Capstone Project Proposal

Project Title: Food Temperature Analysis and Forecasting

Domain and Context:

The project is an application of statistical analysis and Time Series modeling in the Food Industry. The Client is a chain of restaurants where food storage within a safe range of temperature is a mandate.

In USA, FDA (Food and Drug Administration) standards legally require restaurants to record the food storage temperature at specified time intervals to ensure that the food is always maintained within a safe range in order to prevent it from going bad. So far, this process has been carried out manually by employees at restaurants and storage facilities across the country.

As a result of the advent of IOT and Big Data, it is now possible to track food temperature and other parameters more accurately and conveniently using IOT based Smart Sensor devices. Hence, it is now feasible to analyze temperature fluctuations between readings and to study the effects of such environmental dynamism on food quality.

Time Series forecasting models can be used to predict the temperature of stored edibles, to retain their freshness and prevent undesirable outcomes such as food spoilage.

Problem Statement/Definition:

The end goal of this project is twofold; Objective A is to be addressed by the *Interim Report Stage* and Objective B by the *Final Report Stage*. 100% of the project is to get completed before *Presentation Stage*.

Objective A:

The Client requires a thorough statistical comparison of manual temperature data against IOT Sensor data. The intention is to demonstrate that the latter is more reliable in detecting and preventing food spoilage. In statistical terms this can be expressed using Null and Alternate Hypotheses:

***H_0**: Manual temperature readings should suffice in the maintenance of food safety standards and to prevent spoilage. (Sensors do not provide any significant advantage in this regard).*

***H_1**: Temperature measurement using sensors can significantly aid in detecting and preventing food spoilage.*

Should the Client's initial assumption be correct, the study should indicate that manual measurement isn't reliable in maintaining optimal food safety standards. The Client's intention is to utilize the results of this research to empirically and statistically demonstrate to insurance companies the benefits and importance of *IOT Smart Sensors*, generating business value in the process. The report generated for the Client is to have a thorough analysis and conclusion for each of the questions below with an emphasis on visual communication.

The aforementioned statistical study is to be conducted in the context of the parameters listed below:

- What is the average temperature of each food item tracked?
- What percentage out of range is the temperature on average?

- What is the percentage out of temperature during peak times?
- What percentage of temperature readings are manually recorded inaccurately per day on average?
- How long on average does a food item stay out of temperature?
- *How long does an employee spend probing food and recording temperatures on average and total per day?**
- *How many items of food should have been discarded per day and how many were not?**
- *How many customers were served food that contained ingredients that were out of temperature?**

*These are additional questions and are subject to Client's access to data. The Client informed that they are currently working on acquiring the relevant data. Should that occur, these points would get included in the scope of Objective A.

Objective B:

The objective is the implementation of a temperature forecasting Time Series model for food items in storage.

Until the interim stage of the project, the team's efforts would be solely focused on Objective A. As stated earlier, a detailed report on Objective B is to emerge in the Interim Stage Report of the project.

Literary Survey:

1.) Meiyu Wen, Dandan Che, Jean-Pierre Niyigena, Ruohan Li, Qingshan Jiang, 2019 "A Prediction Model of IoT Data Using Long Short-Term Memory Neural Network" - ISAICN

2.) Alexandru Popa, Mihaela Hnatiuc, Mirel Paun, Oana Geman, D. Jude Hemanth, Daniel Dorcea, Le Hoang Son and Simona Ghita ,2019 "An Intelligent IoT-Based Food Quality Monitoring Approach Using Low-Cost Sensors" - symmetry

3.) Roberto Casado-Vara, Paulo Novais, Ana Belen, Javier Prieto, Juan Manuel Corchado, 2018 "Distributed Continuous-Time Fault Estimation Control for Multiple Devices in IoT Networks"- IEEE

4.) V.Sandeep, K.Lalith Gopal, S.Naveen, A.Amudhan, L. S. Kumar, 2015 "Globally Accessible Machine Automation Using Raspberry Pi Based on Internet of Things" - ResearchGate

5.) Pedro Lima Monteiro, Massimiliano Zanin, Ernestina Menasalvas Ruiz, João Pimentão, Pedro Alexandre da Costa Sousa, 2018 "Indoor Temperature Prediction in an IoT Scenario" - sensors

6.) M.K. Nallakaruppan, U. Senthil Kumaran, 2019 "IoT based Machine Learning Techniques for Climate Predictive Analysis" - IJRTE

7.) Marco Ferrari, Msrco Bau, Vittorio Ferrari, 2014 "Resonant Piezo-layer (RPL) Sensors with Contactless Interrogation for Food Monitoring from Outside Sealed Packages" - ScienceDirect

8.) A. Flammini, E. Sisinni, 2014 "Wireless Sensor Networking in the Internet of Things and Cloud Computing Era" - ScienceDirect

9.) Ganesan Sangeetha, Muthuswamy Vijayalakshmi, "Role of Smart Sensors in Minimizing Food Deficit by Prediction of Shelf-Life in Agricultural Supply Chain" - Principles of Internet of Things (IoT) Ecosystem: Insight Paradigm by Sheng-Lung Peng, Souvik Pal, Lianfen Huang (pp 153-175) - Springer (Vol. 174)

10.) BEN ABDEL OUAHAB IKRAM, BOUDHIR ANOUAR ABDELHAKIM, ASTITO ABDELALI, BASSAM ZAFAR, BOUHORMA MOHAMMED, 2019 "Deep Learning architecture for temperature forecasting in an IoT LoRa based system" - ResearchGate

11.) Ganjar Alfiana, Muhammad Syafrudinb, Umar Farooqc, Muhammad Rifqi Ma'arifd, M. Alex Syaekhonib, Norma Latif Fitriyanib, Jaeho Leea, Jongtae Rheeb, 2020 “Improving efficiency of RFID-based traceability system for perishable food by utilizing IoT sensors and machine learning model” - Food Control (Elsevier)

12.) Nevena Golubovic, Rich Wolski, Chandra Krintz, Santa Barbara, Markus Mock, 2019 "Improving the Accuracy of Outdoor Temperature Prediction by IoT Devices" - IEE International Congress on IoT

13.) Hairong Yan, Michele Luvisotto, Kai Lu, Siyu Jiang, 2019 "Temperature prediction intelligent system based on BP neural network in wireless industrial IoT" - IEE Conference on Smart IoT

14.) Siddhartha Bhandari, Neil Bergmann, Raja Jurdak, Branislav Kusy,2017 "Time Series Analysis for Spatial Node Selection in Environment Monitoring Sensor Networks" - sensors

15.) Melanie Swan, 2012 "Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0 - JSAN

Journal of Sensor and Actuator Networks (MDPI)

16.) Radu Manuca, Robert Savit, 1996 "Stationarity and nonstationarity in time series analysis" - Physica D

Dataset:

The data consists of various food items, location information and their respective IOT temperature readings. Some of the food items have fields for both internal and ambient temperature data while others have only one reading. Manual data is to be provided by the Client in a separate file for cross comparison. The data would be recorded by IoT Sensors, subsequently get collected in a data file and transferred to the team along with the corresponding manual temperature data file.

Normally at least 1 full year of data from the source would be required for training the model. However due to the Client's constrains, the data would be collected from multiple locations for smaller durations and put together. Statistical checks and corrections are to be implemented in order to ensure that difference in location doesn't adversely affect model training. It is likely that Manual Data could either be a scanned document or of a csv /xlsx format. In either case it is the team's responsibility to see to it that all data in converted to the appropriate format.

It is important that the data is paired in such a manner or else one runs the risk of an invalid comparison. To achieve this, the containing folder's nomenclature should reflect the date of data acquisition, from both sensors and manual methods. The Client has assured that the entire data would be dispensed to the team in chunks as doing it all at once is not cost effective for them. The team expects the total data to be made available in fractions within a reasonable time period post the commencement of the project.

Sample Data:

Sample Data currently available and provided by the client consists of food temperatures of two to five food items (based on availability) for the period of 4th November 2019 to 6th February 2020.the food temperatures as recorded as 1min, 5 min, 10min 15min and 1hour aggregates and also temperatures recorded without any aggregation.

Tentative list of Algorithms:

- Dickey Fuller Test of Stationarity
- Auto-Regressive Time Series Model
- Moving Average Time Series Model
- ARIMA Time Series Model
- Linear Regression Model
- Lasso Linear Model
- KNN Regression Model
- Decision Tree Regression Model
- Random Forest Regression Model
- Support Vector Regression Model
- LSTM Model (Long short Term Memory)

References:

1.) Meiyu Wen, Dandan Che, Jean-Pierre Niyigena, Ruohan Li, Qingshan Jiang, 2019 "A Prediction Model of IoT Data Using Long Short-Term Memory Neural Network" – ISAICN

2.) M.K. Nallakaruppan, U. Senthil Kumaran, 2019 "IoT based Machine Learning Techniques for Climate Predictive Analysis" – IJRTE

3.) Ganesan Sangeetha, Muthuswamy Vijayalakshmi, "Role of Smart Sensors in Minimizing Food Deficit by Prediction of Shelf-Life in Agricultural Supply Chain" - Principles of Internet of Things (IoT) Ecosystem: Insight