

# On-Care: Applying Predictive Analytics in e-Commerce System

Ace Zheridan S. Gutierrez and Alexander A. Hernandez

**Abstract**— Using massive data to be more competitive amongst other companies is an approach to business through the integration of information and communication technology, more specifically data analytics. This paper describes a development of a web application implementing predictive analytics which can help stakeholders to improved sales and customer relation. Having an online system that monitors the current number of stocks and maintaining the availability of products provides the stake-holders the edge against other competitors through efficient product predictive analytics. The results of the evaluation show that the software is highly acceptable among the end users. The study concludes that predictive analytics is relevant for recent developments in e-commerce systems. This study also presents practical and research recommendations.

**Keywords**— Predictive Analytics, Online Ordering, Linear Regression, Forecasting, Demand Management, Supply-chain.

## I. INTRODUCTION

Large volume of data has been acquired in the past two years [18] and many are now using data to win competitively [12] but having the right kind of data is not all that matters but as importantly is the development of analytics tools that focus on business outcomes [2] [31]. Methods of data science and applications that can manage this large amount of invaluable data have been developed in the form of predictive analytics [33]. Siegel assert that predictive analytics is the science that unleashes the power of data [27]. Predictive analytics include statistical models and other empirical methods that are aimed at creating empirical predictions [26]. Predictive analytics can properly as a core element in companies' efforts to improve performance [2] [11] [27]. Thus, motivates most organization to implement data analytic functions such as predictive analytics [12].

In the paradigm of modern business management individual businesses are referred to as supply chains [17]. Today, competition is not based on brand versus brand but rather on supply chain versus supply chain [16]. Success of business is determined by the management's ability to associate different business relationships [17] "Supply Chain Management is the integration of key business processes from end user through original suppliers that provides products, services, and information that add value for customers and other stakeholder" [17]. Supply chain's focus is still the customer

thus, accurate and timely responses to customer's demands are critical to effective supply Chain Management [17]. Thus, predicting future customer demand can maximize sales and marketing effectiveness [28]. In the different business processes, forecasting is a crucial feature that can contribute to the efficiency of the business functions [19]. Decreased operation costs, improve customer service, increased sales and reductions in inventory are some of the improvements that a consistent, systematic, and appropriate forecasting can offer which in turn has a positive effect on the returns on a shareholder value [20].

## II. RELATED WORKS

### A. Current state of Sales and Challenges

As companies becomes internationally competitive, supply chains face issues and challenges such as; increasing demands to reduce costs, increase quality, improve customer service and ensure continuity of supply [9] [21]. Additionally, the increase in turbulence and uncertainty of the marketplace [6] [8] makes the demands of the industrial sector more volatile. Thus, the increase in vulnerability of supply chains to disruption has a significant and negative impact on the financial aspects of the affected entities [7] and as a result, life cycle demands, are difficult to predict [4] [32].

The impact of unplanned and unforeseen events in supply chains can have severe financial effects [7] across the network as a whole. Research in North America suggests that when companies experience disruptions to their supply chains, the impact on their share price once the problem becomes public knowledge can be significant. The research suggests that companies experiencing these sorts of problems saw their average operating income drop 107 percent, return on sales fall 114 percent and return on assets decrease by 93 percent [4].

### B. Predictive Analytics using Regression models

Most companies are forecast driven [5]. The focus of some executives is the impact of forecasting on shareholder value. This impact can be clearly seen as an increased in sales through improved operational plans and customer service [18]. Accurate demand forecasting leads to an efficient operations and high level of customer service [14] [19]. Improved business performance can be achieved through implementations of perfect data analytic functions such as predictive analytics [11] [2] [27]. As stated by Waller and Fawcett (2013), Forecasting is about predicting the future, and predictive analytics add questions regarding what would happened in the past, given different conditions [33].

Manuscript received Aug. 31, 2017

Ace Zheridan S. Gutierrez is with Technological Institute of the Philippines.  
Alexander A. Hernandez is with Technological Institute of the Philippines.

Predictive analytics as referred by Shmueli & Koppius is the process of building and assessing of a model aimed at making empirical predictions that includes two components as follows:

a. Empirical predictive models designed for future observations or scenarios such as data mining.

b. Methods for evaluating the predictive power of a model.

Regression is a data mining function that predicts a number [1] [13] [14] [29]. When we are examining the relationship between a quantitative outcome and a single quantitative explanatory variable, simple linear regression is the most commonly considered analysis method [23] [22] [25] [30].

### III. METHODOLOGY

This project uses agile methodology that would provide the highest client satisfaction by developing the software incrementally. Through early and continuous delivery of quality software, the specification, development and objectives of the project is achieved. To develop the software, this study use Bootstrap, php, javascript, MySQL, and other third-party technologies. The study used the insights and feedback of the following users: Manager, Staff, and Medical Representatives. In evaluating the quality of the software, this study use ISO-9126 criteria during the test phase of the development. Using a Likert Scale with the interpretations of highly acceptable, acceptable, moderately acceptable, slightly acceptable, and not acceptable, the study measured the level of acceptance or disagreement of the user with regards to the software evaluation criteria. The results of the evaluation are presented in the next chapter of the paper.

To apply predictive analytics in forecasting the number of stocks to be maintained by the stake holders, Linear regression algorithm is used. The structural component of the model in linear regression algorithm, specifies the relationships between explanatory variables and the mean of the outcome variables [25]. In simple regression analysis, one seeks to measure the statistical association between two variables, X and Y. Regression analysis is generally used to measure how changes in the independent variable, X, influence changes in the dependent variable, Y. Regression analysis shows a statistical association or correlation among variables [10]. Correlation is a statistic that measures the strength and direction of a linear relationship which can either be positive or negative association. Correlation Coefficient measures the strength and direction of the linear relationship between response and prediction variable. It is a number between -1 and +1. Higher the absolute value of r, stronger is the linear relationship between the variables. Negative value of correlation coefficient represents negative relation between response and prediction variable, means that if X increases then Y decreases and vice versa. The square of the correlation coefficient denotes how much response variable can be explained from the prediction variable. After establishing that there is a strong linear relationship between response variable and predictor, we can now estimate the regression coefficients. This is equivalent to finding the equation of a straight line that best fits the points on

the scatter diagram of response variable versus the predictor variable. One of the various methods is known as “least square method”. It gives the equation of a straight line, which minimizes the sum of squares of the vertical distances from each point to the line [3] [10].

The formula for the correlation coefficient is:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}}$$

And the coefficient of determination,  $r^2$ , is the percentage of the variance of the dependent variable explained by the independent variable. The formula for the coefficient of determination is as follows:

$$r^2 = (r)^2$$

Using least squares method, the values of slope m and y-intercept c can be given as:

$$\text{Slope } m = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

$$\text{And y-intercept } c = \bar{y} - m * \bar{x}$$

Where  $\bar{x}$  = mean value of x-values,  $\bar{y}$  = mean of y-values

A linear relationship between two variables can be depicted in a general form of an equation which is as follows:

$$y = mx + c$$

To help the company to have a consistent level of stocks, predictive analytics is applied to the system. Historical data of the company provides a valuable resource in determining the forecast value of stocks. Using the correlation between the number of stocks as the response variable and a predictor that can be either the number of orders or the number of sales. We measure the strength of the relation between the two variables to determine the viability of the relationship. Using the least square method, we can determine how much response variable can be explained by the prediction variable. After determining a strong relationship between the two variables, we can now apply linear regression algorithm to create a model based on the historical data of the company.

Linear regression algorithm works by estimating coefficients for a line or hyperplane that best fits the training data. It is a very simple regression algorithm, fast to train and has great performance by using output variable of the data, which is a linear combination of the inputs. The very first thing the researcher did was to determine the strength of linear relationship between “the number of stocks as the response variable and the number of orders as the predictor”, “or the number of stocks as the response variable and the number of sales as the predictor”, using correlation coefficients. After having a strong positive relationship, we can apply the algorithm. In linear regression, we use the equation  $y=mx+b$ , wherein y is the value to be predicted, m is the slope, x is the

predictor, and b as the y-intercept. Using the system, we can find the values for each variable for us to forecast.

#### IV. RESULTS AND DISCUSSION

The next part of this section shows On-Care: An Online Ordering System for Neo Care Philippines using Predictive Analytics. Fig. 1 presents the system architecture for the Online Ordering system of Neo Care Philippines which is divided into between two user types. The modules for the Manager or Staff include Account Module, Stock Module, Inventory Module, Order Module, and Reports Module while the modules for the Medical Representative include Inventory module, and Order Module.

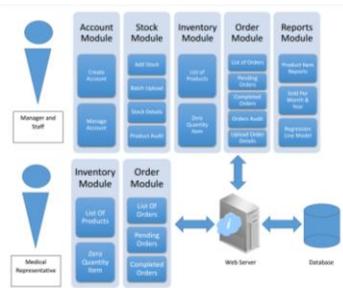


Fig. 1. On-Care: Online Ordering System with Predictive Analytics Architecture

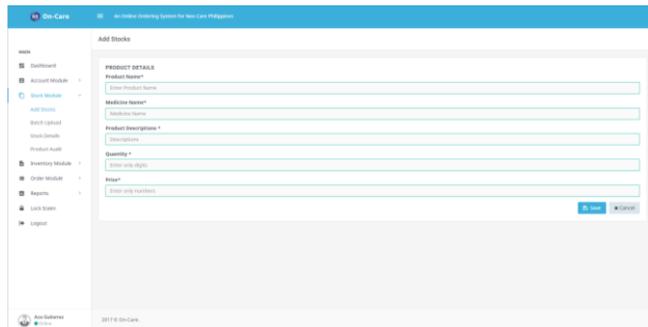


Fig. 2. Stock Module - Manager

Fig. 2 shows the Stock module of the system which contains Add Stocks, Batch Upload, Stock Details, and Product Audit. Through this module, the manager can view the details and add products to the system by providing the necessary details such as; Product name, Medicine name, Product description, Quantity, and Price.

Under the stock module, the manager can view the activities done under the Stock Module through the Product Audit page that can help verify the activities done through the system. The system creates a log of user activities which then can help the user check for previous transactions. Using Batch upload, the manager can add series of products to the system all at once. Batch upload feature lets the user import product data efficiently in a simple manner thus minimize the time for data encoding to the system.

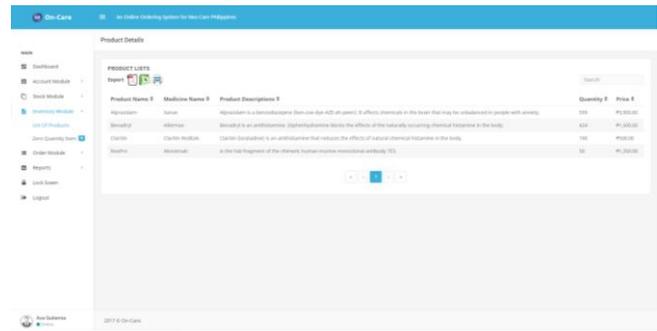


Fig. 3. Inventory Module – Manager

In Fig. 3, it shows the list of available products that is registered to the system. The page displays the basic information about the product such as Product name, Medicine name, Product Description, current Quantity, and Price. Under the Zero Quantity Item page, list of products that has zero quantity are displayed. A small notification icon is also displayed to show the number of products that has zero quantity. The notification provided by the system about zero quantity items can easily be seen by the user for immediate actions.

The information provided by this module can help the manager in decision making and planning specially with product management. The system monitors the current number of stocks and provides notifications to the manager if a certain item has low or zero quantity of items. Thus, help the manager in maintain a significant level of inventory.

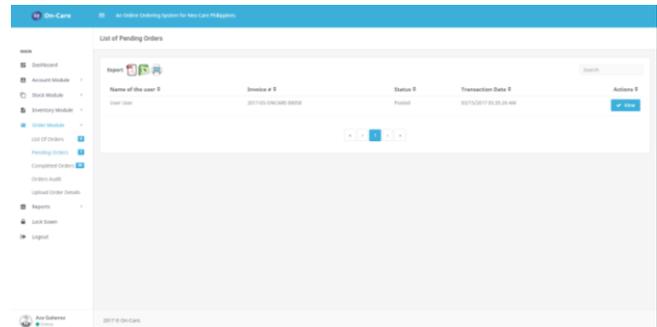


Fig. 4. Order Module - Manager

Fig. 4 shows the list of products that the medical representative wants to order. Under the List of Orders page, the on-going transaction of the medical representative can be seen. And when the medical representative posted his/her order, it will now be seen under the Pending Orders page. After the company finished assembling the order, it will now go to the Completed Orders page, and a notification will be sent to the Medical Representative account notifying him/her that the order is finished and ready for pickup. Through the module, the user can monitor the progress of each transactions made. The series of processes done under this module, from List of Orders, Pending Orders to Completed Orders, can eliminate ambiguity with regards to managing orders.

This module provides the company an online and paper-less transaction processing compared to the old system that uses

manual processing. Through the web application, the user can manage transactions or orders on the go.

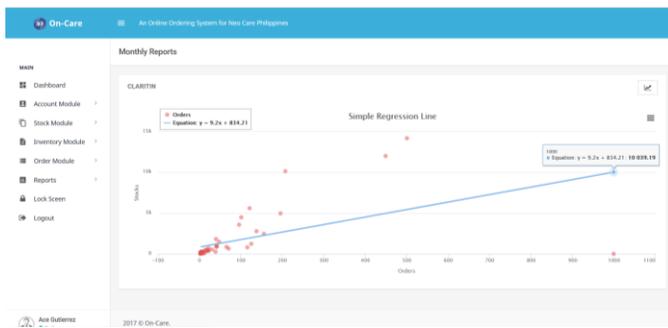


Fig. 5. Reports Module – Regression Line Model

In Fig. 5, a four-month report is shown by the system with the corresponding information about the number of stocks, number of orders, and number of forecasted stocks. The Reports module contains the analytics of the system. Using simple linear regression algorithm, and transactional data done through the system, the system can forecast the number of stocks needed by the company to maintain a consistent level number of stocks available.



Fig. 6. Reports Module

In Fig. 6, a four-month report is shown by the system with the corresponding information about the number of stocks, number of orders, and number of forecasted stocks. The Reports module contains the analytics of the system. Using simple linear regression algorithm, and transactional data done through the system, the system can forecast the number of stocks needed by the company to maintain a consistent level number of stocks available.

Using the forecast report provided by the system, the system helps the manager in planning for the next four months on how much products the company has import to maintain a consistent level of available stocks.

TABLE 1. SUMMARY OF SOFTWARE EVALUATION

Criteria	Mean	Interpretation
1. Functionality	4.67	Highly Acceptable

2. Reliability	4.76	Highly Acceptable
3. Usability	4.60	Highly Acceptable
4. Efficiency	4.82	Highly Acceptable
5. Maintainability	4.35	Moderately Acceptable
6. Portability	4.24	Moderately Acceptable

**Overall Weighted Mean 4.57 Highly Acceptable**

In summary, the software evaluation of On-care: An Online Ordering System for Neo Care Philippines using Predictive Analytics indicates a highly positive response. The evaluation shows that the system is highly functional (4.67), highly reliable (4.76), highly usable (4.60), highly efficient (4.82), moderately maintainable (4.35), and moderately portable (4.24). Hence, the software evaluation receives an overall rating of 4.57 with an interpretation of highly acceptable. The results show that the web application has certainly achieved the goal of the study of creating an online web ordering system that can forecast and provide a report for decision making.

## V. CONCLUSION

This research aims to provide an Online Ordering System with Predictive Analytics for Neo Care Philippines that can improve order management of the company through the web application and provide a forecast report that can aid in decision making. The software evaluation also shows the importance of having an online application that can aid the company in achieving its goals through fast and efficient transaction handling. However, this research has also recommendations to improve the findings of the study including; (a) implement a much complex regression algorithm; (b) acquire more transactional data; (c) continuous system upgrades and adding new features through user feedback; and (d) adapting a delivery and payment module.

## REFERENCES

- Almazyad, A. S., Ahamad, M. G., Siddiqui, M. K., & Almazyad, A. S. (2010). Effective hypertensive treatment using data mining in Saudi Arabia. *Journal of clinical monitoring and computing*, 24(6), 391-401. <https://doi.org/10.1007/s10877-010-9260-2>
- Barton, D., & Court, D. (2012). Making advanced analytics work for you. *Harvard business review*, 90(10), 78-83.
- Bhatia, N. (2009). *Linear Regression: An approach for forecasting*.
- Christopher, M. (2016). *Logistics & supply chain management*. Pearson UK.
- Christopher, M. (2000). The agile supply chain: competing in volatile markets. *Industrial marketing management*, 29(1), 37-44. [https://doi.org/10.1016/S0019-8501\(99\)00110-8](https://doi.org/10.1016/S0019-8501(99)00110-8)
- Coleman, L. (2006). Frequency of man-made disasters in the 20th century. *Journal of Contingencies and Crisis Management*, 14(1), 3-11. <https://doi.org/10.1111/j.1468-5973.2006.00476.x>
- Craighead, C. W., Blackhurst, J., Rungtusanatham, M. J., & Handfield, R. B. (2007). The severity of supply chain disruptions: design characteristics and mitigation capabilities. *Decision Sciences*, 38(1), 131-156. <https://doi.org/10.1111/j.1540-5915.2007.00151.x>
- Elkins, D., Handfield, R. B., Blackhurst, J., & Craighead, C. W. (2005). 18 ways to guard against disruption. *management*, 1(1), 10-11.
- Goebel, D. J., Marshall, G. W., & Locander, W. B. (2003). Enhancing purchasing's strategic reputation: Evidence and recommendations for future research. *Journal of Supply Chain Management*, 39(1), 4-14. <https://doi.org/10.1111/j.1745-493X.2003.tb00150.x>
- Guerard Jr, J. B. (2013). *Introduction to financial forecasting in investment analysis*. Springer Science & Business Media. <https://doi.org/10.1007/978-1-4614-5239-3>

- [11] Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154, 72-80.  
<https://doi.org/10.1016/j.ijpe.2014.04.018>
- [12] Hopkins, M. S., LaValle, S., & Balboni, F. (2010). 10 Insights: A First Look at The New Intelligent Enterprise Survey. *IT Management Select*, 14.
- [13] Hu, J. The Overview of Data Mining Applied in the Department of Public Security. In 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI).
- [14] Hussain, M., Shome, A., & Lee, D. M. (2012). Impact of forecasting methods on variance ratio in order-up-to level policy. *The International Journal of Advanced Manufacturing Technology*, 59(1), 413-420.  
<https://doi.org/10.1007/s00170-011-3485-1>
- [15] Jadhav, M. V., & Chandurkar, S. (2014). Mining: Student Database. *International Journal of Innovative Research and Development* | ISSN 2278-0211.
- [16] Ketchen, D. J., & Hult, G. T. M. (2007). Bridging organization theory and supply chain management: The case of best value supply chains. *Journal of Operations Management*, 25(2), 573-580.  
<https://doi.org/10.1016/j.jom.2006.05.010>
- [17] Lambert, D. M., & Cooper, M. C. (2000). Issues in supply chain management. *Industrial marketing management*, 29(1), 65-83.  
[https://doi.org/10.1016/S0019-8501\(99\)00113-3](https://doi.org/10.1016/S0019-8501(99)00113-3)
- [18] Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- [19] McCarthy, T. M., & Golobic, S. L. (2002). Implementing Collaborative Forecasting to Improve Supply Chain Performance.
- [20] Mentzer, J. T. (1999). The Impact of Forecasting on Return on Shareholder's Value. *The Journal of Business Forecasting Methods & Systems*, 8-12
- [21] Pearson, J. N., Ellram, L. M., & Carter, C. R. (1996). Status and recognition of the purchasing function in the electronic industry. *Journal of Supply Chain Management*, 32(2), 30.  
<https://doi.org/10.1111/j.1745-493X.1996.tb00223.x>
- [22] Polycarp, O. O. (2015). Application of Linear Mixed Effects Model on Hierarchical Data (KCPE Examination Scores) (Doctoral dissertation).
- [23] Radzevičius, A., Viškelis, J., Karklelienė, R., Juškevičienė, D., & Viškelis, P. (2016). Determination of tomato quality attributes using near infrared spectroscopy and reference analysis. *Žemdirbystė (Agriculture)*, 103(1), 91-98.  
<https://doi.org/10.13080/z-a.2016.103.012>
- [24] Richards, N. M., & King, J. H. (2013). Three paradoxes of big data. *Stan. L. Rev. Online*, 66, 41.
- [25] Seltman, H. J. (2015). *Experimental design and analysis*. Pittsburgh: Carnegie Mellon University, 428.
- [26] Shmueli, G., & Koppius, O. R. (2011). Predictive analytics in information systems research. *Mis Quarterly*, 553-572.
- [27] Siegel, E. (2013). *Predictive analytics: The power to predict who will click, buy, lie, or die*. John Wiley & Sons.



[28] Singh, N., Olasky, S. J., Cluff, K. S., & Welch Jr, W. F. (2006). U.S. Patent No. 7,080,026. Washington, DC: U.S. Patent and Trademark Office.

[29] Singh, N., Raw, R. S., & Chauhan, R. K. (2012). Data mining with regression technique. *Journal of Information Systems and*

*Communication*, 3(1), 199.

[30] Teshome, D. M. Rural Household's Exposure to Meteorological Drought in the Abay Valley of Simada Woreda, Upper Blue Nile Basic of Ethiopia.

[31] Van der Aalst, W. M. (2014). Data scientist: The engineer of the future. In *Enterprise Interoperability VI* (pp. 13-26). Springer International Publishing.

[https://doi.org/10.1007/978-3-319-04948-9\\_2](https://doi.org/10.1007/978-3-319-04948-9_2)

[32] Wagner, S. M., & Bode, C. (2008). An empirical examination of supply chain performance along several dimensions of risk. *Journal of business logistics*, 29(1), 307-325.

<https://doi.org/10.1002/j.2158-1592.2008.tb00081.x>

[33] Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77-84.

<https://doi.org/10.1111/jbl.12010>

**Ace Zheridan S. Gutierrez** is graduate student at the Technological Institute of the Philippines since 2015. He is the Program Chairperson of the Information and Technology Department of St. Bridget College. Born in the town of Batangas on August 29, 1988, Philippines. He is a graduate of BS in Information and Technology at St. Bridget College, Batangas, Philippines on March 2015. He is currently a full-time probationary faculty member of St. Bridget College and mostly teaches programming courses. He is also a mid-level administrator of the school acting as the IT Program Chairperson.



**Alexander A. Hernandez** is an associate professor at the College of Information Technology Education of the Technological Institute of the Philippines. He is a graduate of Doctor in Information Technology at the De La Salle University in the Philippines. He holds an MSc and BSc in Information Technology from Technological Institute of the Philippines.

His research interests are in Data Mining, Enterprise and Decision Support Systems, Sustainability and Green Computing. Dr. Hernandez has published in *International Journal of Green Computing*, *International Journal of Enterprise Information Systems*, *International Journal of Social Ecology and Sustainable Development*, *Journal of Cases in Information Technology*, *International Journal of Information and Communication Technology Education*, *International Journal of Asian Business and Information Management*, *International Journal of Sociotechnology and Knowledge Development*, and presented papers at various international conferences and workshops, among others.