Transactions on Affective Computing*, *6*(3), 223-235.

[10] Scellato, S., Noulas, A., & Mascolo, C. (2011, August). Exploiting place features in link prediction on location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1046-1054).

[11] Guille, A., & Hacid, H. (2012, April). A predictive model for the temporal dynamics of information diffusion in online social networks. In *Proceedings of the 21st international conference on World Wide Web* (pp. 1145-1152).

[12] Bulla, P. . "Traffic Sign Detection and Recognition Based on Convolutional Neural Network". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 4, Apr. 2022, pp. 43-53, doi:10.17762/ijritcc.v10i4.5533.

[13] Adedoyin-Olowe, M., Gaber, M. M., & Stahl, F. (2013). A survey of data mining techniques for social media analysis. *arXiv preprint arXiv:1312.4617*.

[14] Benevenuto, F., Rodrigues, T., Cha, M., & Almeida, V. (2009, November). Characterizing user behavior in online social networks. In *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement* (pp. 49-62).

[15] Banerjee, A., & Ghosh, J. (2000). Concept-based clustering of clickstream data.

[16] Beutel, A. (2016). User behavior modeling with large-scale graph analysis. *Computer Science Department, Carnegie Mellon University*.

[17] Dimitrov, A. G., Lazar, A. A., & Victor, J. D. (2011). Information theory in neuroscience. *Journal of computational neuroscience*, *30*(1), 1-5.

[18] Yu, S., & Kak, S. (2012). A survey of prediction using social media. *arXiv preprint arXiv:1203.1647*.

[19] Xiong, R., & Donath, J. (1999, November). PeopleGarden: creating data portraits for users. In *Proceedings of the 12th annual ACM symposium on User interface software and technology* (pp. 37-44).

[20] Graells-Garrido, E. (2014, April). Enhancing web activities with information visualization. In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 39-44).

[21] Frau, S., Roberts, J. C., & Boukhelifa, N. (2005). Dynamic coordinated email visualization. In *WSCG05-13th International Conference on Computer Graphics, Visualization and Computer Vision'2005* (pp. 187-193).

[22] Ghazaly, N. M. . (2022). Data Catalogue Approaches, Implementation and Adoption: A Study of Purpose of Data Catalogue. International Journal on Future Revolution in Computer Science &Amp; Communication Engineering, 8(1), 01–04. https://doi.org/10.17762/ijfrcsce.v8i1.2063

[23] Koehn, D., Lessmann, S., & Schaal, M. (2020). Predicting online shopping behaviour from clickstream data using deep learning. *Expert Systems with Applications*, *150*, 113342.

[24] De Bruyn, P. C. (2021). Predicting behavioral profiles of online extremists through linguistic use of social roles. *Behavioral Sciences of Terrorism and Political Aggression*, *13*(4), 295-319.

[25] Wu, Y., Jiang, Q., Ni, S., & Liang, H. E. (2021). Critical Factors for Predicting Users' Acceptance of Digital Museums for Experience-Influenced Environments. *Information*, *12*(10), 426.

[26] 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), 1-15.

[27] Agarwal, D. A. . (2022). Advancing Privacy and Security of Internet of Things to Find Integrated Solutions. International Journal on Future Revolution in Computer Science &Amp; Communication Engineering, 8(2), 05–08. https://doi.org/10.17762/ijfrcsce.v8i2.2067

[28] Setia, S., Jyoti, V., & Duhan, N. (2020). HPM: A Hybrid Model for User Behavior Prediction Based on N-Gram Parsing and Access Logs. *Scientific Programming*, *2020*.

[29] Safara, F. (2020). A Computational Model to Predict Consumer Behaviour During COVID-19 Pandemic. *Computational Economics*, 1-14.

[30] Yuvaraj, N., Srihari, K., Dhiman, G., Somasundaram, K., Sharma, A., Rajeskannan, S., ... & Masud, M. (2021). Nature-Inspired-Based Approach for Automated Cyberbullying Classification on Multimedia Social Networking. Mathematical Problems in Engineering, 2021.

[31] Vinayak, B., Lee, H. S., & Gedem, S. (2021). Prediction of Land Use and Land Cover Changes in Mumbai City, India, Using Remote Sensing Data and a Multilayer Perceptron Neural Network-Based Markov Chain Model. *Sustainability*, *13*(2), 471.

[32] Natarajan, Y., Kannan, S., & Mohanty, S. N. (2021). Survey of Various Statistical Numerical and Machine Learning Ontological Models on Infectious Disease Ontology. Data Analytics in Bioinformatics: A Machine Learning Perspective, 431-442.

[33] Neher, E., & Taschenberger, H. (2021). Non-negative matrix factorization as a tool to distinguish between synaptic vesicles in different functional states. *Neuroscience*.

[34] Raja, R. A., & Kousik, N. V. (2021). Analyses on Artificial Intelligence Framework to Detect Crime Pattern. Intelligent Data Analytics for Terror Threat Prediction: Architectures, Methodologies, Techniques and Applications, 119-132.

[35] Zhu, W., Lan, C., Xing, J., Zeng, W., Li, Y., Shen, L., & Xie, X. (2016, March). Co-occurrence feature learning for skeleton based action recognition using regularized deep LSTM networks. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 30, No. 1).

[36] P. Modiya and S. Vahora, “Brain Tumor Detection Using Transfer Learning with Dimensionality Reduction Method”, Int J Intell Syst Appl Eng, vol. 10, no. 2, pp. 201–206, May 2022.

[37] Rathish, C. R., and A. Rajaram. "Efficient path reassessment based on node probability in wireless sensor network." International Journal of Control Theory and Applications 34.2016 (2016): 817-832

[38] S Rahamat Basha, Chhavi Sharma, Farrukh Sayeed, AN Arularasan, PV Pramila, Santaji Krishna Shinde, Bhasker Pant, A Rajaram, Alazar Yeshitla, "Implementation of Reliability Antecedent Forwarding Technique Using Straddling Path Recovery in Manet," Wireless Communications & Mobile Computing (Online), vol. 2022, 2022.

[39] Rathish, C. R., and A. Rajaram. "Hierarchical Load Balanced Routing Protocol for Wireless Sensor Networks." International Journal of Applied Engineering Research 10.7 (2015): 16521-16534.

[40] D. N. V. S. L. S. Indira, Rajendra Kumar Ganiya, P. Ashok Babu, A. Jasmine Xavier, L. Kavisankar, S. Hemalatha, V. Senthilkumar, T. Kavitha, A. Rajaram, Karthik Annam, Alazar Yeshitla, "Improved Artificial Neural Network with State Order Dataset Estimation for Brain Cancer Cell Diagnosis", BioMed Research International, vol. 2022, 10 pages, 2022.

[41] P. Ganesh, G. B. S. R. Naidu, Korla Swaroopa, R. Rahul, Ahmad Almadhor, C. Senthilkumar, Durgaprasad Gangodkar, A. Rajaram, Alazar Yeshitla, "Implementation of Hidden Node Detection Scheme for Self-Organization of Data Packet", Wireless Communications and Mobile Computing, vol. 2022, 9 pages, 2022. https://doi.org/10.1155/2022/1332373.

[42] M. . Parhi, A. . Roul, B. Ghosh, and A. Pati, “IOATS: an Intelligent Online Attendance Tracking System based on Facial Recognition and Edge Computing”, Int J Intell Syst Appl Eng, vol. 10, no. 2, pp. 252–259, May 2022.

[43] M. Dinesh, C Arvind, S.S Sreeja Mole, C.S. Subash Kumar, P. Chandra Sekar, K. Somasundaram, K. Srihari, S. Chandragandhi, Venkatesa Prabhu Sundramurthy, "An Energy Efficient Architecture for Furnace Monitor and Control in Foundry Based on Industry 4.0 Using IoT", Scientific Programming, vol. 2022, Article ID 1128717, 8 pages, 2022. https://doi.org/10.1155/2022/1128717.