Multi Objective Evaluator Model Development for Analyze the Customer Behavior

Mr.R.Krishnamoorthy
Department of computer science and engineering
Bharath Institute of Higher Education and Research
Chennai,India
r.krishcse@gmail.com

Dr.K.P.Kaliyamurthie
Department of computer science and engineering
Bharath Institute of Higher Education and Research,
Chennai,India
kpkaliyamurthie@gmail.com

BSH Shayeez Ahamed
Department of computer science and engineering
Madanapalle Institute of Technology and Science
Madanapalli, India
shayeezahamedbsh@mits.ac.in

Nimmala Harathi
Department of Electronics and communication engineering
School of Engineering,
Mohan Babu University (Erstwhile Sree Vidyanikethan Engineering College)
Tirupati, India
nimmalaharathi@vidyanikethan.edu

Dr.R.Senthamil Selvan
lDepartment of Electronics and communication engineering
Annamacharya Institute of Technology and Sciences
Tirupati, India
selvasenthamil2614@gmail.com

2023 3rd International Conference on Advancement in Electronics & Communication Engineering (AECE) | 979-8-3503-3072-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/AECE59614.2023.10428189

Abstract— Predicting consumer behaviour elucidates demographics, tastes, and underlying patterns. By keeping tabs on consumer activity, businesses may learn more about their wants and requirements, allowing them to provide better suggestions and increase conversion rates. Numerous factors affect consumer choices, including customer economics, buyer segmentation, and product quality. The most pressing problem that needs fixing is extracting actionable insights from these massive data sets to predict customers' actions. Author presented the multi objective evolutionary method to considerably improve the precision of consumption-related predictions as part of a cutting-edge quantitative research approach for predicting and analyzing customers' consumption habits. As the foundation for the whole prediction model, the data is first compiled based on customer preference and behaviours. As part of the data preparation, min-max normalization eliminates extraneous or irrelevant information. The author uses the Word2vec model for extraction of features, and we adopt boosting ant colony optimization (BACO) for feature selection. Multi-objective evolutionary algorithms (MOEAs) are used to make the predictions. The performance of the proposed system is evaluated, and its metrics are compared with those of well-established methods. The results indicate that the proposed multi objective evolutionary algorithm (MOEA) technique outperforms traditional Machine Learning (ML), excessive gradient boosting (XGB), Artificial intelligence (AI), and naive bayes (HNB) algorithms in various performance metrics. These metrics include accuracy (96 per cent), prediction quality (97per cent), accuracy (98 %), Recall (94 %), F1 score (99 %), and forecast time fifty seconds). Therefore, the results indicate that the regression model is viable and can be maintained over time. The proposed system for predicting consumption behaviour has shown its effectiveness in enhancing profitability.

Keywords— Consumer behavior, Word2vec Model, MOEA, Machine Learning, Naïve Bayes algorithm.

I. INTRODUCTION

A plethora of commercial websites provide extensive information about events, contacts, and opinions. Knowing how customers respond may help with research into demographics, product-market fit, and other areas. As a result, it's instructive to examine how to construct consumption patterns based on information about the buying habits of various types of customers. Several methods, including texts mined, theory of statistics, association analysis, and visualization, are required to forecast consumer behaviour, extract relevant information, and analyze the impact of users. Users' purchasing preferences and habits are extracted from e-commerce platform activity data, and the probability of users' future payment behaviour is calculated using this method of behaviour prediction. Predictive results may be used in various contexts, such as product recommendation and ad placement. Predicting consumer behaviour reveals insights into consumers' personalities, preferences, and financial limitations[1,2]. Businesses may improve their services, make more informed recommendations, and expand their market share by closely monitoring consumer behaviour. Product quality, buyer groupings, and customer economics are just a few aspects influencing consumer behaviour[3,4]. The most pressing issue is how to efficiently mine these large data sets for insights into customer behaviour. Company growth may be predicted and analyzed using cutting-edge quantitative research techniques if consumers' actions can be better predicted. It's a way to sift through this massive data set and extract relevant insights for future planning[5,6].

Uneven data distribution has exacerbated classification problems in practical settings. Consumers shop online by doing related tasks, such as researching products and making purchases, in a computer-mediated setting. The buying decision is the integration and harmony of consumer needs, incentives, behaviour, and memory. Mining communication networks may meet the prediction goal for the likely connections underlying consumer behaviours. The result was a rapid and effective implementation of a customer prediction technique for grouping consumers into distinct subsets[7,8]. It also facilitated mapping variations between these subsets and comparing consumer behaviour across different markets. Keeping a close eye on the rapid development of computer Internet advancements and Internet service, the company has emerged as one of the critical channels individuals consume regularly. Meanwhile, a pressing problem is getting actionable insights from massive datasets, including information about users' consumption habits. Many companies need help with this issue, particularly when accurately assessing and anticipating user behaviour, building customer profiles, grouping users into categories based on their behaviours, and making personalized suggestions for marketing and product purchases[9,10].

The use of consumption patterns prediction is shown in Fig. 1. This graphic examination of the various aspects that impact consumers' buying choices is meant to address the issues facing e-commerce. From a practical aspect, it is vital to provide specialized guidance to relevant businesses to improve their operational profitability. To completely grasp the user's impact on the company[11], it employs a model of logic to predict the user's purchasing behaviour. Due to the rapid acceleration of loss of biodiversity, economic change, and associated issues, it is vitally important for people to adapt their consumer behaviour to be more responsible for offering secure and healthy living situations for both future generations. Studies have shown that people purchase and use things and services that their immediate environment can replenish, manage, or compost; nonetheless, most people continue to see the economy as primarily concerned with producing and consuming goods. The current consumer culture has to shift with consumer prediction to support the transition to the circular economy. Integrated resource action plans are only theoretical tools, and they won't be able to alter the current extraordinary growth pattern. Although thoughts on sustainable consumption habits have been around for a while, additional research on the core concepts is still necessary due to the complexity and wide range of reasons for the occurrence[12].

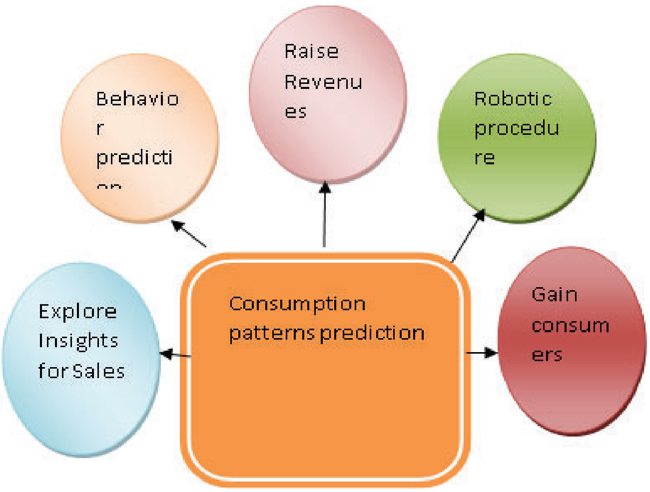

Fig. 1. Expenditure patterns forecast

A thorough assessment of market segments and customer demands is required to evaluate consumer habits for developing new products, new attitudes, and societal psychology. After thoroughly researching customer habits, businesses need to implement a reliable, effective, and flexible marketing strategy that guarantees profits and sales. It succeeded through careful planning and using a computer model to establish an objective market segment aim. The promise's impact on customers' behaviour is evaluated, and any barriers to consumers' capacity to put their ideals into action are eliminated as part of this consumption durability aspect. To address these shortcomings and improve upon previous efforts, we proposed a multi objective evolutionary algorithm to predict customer behaviour better[13,14,15].

II. Methodology Proposal

Consumers' spending habits can be more accurately predicted, boosting corporate profits. Therefore, we proposed a multi-objective evolutionary technique whereby the Word2vec form is used for feature extraction, and the best features for enhancing consumption behaviour prediction are selected using boosted ant colony optimization (BACO). The planned workflow is shown in Fig. 2, and this section provides a detailed explanation of that flow.

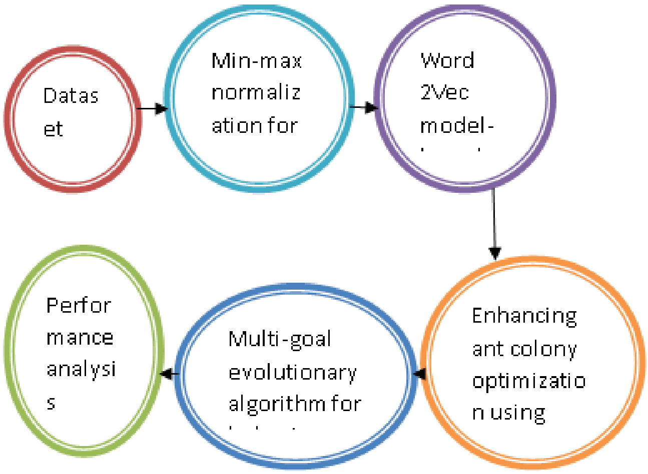

Fig. 2. Planned Workflow

A. *Collection of Data*

The data was collected through online platforms, such as QQ, email, from Chinese customers who have engaged in a maximum of one Online shopping carnival (OSC) within the past 3 years. Consequently, a survey on ease was employed to collect data from customers residing in Changchun and Jilin town, which are stage 2 and 3 cities in the northeastern region of China. The survey aimed to examine the extent of social marketplace utilization, specifically in the context of online networking. However, it was suggested that they expand their geographic scope when studying OSC behaviour since the analysis only included four locations from the level 1 group related to social marketplace use. Were 360 questionnaires initially collected, but after weeding out the non-respondents, only 310 usable surveys remained.

B. *Min-Max Normalization for Data Preprocessing*

The initial step of any inquiry should be data preparation since here is where the data quality is evaluated about the accuracy of each prediction model. As a result, it becomes hard to locate reliable discoveries and make accurate predictions using data analysis; it is necessary to adjust the underlying data. Min-max normalization is widely use as a method of data normalization. All values are transformed to decimals ranging from 0 to 1, except the minimum and maximum for each characteristic, which are changed to a 0. The whole y data set may be represented by a single number between 0 and 1. Denoting the distinction between the highest and lowest numbers as the denominator sets up a data range. By subtracting the lowest rate of every A element from each y factor, it is possible to represent every element as a value between 0 and 1 for the numerator. By subtracting the highest value of each y element from every y part, as shown by equation (2)'s displacement and reversed min-max normalization, it is possible to create a significant value close to "1" and a smaller number close to "2" about the numerator. It is necessary to transform measured data from one scale to another to correlate with the subsequent distribution of the adjusted values. By subtracting the lowest value from the data set, min-max normalization creates two values spanning the range between the highest and lowest.

$$A^{*} = \frac{A-MIN(A)}{RANGE(A)} \qquad (1)$$

Where min (A), max (A), and range (A) all stand for minimum, maximum, and range (from), respectively, of A.

$$A^{*} = \frac{[A-MIN(A)]}{[MAX(A)-MIN(A)]^{a}} \qquad (2)$$

The denominator power constant is a.

C. Word2Vec Model-Based Feature Extraction

When a dataset already has the essential information, feature extraction may be used to convert the data into a numerical characteristic. Computerized feature extraction is an alternative to artificial intelligence that does not involve modifying the raw data but instead utilizes techniques or neural networks to automatically and without human intervention extract features from the data. This strategy may be effective in expediting the process of moving from data collection to developing AI systems. Word2Vec was used to extract features from the data. Following these first steps, the Word2Vec model is built. Word2Vec is a two-layer, deeper neural system that reconstructs data contexts using natural language processing techniques. Word2vec takes a collection of purchase records as input and output a vector set based on the words found within. Each unique term in the corpus of words is assigned a vector, and the whole thing is fed into a vector space with many hundreds of dimensions. By determining the vector value of each word and gauging the semantic gap between words, the Word2vec method improves the likelihood of correctly estimating the word's context or nearby words. It identifies features of the data by assigning them a value of zero. Terms with comparable semantic links will be substituted for the original value. Word2Vec separates the ham and spam keywords from the data in each text. Two new characteristics are generated when these values are added inside their respective classes. Word2vec's computed distributed vector, as illustrated in equation (3), is used to represent organizational data. Having relevant consumer data located nearby is the primary advantage of distributed representations since it facilitates the generalization of observed patterns and yields a more precise model estimate. The purpose of Word2Vec training is to generate term vector representations that are superior at identifying their context in the same content.

$$\frac{1}{s}\sum_{S=1}^{S} q \sum_{I=-K}^{I=K} LOGQ(X_{S=I} \ |X_{s}|\text{Xs}) \qquad (3)$$

D. The Use of Boosted Ant Colony Optimization (BACO) for Feature choice.

Boosted ant colony optimization (BACO) is a natural-inspired algorithm that forages for food much as ants do. BACO is more logical than competing algorithms due to its ability to parallelize while reducing process dependency and giving opinion on the activities of insects in the search space. A pheromone trail and heuristic data are taken into account by BACO to make statistical decisions. The BACO adjusts the concentration of pheromones at every given feature as the user moves along a route.

E. Use of Multi-Objective Evolutionary Algorithm to Predict Consumption Behavior

It has been shown that using a multi objective evolutionary algorithm (MOEA) to forecast customer spending increases revenue and profits. Prediction algorithms may be divided into two broad categories: conventional gradient-based techniques and modern gradient-free straight approach. The multiple-objective evolutionary method is one of the more established prediction methods; it determines the optimal next step based on iteratively changing a nonlinear goal variable. The effectiveness of this technique is very responsive to the first parameters provided. Assuming the objective and constraint functions are differentiable, convergence to the optimal prediction is achieved. The multiple-objective evolutionary method effectively makes predictions with interrupted or non-differentiable variables. Since the relationship between the features of consumption and one another generates a continuous function, this method is well suited for consumption prediction inside the firm. The solution to business worries about consumption forecasting may lie in this method. The MOEA algorithm uses concepts from evolutionary theory and the process of replication to arrive at a prediction. In this method, the processes of replication, selection, crossover, and mutation are all essential. The MOEA algorithm's fundamental operations for forecasting consumption are shown in Fig. 3.

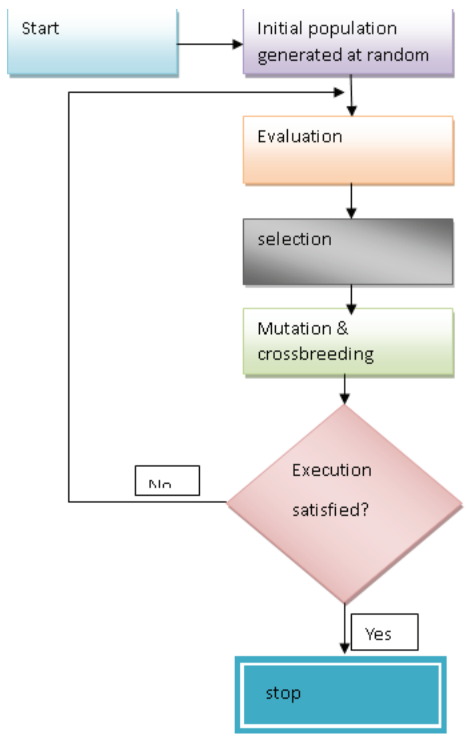

Fig. 3. Steps of MOEA

The MOEA's process is shown in Algorithm 1. Population increase is the initial step in creating several workable answers to the problem. The next thing to do is evaluate the fitness function of predictions, where predictions are the ideal solution that has to be improved. In this evaluation, we choose the most efficient methods for producing the next population. Mate selection follows a fitness evaluation, so the selected prediction may undergo a genetic crossover. The current population is also replaced with a brand new one. This process is repeated until the conditions for ending the consumption forecast are met. Current methods use obtained data as building blocks for a suitable predictive or explanatory model. The hyperparameters of multi-objective evolutionary algorithms are optimized while simultaneously seeking to balance conflicting performance goals. It finds the optimum result for the given task and satisfies all constraint by accurately determining the inputs to the goal function. Mutation procedures attempt to avoid local minima by preventing data populations from developing excessively similar to one another, which delays or even blocks converging to the global optimum. Crossing two original datasets together creates novel solutions by swapping part or all of the data. It's more likely to happen.

III. Result

This study aims to investigate the consumption of consumers prediction (MOEA) using a multiple-objective evolutionary algorithm. This article uses data on the preferences of Chinese customers who have take part in at least one OSC (online shopping carnival). In this part, we analyze how well consumer consumption forecasts work. The most critical metrics are prediction time, prediction quality, prediction accuracy, prediction recall, and prediction precision. The effectiveness of the method of evaluation and improvement (MOEA) is evaluated using these measures. Machine learning (ML), artificial learning (AL), eXtreme gradient boost (XGB), and hybrid naive Bayes (HNB) were among the conventional methods whose outcomes were compared.

A. Accuracy

The precision of a model is the degree to which its predictions align with reality. Incorporating information about customers from several sources, the suggested technique generates predictions according to customer preferences by spotting trends and foreseeing potential outcomes. The proposed technique is proven to improve upon the previous method in accurately predicting consumers' consumption. Accuracy in current methods for predicting consumption is shown in Fig. 4, and the suggested system is labeled. The suggested system achieves 96% accuracy, whereas ML gets 52%, XGB gets 76%, AL gets 86%, and HNB gets 66%.

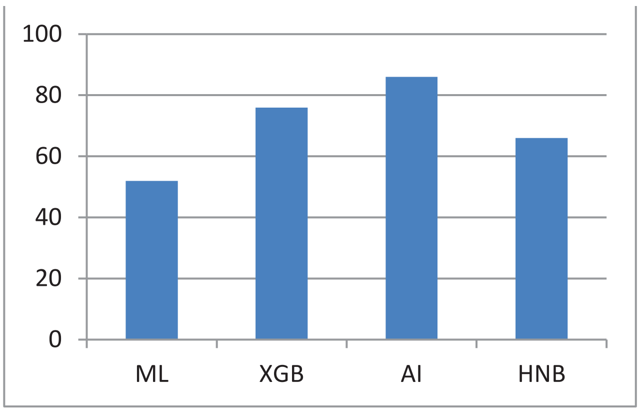

Fig. 4. Accuracy in current methods for predicting consumption.

B. Prediction Quality

The proposed method efficiently and correctly determines client preferences for favourable, overwhelming, reasonable ideas, horrible, and considerably negative things, which may be used to forecast how individuals would react when purchasing. When using past data to estimate future performance, the forecasting quality metric will show how well the system works. Fig. 5 shows the interpreted prediction quality. The suggested system achieves 98% prediction accuracy, whereas ML and XGB achieve 68%, AI 76%, and HNB 59%. Therefore, the efficiency of the suggested system is more significant.

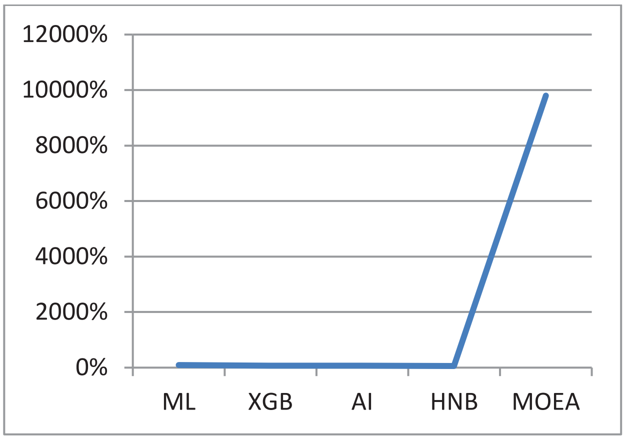

Fig. 5. Interpreted Prediction Quality Curve

C. Precision

Precision is the average likelihood of a correct forecast of a consumer's preferences. Precision is the degree to which actual consumption choices over a wide range of purchases may be anticipated using the suggested technique. Precision (or positive predictive value) refers to the percentage of correct ideas among the retrieved occurrences. Accuracy is the hallmark of quality, and this concept may be defined. The accuracy of current and proposed methods are compared in Fig. 6. The suggested work would be far more precise than current approaches. Existing systems achieve the following levels of accuracy in their consumption predictions: ML 56%, XGB 72%, AI 76%, and HNB 91%, whereas the proposed method achieves 97%. As a result, the suggested system offers the highest degree of performance.

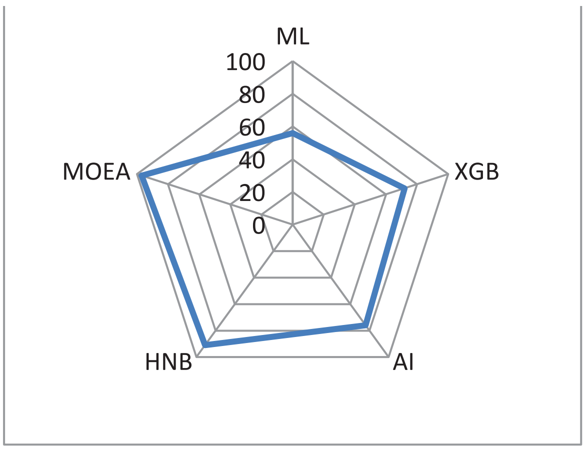

Fig. 6. Accuracy of Current and Proposed Methods

D. Recall

Fig. 7 illustrates a recollection of both suggested and current approaches. The recall measures how many of the necessary events were retrieved. Remember that the sensitivity is the same thing as the actual positive rate. The suggested technique has the most excellent recall of all current methods. Recall levels for behaviour prediction using current systems are as follows: ML: 85%, XGB: 66%, AI: 77% and HNB: 59%; the suggested system, on the other hand, achieves 94%. Recall indicates the effectiveness of the planned action.

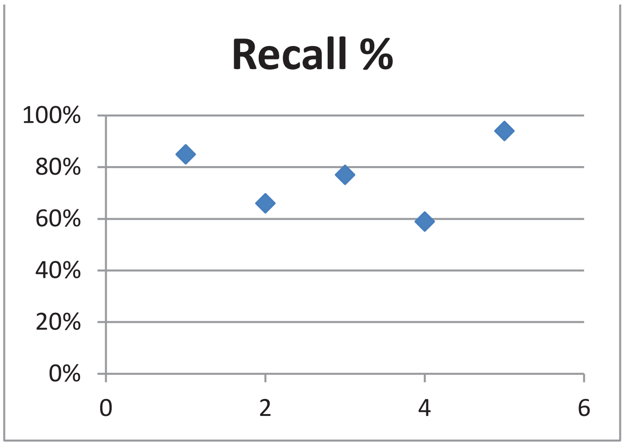

Fig. 7. Recollection of both Suggested and Current Approaches

E. F1 Score

Existing and suggested methods' F1 scores are shown in Fig 8. The F1-score is a single statistic that takes a system's harmonic means for clarity and memory and averages them together. The focus here is on drawing comparisons among the two systems. A advanced F1 -score indicates improved system performance. Fig. 7 shows that whereas ML and XGB get 68%, AI get 65%, and HNB get 85% of the F1-score, the suggested system gets 98%. It indicates the superior performance of the suggested system.

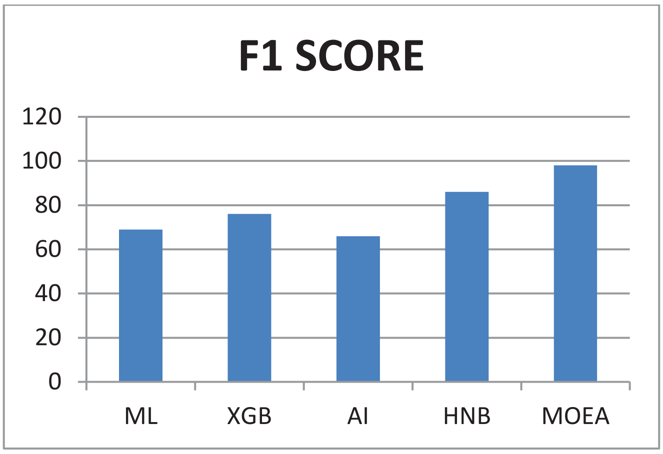

Fig. 8. Existing and suggested methods' F1 scores

F. Prediction Time

Prediction times for known and new methods are shown in Figure 10. The prediction time is when a system is expected to provide a forecast. Fig. 9 shows that. In contrast, ML's prediction time reaches 94 seconds, XGB's reaches 86 seconds, AI's reaches 76 seconds, HNB's reaches 66 seconds, and the proposed system reaches 51 seconds.

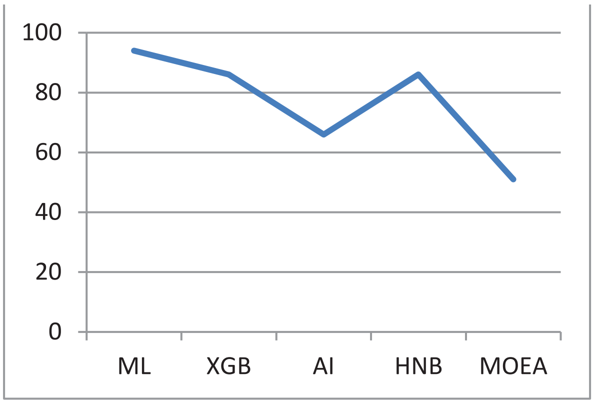

Fig. 9. Prediction Time Analysis chart

It is well known that the suggested system's forecast time is much shorter than the current methods. Therefore, this provides strong evidence that the suggested approach will work effectively in practice. Table.2 demonstrates the results of a similarity between the currently used approach and the one being suggested.

TABLE I. SIMILARITY BETWEEN THE CURRENTLY USED APPROACH AND THE ONE BEING SUGGESTED.

	ML	XGB	AI	HNB	MOEA
ACCURACY%	52	76	86	66	96
QUALITY%	98	68	76	59	98
PRECISION%	56	72	76	91	97
RECALL%	85	66	77	59	94
F1-SCORE%	69	76	66	87	99
TIME(S)	94	86	76	66	51

IV. DISCUSSION

Machine learning (ML) is applied to analyzing user behaviour on social media platforms, using input from a small set of metrics, specific needs, and user preferences. Training machine learning on large data sets is necessary and should lead to more accurate predictions at a lower level. To better predict consumer behaviour and help businesses increase sales and income, the eXtreme Gradient Boosting (XGB) model has been developed. It needs to function more effectively on sparse and unstructured data. The author argues that artificial intelligence (AI) should be used because it can help forecast how customers will feel about a company and how they will react to a product's usefulness. Prediction requires more time. Hybrid naive Bayes (HNB) was designed to categorize customer patterns while they buy things, and the author claims it will be a significant advancement in consumer behaviour research. Naive Bayes cannot learn to predict patterns since it assumes that each characteristic exists in isolation. Therefore, these limitations in predicting consumer behaviour are addressed by the suggested MOEA model.

V. CONCLUSION

The combination of utilization prediction with the company will outcome in significant shifts in revenue and growth for business organization in an era where customer prediction is one of the revolutionary technological aspects. The ability to accurately forecast customer behaviour has allowed for a quantitative characterization of all aspects of the company, improving productivity, precision, and expertise. This chapter suggested using a multi objective evolutionary model (MOEA) to predict future client consumption for optimal company performance. The results reveal that the recommended MOEA technique outperforms the standard Machine learning, XGB, AI, and HNB algorithm approaches in terms of accuracy (96%), accuracy of forecast (98%), accuracy (99%), recall (94%), F1 -score (99%), and time to prediction (51s). It may not be easy to grasp and understand some of the process notions in the proposed technique. Successful company growth may depend on better customer behaviour prediction, which may be the focus of future studies on the issue. The use of evolutionary algorithm approaches to enhance performance measures and the implementation of consumer prediction in business and economic sectors is something to consider for the future.

REFERENCE

[1] Nurjanni, Kartina Puji, Maria S. Carvalho, and Lino Costa. "Green supply chain design: A mathematical modeling approach based on a multi-objective optimization model." *International journal of production economics* 183 (2017): 421-432.

[2] Hwang, C-L., and Abu Syed Md Masud. *Multiple objective decision making—methods and applications: a state-of-the-art survey*. Vol. 164. Springer Science & Business Media, 2012.

[3] Williams, Huw CWL. "On the formation of travel demand models and economic evaluation measures of user benefit." *Environment and planning A* 9.3 (1977): 285-344.

[4] Pavan, KV Sai, et al. "Improvement of Delamination spread Model to gauge a dynamic disappointment of interlaminar in covered Composite Materials &to forecast of Material Debasement." PalArch's Journal of Archaeology of Egypt/Egyptology 17.9 (2020): 6551-6562.

[5] Govindan, Kannan, et al. "Multi criteria decision making approaches for green supplier evaluation and selection: a literature review." *Journal of cleaner production* 98 (2015): 66-83.

[6] Vigneron, Franck, and Lester W. Johnson. "A review and a conceptual framework of prestige-seeking consumer behavior." *Academy of marketing science review* 1.1 (1999): 1-15.

[7] Brauers, Willem Karel, and Edmundas Kazimieras Zavadskas. "The MOORA method and its application to privatization in a transition economy." *Control and cybernetics* 35.2 (2006): 445-469.

[8] Tzeng, Gwo-Hshiung, Cheng-Hsin Chiang, and Chung-Wei Li. "Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL." *Expert systems with Applications* 32.4 (2007): 1028-1044.

[9] G. N. R. Prasad, L. K. Kanulla, V. Ijjagiri and S. S. C. Mary, "Implementation and Health Monitoring System of Vehicle by using IoT and Cloud Computing," 2022 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, 2022, pp. 518-521, doi: 10.1109/ICECA55336.2022.10009432.

[10] Götz, Oliver, Kerstin Liehr-Gobbers, and Manfred Krafft. "Evaluation of structural equation models using the partial least squares (PLS) approach." *Handbook of partial least squares: Concepts, methods and applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. 691-711.

[11] Lupo, Toni. "A fuzzy framework to evaluate service quality in the healthcare industry: An empirical case of public hospital service evaluation in Sicily." *Applied Soft Computing* 40 (2016): 468-478.

[12] Kumar, Abhishek, et al. "A review of multi criteria decision making (MCDM) towards sustainable renewable energy development." *Renewable and Sustainable Energy Reviews* 69 (2017): 596-609.

[13] Devi, P. Anjani, et al. "FATIGUE AND DYNAMIC ANALYSIS OF CONTROL ARM." NOVYI MIR 6.6 (2021): 234-246.

[14] Moutinho, Luiz. "Consumer behaviour in tourism." *European journal of marketing* 21.10 (1987): 5-44.

[15] Singh, C., Rao, M.S.S., Mahaboobjohn, Y.M., Kotaiah, B., Kumar, T.R. (2022). Applied Machine Tool Data Condition to Predictive Smart Maintenance by Using Artificial Intelligence. In: Balas, V.E., Sinha, G.R., Agarwal, B., Sharma, T.K., Dadheech, P., Mahrishi, M. (eds) Emerging Technologies in Computer Engineering: Cognitive Computing and Intelligent IoT. ICETCE 2022. Communications in Computer and Information Science, vol 1591. Springer, Cham. https://doi.org/10.1007/978-3-031-07012-9_49