Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020)
IEEE Xplore Part Number: CFP20N67-ART; ISBN: 978-1-7281-5374-2

Particle Swarm Optimized Neural Network for Predicting Customer Behaviour in Digital Marketing

Dr N Srikanth Reddy
Assistant Professor
School of Management
Presidency University
Itigalapura, Bangalore 560064
srikanthreddyn36@gmail.com

Abstract*-In the digital marketing world, the analysis of online customer's behaviors is considered as one of the key applications that could encourage focusing on more innovative advertisements. Such applications rely a lot on the online customer data analysis for tuning the online ads return on investment. Since the machine learning algorithm called Neural Network (NN) has been certified as the well-performing classifier to handle the complex tasks in different fields, this paper plans to perform the prediction of customer's behavior on digital marketing. Here, the data collection will be done manually based on the knowledge gained from different data formats concerning various online ads like display advertising, social media, native advertising, video ad etc. Further, the data normalization is performed for removing the redundant or duplicate data, which is followed by optimal trained NN for prediction. A well-performing optimization algorithm termed Particle Swarm Optimization (PSO) is used for training the NN and it is compared over several other machine learning algorithms for validating the effectiveness of the proposed model. PSO is a computational method that optimizes a problem by iteratively trying to improve a candidate solution with regard to a given measure of quality. In this paper, the accuracy of the developed PSO-NN was acquiring the best results when compared to conventional NN at any number of hidden neurons. On considering the number of hidden neurons as 25, the accuracy of the introduced PSO-NN was 4.2% advanced than NN. Therefore, it is confirmed that the introduced PSO-NN is performing well in customer behaviour prediction under digital marketing.

Keywords-Digital Marketing; Machine Learning; Neural Network; Optimal Training; Particle Swarm Optimization

Nomenclature

Abbreviations	**Descriptions**
NN	Neural Network
PSO	Particle Swarm Optimization
IoT	Internet of Things
ANN	Artificial Neural Networks
RNN	Recurrent Neural Network
FPR	False Positive Rate
CNN	Convolution Neural Networks
NPV	Negative Predictive Value
DMC	Digital Marketing Communication
FDR	False Discovery Rate
HHS	Health and Human Services
FNR	False Negative Rate
BTM	Biterm Topic Model
B2B	Business 2 Business
CRF	Conditional Random Fields
MCC	Matthew's Correlation Coefficient
ELM	Elaboration Likelihood Model
SEO	Search Engine Optimization
VAR	Vector Autoregressive model
SEM	Search Engine Marketing

I. INTRODUCTION

The rapid growth of communication and information techniques both in public and private sectors in past decades have started the development of a novel digital marketing environment. A massive amount of data is being generated as of now because of the proliferation of information technology. It is observed that "2.5 quintillion bytes" of information is generated daily, which improves the onset of IoT [1]. In addition, it has also been anticipated that over the past two years, 90 percent of the global data has been produced [2]. Based on the quality of knowledge produced, accessed, and employed, the timely and accurate business decisions are taken. Therefore, by introducing new data-oriented approaches to marketing strategy, rapid technological growth, and its barrier-free global distribution generate prospects for financial benefits [3]. Efficient and effective approaches are in demand to retrieve valuable information from the data sets. Data sets with constant changes are generated in various areas everyday [25]. RNNs have the ability to learn temporal sequence by processing arbitrary input sequences using internal memory units [26].

Digital marketing has emerged as a natural response for businesses to exploit and gain from an effective accumulation of customers on the network [4]. The digital marketing is employed as a part of their marketing mechanisms and deployment programs are used in different organizations namely "businesses, hospitals, schools, professional associations, councils, and NGOs" [5]. Few of them have the ability to work on their renowned E-commerce website, still, the internet is used as the channel in many cases in their communication mechanism. In general, these organizations accomplish the client roles or marketers often termed as brands. Many organizational groups also function in the field of digital marketing. For the organizations present in the first set, digital agencies generate and introduce marketing models and employ digital marketing as part of their own marketing mechanism. In third group, organizations such as media are employed with digital agencies for target audience communication.

In machine learning, the new developments specifically in ANNs have demonstrated the best outcomes in learning the hidden patterns automatically from the data [6]. The authors in [7] have seen the evolving implementations in the last century with deep NN learning models in language processing [8] [9] namely "speech recognition [10], video sequence analysis [11], time series predictions [12]" or even designing the customer information that is clicked by them on website for recommendation [3]. The machine learning researchers are motivated by these applications. In recent times, the effective implementation of NNs to such challenges has revealed remarkable outcomes and has motivated many researchers for identifying more strategies for applying NNs to similar concerns. Moreover, in order to improve performance experimentally, optimal trained NN is performing well [6]. The performances are compared over traditional machine learning models that made probable for enabling the implementations in different datasets. A vast amount of data is collected in the field of digital marketing. This propriety allows novel implementations in online advertising [13] [14] and the recommended system [15] [16].

The main contribution of this paper is given below.

- To develop an innovative model for digital marketing that is concentrated on analyzing the behaviors of online customers on diverse advertisements.
- To implement the optimal trained NN with the integration of a meta-heuristic algorithm called PSO that optimizes the weight to attain minimum error difference between the predicted and target output.

The present paper is modeled as shown in the below manner: Section II specifies the literature review and features and challenges of existing prediction models. The improved machine learning algorithm for predicting customer behavior on digital marketing is depicted in Section III. Section IV speaks about the optimal trained NN for predicting customer behaviour in digital marketing. The results and discussions related to this paper are shown in Section V. Finally, the entire paper's conclusion is shown in Section VI.

II. LITERATURE REVIEW

A. Related Works

In 2018, Cui *et al*. [7] have implemented NN to the digital marketing environment in order to reach the potential customers. Initially, with the structure of NN, the online behaviour of the customer was modelled. In the search engine, collecting the user's information starting from the search keywords for the landing page and various succeeding pages was done until the user leaves the site. By using RNN, and CNN, the entire journey of the visited pages were modelled, which provided the semantic keyword used by the user for searching and visited pages. In order to analyze the conversion rates of every user that was visited further, “Monte Carlo simulation” was employed.

In 2019, Miklosik *et al*. [17] have concentrated on the usage and selection of machine learning-driven analytical tools with three different sets namely "media companies, advertisers, and advertisers". The samples on these organizations functioning in Slovakia were analyzed by performing quantitative and qualitative research. The authors have screen lighted some of the observations like (a) during the creation and deployment of the marketing mechanisms, the analytical tools play a significant role, (b) required more knowledge on developing methods like artificial intelligence and machine learning algorithms, (c) the effective implementation of machine learning tools, and (d) the usage of machine learning-driven analytical tools and the low-level of adoption in marketing management. In order to assist the organizations, a framework that includes process maps and enablers was introduced for recognizing the opportunities and executing the projects successfully, which were set in the deployment and analytical machine learning tools were used in digital marketing.

In 2019, Kim *et al*. [18] have aimed at recognizing the influential cited operations in DMC for defining the present status of the analysis on DMC, and for denoting the scope of the shape of influential works. The evaluation of the articles published in a 12 year time in core “DMC-related journals” was done. This evaluation has examined “5865 citations of 141 digital-related articles in the targeted journals” in the provided publications with the analysis of co-citation and citation. Later, the authors have recommended thematic insights and the academic and practitioner suggestions, which have provided the best results in the effective DMC creation.

In 2018, Mackey *et al*. [19] have intended for introducing and deploying a method with machine learning for accurate marketing detection and opioids sale by illicit online sellers using Twitter as a participation section in " HHS Opioid Code-a-Thon event". In Code-a-Thon, the tweets were gathered from the "Twitter public application program interface" for the combination of typical "prescription opioid keywords". During the competition of Code-a-Thon, “an unsupervised machine learning-based model” was introduced for acquiring the summary of tweet contents for segregating those sets linked with criminal online marketing and BTM was employed for sale. Once, the related tweets were separated, hyperlinks linked with these tweets were analyzed for evaluating the illegal online seller features. The analysis was done on 213,041 tweets in a period of Code-a-Thon, which included "keywords codeine, percocet, vicodin, oxycontin, oxycodone, fentanyl, and hydrocodone".

In 2019, Vieria *et al*. [20] have introduced and experimented with the method of digital echoverse in an emerging market B2B context with VAR modelling for evaluating an individual 132-week dataset from a Brazilian hub firm functioning in marketplace. In developing markets, an empirical evidence assisting the conceptual model. For arising markets, the significance of the market development model was underscored; the results have demonstrated that renowned media and digital inbound marketing taken part in customer acquisition. By deserved social media complement renowned media, but not paid media, impressions were generated that highlighted the concept that during the digital sources echoverse might remain similar in all countries, and its components exerted a specific influence pattern in a growing context of the market.

B. Review

Although there are various machine learning algorithms for predicting the customer behaviour in digital marketing, still there are confronts with the conventional models so that a new model needs to be developed for effective prediction of the customer behaviour in digital marketing. Few defects of conventional machine learning algorithms in digital marketing customer behaviour prediction are shown in Table I. NN [7] is used for modelling the behaviour of the customer when visited online, and it has high efficiency. But, it is hardware dependent. Machine learning [17] is based on CRFs, which is an improved "statistical graphical model" for sequential data, and it is employed to recognize the pairs of replicated ads, which denote the similar hosting unit. However, it requires more amounts of training data. In social media, ELM [18] impacted the DMC domain as a discipline, which produced theoretical implications. But, it is on the basis of the assumption that the attitudes formed by central route processing will be more robust and hard. BTM [19] is used to identify the patterns and themes in corpora of short texts like tweets that have been used for examining the behaviour of prescription drug abuse and online marketing in earlier, and it is applied on a small dataset, which produced clear signals of specific selling argument word settings with words like "buy," "online," "cheap," "free," "shipping,". Yet, it needs to enhance performance. VAR [20] is used for providing many directional pathways between various variables, and the stable VAR has the ability to explain more than one variable. However, it is very complex to evaluate complex organizations and portfolios. Hence, the above mentioned challenges might motivate the researchers to develop an efficient approach for customer behaviour prediction.

TABLE I. FEATURES AND CHALLENGES OF EXISTING MACHINE LEARNING ALGORITHMS FOR PREDICTION IN DIGITAL MARKETING

Author [citation]	Methodology	Features	Challenges
Cui *et al*. [7]	NN	• It is used for modelling the behaviour of the customer when visited online. • It has high efficiency.	• It is hardware dependent.
Miklosik *et al*. [17]	Machine Learning	• These are models are based on CRFs, which is an improved "statistical graphical model" for sequential data. • It is used to recognize the pairs of duplicate advertisements, which denotes the similar hosting unit.	• It requires more amount of training data.
Kim *et al*. [18]	ELM	• In social media, it impacted the DMC domain as a discipline, which produced theoretical implications.	• This model is on the basis of assumption that the attitudes formed by central route processing will be more robust and hard.
Mackey *et al*. [19]	BTM	• It is used to identify the patterns and themes in corpora of short texts like tweets that have been used for examining the behaviour of prescription drug abuse and online marketing in earlier. • It is applied on small dataset, which produced clear signals of specific selling argument word settings with words like "buy," "online," "cheap," "free," "shipping,".	• Need to enhance the performance.
Vieria *et al*. [20]	VAR	• It is used for providing many directional pathways between various variables. • The stable VAR has the ability to explain more than one variable.	• It is very complex to evaluate complex organizations and portfolios.

III. IMPROVED MACHINE LEARNING ALGORITHM FOR PREDICTING CUSTOMER BEHAVIOR ON DIGITAL MARKETING

A. Developed Model

Digital marketing is the product or services marketing with the digital techniques especially on internet that consists of "mobile phones, display advertising, and any other digital medium". The term digital marketing represents the data-driven marketing. At present, the digital marketing has modified the way of business and brands that employ the new techniques for marketing. Since customers are using digital devices rather than viewing the digital marketing applications, physical shops are becoming highly significant and effective, as well as the digital platforms are rapidly merged into marketing plans and daily life. Moreover, digital marketing methods namely SEM, content automation, influence marketing, campaign marketing, and SEO were explored in the earlier contributions. By using tablets, smart devices, gaming consoles, and each and every application, service, and channel accessible by these devices, customers are connected to the internet at any time. Thus, the prediction of the customer behaviour is highly important so that a new model needs to be developed in digital marketing. The diagrammatic representation of proposed customer behaviour prediction in digital marketing is shown in Fig. 1.

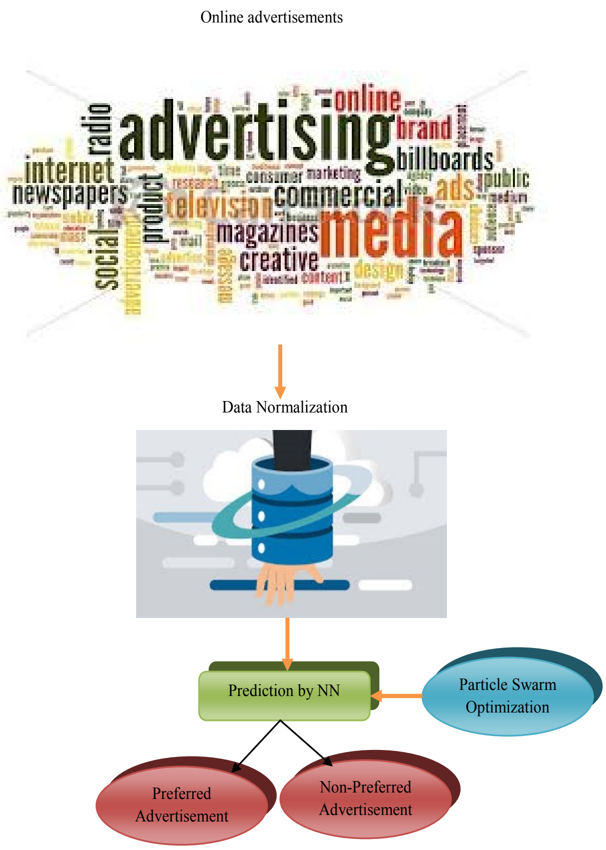

Fig. 1. Proposed model for predicting customer behaviour in digital marketing

The developed customer prediction model in digital marketing includes three phases such as "Data collection, Data Normalization, and Prediction". Initially, the customer behaviour data of different advertisements are manually or synthetically created, which involves diverse ads like display advertising, social media, native advertising, video ad etc. The dataset is given as input to the normalization phase that scales the data in a certain range. Once the data is normalized, it is given as input to the classifier NN, in which the training algorithm is enhanced by PSO. Hence, the PSO optimizes the weight of NN for minimizing the error difference in between the predicted and the actual output. Thus, the final predicted outcome provides the result as the concerned advertisement is preferred by the customer or not.

B. Data Normalization

Data normalization is a process for database organization, which is more significant that normalizes the database for reducing the data redundancy and for ensuring in each table, only the relevant information is maintained. The goal of data normalization is to change the values of numeric columns in the dataset to a common scale, without distorting differences in the ranges of values. For machine learning, every dataset does not require normalization. It is required only when features have different ranges. This also averts the problems that occurred when the operations like "insertion, deletion, and update" are performed in the databases. Here, the data is scaled in between 0 and 1 for obtaining the normalized data. The benefits of data normalization include, Searching, sorting, and creating indexes faster, since tables are narrower, and more rows fit on a data page. The drawbacks such as, that it can make the database smaller by eliminating redundant data, by doing this, the data will be easier to manage and saves us more space of storage, delete anomalies that will cause an error in the systems.

IV. OPTIMAL TRAINED NEURAL NETWORK FOR PREDICTING CUSTOMER BEHAVIOR ON DIGITAL MARKETING

A. Neural Network

The behaviour of the customer is predicted in digital marketing using NN. In many implementations, NN [21] is developed because of its improved performance. This majorly includes three layers such as "input, hidden, and output layers". The input normalized data is termed as FS_{inp}, the input neurons are given by I, hidden neurons are denoted as H, and the output neurons are indicated by O. In this, the resultant of the hidden layer is computed by Eq. (1).

$$\overline{R}^{(R)} = Af\left(\widetilde{W}^{(R)}_{(\hat{A}H)} + \sum_{I=1}^{INP(Nu)} \widetilde{W}^{(R)}_{(IH)} FS_{inp}\right) \tag{1}$$

In the above equation, the activation function is denoted as Af, the bias weight of the hidden neuron is given by $\widetilde{W}^{(R)}_{(\hat{A}H)}$, the count of input neurons is defined by $INP(Nu)$, and the term $\widetilde{W}^{(R)}_{(IH)}$ refers to the weight from input to the hidden neuron. In addition, the network's entire result is calculated by Eq. (2).

$$\hat{B}_O = Af\left(\widetilde{W}^{(B)}_{(\hat{A}O)} + \sum_{H=1}^{ON(Nu)} \widetilde{W}^{(B)}_{(HO)} \overline{R}^{(R)}\right) \tag{2}$$

In Eq. (2), the bias weight of the output neuron is given by $\widetilde{W}^{(B)}_{(\hat{A}O)}$, and the weight from the hidden neuron to the output neuron is represented as $\widetilde{W}^{(B)}_{(HO)}$. The weight $W_w^{NN} = \left\{\widetilde{W}^{(R)}_{(\hat{A}H)}, \widetilde{W}^{(B)}_{(\hat{A}O)}, \widetilde{W}^{(R)}_{(IH)}, \widetilde{W}^{(B)}_{(HO)}\right\}$ is selected optimally for providing better training to NN by PSO.

B. *Weight Optimization and Objective Function*

The main objective of the developed customer behaviour prediction in digital marketing is to minimize the error difference in between the actual and the predicted result. Predicting customer behaviour is an uncertain and difficult task. Thus, developing customer behaviour models requires the right technique and approach. Once a prediction model has been built, it is difficult to manipulate it for the purposes of the marketer, so as to determine exactly what marketing actions to take for each customer or group of customers. Achieving customer satisfaction is no longer satisfied with a simple listing of marketing contacts, but wants detailed information about customers, past purchase as well as prediction of future purchases. In this, only objective function for NN classifier is provided. NN is used for predicting the behaviour of the customer, thus the main intent of the developed approach is to minimize the error difference between the forecasted and the actual outcome. The numerical formula for measured error is expressed in Eq. (3).

$$ME1 = \underset{W_w^{NN}}{\arg\min} \sum_{O=1}^{B(Nu)} \left| B_O - \hat{B}_O \right| \tag{3}$$

In Eq. (3), the predicted result is given by $\hat{B}_O$ and the actual result is denoted as B_O. The error variance among the actual and the forecasted outcomes must be minimum and it is acquired by weight optimization, which is done by the proposed PSO. The mathematical equation of the objective function is denoted in Eq. (4).

$$Obf = Min(ME1) \tag{4}$$

In the above equation, the error variance of NN is given by $ME1$. The developed PSO-NN updates the weight function of NN in order to attain the objective function of the proposed customer behaviour prediction model. The solution encoding of the developed customer behaviour prediction model is shown in Fig. 2.

Fig. 2. Solution encoding of weight optimization for NN

In Fig. 2, the total number of weight applied to NN is represented as N_{nn}.

C. *Particle Swarm Optimization*

PSO [22] is a population-based method designed to replicate the behaviour of the birds group. In search space, the particles are placed at random and they travel in random directions. In general, the direction of the particle is changed because it moves in the direction of the past best location of itself and its gazes, searches in the neighbourhood region and detecting more number of best locations concerning the fitness measure $ft : Po^n \rightarrow Po$. Consider the particle's position as $PP \in Po^n$, and the velocity is denoted as vl. Both the position and velocity are chosen at random and updated the solutions. In order to update the velocity of the particle, the mathematical formulation given in Eq. (5) is used.

$$vl(it+1) = \omega vl(it) + \phi_c st_c\left(P\hat{P}(it) - PP(it)\right) + \phi_d st_d\left(PP_{gb}(it) - PP(it)\right) \tag{5}$$

In the above equation, the best solution of individual particle is denoted as $P\hat{P}(it)$, whereas the global best solution from the whole iterations is given by $PP_{gb}(it)$. The weight of inertia is termed as $\omega \in Po$ that holds the recurrence quantity existing in the particle's velocity. By using the stochastic variables st_c, st_d the earlier best solutions are weighted and $\phi_c, \phi_d \in Po$ are the user determined behavioural parameters. In search space, for moving the particle to another position, the current location of the particle is added with the velocity, which is mathematically represented in Eq. (6).

$$PP(it+1) = PP(it) + vl(it+1) \tag{6}$$

The pseudo code of the conventional PSO algorithm is shown in Algorithm 1.

Algorithm 1: Pseudo code of conventional PSO [22]

Initialization of particles by the velocity and positions
Until the termination criteria is met (Perform fitness evaluation), repeat
for each particle with velocity vl, and position PP
    Particle's velocity is updated by Eq. (5)
    Velocity boundaries are enforced
    Particle is moved to its new location by Eq. (6)
    Search space boundaries are enforced on the particle's position by moving it to the previous value if the boundary value is increased
    If fitness function is improved
        Update particle's best position
    end if
end for

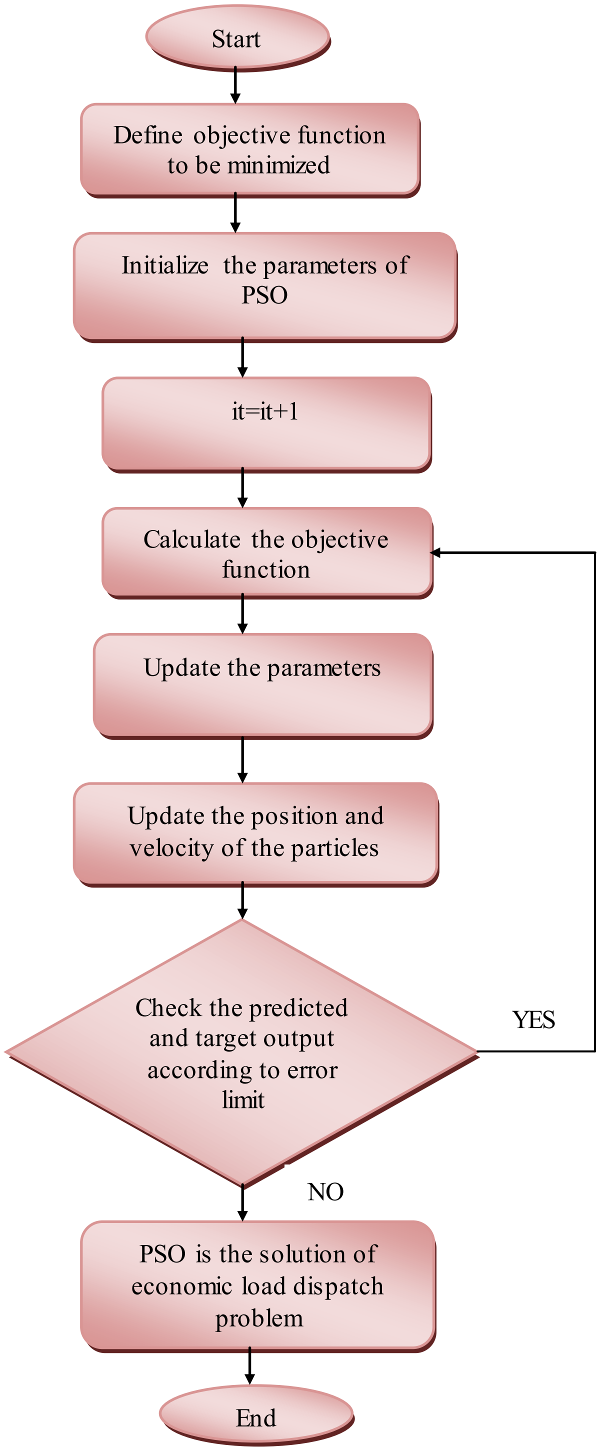

Fig .3. Flowchart of Particle Swarm Optimization

V. Results and Discussions

D. Experimental Setup

E. The developed customer behaviour prediction model in digital marketing was implemented using Python, and the analysis was performed. The datasets considered for analyzing the performance of the proposed model was "Customer behaviour dataset, synthetically created". The proposed PSO-NN was compared over conventional NN [21], KNN [23], and SVM [24] by concerning the evaluation metrics as "accuracy, sensitivity, specificity, precision, FPR, FNR, NPV, FDR, F1 score, and MCC.

F. Performance Metrics

In this paper, ten performance measures are taken into consideration to evaluate the performance.

(a) Accuracy: "ratio of the observation of exactly predicted to the whole observations". Here t^{Po}, t^{Ne} denotes the true positives, and true negatives, respectively. Moreover, f^{Po} and f^{Ne} are the false positives and false negatives, respectively.

$$Acc = \frac{t^{Po} + t^{Ne}}{t^{Po} + t^{Ne} + f^{Po} + f^{Ne}} \tag{7}$$

(b) Sensitivity: "the number of true positives, which are recognized exactly".

$$Sen = \frac{t^{Po}}{t^{Po} + f^{Ne}} \tag{8}$$

(c) Specificity: "the number of true negatives, which are determined precisely".

$$Spe = \frac{t^{Ne}}{f^{Po}} \tag{9}$$

(d) Precision: "the ratio of positive observations that are predicted exactly to the total number of observations that are positively predicted".

$$\Pr e = \frac{t^{Po}}{t^{Po} + f^{Po}} \tag{10}$$

(e) FPR: "the ratio of count of false positive predictions to the entire count of negative predictions".

$$FPR = \frac{f^{Po}}{f^{Po} + t^{Ne}} \tag{11}$$

(f) FNR: "the proportion of positives which yield negative test outcomes with the test".

$$FNR = \frac{f^{Ne}}{t^{Ne} + t^{Po}} \tag{12}$$

(g) NPV: "The probability that subjects with a negative screening test truly don't have the disease".

$$NPV = \frac{f^{Ne}}{f^{Ne} + t^{Ne}} \tag{13}$$

(h) FDR: "The number of false positives in all of the rejected hypotheses".

$$FDR = \frac{f^{Po}}{f^{Po} + t^{Po}} \tag{14}$$

(i) F1 score: "The harmonic mean between precision and recall. It is used as a statistical measure to rate performance".

$$F1score = \frac{Sen \bullet \mathrm{Pre}}{\mathrm{Pre} + Sen} \tag{15}$$

(j) MCC: "correlation coefficient computed by four values".

$$MCC = \frac{t^{Po} \times t^{Ne} - f^{Po} \times f^{Ne}}{\sqrt{\left(t^{Po} + f^{Po}\right)\left(t^{Po} + f^{Ne}\right)\left(t^{Ne} + f^{Po}\right)\left(t^{Ne} + f^{Ne}\right)}} \tag{16}$$

G. Performance Analysis

The performance analysis of the improved PSO-NN and the conventional NN by varying the count of hidden neurons for diverse performance metrics is depicted in Fig. 3. In Fig. 3 (a), the accuracy of the developed PSO-NN is acquiring the best results when compared over conventional NN at any number of hidden neurons. When the number of hidden neurons is considered as 25, the accuracy of the introduced PSO-NN is 4.2% advanced than NN. The specificity of the developed PSO-NN is acquiring the true negative observations accurately and it is shown in Fig. 3 (c). It is 2.1% better than NN at 5 numbers of hidden neurons. Moreover, the precision of the suggested PSO-NN is producing the best performance when considering any number of hidden neurons, which is shown in Fig. 3 (d). By considering the number of hidden neurons as 10, the precision of the recommended PSO-NN is 3.1% progressed than NN. The FPR of the presented PSO-NN is 84% superior to NN at 20 numbers of hidden neurons, which is shown in Fig. 3 (e). In addition, the NPV of the proffered PSO-NN is exactly predicting the negative values from the whole values and it is depicted in Fig. 3 (g). The NPV of the offered PSO-NN is 5.2% advanced than NN. The overall classification analysis of the proposed PSO-NN and the traditional machine learning algorithms is shown in Table II. In Table II, the accuracy of the recommended PSO-NN is producing the best outcomes when compared over other machine learning classifiers. Therefore, the accuracy of the suggested PSO-NN is 4.2% progressed than KNN, 2.8% progressed than SVM, and 1.3% progressed than NN. Moreover, the precision of the proposed PSO-NN is 2.1% excellent than KNN, 1.6% excellent than SVM, and 3.5% excellent than NN. Hence, it is confirmed that the developed PSO-NN is performing well for predicting the customer behaviour in digital marketing field than existing NN.

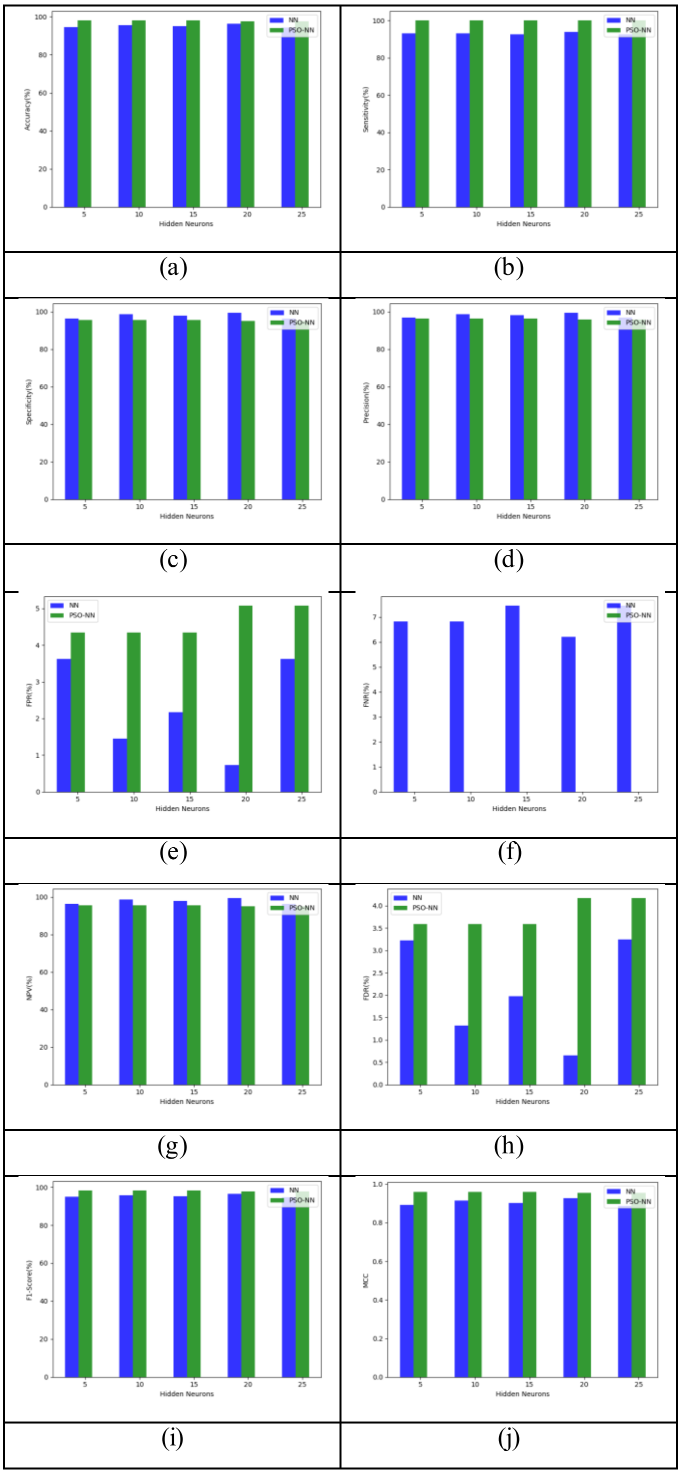

Fig . 4. Performance analysis of the proposed and the conventional machine learning algorithms for customer behaviour prediction in digital marketing with respect to the performance measures "(a) accuracy, (b) sensitivity, (c) specificity, (d) precision, (e) FPR, (f) FNR, (g) NPV, (h) FDR, (i) F1 score, and (j) MCC"

TABLE II. OVERALL PERFORMANCE ANALYSIS OF THE PROPOSED AND THE EXISTING MACHINE LEARNING ALGORITHMS FOR CUSTOMER BEHAVIOUR PREDICTION

Performance Measures	KNN [23]	SVM [24]	NN [21]	PSO-NN
Accuracy	0.936455	0.949833	0.963211	0.976589
Sensitivity	0.900621	0.931677	0.937888	1
Specificity	0.978261	0.971014	0.992754	0.949275
Precision	0.97973	0.974026	0.993421	0.958333
FPR	0.021739	0.028986	0.007246	0.050725
FNR	0.099379	0.068323	0.062112	0
NPV	0.978261	0.971014	0.992754	0.949275
FDR	0.02027	0.025974	0.006579	0.041667
F1-Score	0.938511	0.952381	0.964856	0.978723
MCC	0.876322	0.900425	0.928014	0.953794

V. CONCLUSION

A new method was introduced in this paper for predicting the customer's behaviour on digital marketing by optimal trained NN. In this, the data was gathered from various formats regarding diverse advertisements such as "display advertising, social media, native advertising, video ad etc". For eliminating the unwanted information, the data normalization was done. Later, the optimal trained NN was done for predicting the behaviour of the customer. This was performed by the proposed PSO. From the analysis, the accuracy of the developed PSO-NN was acquiring the best results when compared to conventional NN at any number of hidden neurons. When the number of hidden neurons was considered as 25, the accuracy of the introduced PSO-NN was 4.2% advanced than NN. Hence, it is confirmed that the introduced PSO-NN is performing well in customer behaviour prediction under digital marketing. In the future, the significance of deep learning could be applied for better mining of customer behaviour prediction.

REFERENCES

[1] J. Yeo, S. Hwang, s. kim, E. Koh and N. Lipka, "Conversion Prediction from Clickstream: Modeling Market Prediction and Customer Predictability," IEEE Transactions on Knowledge and Data Engineering, vol. 32, no. 2, pp. 246-259, 1 Feb. 2020.

[2] M. Papoutsoglou, A. Ampatzoglou, N. Mittas and L. Angelis, "Extracting Knowledge From On-Line Sources for Software Engineering Labor Market: A Mapping Study," IEEE Access, vol. 7, pp. 157595-157613, 2019.

[3] L. Zhang, R. Cushing, L. Gommans, C. De Laat and P. Grosso, "Modeling of Collaboration Archetypes in Digital Market Places," IEEE Access, vol. 7, pp. 102689-102700, 2019.

[4] F. Liang, W. Yu, D. An, Q. Yang, X. Fu and W. Zhao, "A Survey on Big Data Market: Pricing, Trading and Protection," IEEE Access, vol. 6, pp. 15132-15154, 2018.

[5] Y. Zhang, Y. Xu, Y. Sun, Q. Wu and K. Yao, "Energy Efficiency of Small Cell Networks: Metrics, Methods and Market," IEEE Access, vol. 5, pp. 5965-5971, 2017.

[6] LeCun Y, Bengio Y, and Hinton G, "Deep learning", Nature, vol. 521, no. 7553, pp. 436-444, 28 May 2015

[7] Yanwei Cui, Rogatien Tobossi, and Olivia Vigouroux, "Modelling customer online behaviours with neural networks: applications to conversion prediction and advertising retargeting", arXiv:1804.07669, April 2018..

[8] Nal Kalchbrenner, Edward Grefenstette, and Phil Blunsom, "A Convolutional Neural Network for Modelling Sentences", Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 655–665, 8 Apr 2014.

[9] Xiang Zhang, Junbo Zhao, and Yann LeCun, "Character-level Convolutional Networks for Text Classification", arXiv:1509.01626, pp. 649-657, 2015.

[10] G. S. Meltzner, J. T. Heaton, Y. Deng, G. De Luca, S. H. Roy and J. C. Kline, "Silent Speech Recognition as an Alternative Communication Device for Persons With Laryngectomy," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 25, no. 12, pp. 2386-2398, Dec. 2017.

[11] L. A. Thomaz, E. Jardim, A. F. da Silva, E. A. B. da Silva, S. L. Netto and H. Krim, "Anomaly Detection in Moving-Camera Video Sequences Using Principal Subspace Analysis," IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 65, no. 3, pp. 1003-1015, March 2018.

[12] C. Lin, Y. Hsieh, F. Cheng, H. Huang and M. Adnan, "Time Series Prediction Algorithm for Intelligent Predictive Maintenance," IEEE Robotics and Automation Letters, vol. 4, no. 3, pp. 2807-2814, July 2019.

[13] M. Pärssinen, M. Kotila, R. Cuevas Rumin, A. Phansalkar and J. Manner, "Is Blockchain Ready to Revolutionize Online Advertising?," IEEE Access, vol. 6, pp. 54884-54899, 2018.

[14] A. Arrate Galán, J. González Cabañas, Á. Cuevas, M. Calderón and R. Cuevas Rumin, "Large-Scale Analysis of User Exposure to Online Advertising on Facebook," IEEE Access, vol. 7, pp. 11959-11971, 2019.

[15] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk, "Session-based Recommendations with Recurrent Neural Networks", arXiv:1511.06939, 29 Mar 2016.

[16] R. Wang et al., "TaxiRec: Recommending Road Clusters to Taxi Drivers Using Ranking-Based Extreme Learning Machines," IEEE Transactions on Knowledge and Data Engineering, vol. 30, no. 3, pp. 585-598, 1 March 2018.

[17] A. Miklosik, M. Kuchta, N. Evans and S. Zak, "Towards the Adoption of Machine Learning-Based Analytical Tools in Digital Marketing," IEEE Access, vol. 7, pp. 85705-85718, 2019.

[18] JuranKim, SeungmookKang, and Ki HoonLee, "Evolution of digital marketing communication: Bibliometric analysis and network visualization from key articles", Journal of Business Research, Available online 6 November 2019.

[19] Mackey T, Kalyanam J, Klugman J, Kuzmenko E, and Gupta R, "Solution to Detect, Classify, and Report Illicit Online Marketing and Sales of Controlled Substances via Twitter: Using Machine Learning and Web Forensics to Combat Digital Opioid Access", J Med Internet Res, vol. 20, no. 4, 27 April 2018.

[20] Valter Afonso Vieira, Marcos Inácio Severo de Almeida, Raj Agnihotri, Nôga Simões De Arruda Corrêa da Silva, and S. Arunachalam, "In pursuit of an effective B2B digital marketing strategy in an emerging market", Journal of the Academy of Marketing Science, vol. 47, pp. 1085-1108, 2019.

[21] F. Fernández-Navarro, M. Carbonero-Ruz, D. Becerra Alonso and M. Torres-Jiménez, "Global Sensitivity Estimates for Neural Network Classifiers," IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 11, pp. 2592-2604, Nov. 2017.

[22] M.E.H.Pedersen, and A.J.Chipperfield, "Simplifying Particle Swarm Optimization", Applied Soft Computing, vol.10, no.2, pp.618-628, March 2010.

[23] YewangChen, XiaoliangHu, WentaoFan, LianlianShen, ZhengZhang, XinLiu, JixiangDu, HaiboLi, YiChen, and HailinLi, "Fast density peak clustering for large scale data based on kNN", Knowledge-Based Systems, Available online 3 July 2019.

[24] ShuangYu, Kok KiongTan, Ban LeongSng, Shengjin, and Alex Tiong HengSia, "Lumbar Ultrasound Image Feature Extraction and Classification with Support Vector Machine", Ultrasound in Medicine & Biology, vol.41, no.10, pp.2677-2689, October 2015.

[25] Shi, Yong, and Arialdis Japa, "Multi-dimensional analysis on data sets for information retrieval", In 2017 International Conference on Intelligent Sustainable Systems (ICISS), IEEE, pp. 157-162, 2017.

[26] Raj, J. S., and Ananthi, J. V., " Recurrent Neural Networks and nonlinear prediction in support vector machines" Journal of Soft Computing Paradigm (JSCP), vol.1, no.1, pp. 33-40, 2019.