



# A Computational Model to Predict Consumer Behaviour During COVID-19 Pandemic

Fatemeh Safara<sup>1</sup>

Accepted: 25 October 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

The knowledge-based economy has drawn increasing attention recently, particularly in online shopping applications where all the transactions and consumer opinions are logged. Machine learning methods could be used to extract implicit knowledge from the logs. Industries and businesses use the knowledge to better understand the consumer behavior, and opportunities and threats correspondingly. The outbreak of coronavirus (COVID-19) pandemic has a great impact on the different aspects of our daily life, in particular, on our shopping behaviour. To predict electronic consumer behaviour could be of valuable help for managers in government, supply chain and retail industry. Although, before coronavirus pandemic we have experienced online shopping, during the disease the number of online shopping increased dramatically. Due to high speed transmission of COVID-19, we have to observe personal and social health issues such as social distancing and staying at home. These issues have direct effect on consumer behaviour in online shopping. In this paper, a prediction model is proposed to anticipate the consumers behaviour using machine learning methods. Five individual classifiers, and their ensembles with Bagging and Boosting are examined on the dataset collected from an online shopping site. The results indicate the model constructed using decision tree ensembles with Bagging achieved the best prediction of consumer behavior with the accuracy of 95.3%. In addition, correlation analysis is performed to determine the most important features influencing the volume of online purchase during coronavirus pandemic.

**Keywords** Coronavirus disease (COVID-19) · Machine learning · E-commerce · Consumer behavior · Prediction model · Bagging · Boosting

---

✉ Fatemeh Safara  
fsafara@iiu.ac.ir; fsafara@yahoo.com

<sup>1</sup> Department of Computer Engineering, Islamic Azad University, Islamshahr Branch, Islamshahr, Iran

## 1 Introduction

Coronavirus disease (COVID-19) is a severe respiratory syndrome, which first recognized in Wuhan, China in December 2019 (Kucharski et al. 2020). The initial symptoms of COVID-19 are dry cough, fever, and tiredness which are common for a number of respiratory diseases. Then, aches and pains, sore throat, loss of smell and taste would be presented. In the next stages of the disease, serious symptoms such as pneumonia and severe respiratory diseases and even heart failure would be occurred. Several research papers are published on COVID-19 from different views such as the transmission speed, drug and vaccine production, general health and psychological problems, however, consumer behaviour has received less attention.

With the advent of internet, knowledge-based economies such as online shopping was emerged long time ago. Machine learning methods could be used to extract implicit knowledge from the online shopping sites' logs (Ahmadi 2020). Industries and businesses use the knowledge to better understand the consumer behavior, and opportunities and threats correspondingly. However, with the outbreak of COVID-19, online shopping attracted more attention. We have to stay at home and observe self-isolation and social distancing because the transmission speed of the disease is very high. In particular, for elderly and people with background disease such as respiratory diseases, heart (Ivanov 2020) conditions, hypertension, diabetes, multiple sclerosis (MS), and cancer, we have to pay special attention and provide specific facilities. These conditions considerably affect the consumers shopping behavior.

Nowadays, big projects such as Siri Apple, Echo Amazon, Google, Facebook, and Microsoft employ machine learning techniques to automatically analyze their consumers shopping behavior and provide appropriate service for them (Wu et al. 2019). However, using machine learning methods is not restricted to projects with huge research budgets. Small and Medium Enterprises (SMEs) could benefit from these methods to improve their efficiency as well. Beside the spread of COVID-19, different important reasons could be presented to show the appropriateness of machine learning methods for SMEs (Zeng et al. 2019). For example, considering urban traffic management issues, using online tools for online shopping and online marketing is on the rise (McDonald et al. 2019). In addition, online catalogs help price and quality comparison (Ballestar et al. 2019). The required time for buying a product could be reduced considerably as well. Recognizing the consumer behaviors and providing proper service in a very short time is an important issue in any e-commerce application (Souri et al. 2019). Sometimes users try to buy a product, however, they cannot add the selected product to their shopping cart or they have problem with paying for it. Therefore, consumers leave the website without buying. This is an important issue for many e-commerce retailers (Vanneschi et al. 2018).

In a successful e-commerce application, understanding consumer behavior requires identifying reasons that encourage consumers to buy from the websites, however, identifying consumer behavior and encouraging factors is difficult

(Agrawal et al. 2018). Machine learning techniques use data analytic tools to discover behavioral patterns of consumers (Patil and Rao 2019). Increasing consumer satisfaction with online shopping, and improving the accuracy of consumer need prediction have always been the main challenges in e-commerce. Machine learning and data mining methods such as classification, clustering, association rule mining, and pattern recognition techniques have been effectively employed in previous researches for predicting influential factors on e-commerce (Khodabandehlou and Rahman 2017). Problems raised with COVID-19 added the importance of paying attention to consumer satisfaction during online shopping. Therefore, a model to predict consumer behaviour in online shopping is proposed using machine learning techniques. Contributions of the method proposed could be summarized as follows:

- A predictive model is presented to anticipate consumer behavior in online shopping in an e-commerce environment in COVID-19 era.
- The accuracy of predicting consumer behaviors in online shopping is increased through Bagging and Boosting of the classifiers.

The rest of the paper is organized as follows: Sect. 2 provides related works published about the prediction of customer behavior. Section 3 presents the proposed method for predicting consumer behavior. Section 4 provides experimental data, experimental setup and results. Finally, Sect. 5 concludes the paper.

## 2 Related work

In this section, a number of research studies reported before on consumer behavior are reviewed. A thorough review has been conducted by Maria Nicola et al. (2020) on socio-economic implications of COVID-19 on world economy, including agriculture, petroleum and oil, manufacturing industry, real estate and housing sector, education sector, financial industry, pharmaceutical industry, sports industry, hospitality, tourism and aviation, family dynamics, domestic violence and home video-gaming. COVID-19 has influenced the people life from different aspects, in particular, in their shopping behaviour. Increased demands for essential goods such as food, groceries, and healthcare services encountered retailers and service providers with more chances for serving consumers at home (Change 2020). Understanding the behaviour of consumers for food stockpiling is studied by Wang et al. (2020). They showed that the scale of food reserve is almost twice during COVID-19. In addition, they concluded that the women with higher education level and higher income level are more willing to buy more food and in particular more fresh food. On the other hand, demands for non-necessary goods such as clothes, shoes, and home appliances are dropped, and caused the retailers to be faced with decrease in their sales 2019 (Pantano et al. 2020). Therefore, there is a need for models to predict consumer behaviors to plan for surviving in this market accordingly.

Two models are presented by Yoon (2020) to predict GDP growth in Japan including gradient boosting and random forest. Between two models studied, the

gradient boosting was more accurate than random forest model. He claims that machine learning models are encouraging techniques in macroeconomic forecasting. A theoretical mapping is proposed in Barua (2020) to show the relationship between macroeconomic impact of the COVID-19 and affected economies. AD-AS model is introduced that considers essential features required in determining policies to cope with such economic shocks. They conclude that any action would be taken before economic depression occurs, because it would be too late if we decide to wait for COVID-19 to end and then think about its problems and solutions.

A formal modeling is performed in Souri et al. (2019a, b) to investigate the correctness of Knowledge Creation Process (KCP) in the social networks and social systems. Human behaviour in social relationships is studied and the correctness of the proposed user relationship management approach is discussed.

In Büşra Altunan et al. (2019) a dataset of Turkish statistical community is used to determine the default risk of individuals regarding to their characteristics. The dataset includes 22,745 samples with 14 attributes. Several classifiers were examined including j48, Bayesian network, Naive Bayesian, artificial neural network, logistic regression, and random forest. Their results were compared based on accuracy, performance, ROC curve and mean squared error. Finally, the most appropriate classification algorithm for risk detection is determined. Experiments performed on a real dataset indicate that consumer priorities for specific products plays a significant role in making decision for the e-commerce application (Souri et al. 2020).

Qiu et al. (2015), proposed a framework to predict the behavior of consumers in an e-commerce environment. The framework, which is called the “consumer modeling process”, is a two-step process. Relationships of products are examined and employed to predict consumer incentives, namely the construction of a candidate product set. Then, consumer priorities for the product characteristics are learned and ultimately used to identify the candidates. When a product is shipped to a consumer behavior system, the program can return products that are likely to be purchased for that consumer in the upcoming purchases. In another study conducted by Chinchilla and Ferreira (2016), the results described the use of machine learning methods to investigate the behavior of a fashion company’s consumers on the Instagram. Clustering methods and associative rules are used in the cross-industry standard process to extract descriptive models. They concluded that appropriate information extracted using the proposed descriptive model could be used in marketing strategies.

A prediction model of supply chain is presented by Ivanov (2020), which is based on the simulation data. He determined specific characteristics that form epidemic coronavirus as a unique supply chain distraction risk. He also showed the effectiveness of simulation results to inspect and anticipate the influence of the disease on supply chain profits. Two software are employed: anyLogistix simulation and optimization software. Both temporary and permanent effects of COVID-19 on supply chain and management insights are considered. From the simulation results, Ivanov revealed that scheduling of closing and opening of the services at diverse levels might be an important factor determining the influence of COVID-19 on supply chain performance. Zhang et al. (2020) proposed a novel two-stage ensemble machine learning model to predict stuck price. They ensemble support vector regression and adaptive neurofuzzy inference system. They

studied stock exchange of Shanghai and Shenzhen during 2012 to 2017. Their experiments demonstrate that the proposed model has superior prediction results than single and two stage models studies and compared with their model. As can be concluded from the researches reviewed above consumer behavior is an important factor in determining the future success of e-commerce applications and predicting consumer behavior could be of great help for both retailers and supply chain industry managers.

### 3 Methods

Stability of a long-term sale depends heavily on consumer behavior and his/her needs, instead of having just great products. In particular, in the current situation of COVID-19 pandemic, consumer behavior is changing every day. Therefore, predicting consumer behavior could be of crucial importance for future business planning. An e-commerce company with the capability of anticipating consumers' shopping behavior will gain several advantages such as increasing consumer purchase rates, increasing sales and consumer satisfaction, and greater competition (Souri et al. 2017).

In this paper, two approaches are taken to predict consumer behavior: statistical approach and machine learning approach. In statistical approach, correlation between different features are calculated and analyzed. In machine learning approach, a predictive model is proposed to anticipate consumer behavior in online shopping. The research process followed in this paper is illustrated in Fig. 1. The methods, are examined to predict consumer behavior based on the data collected from online purchase at DigiKala site ([www.digikala.com](http://www.digikala.com)). DigiKala is one of the largest and successful online shopping sites in middle east. The well-known classification methods of machine learning are compared and the best practice with the highest accuracy is identified.

In the rest of this section, first data preprocessing is presented. Then, correlation analysis explained briefly. After that, the proposed machine learning method for predicting consumer behavior is presented. At the end of the section, evaluation criteria employed to assess the proposed method are provided.

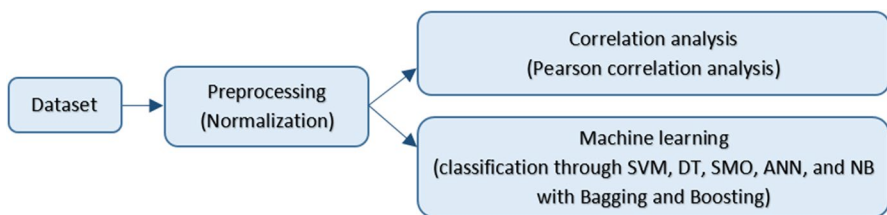


Fig. 1 Research process

### 3.1 Preprocessing

In the first step, the data should be preprocessed to reduce the implementation time and improve the results. For this purpose, we normalized the data so that the attributes are normalized as follows. The general formula is:

$$newValue = \frac{originalValue - oldMin(newMax - NewMin) + newMin}{oldMax - oldMin} \quad (1)$$

In this paper, the desired range for normalization is considered as [0, 1], therefore, the Eq. (1) would be changes to the following equation:

$$newValue = \frac{originalValue - oldMin}{oldMax - oldMin} \quad (2)$$

### 3.2 Correlation Analysis

The data employed in this paper is constructed of 11 features where the “Effective” that is the 11th features, is the impact of COVID-19 on consumer purchase volume (the structure of the dataset is explained in Sect. 4.1 in detail). After normalizing the data, we calculated the correlation between the features to find the feature that has higher correlation with “Effective” feature. Pearson correlation is calculated from the following equation:

$$\rho_{X,Y} = corr(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (3)$$

where  $\rho_{X,Y}$  is the Pearson correlation between  $X$  and  $Y$  that are two variables we are interested to calculate their correlation.  $\mu_X$  and  $\mu_Y$  are expected values of  $X$  and  $Y$ , and  $\sigma_X$  and  $\sigma_Y$  are standard deviations of  $X$  and  $Y$ . Standard deviation should be finite and positive to make Pearson correlation valid.  $E$  is expected value operation,  $cov$  is covariance operator, and  $corr$  is correlation coefficient.

The feature “Effective” accepts two values: ‘yes’ and ‘no’. The value ‘yes’ indicates that COVID-19 influenced consumers’ purchase behavior and the value ‘no’ indicates that COVID-19 did not influence consumers’ purchase behavior. The correlation results are provided in Sect. 4.2.

### 3.3 Classification

Classification methods are usually used to design both description and prediction models. In this research study, we examined five classification methods namely: support vector machine (SVM), decision tree (DT), Sequential Minimal Optimization (SMO), artificial neural network (ANN), and Naïve Bayes (NB). Ensemble meta-algorithms, Boosting and Bagging are also examined with the above classifiers to improve the accuracy of the proposed method. Ensemble methods use a number of

classification algorithms to achieve a better predictive performance than each of the classification algorithms. Boosting and Bagging are ensemble methods that convert weak classifiers to strong classifiers through reducing bias and variance. Bagging is a statistical approximation method that uses a statistical quantity, such as an average, to be estimated from multiple random samples of your data. It is a beneficial method when there is a small amount of data and we are interested in a stronger approximation of a statistical measure. Boosting works as Bagging, however uses weighted average. In training stage, they generate  $N$  training datasets from the original dataset through random sampling with replacement. In Boosting, the samples are weighted and some of the samples has more chance to influence the classification results. However, in Bagging all the samples has same chance to be participated in training process. Bagging improves the accuracy of the weak classifiers via training the learners in parallel. On the other hand, Boosting trains weak classifiers sequentially, where each classifier tries to improve the results of its predecessor. Bagging and Boosting reduce the bias and variance of a single classifier and improve the accuracy of prediction. In order to predict the consumer behavior, individual classifiers and their ensembles with Bagging and Boosting are examined in this paper, and the results are provided in Sect. 4.2.

### 3.4 Evaluation Criteria

The results of applying the model on the dataset are evaluated through accuracy, precision, recall, and F-Measure criteria calculated from confusion matrix values. A confusion matrix for a typical two-value classification problem is presented in Table 1.

Four important criteria to evaluate classification performance are accuracy, precision, recall, and F-Measure. Accuracy is one of the important classification evaluation criteria that could be calculated from Eq. 4:

$$AC = \frac{TN + TP}{TN + FN + TP + FP} \quad (4)$$

Precision and recall refer to quality and quantity of the results. In other words, precision indicates the ratio of the relevant results to irrelevant ones, and recall indicates

**Table 1** A typical confusion matrix for a binary classification problem

|                                       | Predicted values<br>(as predicted by the proposed algorithm) |                      |
|---------------------------------------|--|----------------------|
|                                       | Positive   | Negative             |
| Actual values ( as is in the dataset) |  |                      |
| Positive                              | TP<br>True Positive  | FN<br>False Negative |
| Negative                              | FP<br>False Positive   | TN<br>True Negative  |

the overall percentage of relevant results retrieved. Precision and recall are measured through the Eqs. 5 and 6, respectively.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

F-Measure deals with the accuracy of a test and is calculated based on precision and recall as follows:

$$F - Measure = \frac{Precision * Recall * 2}{Precision + Recall} \quad (7)$$

F-measure is the harmonic mean of the precision and recall of the model developed. The maximum value of F-measure is 1, which means the best precision and recall. Its minimum value is 0, which means that one of the precision and recall is 0.

## 4 Results

In this section, first a description of the dataset is provided. Then, the results of the proposed method on the consumers' online shopping dataset is presented. The results are obtained and compared based on the evaluation criteria explained in Sect. 3.4.

**Table 2** Description of the dataset records

|    | Feature              | Description  |
|----|----------------------|--|
| 1  | Gender               | Male or Female   |
| 2  | Education level      | Defined in five stages:<br>No / Diploma / Bachelor / Master / PhD                                    |
| 3  | Job                  | Categorized in seven classes:<br>Housewife/Student/Employee/Self-employment/Teacher/Manager/Academic |
| 4  | Age                  | 18–60 years  |
| 5  | Diabetes disorder    | Diabetes disorder is labeled with No/Yes   |
| 6  | Respiratory diseases | Respiratory diseases is labeled with No/Yes  |
| 7  | Cancer               | Cancer diseases or Improved status is labeled with No/Yes  |
| 8  | MS diseases          | Multiple sclerosis (MS) diseases is labeled with No/Yes  |
| 9  | NP-BP                | Number of online purchase in 2 months before pandemic<br>(20 January 2020 – 20 March 2020)           |
| 10 | NP-AP                | Number of online purchase in 2 months after pandemic<br>(20 March 2020 – 20 May 2020)                |
| 11 | Effective            | The effect of COVID-19 on number of online purchase during the era                                   |



## 4.1 Experimental data and experimental setup

The dataset is obtained from DigiKala online shopping site ([www.digikala.com](http://www.digikala.com)). Each record of the dataset is composed of 11 features, where the 11th feature with is the class label. The decryption of the data is presented in Table 2. Gender, age, education level, and job are the demographic features. Next four features are about probable consumers' diseases. The last two features are about the number of purchase in two months before and after the COVID-19 pandemic (before and after 20 March 2020).

The computer used in this experiment is a CPU Intel(R) Core (TM) i74700 HQ CPU @ 2.40 GHz. RAM is 8.00 GB. The system is 64-bit Windows 8 operating system. Correlation analysis is performed in Python anaconda Ver 2020.20 with seaborn library Ver 0.10.0, and machine learning models are constructed and evaluated in Weka Ver 3.8.1. Anaconda is an open source and free version of Python and R programming language. It is widely used for research computing with machine learning algorithms. It works with almost all operating systems including Windows and Linux.

## 4.2 Experimental Results

Correlation analysis calculates the statistical relationship between two variables. The correlation analysis explained in Sect. 3.2 is applied on the dataset and the results in Fig. 2 is obtained. Among the demographic features, Gender has the least correlation and Age has the most correlation with Effective feature. This

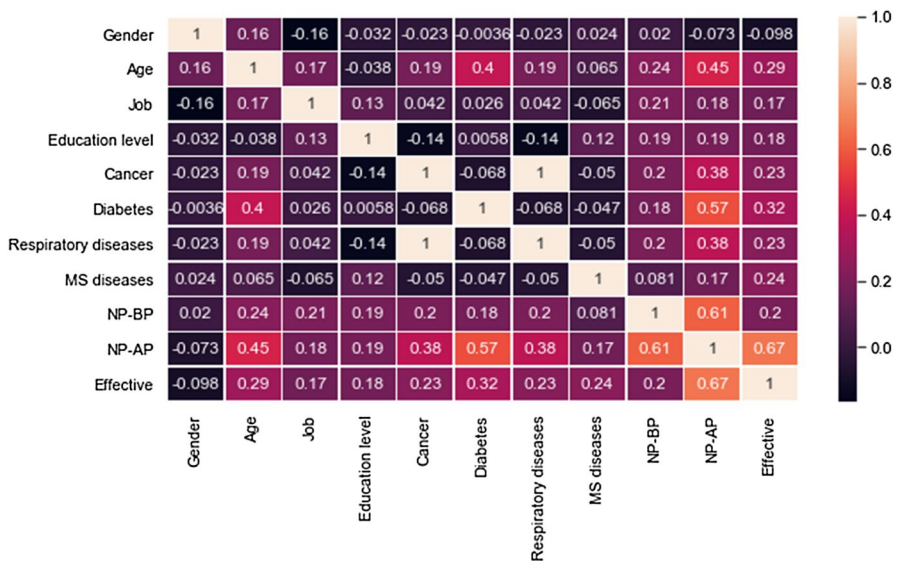


Fig. 2 Correlation between all the features

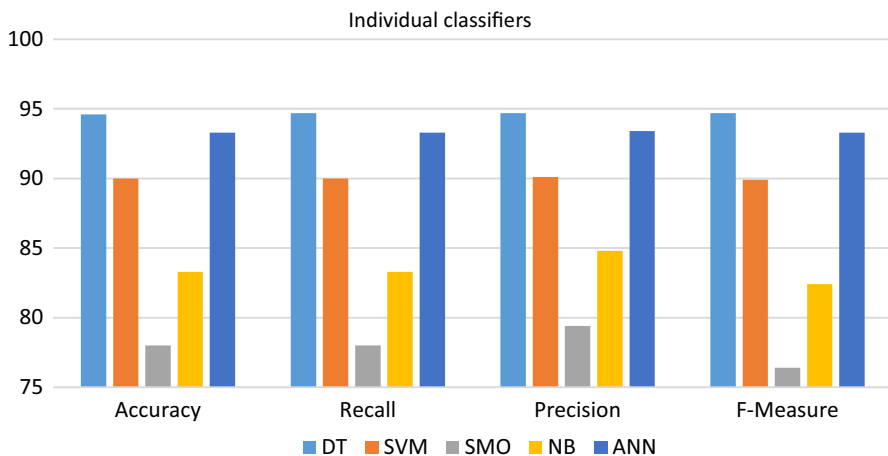
indicates that Gender feature has the least impact on increasing the number of purchases during the COVID-19 era. As can be concluded, the older the Age, the greater the interest in online shopping.

Among the four diseases studied, Cancer, Diabetes, Respiratory disease, and MS, three of them has almost the same correlation with Effective feature; with the correlation of 0.23, 0.23, and 0.24. This means they have almost the same influence on increasing the number of purchase in an online shopping. However, Diabetes has 0.32 correlation with Effective features that means people with Diabetes has more online purchases than people with the other diseases.

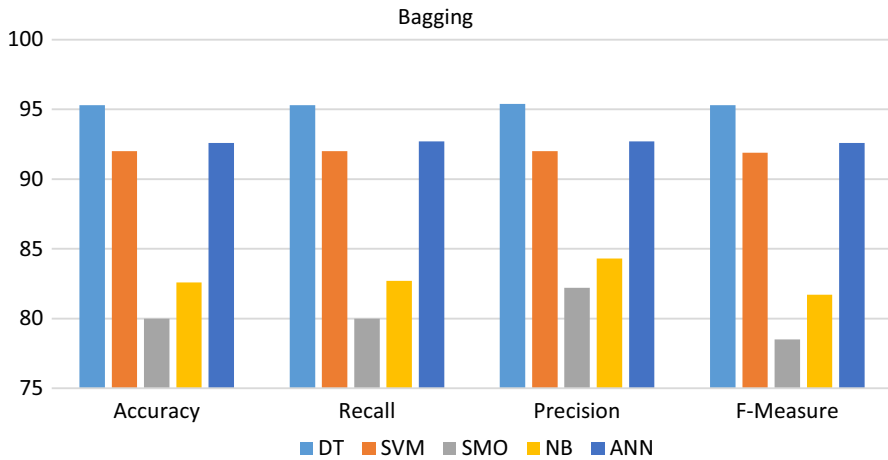
As for NP-BP and NP-AP features, during the four months of study, number of purchase before 20 of March has a weak correlation with Effective feature. However, in the next two months the number of purchase are increased and the correlation of 0.67 between NP-AP and Effective features is obtained.

After correlation analysis, with the aim of predicting consumers' online shopping behavior, five machine learning models including DT, SVM, SMO, NB, and ANN, are created and their results are compared. First, the machine learning classification methods examined individually. Then, individual classifiers and their ensembles with Boosting and Bagging are experienced. tenfold cross validation is performed for validating the results. the results of classification through each single classifier are presented in Fig. 3. Among the five classifiers examined, DT has gained the higher accuracy and SMO presented the lower accuracies. The same explanation is true for recall, precision and F-Measure.

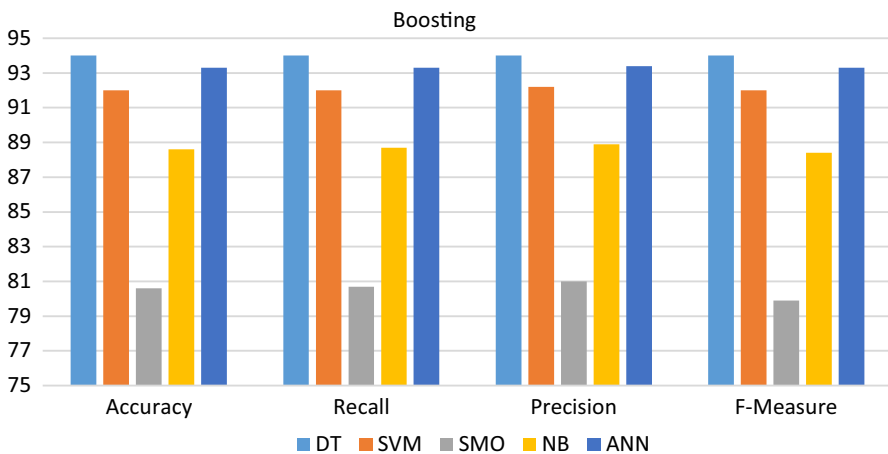
In supervised learning, Bagging tries to improve the accuracy of the weak classifiers through training the learners in parallel. However, Boosting trains weak classifiers sequentially, where each classifier tries to improve the results of its predecessor. Bagging and Boosting reduce the bias and variance of a single estimate and improve the accuracy of prediction.



**Fig. 3** Percentage of classification results obtained from individual classifiers



**Fig. 4** Percentage of classification results obtained from each of the classifiers ensembles with Bagging



**Fig. 5** Percentage of classification results obtained from each of the classifiers ensembles with Boosting

In our experiments, Bagging improved the results of three classifiers, DT, SVM, and SMO, while depressed the results of the other two classifiers, NB and ANN. Overall, the best classification results are obtained through DT ensembles with Bagging. Figure 4, presents the results of Bagging ensembles.

Boosting does not improve the results of neither single DT nor DT ensembles with Bagging. However, the results of SVM, SMO and NB is improved via Boosting for all criteria. For ANN, Bagging achieved same results as individual ANN (Fig. 5).

Although, we expect Bagging and Boosting increase the accuracy of the base classifier, in some cases single classifier has achieved same or even better results. For example, in the case of NB, single NB obtained better results than

**Table 3** The results of individual classifiers and their ensembles with Bagging and Boosting

|                             | Classifier | Accuracy (%) |
|-----------------------------|------------|--------------|
| Single classifiers          | DT         | 94.6         |
|                             | SVM        | 90           |
|                             | SMO        | 78           |
|                             | NB         | 83.3         |
|                             | ANN        | 93.3         |
| <i>Ensemble classifiers</i> |            |              |
| Bagging                     | DT         | 95.3         |
|                             | SVM        | 92           |
|                             | SMO        | 80           |
|                             | NB         | 82.6         |
|                             | ANN        | 92.6         |
| Boosting                    | DT         | 94           |
|                             | SVM        | 92           |
|                             | SMO        | 80.6         |
|                             | NB         | 88.6         |
|                             | ANN        | 93.3         |

NB ensembles with Bagging, with 83.3% and 82.6% accuracy, respectively. With ANN, Bagging achieved even worst accuracy than single ANN, with 93.3% and 92.6% accuracy. In addition, ANN ensembles with Boosting achieved same results as single ANN with 93.3% accuracy. Overall, DT ensembles with Bagging achieved the best results among the 15 variations of classifiers ensembles with Bagging and Boosting with 95.3% accuracy, as summarized in Table 3.

## 5 Conclusion

The impact of COVID-19 on consumer behavior is research in this paper. A prediction model is proposed to anticipate consumer behavior in online shopping in the COVID-19 pandemic. Five classification models are examined, where Decision tree achieved the best result of 94.6% accuracy. Then, ensembles meta-algorithms, Bagging and Boosting, are used to improve the accuracy of the classifiers. Again, decision tree obtained best results, where DT ensembles with Bagging gained superior results of 95.3% accuracy in compare to other base classifiers ensembles with Bagging and Boosting examined. In addition, correlation analysis is performed to find the most influential features on predicting consumer behavior. Age and Diabetes have found as features that have more impact on consumers' online purchase behavior. As future research, meta-heuristic algorithms could be examined to gain higher accuracies in consumer behavior prediction. Different classifiers could also be used for the prediction for better accuracy. In addition, a dataset with more features affecting from COVID-19 pandemic could be employed to construct prediction models.

## References

- Agrawal, S., Das, A., Gaikwad, A., & Dhage, S. (2018). Customer churn prediction modelling based on behavioural patterns analysis using deep learning. In *2018 International conference on smart computing and electronic enterprise (ICSCEE)* (pp. 1–6). IEEE.
- Ahmadi, M. (2020). A computational approach to uncovering economic growth factors. *Computational Economics*, 1–26.
- Ballestar, M. T., Grau-Carles, P., & Sainz, J. (2019). Predicting customer quality in e-commerce social networks: A machine learning approach. *Review of Managerial Science*, 13(3), 589–603.
- Barua, S. (2020). *Understanding Coronanomics : The economic implications of the coronavirus ( COVID-19 ) pandemic*, 1–44.
- Büşra Altun, Ebru D. Arslan, Merve Seyis, & Merve Birir, F. Ü.-Y. (2019). A Data Mining Approach to Predict E-Commerce Customer Behaviour. In *The International Symposium for Production Research*, 29–43.
- Change, P. M. (2020). *How the COVID-19 Pandemic May Change the World of Retailing*. 3–5. <https://doi.org/https://doi.org/10.1016/j.jretai.2020.04.002>
- del Chinchilla, L. C. C., & Ferreira, K. A. R. (2016). Analysis of the behavior of customers in the social networks using data mining techniques. *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2016, 623–625. <https://doi.org/10.1109/ASONAM.2016.7752301>
- Ivanov, D. (2020). Predicting the impacts of epidemic outbreaks on global supply chains : A simulation-based analysis on the coronavirus outbreak. *Transportation Research Part E*, 136(March), 101922. <https://doi.org/10.1016/j.tre.2020.101922>
- Khodabandehlou, S., & Rahman, M. Z. (2017). Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behavior. *Journal of Systems and Information Technology*.
- Kucharski, A. J., Russell, T. W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., & Eggo, R. M. (2020). *Articles Early dynamics of transmission and control of COVID-19 : a mathematical modelling study*. 553–558. [https://doi.org/https://doi.org/10.1016/S1473-3099\(20\)30144-4](https://doi.org/https://doi.org/10.1016/S1473-3099(20)30144-4)
- McDonald, N., Yuan, Q., & Naumann, R. (2019). Urban freight and road safety in the era of e-commerce. *Traffic Injury Prevention*, 20(7), 764–770.
- Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-jabir, A., Iosifidis, C., & Agha, R. (2020). The socio-economic implications of the coronavirus pandemic ( COVID-19): A review. *International Journal of Surgery*, 78(April), 185–193. <https://doi.org/10.1016/j.ijsu.2020.04.018>
- Pantano, E., Pizzi, G., Scarpi, D., & Dennis, C. (2020). Competing during a pandemic ? Retailers ' ups and downs during the COVID- 19 outbreak. *Journal of Business Research*, 116(May), 209–213. <https://doi.org/10.1016/j.jbusres.2020.05.036>
- Patil, M., & Rao, M. (2019). Studying the contribution of machine learning and artificial intelligence in the interface design of e-commerce site. In *Smart intelligent computing and applications* (pp. 197–206). Springer.
- Qiu, J., Lin, Z., & Li, Y. (2015). Predicting customer purchase behavior in the e-commerce context. *Electronic Commerce Research*, 15(4), 427–452. <https://doi.org/10.1007/s10660-015-9191-6>
- Souri, A., Asghari, P., & Rezaei, R. (2017). Software as a service based CRM providers in the cloud computing: challenges and technical issues. *Journal of Service Science Research*, 9(2), 219–237.
- Souri, A., Ghafour, M. Y., Ahmed, A. M., Safara, F., Yamini, A., & Hoseyninezhad, M. (2020). A new machine learning-based healthcare monitoring model for student's condition diagnosis in Internet of Things environment. *SOFT COMPUTING*.
- Souri, A., Nourozi, M., Rahmani, A. M., & Navimipour, N. J. (2019). A model checking approach for user relationship management in the social network. *Kybernetes*.
- Souri, A., Rahmani, A. M., Navimipour, N. J., & Rezaei, R. (2019). Formal modeling and verification of a service composition approach in the social customer relationship management system. *Information Technology & People*, 32(6), 1591–1607. <https://doi.org/10.1108/ITP-02-2018-0109>
- Vanneschi, L., Horn, D. M., Castelli, M., & Popovič, A. (2018). An artificial intelligence system for predicting customer default in e-commerce. *Expert Systems with Applications*, 104, 1–21.
- Wang, E., An, N., Gao, Z., Kiprop, E., & Geng, X. (2020). Consumer food stockpiling behavior and willingness to pay for food reserves in COVID-19. *Food Security*, 12(4), 739–747. <https://doi.org/10.1007/s12571-020-01092-1>

- Wu, X., Zhang, L., & Lan, T. (2019). Scenario based e-commerce recommendation algorithm based on customer interest in Internet of things environment. *Electronic Commerce Research*. <https://doi.org/10.1007/s10660-019-09339-6>
- Yoon, J. (2020). Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach. *Computational Economics*. <https://doi.org/10.1007/s10614-020-10054-w>
- Zeng, M., Cao, H., Chen, M., & Li, Y. (2019). User behaviour modeling, recommendations, and purchase prediction during shopping festivals. *Electronic Markets*, 29(2), 263–274.
- Zhang, J., Li, L., & Chen, W. (2020). Predicting stock price using two-stage machine learning techniques. *Computational Economics*. <https://doi.org/10.1007/s10614-020-10013-5>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations