PREDICTING A SERVICE REQUESTFORECASTING OF REGIONAL STANDARDS & TESTING LABORATORY FOR FOOD AND QUALITY ASSURANCE

[1]JAYSON T. GUIA, [2]JASMIN DE CASTRO – NIGUIDULA

[1,2]Technological Institute of the Philippines
E-mail: [1]guia.jayson.43@gmail.com, [2]jasniguidula@yahoo.com

Abstract - Demand forecasting are crucial for managing supply and when properly applied will help address the surpluses or shortage of materials requested for a certain period of time. This paper explores and addresses the problems of forecasting the services requested of RSTL in the government agency in one of the province of the Philippines to correctly determine the number of materials needed to conduct a test. This involves the activity of generating a forecasting models for each test services to determine the number of supplies and materials that RSTL will purchase. In this research, the researcher used data mining techniques to forecast the number of supplies needed based on the customer's laboratory services request. The researcher also discussed each food test request to provide the reader an insight of the test and why it was needed to test. Database record from year 2014 to 2016 were extracted, group, summarized and presented applying Knowledge Discovery for Data (KDD) methodology to evaluate them. Result of the evaluation shows that the data mining techniques were effectively forecast the trend especially in a seasonal pattern. It was also noticeable that test service harmful to human health when consume and those test that concerned with the shelf life of the product have a high request while low demand for the test request for healthy product contents.

Keywords - Kdd, Trend And Seasonality, Data Mining, Test Services, Shelf Life

I. INTRODUCTION

The world's datum and each lives' information keeps expanding and will never end. [1]These data if properly studied and analyzed can provide informative result to be used in decision making. Analyzing the patterns to data can help discover how the fully utilized them in our own advantage through identifying opportunities using prediction or forecasting[2]. To evaluate the data, data mining methodology must be applied. According to Witten and Frank (2011),data mining integrates data analysis through a series of activities that will interpret the result through evaluation[1]. Algorithm in data mining is one of many approaches that produce and generate useful information [3]. People whose main job are related to economy, especially statisticians and forecasters believe that pattern in data can be identified, filtered that can be used in prediction if properly validated which in turn provide opportunities[4]. Regional Office of the Department of Science and Technology (DOST)[5]houses the Food Safety & Quality Assurance Center, Microbiological to fast track technological capability's development of industry but also to improve the competitiveness of local products and/or services through industry initiatives encouragement to meet not just national but also international standards as well for various products in generated from agriculture and other related products.[5]. Food Safety & Quality Assurance Center and Microbiological Laboratories (FSQAS&ML) requires consumable supplies and materials to conduct However, in their requisition of supplies and materials, laboratory personnel sorted the top 10 in-demand test services per quarter to determine the necessary supplies and materials that will be requested.

This study aims to analyze the number services request provided per month from year 2014 to 2016 for food to forecast the demand in the future periods. Specifically, it will forecast each test request data to determine the number of supplies and materials needed to conduct the test, and utilized them to provide guideline to reduce the deviation between actual and predicted values.

II. RELATED LITERATURE

Food and food products are required to be tested in the laboratory for safety and quality before it will released in the market. Determining the chemical, biological and or physio-chemical contents in food are important because it guarantee the quality and safety of the food products [6]. Ash content in food is the remaining residue after heating usually is less than 5% but some meat based product such as bacon and dried beef may contain as high as 6% and 11.6% respectively [7]. The higher the level of ash content in food can change the pH level in the blood and will be harmful to the organ [8].The higher the level of pH, the better and lesser the acidity in the body [9]. Determining the nutritional facts which is use for food labeling of food products [10].The next test services is the reversal of the latter because it's very harmful to health when consume by any person, the Escherichiacoli (E.coli)&Salmonella.(E.coli)are found in all warm-blooded animals and humans which has a useful function in the intestinal track[11]. However, a pathogenic strain such as Salmonella [12]were transferred to seafood through polluted sewage that ends in the coastal waters which in turn,

infects sea creatures such as shellfish. Thus when consume by human and will cause sickness such as nausea and vomiting, abdominal cramping, fever and watery or bloody diarrhea[12]. Flavor and aroma of many food and food products rich in fats such as milk, butter or cheese depends on the concentration of free fatty acids (FFA). This is the main reason many company requested a conduct of test service to determine the measure accurately that are desirable to their manufactured product[13]. Dairy products applies Standard Plate Count (SPC) test to determine if its quality of milk such as high grade or pastured or in frozen foods. To eliminate the pathogens and prevent food wastes and decays, the SPC for organism (microbes) must be limited[14]. Another test services provided is to determine the number of fungi used or presence in food products. It outwits the harmful effects to human which are useful to food processing including flavors, odors, textures etc. of a certain food products[15]. The most common properties of food and food products is its moisture content which defined the taste, shelf life, texture, etc., with a slight differences may affect the physical content of the product[16]. Food manufacturers arerequired to requestfor the moisture contents of their product to determine the total carbohydrate content of their product[17].The same is said on the water activity which determine the stabilityof the properties which are essentials to the food and food products which usually above 0.95% moisture. However, a high water activity contents increases the growth of microorganism causing spoilage that lessen the food products shelf life[18]. Staphylococcus detection in food production and products is one of the most important test services needed to be conducted. This is due to the threat to a person's health when food contaminated with the said Staphylococcus was consumed. It was usually found on water, soil, and air and even in the skin and nose of a human. Thus it could easily transmitted to food when improper preparation of food or food products are made [19].

Test services mentioned requires supply and materials and its availability of it are essentials to implement the said services.Thus, decision making must be improve to with the proper use of forecasting[20]. Demand forecasting [21], time series forecasting (TSF) [22], moving-average[23], artificial neural network (ANN)[24] algorithms, exponential smoothing[22], Box-Jenkins method[22] casual models like regression[23], artificial intelligence (AI)[21], and genetic algorithm (GA)[21] are some of techniques in forecasting.

One of the key areas in forecasting that attracted many researcher is the demand forecasting (DA). It includes several models such as time series, casual, and econometric models[22]. According to Chodak and Kwaśnicki (2000),[21] effective decision making of a firm rely on DA.With DA, they which they can institute "just-in-time" strategy to monitor their daily transaction such as the flow of materials.

Because of the limitations using traditional forecasting method,Shao, Zhang, et al. (2016) appliedTime Series (TSF) to determine the future values which came from the previous time series that was observed[25].They noted that there were limitations in forecasting in the traditional forecasting methods when applied in the manufacturing and preferred Artificial Neural Network (ANN) [25] algorithms techniques. This technique was also applied by Kochak and Sharma (2015) in a valve manufacturing industry[24]which increased forecasting accuracy and more reliable forecasting technique for this case.

Parikh (2003) compare two models in data mining and concluded that theHybrid Clustering Classification model (HCCM)[26] accurately produce better result than Pure Classification model (PCM). It was also able to provide better demand forecasting[26]. While Chodak and Kwaśnicki (2000)noted that the applicability seasonal in nature does not work well in some data but was useful in the Wroclaw rolls production's decision making[21].

In relation to warehouse, a study conducted by İşlek, and Öğüdücü (2015) noted the difficulty of estimation of product demand as it increases in relation with several distribution warehouses. They applied clustering the warehouse through splitting forecasting model and achieved success in their goal in solving the warehouse problem[27].P. Murray, B. Agard et al. (2015) also applied clustering method but also included segmentation of customers and applied applicable forecasting model to each segment[28].

III. METHODOLOGY

The Knowledge Discovery in Data (KDD) methodology as shown in figure 1 by Jiawei, H., et al. (2012) includes steps from dataset cleaning, evaluation up to presentation[29]. The researcher will follow this process in analyzing the data. Prior to attempting to analyze the case study data, the important stage of evaluating and resolving data quality had to be addressed.[30]

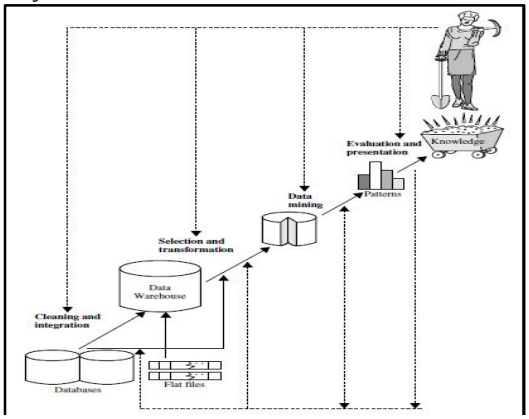

Figure 1. Knowledge Discovery in Data, Jiawei, H., et al. 2012

A three (3) year period dataset of sample analyzed per test activity extracted from year 2014 to 2016 consisting of 4,193 instances of record that include test name, month and year from year 2014 to 2016. Using the process of Knowledge Discovery in Database (KDD) model in figure 1, datasets cleaned through selection of necessary fields from both records and combining and exporting them in Microsoft Excel format. From the excel format, data were selected and transform through trimming, summarizing and group according to test name to get the total number for each quarter and per year.

Using trend and seasonality forecasting as a data mining technique. Many time series data have a common tendency of growing over time, and therefore contain a time trend. When making a causal (changes in X causing changes in Y) inference, we may falsely conclude that change in X caused change in Y. It might be that both variables are trending in the same or opposite directions because of a third unobserved factor. Trend analysis fits a general model to multiple time series data having trend pattern and provides idea to traders about what will happen in the future based on historical data. Trend can be linear, quadratic, or S-curve where Ft is forecast for time period t, t is specified number of time periods, a is intercept of the trend line, b is slope of the line, n is number of periods, and y is value of the time series.[31]

The results were evaluated and discussed in the results and discussion while graphical layout were also presented to help the researcher in presenting using Tableau 10.4 Professional Editionto generate the actual demand, trend line forecast, and seasonal forecast with trend separated by year and quarter.

RESULTS AND DISCUSSION

Table 1 is the summary of test services for food that were tabulated per year and quarter.

Test Name	2014				2015				2016			
	1	2	3	4	1	2	3	4	1	2	3	4
Ash	3	0	1	0	6	0	3	2	4	3	7	0
Drafting of nutrition Facts	4	0	5	0	0	0	0	1	0	0	7	2
Escherichia coli & Coliform Count	83	76	44	47	110	55	140	52	74	75	94	57
FFA	2	0	0	0	3	0	0	1	0	0	1	0
Fungi Count	78	71	43	43	83	57	129	41	72	45	58	10
Moisture	6	0	1	0	6	0	3	1	19	6	7	1
pH	24	2	4	0	14	1	2	0	14	0	0	0
Salmonella	38	30	25	27	33	25	87	36	32	52	75	71
Standard Plate Count	69	76	15	44	110	64	124	52	78	55	38	43
Staphylococcus detection	27	43	23	22	22	12	79	22	16	41	57	19
Water Activity	26	19	4	0	14	12	22	23	73	69	4	17

Table 1. Summary of Record per Year and Quarter for Food Test Dataset

From the total of 50 test services provided, 11 were selected because it provide complete data from 2014-2016.Total per quarter for the last 3 years was computed for each test services conducted were totaled to forecast the demand for the year 2017.

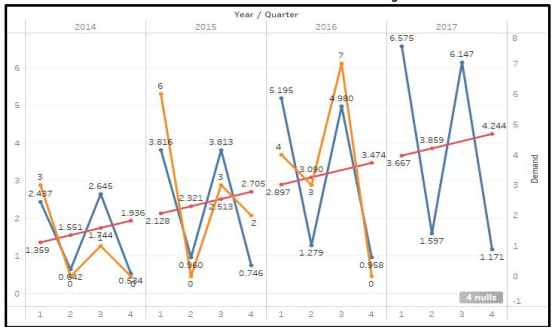

Figure 2 Forecasted service for Ash Test

The overall trend as shown in figure 2 shows that the demand for Ash test service will continue to increase through 2017.

The actual demand represented by orange line shows pattern of high demand every 1st and 3rd quarter of each year proves that the service request were seasonal in nature. This is because of the possibility of high demand for meat based products that need to be tested in preparation for the high demand for holiday season during the 4th quarter of each year. This trend became the basis of forecasted number of request,it is possiblethat the increase of service request to seven (7) and six (6) request for the first and third quarter for the year 2017.

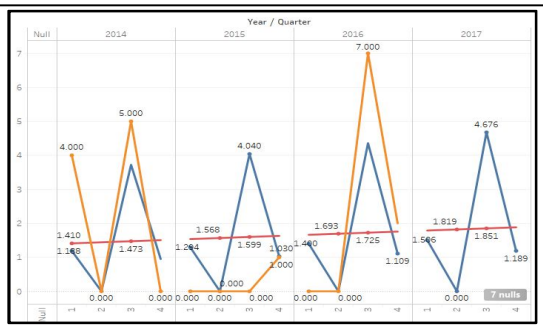

Figure 3 Forecasted service for Drafting of Nutrition Facts Test

Figure 3 shows that the actual demand in comparison with the seasonal demand were almost identical in the year 2014 and 2016 which shows an increase in test request every 3rd. quarter. However, results in 2015 contradicted the actual compare to the forecast service requested. Nonetheless, seasonal forecasting method followed the linear trend of the actual demand to generate a forecast service request for 2017 in which the demand will increase by five (5) on the 3rd quarter. The results also suggest that most of the food products were made during the said period and the same level of nutritional facts must bevalidated before releasing it to the market.

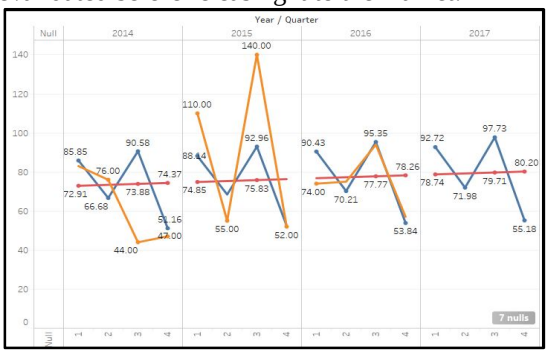

Figure 4 Forecasted service for Escherichia coli & Coliform Count Test

The year 2014 shows a contradicting result between actual services provided and seasonal forecast as shown in Figure 4 but eventually generated a pattern beginning 2015 to 2016. This became the basis of the linear trend forecast which was then became the pattern of seasonal forecast with trend. Figure 4 also shows that the actual and forecasted demand generated the same pattern in 2016 thus the forecasted number of test request for 2017 will be almost the same as in 2016. The results also suggest that there are few sea-based food product are being developed in the CALABARZON area.

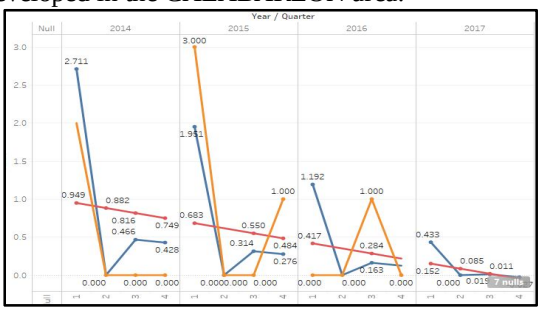

Figure 5 Forecasted service for Free Fatty Acid Test

Seasonality was not evident in figure 5 the year 2014 and 2015 with high request of service in the 1st quarter and no service request in the succeeding quarter. A test request was seen again in the last quarter of 2015 and in the 3rd quarter of 2016. However, the figure 5 shows a decreasing trend of service request that the linear trend forecast was able to trace it. Although the region provides different dairy products, it was ironic that companies requested for FFA level continue to decrease as showed in the projected forecast for 2017. This suggest that companies requested the test from other laboratories or not being tested at all.

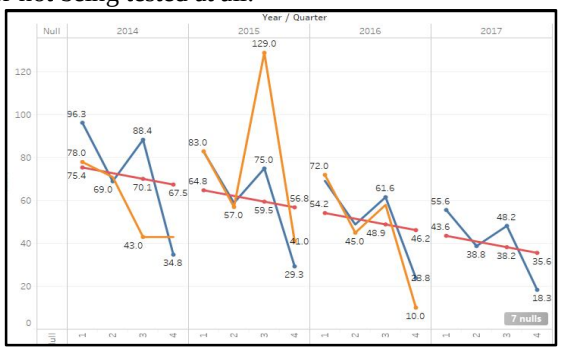

Figure 6 Forecasted service for Fungi Count Test

Figure 6 also shows patternof decreasing demand of service request for the test from 2014 to 2016 though a sudden spike of request was seen in the 3rd quarter of 2015. The contradicting result between actual forecast and seasonal forecast for 2014 was evident, it was able to generate in almost the same result of graph in 2016. However, the test request services of the seasonal forecast trend for 2017 will continue to decrease with test request ranging from approximately up to 66 forecasted request in the 1st quarter to as low as 20 for the year 2017. This was also more evident in the linear trend forecast wherein a downtrend path was generated from 2014 to 2017.Decreasing test requested in the region suggest the possibility that food being tested does not contains fungi necessary for the test or companies have their test by someone else.

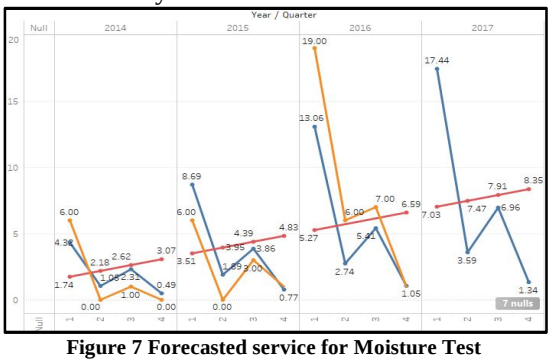

Figure 7 Forecasted service for Moisture Test

Figure 7 shows consistent trend pattern of request that shows seasonality in nature from 2014 to 2016 which shows an increase in the demand every 1st quarter in each year. This shows that the seasonal forecast with trend for the year 2017 will follow the same pattern

which will increase by 18 request for the 1st quarter and 6 request for the 3rd quarter. It was also evident in the line chart generated by the linear trend forecast from 2014 to 2016. Since the moisture test determine the carbohydrate content of the food product, this suggest that there are possibility that many companies in the region focus on high carbohydrate based products.

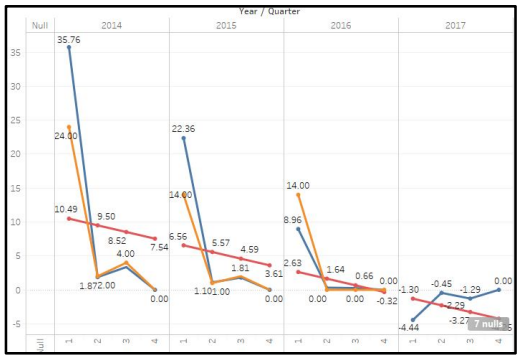

Figure 8 Forecasted service for pH Test

Just like in figure 6 wherein the trend is decreasing, the same as what has happening in figure 8 that shows that pH test request in which a high request of service requested was made in the 1st quarter and low increase in the 3rd quarter of each year from 2014 to 2016. However, it also shows the decline of request for three (3) consecutive years as shown in the line chart generated also from 2014 to 2016. This pattern became the basis of the seasonal forecast with trend that the same will happen in 2017 with one (1) or no request at all. The results also suggest the possibility that companies are less concerned of the pH level of food product they produced.



Figure 9 Forecasted service for Salmonella Test

Although figure 9 show an increasing trend pattern based on the linear trend forecast analysis, actual demand shows that requested services made in 2014 were different in 2015 and 2016 in terms on which quarter in was requested. The actual demand for 2015 and 2016 was the same with each made in the 3rd quarter but at the lower number for 2016. This trend became the basis of generating a seasonal forecast with trend for 2017 wherein it will follows the increase request for the test by 100 for the 3rd quarter of 2017. The trend also suggest the concern of the companies regarding the possibility of salmonella contamination in the raw materials that they used.



Figure 10 Forecasted service for Standard Plate Count Test

Figure 10 shows the same pattern in figure 5 which incidentally are using milk as raw materials,the year 2015 were able to generate an unusual high demand for the test conducted but low in test request in 2014 and 2016. The only noticeable in the figure is the downward trend that the linear trend forecast generated. This trend became the basis of seasonal forecasting with trend for the year 2017 which shows that only in the 1stquarter 2017 will increase the demand for the test request by 80 and will decrease to 39 request in the 4th quarter. Since both FFA and SPC are milk based food products, there is a possibility that the same reason forthe downtrend of FFA test to SPC test.

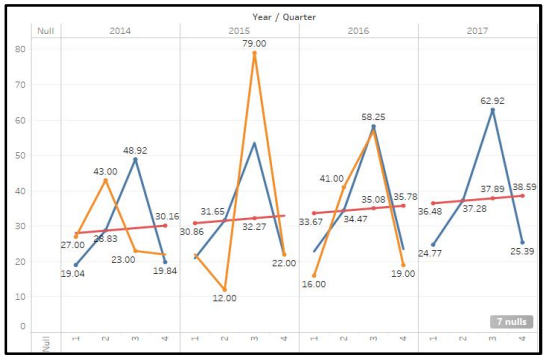

Figure 11 Forecasted service for Staphylococcus detection Test

Figure 11 shows the seasonality of demand based on the actual demand in which high demand occurs every 2nd and 3rd quarter of each year for the last three (3) years of 2014 to 2016. The same sequence was generated by the linear trend seasonality that shows a slightly upward trend from 2014 to 2017. This pattern of seasonality of service requested became the basis to forecast the demand for each quarter of the year 2017 in which from 25 request on the 1st quarter to as high as 60 request on the 3rd quarter before the demand for request will decline again in the 4th quarter. An upward pattern every 3rd quarter of each year suggest that raw materials for food products are also seasonal. However it also suggest that there is a possibility that Staphylococcus are most active or prevalent on the said quarter.

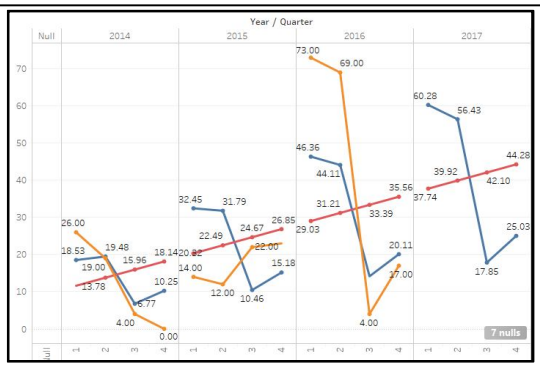

Figure 12 Forecasted service for Water Activity Test

Aside from the presence of harmful bacteria in food product that must be determine, it also important for the company the shelf life of its product improve and monitored thus in figure 12 shows pattern of increasing trend of based on the linear trend forecast generated graph. The year 2014 with small number of services requested is similar to 2016 in terms of service requested made per quarter with high request in the 1st quarter and decrease up to 4th quarter of each year. The 2015 actual demand was different but it does not prevent the seasonal forecast with trend to generate a seasonal pattern which it follows through 2017.

CONCLUSION AND RECCOMMENDATION

Based on the analyzed results the number services request provided per month from year 2014 to 2016 for food, the researcher was able to forecast the future demand using trend with seasonal forecasting technique and Tableau 10.4 Professional Edition as a software tool to apply the said algorithm.

Each test request data comprising of 11 test category for food were evaluated using the said techniques were applied and generated forecast per quarter for the year 2017. Each forecasted demand will be used to determine the number of supplies and materials needed to conduct the test. Though no such thing as accurate or exact forecasted data, this will serve as guidelines to reduce the deviation between actual and predicted values. It was also noticeable that test service provided by the government agency to their client such as those harmful to human when consume and those test that concerned with the shelf life of the product have a high request while low demand for the test request for healthy product contents. Most of the generated result generated in graphical format were somehow near the actual demand specifically if data have a seasonal pattern. The researcher therefore recommend to applied it especially in those test with trend and seasonal pattern. Some forecasted test services were somehow shows not recommendable results, therefore other forecasting method must the applied.

REFERENCE

[1] H. Witten and E. H. M. A. Frank, Data Mining Practical Machine Learning Tools and Techniques, Third Edition, Burlington,: Morgan Kaufmann Publishers, 2011.

[2] H. Witten, E. Frank and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques Third Edition, Burlington,: Morgan Kaufmann Publishers, 2011.

[3] K. A, Data mining in industrial applications and innovation, 2007.

[4] P. Giudici and S. Figni, Applied Data Mining for Business and Industry, Second Edition, West Sussex: A John Wiley and Sons, Ltd., Publication, 2009.

[5] "region4a.dost.gov.ph," Department of Science and Technology CALABARZON, [Online]. Available: http://region4a.dost.gov.ph/programs-and-services/testing-and-calibration/regional-services-and-testing-laboratory.

[6] A. Cifuentes, "Food Analysis: Present, Future, and Foodomics," International Scholarly Research Notices , vol. 2012, p. 16, 2012.

[7] J. McClements, "http://people.umass.edu/~mcclemen/," [Online]. Available: http://people.umass.edu/~mcclemen/581Ash&Minerals.html. [Accessed 15 October 2017].

[8] Pike, Brian, "https://www.livestrong.com," [Online]. Available: https://www.livestrong.com/article/449416-what-is-ash-in-food/. [Accessed 15 October 2017].

[9] P. Shepherd, "https://trans4mind.com/," Trans4mind Ltd, 1997. [Online]. Available: https://trans4mind.com/nutrition/pH.html. [Accessed 15 October 2017].

[10] "http://www.fnri.dost.gov.ph/," [Online]. Available: http://www.fnri.dost.gov.ph/index.php/publications/writers-pool-corner/57-food-and-nutrition/85-what-is-in-your-food-label. [Accessed 15 October 2017].

[11] D. Ward, D. Bernard, R. Collette, D. Kraemer, K. Hart, R. Price and S. Otwell, "Hazards Found in Seafoods," in Hazard Analysis and Critical Control Point Training Curriculum, North Carolina Sea Grant, Raleigh, NC., 1997, p. 177.

[12] P. Feng, S. D. Weagant, M. A. Grant and W. Burkhardt, "https://www.fda.gov," U.S. Food and Drug Administration, [Online]. Available: https://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm064948.htm. [Accessed 15 October 2017].

[13] R. Puchades, A. Suescun and A. Maquieira, "Determination of Free Fatty Acids in Food Flow Injection," Journal of the Science of Food and Agriculture, vol. 66, no. 4, pp. 473-478, 1994.

[14] A. Singh, U. Bajwa and R. K. Goraya, "Effect of Storage Period on the Physicochemical, Sensory and Microbiological Quality of Bakery Flavoured Ice Cream," Journal of Engineering Research and Applications, vol. 4, no. 8, pp. 80-90, 2014.

[15] "http://www.encyclopedia.com," 2003. [Online]. Available: http://www.encyclopedia.com/plants-and-animals/microbes-algae-and-fungi/moneran-and-protistan/fungi. [Accessed 15 October 2017].

[16] D. J. McClements, "http://people.umass.edu," [Online]. Available: http://people.umass.edu/~mcclemen/581Moisture.html. [Accessed 16 October 2017].

[17] S. S. Nielsen, "Determination of Moisture Content," in Food Science Text Series, Springer Link, 2009, pp. 17-27.

[18] M. R. Swain, M. Anandharaj, R. C. Ray and R. P. Rani, "Fermented Fruits and Vegetables of Asia: A Potential Source of Probiotics," Biotechnology Research International, vol. 2014, p. 19, 2014.

[19] Medveďová, Ľ. Valík and A. Studeničová, "The Effect of Temperature and Water Activity on the Growth of Staphylococcus aureus," Czech Journal of Food Science, vol. 27, no. 2, pp. 28-35, 2009.

[20] A. Kochak and S. Sharma, "Demand Forecasting Using Neural Network for Supply Chain Management,"

International Journal of Mechanical Engineering and Robotics Reasearch, vol. 4, no. January 2015, 2015.

[21] G. Chodak and W. Kwaśnicki, "Genetic Algorithms in Seasonal Demand Forecasting," Munich Personal RePEc Archive, pp. 91-99, 11 October 2011.

[22] P. Kumar, M. Herbert and S. Rao, "Demand forecasting Using Artificial Neural Network Based on Different Learning Methods: Comparative Analysis," 2014.

[23] Z. Rakićević and M. Vujošević, "FOCUS FORECASTING IN SUPPLY CHAIN: THE CASE STUDY OF FAST MOVING CONSUMER GOODS COMPANY IN SERBIA," Serbian Journal of Management, 10 January 2015.

[24] A. Kochak and S. Sharma, "Demand Forecasting Using Neural Network for Supply Chain Management," International Journal of Mechanical Engineering and Robotics Research, vol. 4, no. January 1, 2015, 2015.

[25] D. Shao, T. Zhang, K. Mannar and Y. Han, "Time Series Forecasting on Engineering Systems Using Recurrent Neural Networks," in Data Mining and Application, 12th International Conference, ADMA 2016 , Gold Coast, 2016.

[26] B. Parikh, "Applying Data Mining to Demand Forecasting and Product Allocations," 2003.

[27] Ş. G. Ö. İrem İşlek, "A Retail Demand Forecasting Model Based on Data Mining Techniques," in 2015 IEEE 24th International Symposium on Industrial Electronics (ISIE), Buzios, Brazil, June 3-5, 2015.

[28] P. W. Murray, B. Agard and M. A. Barajas, "Forecasting Supply Chain Demand by Clustering Customers," in 2015 IFAC Symposium on Information Control in Manufacturing, Ottawa, Canada, 2015.

[29] J. Han, M. Kamber and J. Pei, Data Mining: Concepts and Techniques, 3rd Edition, Waltham,: Morgan Kaufmann Publishers, 2012.

[30] K. M., Data Mining: Concepts, Models, Methods, and Algorithms, Hoboken, NJ: Wiley-IEEE Press, 2011.

[31] L. Karmaker, P. K. Halder and S. E., "A Study of Time Series Model for Predicting Jute Yarn Demand: Case Study," Journal of Industrial Engineering, no. 27 July 2017, p. 8, 2017.

★★★