

A Predictive Analytics Model for E-commerce Sales Transactions to Support Decision Making: A Case Study

Shereen Morsi¹

¹ Business Information Systems Department, College of Management & Technology
Arab Academy for Science, Technology & Maritime Transport, Cairo, Egypt
Email: shereen_morsi [AT] aast.edu

Abstract—Given the significant growth in electronic commerce, firms are seeking technological innovations and innovative capabilities to deal concurrently with the data' volume generated and gaining insights from it for better decisions. Although recent studies identify predictive analytics as becoming the keystone of all business decision making and a crucial aspect in firms by it is a possible means for driving strategic decisions. Significant inroads into the interrelationships between capabilities and the execution of a pathway to an analytical capability to many Egyptian e-commerce businesses have yet to be made. Therefore, this paper aims to shed light on the importance and the role of using predictive analytics models in the Egyptian e-commerce firms where these tools became dominant resources for gaining valuable knowledge for better decision making by precautionary measures from prediction rates and different applications that have been applied by global e-commerce firms. The aim of the paper was achieved by building a predictive analytics model for sales forecasting by tackling to one of the e-commerce company in Egypt, and the online transaction dataset has been analyzed. The result obtained from the model has been displayed, and some insights extracted from the prediction model have been explained.

Keywords-component; Predictive analytics; Predictive models; Online Transactions; e-commerce; e-commerce in Egypt.

I. INTRODUCTION

The electronic commerce industry is growing at an unprecedented rate all over the world [1]. Where retail e-commerce sales amounted to 3.53 trillion US dollars in 2019, and the revenue of e-retail is expected to grow to 6.54 trillion US dollars in 2022 [2]. The growth is because of believed companies that electronic commerce was considered a vital weapon that the technological innovation invented from internet technology and enabled firms to compete in the global market [3]. In addition, e-commerce uses as a way of providing strategic planning, customer services, reducing cost, and can enhance commercial productivity and efficiency, push the firms towards prosperity and development and create new market opportunities [4]. Therefore, companies are developing competence by using and embedding advances in technologies in the processes of e-commerce.

Although the importance of electronic commerce to organizations, the emergence of web technologies and their applications led to the ever-increasing growth of the volume and varieties of data that result from its transactions [5]. This massive increase in the volume, variety, and velocity of data generated from various e-commerce platforms pose a huge computational challenge for data analysis and intelligence tasks. Whereas the fundamental prerequisite of decision-making processes' efficiency is relying on the actual and right information for monitoring past and current actual information to generate trends. Therefore, the present state of data and change in the e-commerce environment urges a need for different analytical tools to find out the hidden patterns for valuable outcomes.

One of the most crucial applications of technology is predictive analytics, which since the emerging of it has gained popularity rapidly in the Big Data field, and it is becoming the keystone of all business decision making [6]. Predictive analytics is one of data analytics' category and are using for forecasting future outcomes by interpreting and analyzing historical data to uncover the relationships and patterns in these data. Through the applying of statistics algorithms, and machine learning techniques [7]. According to [8] report show that it's estimated to have a growth in the predictive analytics in a global market at a compound annual growth rate (CAGR) of around 25% in 2013 from \$1.7 billion to \$5.24 billion in 2018 and projected to reach approximately \$10.95 billion by 2022. This means that more organizations every year are turning to predictive analytics for increasing their competitive advantage [9].

Correspondingly, predictive analytics is used to detect the relationships and patterns in data for predicting the future by analyzing the historical data and taking better preventive decisions. Where the application of predictive analytics models are based on machine learning for analyzing the customer's reviews and click-through actions, preferences and past purchase history, enhance sale prospects and discover the reasons for customer churn, in real-time[10]. Also, many retailers around the world are using predictive analytics for price optimization, merchandise planning, and analyzing the promotional events' effectiveness for determining which offers are more appropriate to consumers [9]. Predictive analytics

capability – that is, the ability to compete in analytics through organizing, standardizing, and manipulating data – is an essential determinant of an organization's competitive advantage in today's rapidly changing business environment [7].

Therefore, many organizations in various industries use predictive analytics for achieving many desired goals, such as determining customer purchases, promote cross-sell opportunities, forecast inventory, and manage resources, reducing the risk [11]. In addition, predictive analytics models can help businesses attract keep and grow their most profitable customers and enable organizations to function more efficiently to drive decisions and actions [12].

There is no doubt that many global e-commerce firms such as Google, Amazon, e-Bay, Netflix, Flipkart have applied various types of predictive analytics models (machine learning algorithms) and gained secured compelling competitive advantages matchless competitiveness in the market by integrated these models in business intelligence by injecting it in their value chain for staying on top of their game [6, 7].

Even though many organizations from varied sectors have applied the predictive analytics, many electronic commerce companies have been slow for adopting the predictive analytics approach due to many reasons such as IT infrastructure does not address big data and analytics, unaffordability, shortage of skills, and complexity [13, 14]. As the data analytics market progresses and predictive analytics is an essential technique while dealing with a massive amount of data, more affordable and easy-to-use solutions are being developed, which can be integrated by e-commerce firms on a variety of platforms. Therefore, the paper first, discussing the predictive analytics fields, displays types of predictive analytics models and their importance of use in the e-commerce field. Second displays the current state of e-commerce in Egypt. Then, a predictive Model for e-commerce sales prediction proposed and deployed by using the Python package, the building of the model was done by tackling one of the e-commerce company in Egypt and analyzed a dataset of its online transaction for selling different product categories. The obtained results were evaluated on the basis of prediction accuracy. The results attained by the proposed model were promising.

II. LITREATURE REVIEW

A. Predictive Analytics

Predictive analytics is defined as a type of data analytics that examines data for answering the question of "What is likely to happen?" by using forecasting, regression analysis, pattern matching, multivariate statistics, and predictive modeling [15]. Moreover, Kopp [16] defined predictive analytics as to the identification of the event before it takes place by using big data. The predictive analytics applications depend on robust data mining, machine learning, artificial intelligence (AI), and deep learning algorithms [17, 18].

Predictive analytics has a long tradition in computer science and information systems literature and has been demonstrated to be applicable to various domains, such as sales forecasting, strategic sales management [20; 19], recommender systems [21] & fraud detection [22]. Predictive analytics models build by using quantitative techniques, generally machine learning algorithms, to build predictive models [23]. It is an essential part of the business intelligence [24] and decision support systems literature.

Predictive modeling uses predictive models that rely on capturing relationships between predicted variables and explanatory variables from past occurrences, exploiting them for analysis, and to predict the outcome [25]. The predictive models' objective are to assess the likelihood that a similar unit in a different sample will exhibit the performance. Predictive models often perform calculations during real time of transactions, such as evaluating the opportunity of a given customer or transaction, in order to guide a decision. In this context, Lovemen [26], CEO and president of Caesar's Entertainment, stated that: "The best way to engage in ... data-driven marketing is to gather more and more specific information about customer preferences, run experiments and analyses on the new data, and determine ways of appealing to [casino game] players' interests. We realized that the information in our database, coupled with decision science tools that enabled us to predict individual customer's theoretical value to us, would allow us to create marketing interventions that profitably addressed players' unique preferences".

This means the predictive models help firms in preparing revenue budgets. The preparation of the revenue budget assists firms in recognizing future sales patterns from data of past sales. Also, its impact helps in better determination and forecasting requirements of inventory which leads to avoiding loss of the customers [27]. Also, predictive models help firms' in predicting risk, tendency, and in attaining better revenues by enhancing their key metrics making strategic corrections and by making accurate predictions from structured and unstructured data[28].

B. Predictive Models Techniques

During the predictive modeling process, the predictive models are creating for discovering the patterns between explanatory variables and dependent variables, and predicting an outcome [11, 29]. Many techniques are using:-

- **Regression:** Regression models are considered the backbone of predictive analytics. Where the focus of regression models is relying on establishing a mathematical equation that represents the interactions between dependent and independent variables, or predictors (numerical and categorical variables). Many various regression models can be applied like a linear regression model, discrete choice models, Logistic regression, Multinomial logistic regression, Time series models.
- **Classification:** classification models are built to answer the questions that required yes or no (like, is

this customer about to churn? or is this a fraudulent transaction?” by providing broad analysis which helps in the guidance of a decisive variable (target or class) and the construction of the model is based on decisive or numerical variables (predictors or attributes).

- Clustering: It is one of the unsupervised machine learning methods. The clustering model sorts' data into separate, each cluster contains similar attributes. This process helps in find out the unknown relationships in a dataset such as using a cluster model for separating customers into similar groups based on common characteristics and devise strategies for each group at a larger scale.
- Association rules: Association rules are a rule-based machine learning method, which is using to find essential associations in the observations [30]. This means to find all item sets that have greater support than minimum support and then using the large item sets to generate the desired rules that have confidence greater than the minimum confidence [31]. One of application of association rules is market basket analysis which is a modeling technique for discovering regularities between products in large-scale transaction data [32] and can be described simply as if a customer buys a specific set of items, maybe probably buy another set of items.

C. E-Commerce in Egypt

Egypt started the investment in the ICT sector by building in its infrastructure since 1985 as an essential tool for economic development and as the base for e-commerce infrastructure establishment. Through the previous decades, the ICT sector became a significant platform for socioeconomic development and the business sector. According to [33], the Internet in Egypt is playing an increasingly different role in every aspect of life, since the number of Egyptians using the internet has risen every year along the last decade. In 2017, the number of internet users increased from 0.65 million users in 2000 to 48 million users. Also, internet penetration grew from 1.01% in 2000 to 35% in 2017, with an average annual growth rate of 64% during such period.

On the other hand, almost one- third of Egyptian private enterprises connected to the internet; the usage of the internet was increased by 25% from 2016 to 2017. Furthermore, about 45% of private enterprises in urban areas use the internet compared to 20% in rural areas [34]. The gap between rural and urban areas is due to the nature of working conditions of both regions; companies in urban areas depend mainly upon ICT in their work, while in rural areas companies specialize in agriculture, fishing, and other activities that don't rely mainly on ICT to function [35].

Currently, many current and new business firms have employing e-commerce systems by relying on various technologies and the Egyptian e-commerce B2C turnover grew by 22% to USD 5 billion. From 48 million online users

of the total population who were aged 15 and older, about 18 million are actually online shoppers who are buying something online. The average spending per online shopper was USD 540 in 2017. Most online transactions are in the electronics categories, airline tickets, fashion, and entertainment. According to [36], by 2020, Egypt is expected to generate US\$ 1.9 billion in B2C e-commerce volume. This means the Egyptian e-commerce have to grow at least 45 percent annually, which represents an increase to 1.52 per cent of GDP ratio in 2019/20.

Although, e-commerce systems had changed the face of most business functions in Egyptians organizations and enabled online transaction by allowing web-based applications to be established and provide an opportunity for existing systems to automate the process of sales as well as using of online presence for achieving higher volumes of sales and reaching global markets. Except, the use of advances in internet technologies has led to generate large-scale historical and real-time data from various dealing of business transactions. Therefore, these changes required attention from decision-makers in e-commerce firms for using new tools that can be analyzing datasets by analytical tools of various transactions in order to understand and examine emerging patterns for better decision making [37]. Therefore, for Egyptian e-commerce firms to grow, sustain and compete in the global market, a potential way is by creating business value through the use of advanced technologies and relies on prediction models for better strategic planning.

III. DATA ANALYTICS

Based on the classification of the Institute of Operations Research and Management Sciences (INFORMS) organization [38], the analytics can be grouped to descriptive, predictive, and prescriptive. Each type can be viewed independently such as many of business may use only descriptive analytics for providing information on decisions they face. Others may use many different types of analytic to glean insightful information needed to plan and make decisions. This paper tackle to build a prediction model for predicting total purchasing from all products categories for one of an online store as one of the necessary model that any firm has to rely on it in its decision-making processes. This methods may be useful for planning its demand and supply a need for running the business, or as a guide for budgets and targets' planning phase.

A. Predictive Modeling Framework

In this section, we illustrate the detail of the predictive modeling framework that applies to the task proposed. The predictive analytics process passes by six phases as shown in figure 1. In the initial phase (the identification of the problem), the problem is defined with outcomes, objectives, scope, and deliverables from the proposed model. In the next phase (the collection and preparation of the data), data is collected from various sources and is analyzed. The third step is the analysis (analysis of the data and the development of the model) requires strategies for pre-processing such as data cleaning,

transformation, and data modeling so that useful data is extracted for further processing. Subsequently, validate the initial hypothesis using statistical models. The next phase (the deployment, observation and control of the predictive model) is predictive modeling for forecasting the future. Results after implementation can be deployed for using it in the day-to-day decision-making. The last phase is monitoring the model in order to ensure that it is providing the expected outcome.

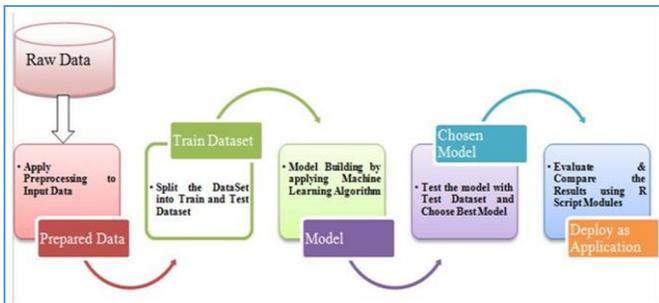


Figure 1. Predictive modeling framework [39].

B. Dataset and Features

For our analysis, we used a sample of the online transactions historical data from one of well-known company that conduct e-commerce in Egypt. In the first step, the dataset was loaded for exploratory data analysis and the importing necessary libraries and reading data from CSV with the help of pandas for doing basic operations. The calculations were conducted using the Python environment, the main packages numpy, pandas, matplotlib, and seaborn. For conducting the analysis, Jupyter Notebook was used. Figure 2 shows the importing of various packages and the loading of data.

```
In [94]: # These are several packages loaded in
import numpy as np
import pandas as pd
import seaborn as sns
from pylab import rcParams
rcParams['figure.figsize'] = 10, 5

In [97]: # reading the data file
df = pd.read_csv('C:\Users\Waseem\Desktop\ecomg.csv')

df.shape
df.head()

Out[97]:
   user_ID  Gender  Age  Work_Years  City  Stay_In_Current_City_Years  Marital_Status  Ps_Category1  Ps_Category2  Ps_Category3  Ps_Category4  Purchase
0  3069042    F    0-17    10    C    4+    0    3.0    4.0    1.0    2.0    9730
1  3248942    F    0-17    10    C    4+    0    6.0    6.0    1.0    3.0    7200
2  3067842    F    0-17    10    C    4+    0    7.0    NaN    10.0    2.0    10422
3  3065442    F    0-17    10    C    4+    0    2.0    14.0    5.0    7.0    9570
4  3265442    M    55+    16    G    4+    0    4.0    NaN    1.0    11.0    7968
```

Figure 2. Dataset loading

Dataset is part of data transactions that stored in the database which is about selling a various collections of products. It contains approximately 10k of online transactions by Egyptian shoppers and each record has 12 different variables which include either numerical or categorical data about several characteristics of customers such as user-id, product-id, gender, age, city, work of years' experience, for a type of product's categories, and total amount of purchase of all products categories as shown in Figure 3. The main goal of using this dataset is to build a prediction model that can predict total purchasing of all the categories of all products based on the personal characteristics of customers.

```
#checking if there is any NULL values present in the current dataset
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10736 entries, 0 to 10735
Data columns (total 12 columns):
User_ID      10736 non-null int64
Product_ID   10736 non-null object
Gender       10736 non-null object
Age          10736 non-null object
Work_Years   10736 non-null int64
City         10736 non-null object
Marital_Status 10736 non-null int64
Ps_Category1 10736 non-null float64
Ps_Category2 7817 non-null float64
Ps_Category3 8836 non-null float64
Ps_Category4 3521 non-null float64
Purchase     10736 non-null int64
dtypes: float64(4), int64(4), object(4)
memory usage: 1066.6+ KB

print("Number of unique users:", len(df.User_ID.unique()))
print("Number of unique products:", len(df.Product_ID.unique()))

Number of unique users: 1460
Number of unique products: 2334

df['Ps_Category1'] = df['Ps_Category1'].fillna(0)
df['Ps_Category2'] = df['Ps_Category2'].fillna(0)
df['Ps_Category3'] = df['Ps_Category3'].fillna(0)
df['Ps_Category4'] = df['Ps_Category4'].fillna(0)
```

Figure 3. Exploring Dataset

Figure 3 represent an overview of the dataset which has 10736 rows and 12 columns as described below:

- User_ID: Unique ID of the user (primary key), and the total number of users is 1460 in dataset.
- Product_ID: Unique ID of the product, products are 2334 in dataset.
- Age: the age of users who done purchased online.
- Gender: the gender of the online users who make the transaction.
- City: refer to user's living city. Cities are categorized into 3 categories 'C: Cairo', 'G: Giza', and 'A: Alex'.
- Marital_Status: refers current status of users which is 0 if the user is not married and 1 is married.
- Ps_Category 1 to _4: it is the category of the products and there are 4 categories and they represent in numbers.
- Purchase: total amount of products categories purchasing.

In addition, as shown from figure 3, the data columns of Ps_category1, 2, 3, & 4 fields contains null values which are required to be cleaned its before starting the data analysis's step.

C. Data Cleaning

In data cleaning step, both Ps_Category1, 2, 3, & Ps_Category4 had some null values which assigned by 0 value for all Nan cases.

```
df['Ps_Category1'] = df['Ps_Category1'].fillna(0)
df['Ps_Category2'] = df['Ps_Category2'].fillna(0)
df['Ps_Category3'] = df['Ps_Category3'].fillna(0)
df['Ps_Category4'] = df['Ps_Category4'].fillna(0)
```

Figure 4. Removing null values

D. Data Analysis

After reviewing the dataset, data analysis step was used for exploring and dig deep in data for finding insights. Some questions were listed for guiding in the analysis phase:-

- What are the characteristics of buyers (gender, age, city)?
- Who is purchase more men or women?
- Which type of product categories are most sold?
- What are the distribution of purchase among age groups and cities?
- What are the distribution of purchase among products categories?
- And which product categories more sell?

Before answering these questions and gain some insights about the dataset of transactions, descriptive analysis were conducted, which is a study of purchasing distributions, data visualization with different pair plots. descriptive analysis is helpful for finding the correlations among variables and determine which variables that drive the sales.

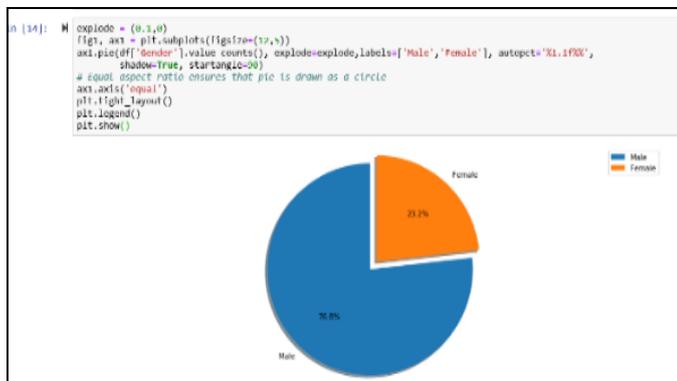


Figure 5. Distribution of percentage of men & women

From figure 5, the percentage of men shoppers is about 76.8% which is more than women 23.2%. This means that the men are willing to buy products online more than women. And in the same time, it could be a good indicator for e-retailer when displaying products or promotions to online shoppers.

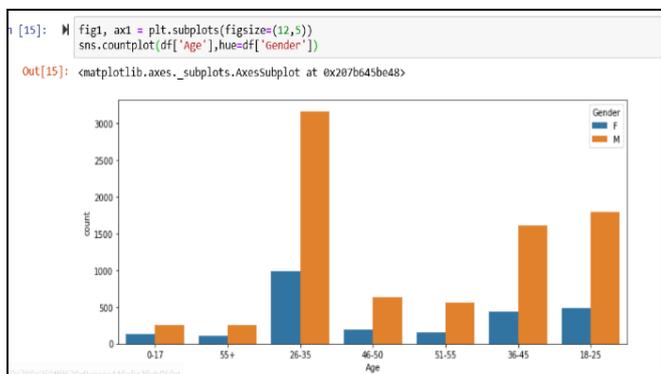


Figure 6. Distribution of Age

Figure 6 displaying the distribution of online shoppers' age. The most purchasing from men is age 26-35 among other age groups.

For dig deep in analysis among variables, the seaborn library and the countplot function were used to plot the total number of male and female in gender, categories of cities, marital status, and products for each category.

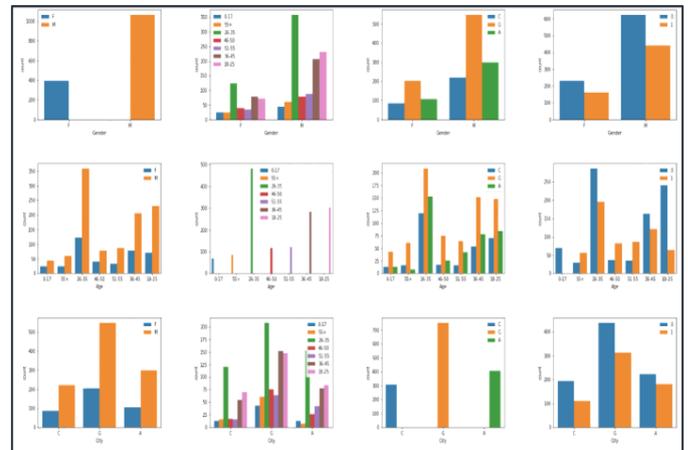


Figure 7. Distribution of Purchase among user's characteristics

As the descriptive analysis enables for seeing the relationship among multiple variables in easy-to-read graphs, so figure 7 & 8 shows the results of the exploratory analysis by depicting the distribution of products purchases with each online consumer characteristics (ex: gender with age, city, amounts, and city with age, and distribution of purchasing by city).

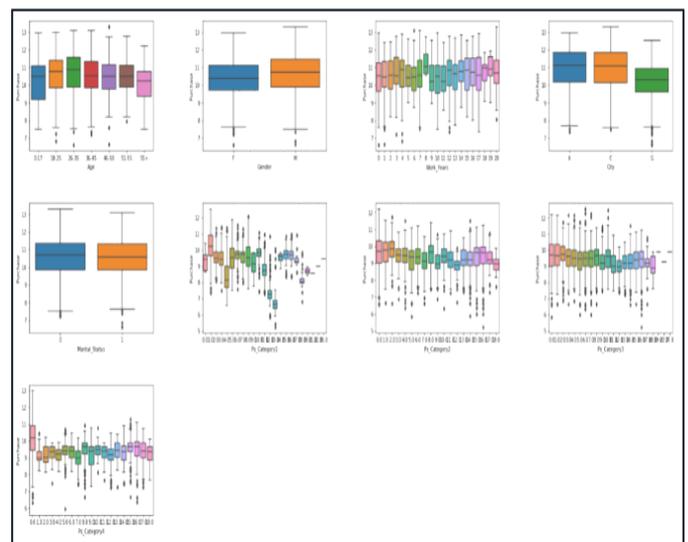


Figure 8. Exploratory analysis

From the above figures 7 & 8, we can conclude that, men are more purchaser than women by 77% for males and 23% for females' buyers. Most of buyers by 45% within the age from 26-35 for both gender. About 38% from buyers are from

Cairo (c) city, whereas, 32% from Alex (A) and 31% from Giza (G).

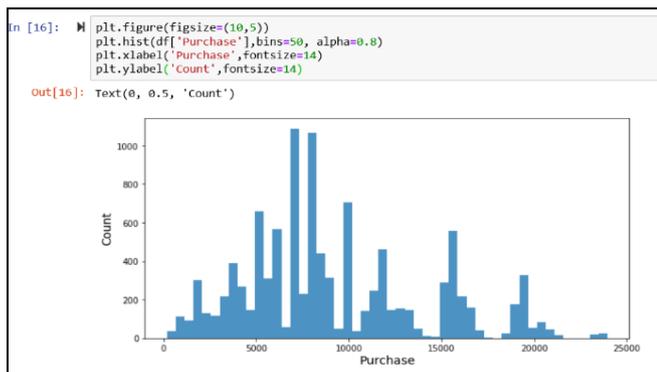


Figure 9. Distribution of purchase

By drawing the heat map for seeing the correlations between variables. It is obvious that independent variables are not highly correlated with the Purchase variable. Also, from figure 10 can conclude that:

- Ps_Category4 has a negative correlation with Ps_Category1.
- Ps_Category1 has a negative correlation with purchase.
- Ps_Category4 has positive correlation with purchase, which it may be due to less prices to these category of products.

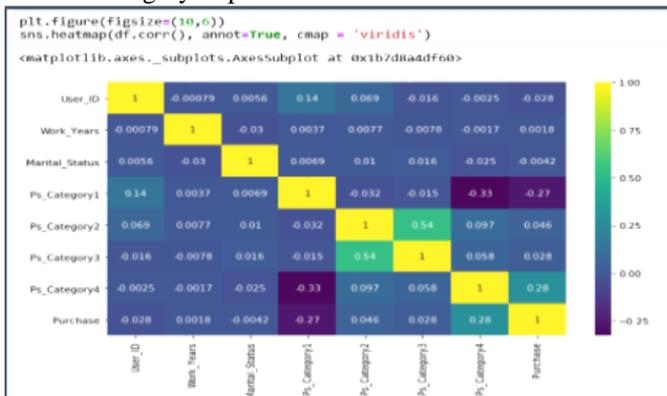


Figure 10. Correlations among all Variables

In the above steps, descriptive analysis has been done on row data. Now one of the machine learning algorithms is applying to predict the value of purchase through customer characteristics data.

E. Prediction model – Linear Regression

Linear regression model is one of machine learning models (supervised learning), and it is an effective tool for predicting quantitative responses. It used to predict the amount of the total purchasing from online users. The linear regression model, like most of the supervised machine learning algorithms, makes a prediction based on the input features. The predicted output values are used for comparisons with desired outputs and an error is calculated. The error signal is

propagated back through the model and model parameters are updating in a way to minimize the error. Finally, the model is considered to be fully trained if the error is small enough.

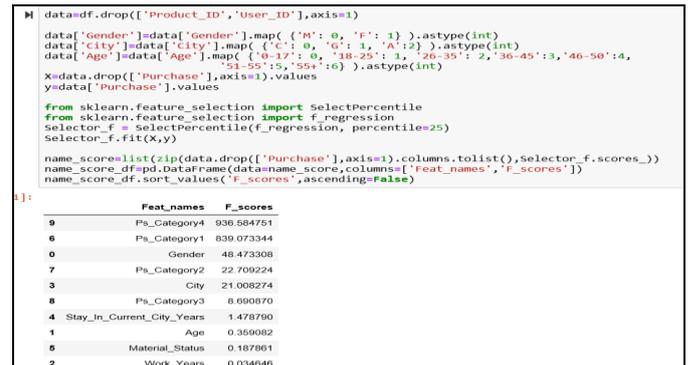


Figure 11. F_Scores for All variables

In this step the sklearn library was used to estimate the performance of machine learning model by calculating the mean absolute error (MAE) and mean squared error (MSE).

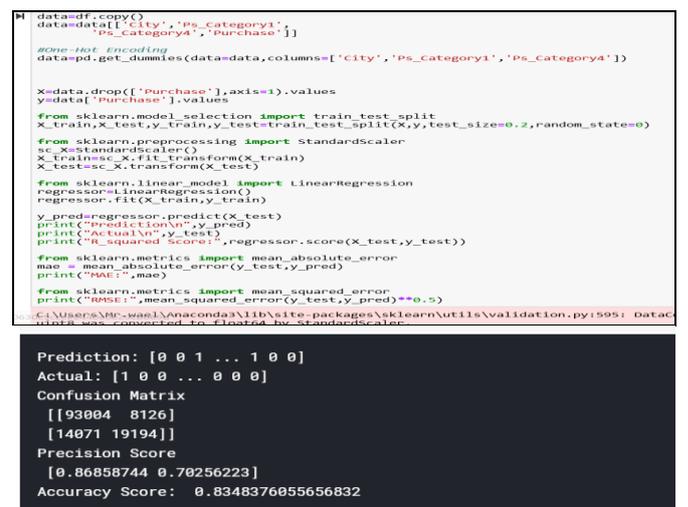


Figure 12. Linear Regression Model & Prediction score

The prediction of sales is a very important part of modern business intelligence [40-42]. Machine learning algorithms can be applied to these tasks. In this paper, the machine learning algorithm for predicting the total purchasing from online customers was applied. And, some exploratory data analysis has been performed for finding interesting relations and trends from the dataset.

IV. DISCUSSION

The purpose of this research was to shed light on the importance and the role of using predictive analytics models in e-commerce firms in Egypt where these tools became dominant resources for gaining valuable knowledge for better decision making. In order to achieve such purpose, predictive analytics was applied to an online transaction dataset obtained through one of an e-commerce company for forecasting categories product sales and the dataset has been analyzed.

Then, the predictive model was built for displaying how it can be used to gain critical insights.

Predictive modeling is the key for providing adaptive decisions, and firms now need to use the most advanced analytics to keep up with the competition. Such models could be useful for the e-commerce firms to make high operational efficiency, improve customer satisfaction, and robust revenue and profit levels.

V. CONCLUSION

With the global explosion in big data in all aspects of the business world, where big Data has brought a radical change in setting up and managing businesses, the transformation needed now from business to become data-driven evidence-based organizations is mandatory. Therefore, the applications and utilization of data analytics tools are considered a vital part of the decision-making processes in all businesses and in e-commerce firms especially. As the exponential growth and speed in the e-commerce industry is booming, which is expected [2] to reach \$4.5 trillion in 2021, where the data volume will continue in an increase in every second. Therefore, data will become a blessing when using it with analytical tools for unearthing hidden patterns, market trends, new opportunities, or solving existing problems or providing other insights. By using data analytics tools such as predictive models, e-commerce industries can peak into their trends, performance, and growth to improve their operation and customer satisfaction, increase revenue, gain new customers, and streamline their operations, and fact-based management to drive decisions and actions.

In This paper an example of predictive model for dataset to online transactions was presented to give a lot of insights for the importance and impact of employing one of the tool of data analytics models in Egyptian e-commerce firms. Moreover, such an analysis is useful for the decision makers in e-commerce companies for gain insights about understanding customers purchase patterns among different categories of products which can assist for example when preparing sales strategy. Where the analysis of big data is very important for e-commerce firms for many perspectives. The significant generation of data amount allows e-commerce firms to make decisions in timely manner where money can be saved and operations can be more optimized. Like, consumer behavior and preferences can be understand by analyzing the big data which includes, customer movement online webs site, transactions, product searches and finally for better decision-making.

REFERENCES

- [1] Haniska Roy, "E-commerce-is-growing-at-an-unprecedented-rate-all-over-the-globe",2019.<https://thenextscoop.com/e-commerce-is-growing-at-an-unprecedented-rate-all-over-the-globe/>
- [2] J. Clement, Global retail e-commerce sales 2014-2023, 2019. <https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/>
- [3] E. M. Agwu, and P. J. Murray, "Empirical Study of Barriers to Electronic Commerce Adoption by Small and Medium Scale Businesses in Nigeria," International Journal of Innovation in the Digital Economy, vol. 6, no. 2, pp. 1-19, 2015.
- [4] K. C. Laudon, and J. P. Laudon, "Management Information Systems: Managing the Digital Firm, 11th ed.," Upper Saddle River, NJ: Pearson Prentice Hall, 2010.
- [5] Beath, C., Becerra-Fernandez, I., Ross, J., & Short, J. " Finding value in the information explosion.", MIT Sloan Management Review, 53, 18–20., 2012.
- [6] H Chen, RHL Chiang, VC Storey – "BUSINESS INTELLIGENCE AND ANALYTICS: FROM BIG DATA TO BIG IMPACT", MIS quarterly, 2012. <https://pdfs.semanticscholar.org/f5fe/b79e04b2e7b61d17a6df79a44faf358e60cd.pdf>
- [7] HodaMoghimi, Stephen Vaughan, Steven McConche, NilminiWickramasinghe." How Do Business Analytics and Business Intelligence Contribute to Improving Care Efficiency?". 2016. 49th Hawaii International Conference on System Sciences.
- [8] Zion Market Research, "Predictive Analytics Market by Software Solutions (Data Mining & Management, Decision Support Systems, Fraud & Security Intelligence, Financial Intelligence, Customer Intelligence and Others) for Customer & Channel, Sales and Marketing, Finance & Risk and Other Applications: Global Industry Perspective, Comprehensive Analysis And Forecast 2016-2022" <https://www.globenewswire.com/news-release/2018/03/02/1414176/0/en/Trends-in-Predictive-Analytics-Market-Size-Share-will-Reach-10-95-Billion-by-2022.html>
- [9] SAS, Predictive Analytics, what it is and why it matters, 2019. https://www.sas.com/en_nz/insights/analytics/predictive-analytics.html
- [10] Kajal Govinda Fegade, Varsha Namdeo & Ravindra Gupta, "Predictive Model for Multiclass Classification of ECommerce Data: An Azure Machine Learning Approach", International Journal of Computer Applications (0975 – 8887) Volume 168 – No.7, June -2017.
- [11] PrasadaBabu, S.HanumanthSastry." Big Data and Predictive Analytics in ERP Systems for Automating Decision Making Process". 2014. IEEE
- [12] Michael Goul, Sule Balkan, Dan Dolk." Predictive Analytics-Driven Campaign Management Support Systems". 2015. Hawaii International Conference on System Sciences.
- [13] Shirley Coleman, Rainer Göb, Giuseppe Manco, Antonio Pievatolo, Xavier Tort-Martorell and Marco Seabra Reis, "How Can SMEs Benefit from Big Data? Challenges and a Path Forward". 2016. Published online by (wileyonlinelibrary.com) DOI: 10.1002/qre.2008. https://www.academia.edu/35085839/How_Can_SMEs_Benefit_from_Big_Data_Challenges_and_a_Path_Forward
- [14] McKinsey Analytics, "Analytics comes of age". McKinsey & Company, 2018. <http://www.rich.co.ke/rcfrbs/docs/Analytics-comes-of-age.pdf>
- [15] Gartner, Information Technology, <https://www.gartner.com/en/information-technology/glossary/predictive-analytics-2>
- [16] Kopp, M., Seizing the big data opportunity, Ecommerce Times, 2013. Available at: <http://www.ecommercetimes.com/story/78390.html>
- [17] Cherif, E., & Grant, D., Analysis of E-business models in real estate. Electronic Commerce Research, 14,1-26, 2013.
- [18] Personali (2018-10-11. Predictive Analytics, <http://www.personali.com/answers/predictive-analytics/>
- [19] Sun, Z.-L., Choi, T.-M., Au, K.-F., and Yu, Y., "Sales forecasting using extreme learning machine with applications in fashion retailing," Decision Support Systems (46:1), pp. 411–419, 2008.
- [20] Choi, T.-M., Hui, C.-L., Liu, N., Ng, S.-F., and Yu, Y. "Fast fashion sales forecasting with limited data and time," Decision Support Systems (59), pp. 84–92, 2014.
- [21] Sahoo, N., Singh, P. V., and Mukhopadhyay, T. "A Hidden Markov Model for Collaborative Filtering." MIS Quarterly (36:4), pp. 1329–1356,2012.
- [22] Abbasi, A., Albrecht, C., Vance, A., and Hansen, J. "Metafraud: a meta-learning framework for detecting financial fraud," Mis Quarterly (36:4), pp. 1293–1327, 2012.
- [23] Shmueli and Koppius, 2011, Predictive analytics in information systems research. MIS Quarterly Vol. 35 No. X/Forthcoming 2011.

- [24] Watson and Wixom, “the Current State of Business Intelligence”, IEEE, *Computer* Volume: 40, Issue:9, Sept. 2007 , DOI: [10.1109/MC.2007.331](https://doi.org/10.1109/MC.2007.331)
- [25] Finlay, Steven (2014). Predictive Analytics, Data Mining and Big Data. Myths, Misconceptions and Methods (1st ed.). Basingstoke: Palgrave Macmillan. p. 237. ISBN 1137379278.
- [26] Davenport, T.H., Harris, J.G., The dark side of customer analytics. *Harvard Business Review* 85, 37, 2007b.
- [27] Mehra, G., (2013). 6 uses of big data for online retailers, Practical Ecommerce. Available at: <http://www.practicalecommerce.com/articles/3960-6-Uses-of-Big-Data-for-Online-Retailers>
- [28] Cherif, E., & Grant, D. Analysis of E-business models in real estate. *Electronic Commerce Research*, 14,1-26, 2013
- [29] Meryem Ouahilal, Mohammed El Mohajir, Mohamed chahhou, BadrEddine El Mohajir. “A Comparative Study of Predictive Algorithms for Business Analytics and Decision Support systems: Finance as a Case Study”. 2016. IEEE.
- [30] Piatetsky-Shapiro, Gregory (1991), Discovery, analysis, and presentation of strong rules, in Piatetsky-Shapiro, Gregory; and Frawley, William J.; eds., *Knowledge Discovery in Databases*, AAAI/MIT Press, Cambridge, MA.
- [31] Fatimetou Zahra Mohamed, " The application of predictive analytics: Benefits, challenges and how it can be improved", *International Journal of Scientific and Research Publications*, Volume 7, Issue 5, 549 ISSN 2250-3153, May (2017).
- [32] Agrawal, R.; Imieliński, T.; Swami, A. (1993). "Mining association rules between sets of items in large databases". *Proceedings of the 1993 ACM SIGMOD international conference on Management of data – SIGMOD '93*. P.207.
[CiteSeerX 10.1.1.40.6984. doi:10.1145/170035.170072](https://doi.org/10.1145/170035.170072). ISBN 978-0897915922.
- [33] Report of MCIT, (2011), “The future of the internet economy in Egypt; a statistical profile”, Egypt, May.
- [34] Measuring the Digital Society in Egypt: Internet at a Glance Statistical Profile, Ministry of Communications & Information Technology, 2015. http://mcit.gov.eg/Upcont/Documents/Publications_1272015000_Measuring_the_Digital_Society_in_Egypt_12_.pdf
- [35] Algawady, Zynab (2005), The impact of e-commerce on developed and developing countries- case study; Egypt and united States, a paper presented in the international conference of globalization, technology and sustainable development, United Arab of Emirates University, Alain, November.
- [36] United Nations Conference on trade and development. ICT Policy review: National E-Commerce Strategy for Egypt. https://unctad.org/en/PublicationsLibrary/dtlstict2017d3_en.pdf
- [37] Ketel, Mark van & Nelson, Tim D., (2007), e-commerce, <http://searchcio.techtarget.com/definition/e-commerce>,
- [38] The Institute of Operations Research and Management Sciences (INFORMS) organization, (www.informs.org).
- [39] Kajal Govinda Fegade, Dr. Varsha Namdeo and Prof. Ravindra Gupta, “Predictive Modelling for E-Commerce Data Classification Tasks: An Azure Machine Learning Approach”, *International Journal of Electrical, Electronics and Computer Engineering, IJEE- ComE* 6(1): 45-50(2017)
- [40] Mentzer, J.T.; Moon, M.A. *Sales Forecasting Management: A Demand Management Approach*; Sage: Thousand Oaks, CA, USA, 2004.
- [41] Efendigil, T.; Önüt, S.; Kahraman, C. A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: A comparative analysis. *Expert Syst. Appl.* 2009, 36 , 6697–6707.
- [42] Machine-Learning Models for Sales Time Series Forecasting. Available from: https://www.researchgate.net/publication/330484523_Machine-Learning_Models_for_Sales_Time_Series_Forecasting