Predicting Purchase Intent: Online Shopping Behavior Learning Through Recurrent Neural Networks

Teena Rathore[1] and Amit Singhal[2] and Manish Tiwari[3]

[1]Geetanjali Institute of Techanical Studies, Udaipur, India.

Abstract

Online purchases are phenomena that are growing rapidly nowadays. We present a neural network to predict purchase intent in an electronic commerce environment. We use trainable vector spaces to model semi-structured and varied input data that comprise categorical and unique instances. In this paper, we compare the traditional machine learning technique with the more advanced deep learning approaches. The convenience of online shopping makes it an emerging trend among consumers. The prevalence of online shopping has sparked the interest of retailers to focus on this area. Therefore, this study determined the relationship between the subjective norm, perceived utility and online shopping behavior while mediated by purchase intention. Purchase intentions are often measured and used by marketing managers as an input to decisions about new and existing products and services. An exploration of the design decision of the model, which includes the sharing of parameters and the omission connections, further increases the accuracy of the model. The results in reference datasets provide classification accuracy within 98% of the state of the art in one and exceed the state of the art in the second without the need for any domain-specific feature engineering / data set in events short and long sequences It is interesting to note that perceived utility also negligibly influences the behavior of online purchases. The discovery also revealed that significant purchase intent positively influences online shopping behavior. In this paper Naïve Bayes (NB), BayesNet, Lazy.LWL and Lazy.IBK algorithms are applied in Weka for predicting purchasing intent online shopping behavior model. There are 18 attributes are used in different algorithms to show terms and error rates with accuracy. Lazy.LWL algorithm gives best results in all algo with minimum error rate.

Keywords : Recurrent Neural Network, WEKA Algorithm, e-commerce, Machine learning.

1. Introduction

In mobile e-commerce, a large data set is available and potential consumers seek information about the product before making purchasing decisions, thus reacting to consumers' buying intentions. Users show different search patterns, that is, the time spent per article, the frequency of search and recurring visits [1]. Click flow data can be used to quantify search behavior using machine learning techniques [5], mainly focused on purchase records. While purchases indicate the final preferences of consumers in the same category, the search is also an essential component to measure intentionality towards a special category.

In the e-commerce domain, merchants can increase their sales volume and profit margin by getting better answers for two questions:

• Which users are more likely to buy (predict purchase (intention)).

• What elements of the product catalog users prefer (to classify the content).

	McKinsey	A.T. Kearney	Affected by shopping intent
Price management	11.1%	8.2%	Yes
Variable cost	7.8%	5.1%	Yes
Sales volume	3.3%	3.0%	Yes
Fixed cost	2.3%	2.0%	No

Table 1: Effect of the improvement of different variables on the operating benefit of [22]. In three out of four categories, knowing more about a user's purchase intention can be used to improve the profits of merchants.

To what extent can merchants realistically increase profits? Table 1 illustrates that traders can improve profits between 2% and 11%, depending on the contributing variable. In the fluid and highly competitive world of online retail, these margins are significant, and understanding a user's buying intent can positively influence three of the four main variables that affect profit. In addition, merchants increasingly depend (and pay advertising to) much larger third-party portals (eg, eBay, Google, Bing, Taobao, Amazon) to achieve their distribution, so any direct measure that the merchant group can use to increase their profits is very necessary

The objective of this document is to identify the activity patterns of certain users that lead to buying sessions and then extrapolate them as templates to predict a high probability of purchase in related websites. The data used consists of approximately 1 million sessions containing user's click data; however, only 3% of the training data consist of purchasing sessions, which makes it a very unbalanced data set. The rest of this document is organized as follows: Section 2 describes the data used in our study and the methods of preprocessing and non-negative matrix factorization for the reduction of dimensionality. Section 3 presents the classification algorithms. Section 4 describes in detail the deep learning algorithms (Deep Belief Networks and Auto-encoders noise elimination stacked) and Section 5 presents the results.

When a merchant is more confident that they are more likely to buy a subset of users, they can use this information in the form of proactive actions to maximize conversion and performance. The merchant can offer a limited time discount, spend more on targeted advertising (and relevant) to re-engage these users, create complementary product packages to push the user to complete their purchase, or even offer their own brand with a lower price. . alternative if the product is considered fungible.

However, there are balances the desire to create ever more accurate models of user behavior online, ie, user privacy and ease of implementation. Users are increasingly reluctant to share personal information with online services, while complex Machine Learning models are difficult to implement and maintain. in a production environment [28]. We surveyed the existing work in this area and we found that well Implementation approaches have a number of factors in common:

• A large investment was needed in the engineering of specific characteristics of the data set, regardless of the implementation of the chosen model.

• The options of models they favor techniques such as gradual reinforcement machines (GBM) [7] and field-aware factoring machines

(FFM) [24] what are well-suited to create representations of Data of click flow semi-structured once good characteristics have been developed [26], [36], [34]. In [29], a class of important entity used the notion of similarity of elements, modeled as a learned vector space generated bywordword2vec [18] and calculated using a metric cosine of couple Standard vector elements. In an electronic commerce context, the elements are more similar if they occur frequently in all user sessions in the corpus and are different if they occur infrequently. The items themselves can be physically different (for example, headphones and batteries), but can often be searched and bought together.

However, in common with other works, [29] it still requires a large investment in feature engineering. The drawback of specific features is that they are linked to a domain, data set or both. In addition, artificial neural networks (ANNs) work well with distributed representations, such as inlays, and ANNs with a recurrence capability to model events over time - Recurrent Neural Networks (RNN) - are suitable for processing sequences and homework of labeled [16] Therefore, our motivation is to build a good prediction model of the user's intention that is not based on the user's private data and that is also easy to implement in a real-world environment. What performance can RNNs achieve with adequate input representation and an end-to-end training regime in the prediction of purchases task of intention? Can this performance be achieved within the restriction of processing only Anonymous session data and continue to be easy to implement in other e-commerce data sets?

2. Literature survey

In order to identify the relevant studies, an electronic search and a We searched the number of index databases of academic journals. so Then, titles and The summaries of the studies. Keywords and phrases used in the literature review were online shopping, online shopping, online consumers, online purchase behavior, online shopping behavior, online consumer behavior and Electronic consumer behavior. Therefore, conceptual documents and articles that use other types of research methods are out of the scope of this study.

In the results of this Bibliographic analysis, many independent variables are identified. While some of the independent variables cited in Table 1 refer to a single article, other independent The variables appear in multiple articles. All the variables are classified according to their similarities.

Hsu , Chaung and Hsu (2014) they study trust from four prospects different : website, provider , initiator of the auction and group members. They find a positive effect of trust in intent to purchase online only on the website, the provider and the group members, but not for an auction starter.
Wu and Lee (2012) they study trust from a different perspective. They investigate the reliability blog on place of the websites and the reliability of the provider's blog. They claim that bloggers have an impact on the consumer's purchase intention. However, they do not find a significant impact of the blog's reliability in the intent to purchase online.

Li et al. (2007, p.272) defines perceived risk as "consumer perceptions of the uncertainty and the adverse consequences of participating in an activity." Everyone Perceived risk studies they have a negative impact on consumers' online shopping intentions. It implies that if consumers think that buying is a big risk on the Internet due to security or privacy issues, their online purchases decrease.

The perception of oneself efficieny It is another dimension, as the previous research shows. According *Wang et al. (2010, p 56),* the self-efficacy he known as " auto -evaluation of the consumer of their abilities to buy online. " If the level of self - affirmation of consumer It is high, your online purchases increase.

Revenue is also analyze in most studies. Consumers who have higher income levels are more likely to buy online than those who have lower income levels (Calik Y Ersoy , 2008; Doolin et al., 2005). Conversely, Clemes and others (2014) establish that consumers who have high income levels do not tend to buy online because they prefer to buy branded products. in retail stores to have a good user experience and get support and service. It is stated that the previous experience of online shopping is positively related to online purchases.

4 OUR APPROACHES

Classic machine learning approaches, such as the GBM, work well and are widely used in e-commerce data, at least in part because the data is structured. GBM is an efficient model since it allows an additive expansion in a set of functions of base or weakness of the students to continually minimize a residual error. A weakness of GBM is the propensity for decision trees too deep or wide fit the training data and, therefore, recorded poor performance validation and test suite because of the high variance [33] , [3. 4]. even though this can be controlled by hyperparameters (namely, tree depth, learning speed, minimum weight to divide a tree node (min_child_weight), and data sub-sampling). GBM also requires a significant effort of feature engineering and, of course, does not process the sequence in order, but consumes a compressed version of it (although it is possible to provide a vector representation of the input sequence as a characteristic).

4.1 Embeddings as item / word representations

Tasks of Natural Language Processing (NLP), such as information retrieval, part of the discourse labeled and fragmentation, operate assigning a probability value for a sequence of words. For this, the language. Models have been developed, defining a mathematical model for capture the statistical properties of words and dependencies between they. Learning good representations of input data is a central task in the design of an automatic learning model that can work well. A the incrustation is a model of vector space where the words become to a low-dimensional vector.

4.2 Recurrent neural networks

Recurrent Neural Networks [27] (RNN) is a specialized class of Neural networks for sequential data processing. A recurring network it is deep in time instead of space and organizes the hidden state vectors hl t on a two-dimensional grid, where t = 1...T it is thought of how time and l =1 ... L It is the depth. All intermediate vectors hl t they are calculated according to hlt -one Yh l - one t Through these hidden vectors, each output Y at some particular time, the step t it becomes an approximation function of all the vectors introduced up to that moment, x one,..., xt [13]

4.2.1 LSTM and GRU. Long-term long-term memory (LSTM) [11] is an extension to colloquial or vanilla RNNs designed to address the Twin problems of disappearance and explosion of gradients during training[19]. Leakage gradients make learning difficult like the right one. The trajectory (downward) of the gradient is difficult to discern, whereas the explosive gradients make the training unstable, both are undesirable results The long-term dependencies on the input data, causing a deep computational graph that you should iterate over the data are the root cause of the disappearance/explosion of gradients. [8] Explain this phenomenon succinctly. Like all deep learning models, RNNs require multiplication by a matrix W. After t steps, this amounts to multiplying by Wt. Thus:

Wt = (Vdiaд(λ)V −1)t = Vdiaд(λ)tV −1 (1)

The eigenvalues (λ) that are not more or less equal to 1 will explode if they are > 1, or they will disappear if they are < 1. Then, the gradients will be scaled by diad (λ)t.

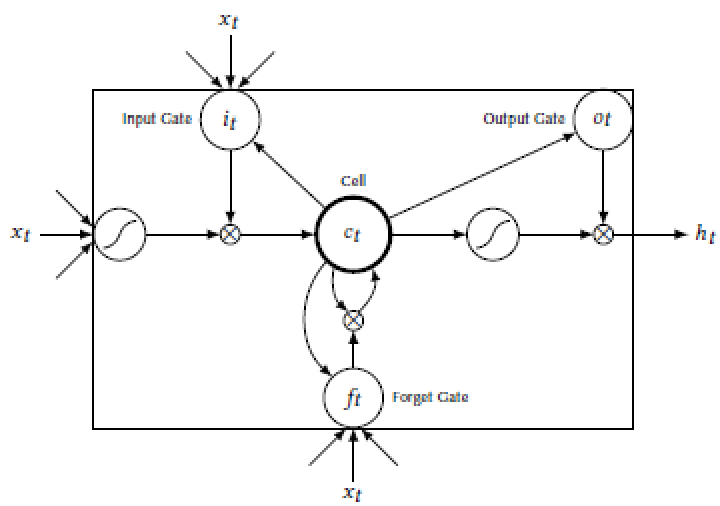

Fig : 1 A single LSTM cell, depicting the hidden and cell states, as well as the three gates controlling memory.

5. IMPLIMENTATION

5.1 Datasets used

Purchase intention of online buyers the data set provides anonymous e-commerce click-through data that is suitable for testing purchase prediction models. The data set consists of vectors of characteristics that belong to 12,330 sessions. The data set was formed so that each session would belong to a different user in a period of 1 year to avoid Any tendency to a specific campaign, day special , user. profile , or period. Online buyers d In the 12,330 sessions in the data set, 84.5% (10,422) were negative class samples that did not end with the purchases, and the rest (1,908) were positive class samples that end with purchases.

The data set consists of 10 numeric attributes and 8 categorical attributes.
The attribute 'Revenue' can be used as a class label.

"Administrative", "Administrative duration", "Informative", "Information duration", "Related to the product" and "Duration related to the product" represent the number of different types of pages visited by the visitor in that session and the time Total invested in each of these categories of pages. The values of these characteristics are derived from the URL information of the pages visited by the user and are updated in real time when a user performs an action. The functions "Bounce Rate", "Output Rate" and "Page Value" represent the metrics measured by "Google Analytics" for each page on the e-commerce site. The value of the "Bounce Index" function for a web page refers to the percentage of visitors that enter the site from that page and then leave ("bounce") without activating any other request to the analysis server during that session. The value of the "Output rate" function for a specific web page is calculated as for all visits of page to the page, the percentage that was the last in the session. The "Page Value" function represents the average value of a web page that a user visited before completing an e-commerce transaction. The "Special Day" feature indicates the proximity of the site visit time to a specific special day (for example, Mother's Day, Valentine's Day) in which the sessions are most likely to end with the transaction. The value of thisattribute is determined considering the dynamics of electronic commerce, such as the duration between the date of the order and the delivery date. For example, to the day by Valentina , this value takes a non-zero value between February 2 and February 12, zero before and after this date unless it is close to another special day, and its maximum value of 1 on February 8. The data set also includes system operations, navigator, region, type of traffic, type of visitor as recurrent or new visitor, a Boolean value that indicates whether the date of the visit is the weekend and the month of the year.

The experimental results were made in this data set using the well-known data extraction software WEKA to show the efficiency of the classification of machine learning algorithms and to justify the proposed algorithm. We have tested three algorithms (JRip, J48, Naïve Bayes (NB)) to show the best results with accuracy and error rate and number of derived rules.

5.2 Experimental Result

We have shown experimental results with the Algorithm Lazy.LWL gives the 87.33% accuracy which is highest then others algorithms. It gives minimum error rate which is 0.16%. We have shown accuracy and error rate comparisons in graphical representation. In which BayesNet, NaiveBayes and Lazy.IBK represents 81.59, 83.64 and 81.53 % accuracy respectively. And Lazy.LWL shows 87.33 % highest percentage of accuracy.

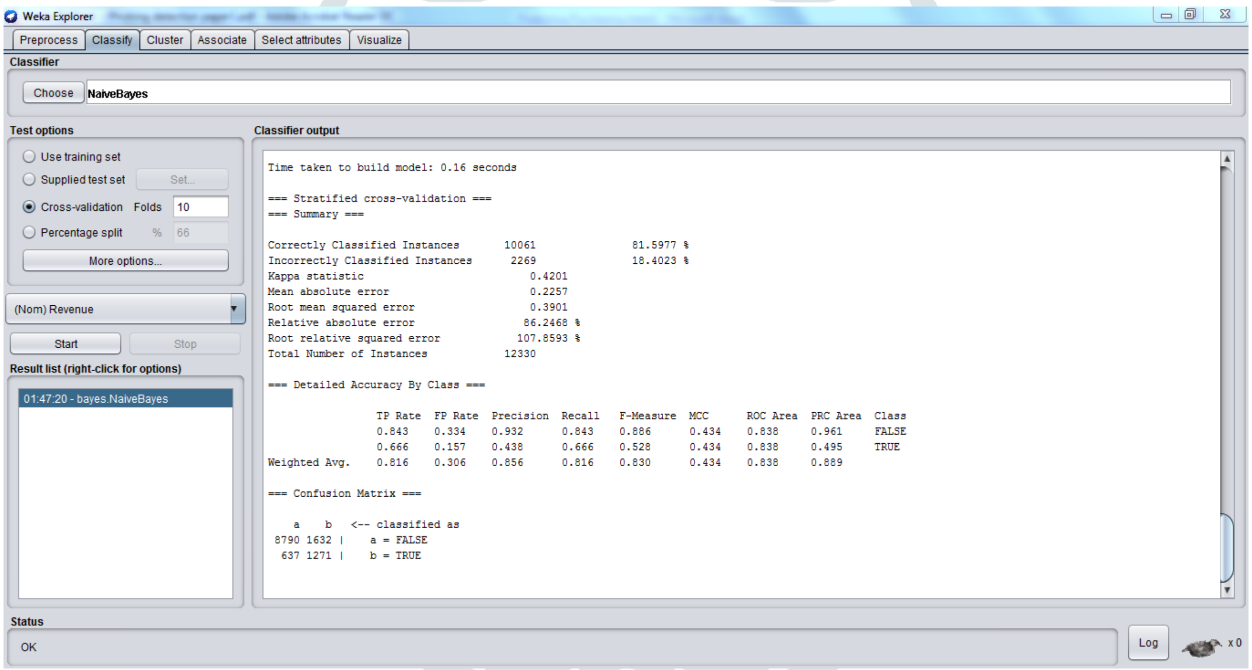

Fig 2: Result with NaiveBayes Algo.

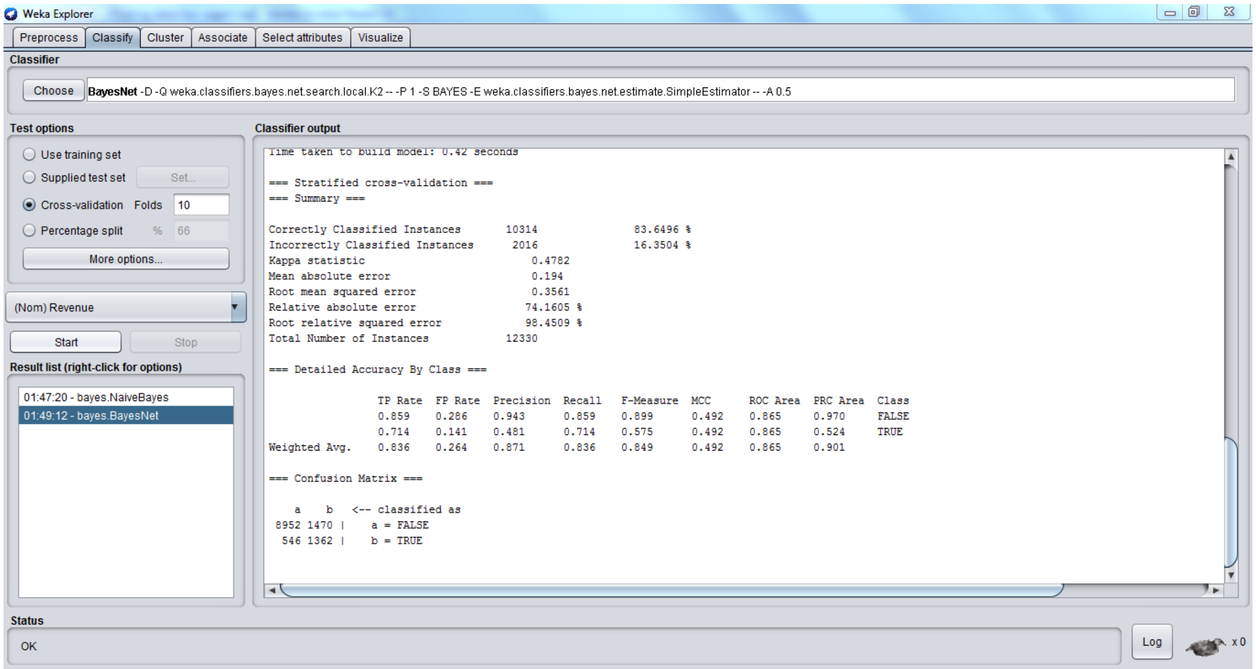

Fig 3: Result with BayesNet Algo.

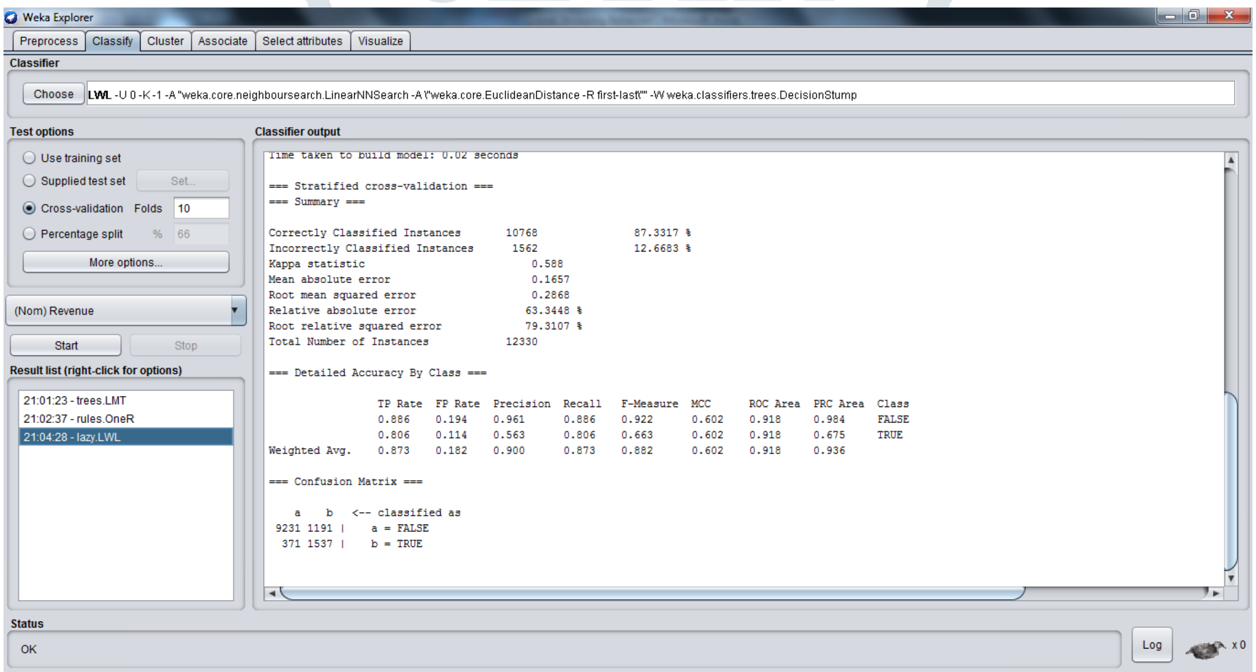

Fig 4: Result with Lazy.LWL Algo.

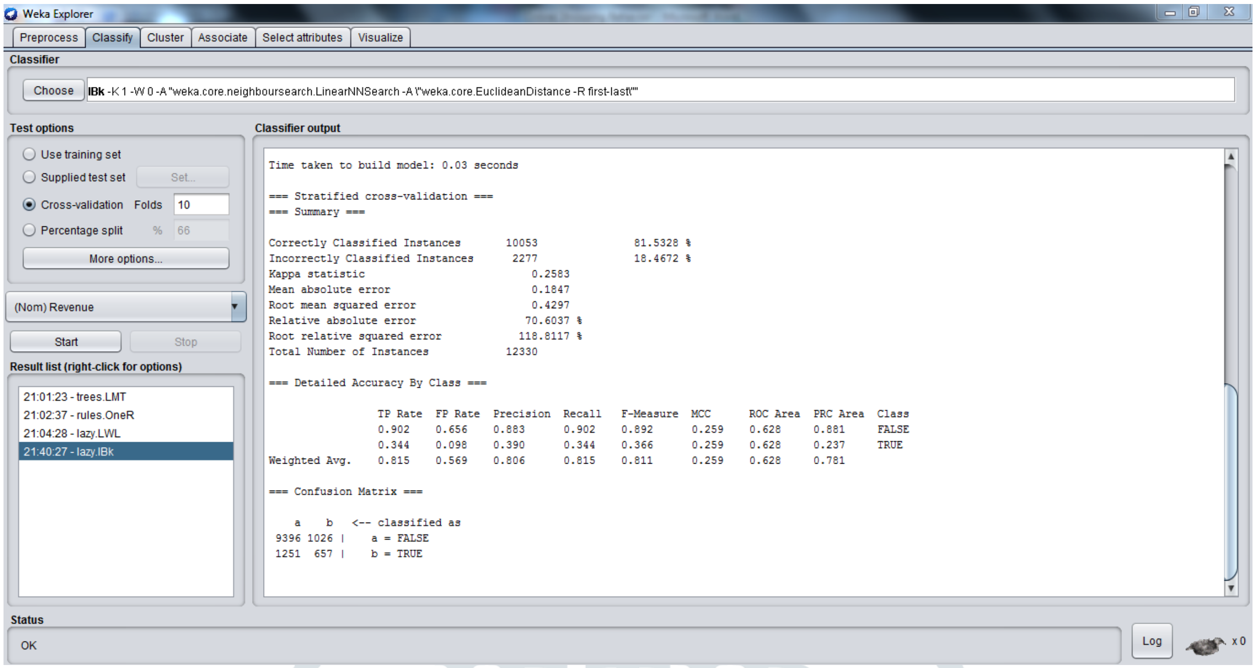

Fig 5: Result with Lazy.IBK Algo.

	WEKA Algorithms	Accuracy Rate %	Error Rate %
Websites Data	NB	81.59	0.22
	Bayes Net	83.64	0.19
	Lazy.LWL	87.53	0.16
	Lazy.IBK	81.53	0.18

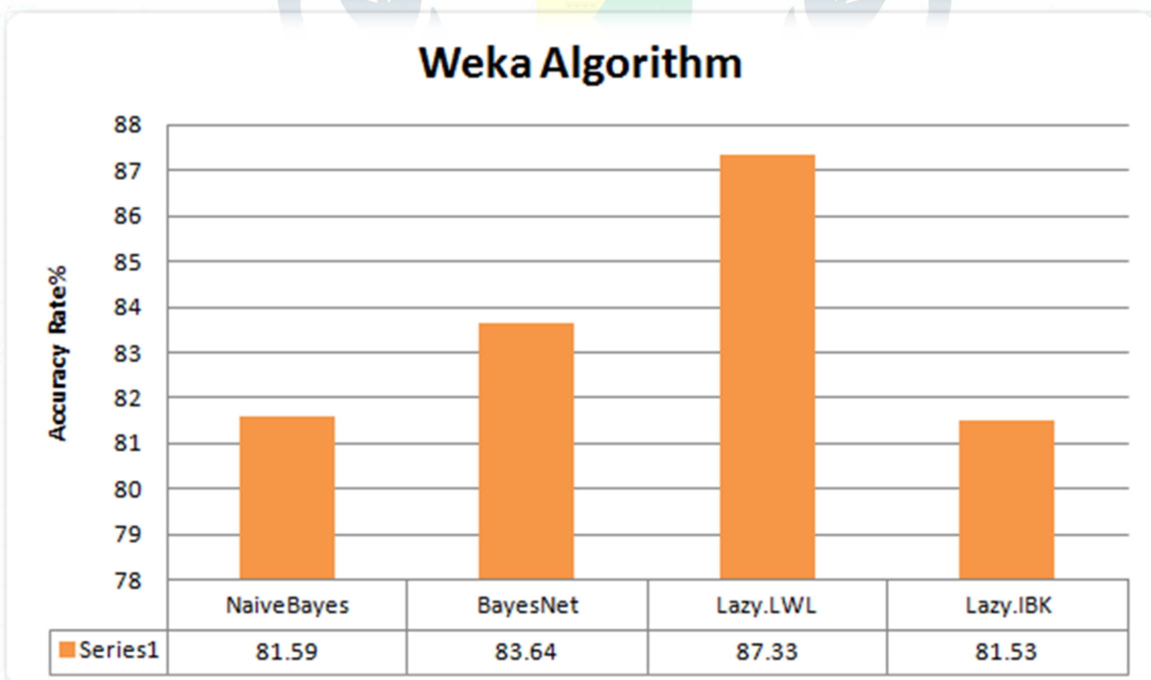

Fig.6 : Graphical Representation of Accuracy Rate %

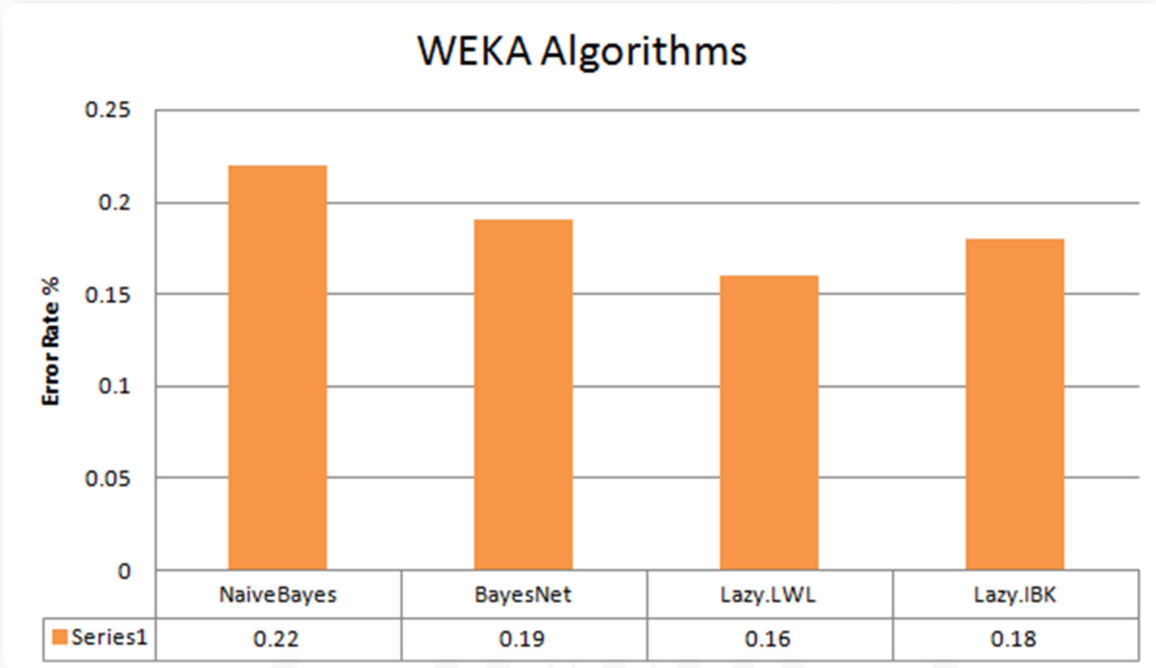

Fig.7 : Graphical Representation of Error Rate %

6. Conclusions and future work

We present a model of recurrent neural network (RNN) that based on Bayes algorithm. Problem of prediction in electronic commerce without using explicit characteristics. The model is simple to implement, generalizes to different datasets comparable yield and can be trained with modest hardware resource requirements. It is promising that closed RNNs without feature engineering can be competitive with Gradient Boosted Machines in short session Lengths and structured data. Choice in the domain of recommendation systems and electronic commerce in general. We create additional work in the input representation (although it continues to avoid feature engineering) it can further improve the results For RNNs both gated and non-gated. A focus area will be to investigate how to share parameters in the hidden layer help RNNs to operate on short sequences of prevalent structured data in electronic commerce. Last , We note that although our approach does not require any feature Engineering is also intrinsically transductive : we plan to investigate Embed generation and maintenance approaches for new Elements/documents invisible to add an inductive capacity to the architecture.

Experiment result show that our approach achieves good accuracy for phishing detection, indicating the effectiveness of the proposed mechanism. The previous study that is analyzed, on the other hand, exposes a performance decline due to the evolution of the phishing ecosystem, while proposed methodology and feature set demon states significant superiority.

In this paper there are three to four algorithms are applied to show the highest accuracy and less error rate. Algorithm Lazy.LWL gives the 87.33% accuracy which is highest then others algorithms. It gives minimum error rate which is 0.16%. We have shown accuracy and error rate comparisons in graphical representation. In which BayesNet, NaiveBayes and Lazy.IBK represents 81.59, 83.64 and 81.53 % accuracy respectively. And Lazy.LWL shows 87.33 % highest percentage of accuracy.

7. ACKNOWLEDGEMENTS

We would like to thank the authors of [26] for making their original Proof of shipping available and the organizers of the original. Challenge in releasing the solution file after the competition ended, Allowing us to make our comparisons.

REFFERENCES

2. Pray Barkan and Noam Koenigstein . 2016. Item2Vec: Incrustation of neuronal elements for him collaborative filtering. CoRR abs / 1603.04259 (2016). arXiv : 1603.04259 http: //arxiv.org/abs/1603.04259
3. Veronika Bogina Y Tsvi Kuflik. 2017. Incorporation of the time of permanence in the recommendations based in sessions with recurrent neural networks. In RecTemp @ RecSys.
4. Tianqi Chen and Carlos Guestrin . 2016 XGBoost : A scalable tree that increase System. In the Proceedings of the 22nd ACM International Conference SIGKDD on Knowledge discovery and data mining (KDD '16). ACM, New York, NY, USA, 785 794 https://doi.org/10.1145/2939672.2939785
5. KyungHyun Cho, Bart van Merrienboer , Dzmitry Bahdanau Y Yoshua Bengio. 2014. In The properties of automatic neural translation: Encoder-decoder approaches. CoRR abs/1409.1259 (2014). arXiv : 1409.1259 http://arxiv.org/abs/1409.1259
6. Xiao Ding, Ting Liu, Junwen Duan Y Jian-Yun Nie. 2015. Consumption of mining users.
7. Intention of social networks using the adaptive domain convolutional Neural network In the procedures of the twenty-ninth Conference of the AAAI on Artificial Intelligence (AAAI'15) .AAAI Press, 2389-2395. http://dl.acm.org/citation.cfm ? id = 2886521.2886653
8. Jerome H. Friedman 2000. Approximation of the greedy function: a machine of improvement of degraded. Annals of Statistics 29 (2000), 1189-1232.
9. I am Goodfellow , Yoshua Bengio , and Aaron Courville . 2016. Deep learning. MIT Press. http://www.deeplearningbook.org.
10. Mihajlo Grbovic , Vladan Radosavljevic , Nemanja Djuric , Narayan Bhamidipati , Jaikit Savla , Varun Bhagwan an Doug Sharp. 2016. E-commerce in your inbox: Product recommendations to scale. CoRR abs / 1606.07154 (2016). arXiv : 1606.07154 http://arxiv.org/abs/1606.07154
11. Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, Y Domonkos Tikk. Recommendations based on session with recurring neural networks. CDR abs/1511.06939 (2015). arXiv : 1511.06939 http://arxiv.org/abs/1511.06939
12. Sepp Hochreiter Y Jürgen Schmidhuber . 1997. Long-term long-term memory. Neural Comput. 9, 8 (November 1997), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735
13. Dietmar Jannach, Malte Ludewig , and Lukas Lerche. 2017. Article based on session recommendation in electronic commerce: about short-term objectives, reminders, trends and descuentosModelado user interaction and adapted by the user 27 (2017), 351-392.
14. Andrej Karpathy , Justin Johnson and Fei-Fei Li 2015 Visualization and understanding
15. Recurrent networks CoRR abs / 1506.02078 (2015). ArXiv : 1506.0207 http://arxiv.org/abs/1506.02078
16. Diederik P.Kingma and Jimmy Ba . 2014. Adam : a method for stochastic optimization. CoRR abs/1412.6980 (2014). arXiv : 1412.6980 http://arxiv.org/abs/1412.6980
17. Yehuda Koren. 2009. Collaborative filtering with temporal dynamics. In proceedings of the 15th International Knowledge Conference ACM SIGKDD Discovery and data mining (KDD '09). ACM, New York, NY, USA UU., 447-456. https://doi.org/10.1145/1557019.1557072
18. Zachary Chase Lipton. 2015. A critical review of recurrent neural networks for sequential learning. CoRR abs/1506,00019 (2015). arXiv:1506,00019 http://arxiv.org/abs/1506.00019
19. Malte Ludewig Y Dietmar Jannach 2018. Evaluation of the Recommendation based on the session Algorithms CoRRabs/1803.09587(2018). arXiv :1803.09587 http://arxiv.org/abs/1803.09587
20. Mikolov, Kai Chen, Greg Corrado and Jeffrey Dean. 2013. Efficient Estimation of representations of words in the vector space. CoRR abs/1301.3781 (2013). arXiv:1301.3781 http://arxiv.org/abs/1301.3781
21. Razvan Pascanu, Tomas Mikolov Y Yoshua Bengio. 2013 In The difficulty of Training in Recurrent Neural Networks.