West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City

PROJECT AHON: AN AGILE TRACKING SYSTEM USING MOBILE EMBEDDED SYSTEM FOR COVID-19 TRACKING

An Undergraduate Thesis

Presented to the Faculty of the

College of Information and Communications Technology

West Visayas State University

La Paz, Iloilo City

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Science in Computer Science

by

Horacio Jason A. Barba V

Joseph Anthony M. Duran

Jovan Earl P. Hinayan

Carlo Adrian R. Reyes

December 2021

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City

Approval Sheet

PROJECT AHON: AN AGILE TRACKING SYSTEM USING MOBILE EMBEDDED SYSTEM FOR COVID-19 TRACKING

An Undergraduate Thesis for the Degree

Bachelor of Science in Computer Science

by

Horacio Jason A. Barba V

Joseph Anthony M. Duran

Jovan Earl P. Hinayan

Carlo Adrian R. Reyes

Approved:

EVANS B. SANSOLIS, Ph.D
Adviser

Ma. Luche Sabayle Ph.D
Chair, Computer Science

MA. BETH S. CONCEPCION, DIT
Dean, CICT

June 2022

Acknowledgment

This research paper will not be possible without the guidance of various people who never failed to support them in every endeavor especially during the conduct of the study.

Horacio Jason A. Barba V

Joseph Anthony M. Duran

Jovan Earl P. Hinayan

Carlo Adrian R. Reyes

June 2022

Horacio Jason A. Barba V, Joseph Anthony Duran, Jovan Earl Hinayan, Carlo Adrian Reyes :" Project Ahon: An Agile Tracking System using Mobile Embedded Systems for COVID-19 Tracing ". Unpublished Undergraduate Thesis, Bachelor of Science in Information Technology, West Visayas State University, Iloilo City, Philippines, April 2022.

TABLE OF CONTENTS

List of Figures

Abstract

The growing complacency of the population during the ongoing COVID-19 pandemic has made the re-emergence of the virus a strong possibility. Currently, 2% of the global population are accounted as falling victims to the virus. This paper proposed a system that can safely and easily detect the emergence of a COVID-19 outbreak. This system employs a multivariate linear regression model which serves as a basis of predictions on the total population of the area, the number of vaccinated people in that area, the occupations the people have in that area, and the location itself. These factors were considered as triggers to an outbreak, such that it is considered as the main factors of the regression algorithm. Prediction of the upcoming outbreaks is computed by the system and is sent through a web application for early warning purposes and information dissemination. Thus, this tool is hoped to mitigate the widespread transmission of the COVID-19 pandemic.

List of Appendices

CHAPTER 1 INTRODUCTION TO THE STUDY

Background of the Study and Theoretical Framework

The COVID-19 virus was first started in China in late 2019 and started to spread around the world by the start of 2020, causing casualties worldwide and health systems to crash. With the US closing its borders in April of 2020 and the other countries following. Movement was then restricted in many countries, putting pressure on the populations' education, health, and economy. Radil (2020)

The outbreak also had a more severe effect on the affected schools that did not have a prepared plan for an online learning platform. Schools and universities had a weak operating performance at the start of 2020 and it was difficult for these universities and colleges to adapt with the academic challenges that coronavirus provided. UNESCO (2020) In the Philippines, classes in all levels were suspended starting march of 2020, medical schools had to suspend face-to-face learning classes, and medical students were forced to be removed from hospitals, clinics, and emergency departments, and had to settle with remote learning instead Hapal (2021). Of all the affected

countries, only Canada, the UK, and the US had the capacity to arrange an online learning program that is still effective for their students. Ali (2020)

There are similarities with how the COVID-19 pandemic has affected the industries now with how other disasters have affected the industries then, back in the events of the 2007-2008 recession, many thought that the impacts would only be localized on the US, but it ultimately ended up affecting the global financial system.de Lusignan (2020) The sudden economic disruption caused by COVID-19 is not only disruptive in one area, but also have accompanying effects that later cause implications in other areas just like how disasters would. And with the current numbers, trends, and industries that it is affecting, it is indicating that the COVID-19 pandemic will still remain as a disaster for the foreseeable future. Phillips (2021)

As an answer to the COVID-19 pandemic, some mitigation strategies have been developed to combat the spread of the virus, since at the start of the pandemic no specific drugs or vaccines were available, the people had to rely on community driven interventions to mitigate the spread of

the virus. Mass gatherings were the first thing to be restricted. Transmissions through respiratory droplets are most common and likely during these events, since population density contributes to a high rate of transmission. Rashid (2008)

Social distancing measures were the next mitigation strategy to be observed, by decreasing the frequency and contact between, population density is then reduced, which leads to ultimately leading to reduced transmission counts. In view of this, schools have been the first to shut down as they are the most densely populated places compared to offices or homes. Russell (2013)

To combat the risk of infections through COVID-19, researchers have experimented with the idea of a digital outbreak detection system as a method for identification and determination of possible contacts. Since manual contact tracing systems are very time consuming and require human intervention, which is avoided during pandemic outbreaks and also accompanied by human inefficiencies such as poor human recollection, results in a contact tracing system that will be incomplete and inefficient. Altmann (2020)

Although digital contact tracing does allow for more efficient and reliable methods in terms of scalability and reach, there are still some factors that allow the method to be debatable. Digital Contact tracing has issues concerning privacy and security, users and governments have access to data stored that should be stored anonymously, but is still susceptible to exploitation from these handlers[9]. User data leakage is still one of the most concerning events that might happen from these systems.

The outbreak of the COVID-19 virus started in China in late 2019 and the impact of the outbreak has spread and been felt throughout the world since the outbreak started. It has caused massive economical and financial damages in many countries, global slowdowns in terms of production, and disruptions in the education and tourism fields.McKibbin (2021)

The first suspected case was investigated on January 22, and by March 1, 633 cases were reported.Amit (2021) The pandemic has also impacted the country in different fields, psychologically, one-fourth of respondents from a study by

Tee reported that they had moderate-to-severe anxiety and one-sixth of them had moderate-to-severe depression and psychological impact.Tee (2020) The Philippines has also suffered in terms of its economic and finance sectors, with a decrease of 1.5% in the country's GDP due to the loss of an active tourism sector in the country due to the pandemic. Vlasov (2021)

In a research from James etal. (2020) He states that quarantine and isolation are crucial components for a successful contact tracing system. Since within 4 days, 80% of cases need to be immediately quarantined or isolated, this makes it so that the contact tracing methods will be reliable and data-confident. According to a new research by Juneau (2020), He builds on this definition, stating that effective contract tracing is associated with better control of COVID-19, with efficacy depending on how fast the contact is traced and quarantined, and how the quarantines are at preventing further transmissions. This helps with the study being both a contact tracing application and a risk awareness system, dissuading

possible visitors to an area and preventing potential transmission.

Objectives of the Study

The study aims to primarily develop a system that can detect an outbreak through multiple factors using the multivariate linear regression algorithm, collect data through satellite machines and process these data into easily readable information, then for these information to be presented in a clear and concise manner in a web application.

Specifically, it aims to:

1. Make a deployable automated contact tracing system able to ping its location to a central database and send all necessary information to it.
2. Developing a central database that can communicate with a deployed contact tracing systems
3. To develop a machine-learning based algorithm that can detect the presence of an outbreak through multiple factors.

Significance of the Study

This research will provide new insights into the technological benefits of creating a contact tracing system specifically utilizing a multivariate linear regression algorithm.

Specifically, the results of this study will be of great help to the following:

Students - The results of this study will benefit students that will have curriculums in Information Technology or Computer Science, promoting the use of Information Technology and Computer Science concepts into real life applications.

Local Government - They are the main beneficiary of this study, providing a contact tracing system with an algorithm that allows for the detection and risk assessment of a certain area.

Epidemiologists - Epidemiologists may use the analysis presented in this study to promote better pandemic coping methods.

Computer Scientists - The findings and the approach of this study helps promote better understanding of the different algorithms available and the different uses of each.

Proponents of the Study - This refers to the students conducting the study. They will find self-fulfillment and gain learnings and skills from this study. This study will help and inspire them to be more innovative and creative in their future endeavors.

Future Researchers - Future researchers will benefit greatly from this research as this will serve as their stepping-stone towards the study of other contact tracing and pandemic detection systems.

Definition of Terms

For better understanding, the following terms were defined conceptually and operationally:

Agile Development(Conceptual) - any development process that is aligned with the concepts of the Agile Manifesto

Agile Development(Operational) - refers to the chosen development process for the study.

Centralized database(Conceptual) - refers to the database that is stored, and maintained locally in a single location.

Centralized database(Operational) - This is the database where all the information gathered by the deployable devices are consolidated and processed.

COVID-19(Conceptual) - refers to the illness caused by a novel coronavirus.

COVID-19(Operational) - is the main source of an outbreak that the study is detecting.

Contact tracing system(Conceptual) - refers to the system responsible for identifying, assessing, and managing people who may have been exposed to a disease.

Contact tracing system(Operational) - is a system that allows the study to track previously logged individuals.

Database(Conceptual) - refers to the collection of data retrieved from different sources.

Database(Operational) - is where all of the gathered data are consolidated and processed for interpretation.

Embedded devices(Conceptual) - refers to the computer devices meant for one or very few specific purposes.

Embedded devices(Operational) - refers to the small and lightweight devices that allow for deployable devices to function.

Infected People(Conceptual) - refers to people that have contracted the COVID-19 infection

Infected People(Operational) - refers to the number of infected people that scanner or database has tracked in an area

Location(Conceptual) - refers to a particular place or position

Location(Operational) - refers to the geographical position the deployable devices are reporting.

Machine-learning Algorithm(Conceptual) - refers to the mathematical model mapping methods used to learn or uncover underlying patterns in the data.

Machine-learning Algorithm(Operational) - refers to the system that allows the system to analyze data on its own without human input.

Mobile Computing(Conceptual) - refers to the variety of devices that allow people to access data and information from wherever they are.

Mobile Computing(Operational) - refers to data computing being done remotely and outside of the main server.

Multivariate linear Regression Algorithm(Conceptual) - is a supervised machine learning algorithm involving multiple data variables for analysis.

Multivariate linear Regression Algorithm(Operational) - refers to the algorithm being used by the study.

Outbreak(Conceptual) - the sudden or violent start of something unwelcome, such as war, disease, etc.

Outbreak(Operational) - refers to the rise of COVID-19 cases in an area

Population(Conceptual) - all the inhabitants of a

particular town, area, or country.

Population - refers to the number of people in an area

RFID(Conceptual) - is a wireless system made up of two components, a tag and a reader

RFID(Operational) - refers to the main identification system used by the study

Satellite Machines(Conceptual) - refers to machines that are deployed from the central server

Satellite Machines(Operational) - refers to the machines being deployed to collect data

Vaccinated People(Conceptual) - refers to the people that have been vaccinated

Vaccinated People(Operational) - refers to the number of vaccinated people the scanner or database has tracked in an area

Web(Conceptual) - a complex system of interconnected elements

Web(Operational) - refers to the system of connected resources the study is using for the study.

Delimitation of the Study

This study focused on developing a contract tracing system that allows for data collection, interpretation, and risk assessment using different variables.

This study's gathered data are limited to theoretical data from publicly available sources, therefore the results are also theoretical.

Among the variables used in the study are variable population, number of vaccinated people in an area, number of health workers in an area, and the location.

As for the population, this was only limited to populations that can be gathered from authenticated publicly available data.

For the number of vaccinated people in an area, this was limited to data given by DOH and/or the LGUs pertaining to the percentage of vaccinated people per area.

And the data regarding the number of health workers were assumed, and limited to the city of Iloilo. The data was publicly given by the Statista Research Department(2021).

Databases, statistical tables, and algorithms were adopted and modified by the researchers to fit into the specific needs of testing our hypothesis.

This study has potential limitations. The effect estimates in the tracing models are subjected to generational and mechanical faults.

They are therefore subject to biases and confounding that may have influenced our model estimates.

Chapter 2 Review of RELATED STUDIES

Review of Existing and Related Studies

In lieu of the growing pandemic in the country, the country has started research on the different outbreak detection systems that the country can implement.

The Philippine National Government has started using an app called StaySafe.ph, a contact tracing app developed and powered by the country's own PLDT-backed Multisys Technologies Corporation, it helps contact tracing efforts by generating heatmaps from user-generated reports. Lacsa (2021)

The Cebuano government has experimented with the use of their own WeTrace contact tracing app developed by a Cebuano scholar from the Department of Science and Technology, which was required to be installed by their governor Gwendolyn Garcia. It utilizes a Bluetooth Low Energy (BTE) communication channel and asymmetric cryptography to encode the message to an intended destination, therefore preserving privacy. Franco (2021)

The government of Butuan City and Zamboanga City has partnered with private organizations CMI Tech, Spring Valley, and telecommunications businesses Globe and Smart to develop the Citizen's Logistics and Early Assessment Reporting Tool (Clear), an online app that enables users to track their state of health on a daily basis and enable contract tracing when necessary, compared to the other monitoring apps, this app relies its risk assessment on the user’s own judgment instead of through an algorithm. The app has a platform and database of information for contact tracing, RED zones identification, and has access to mass testing records when in collaboration with Local Government Units. Technology for COVID-19 Surveillance. (2020)

A study by Jamdade examined the feasibility of an outbreak detection system that is based on the Vector Autoregression Moving-Average with Exogenous Regressors time series algorithm or VARMAX algorithm and its uses as a way to forecast possible transmission movements that could be done in the country, while also discussing current concerns, and challenges that the country may face post-pandemic. Jamdade (2021)

Apple and Google are collaborating for a contact tracing project as well, currently in development, the project by the 2 biggest mobile-phone platforms will be an interoperable system that would allow bluetooth-enabled tracing efforts to work between the two different operating systems, while still maintaining security and privacy protocols. Sharon (2020)

Studies have experimented with the use of a multivariate regression algorithm for outbreak detection systems; multivariate linear regression systems allow for short-term and long-term methods for determining the correlation of each factor that may arise when trying to detect a pandemic outbreak. A study by Bhaskhar suggests using a linear regression algorithm to forecast distribution and growth rate of pandemic cases in certain locations, which determines the effectiveness of the controlling measures that are being done. Bhaskhar (2020)

Digital outbreak prediction for many countries is now top priority to combat the effects of the pandemic.

Although manual outbreak prediction is being used, it is still too inefficient, making it difficult to reliably detect outbreaks before they start to happen, whereas digital outbreak detection systems offer scalability and reliability from a digital approach. Ferretti (2020)

The study aims to primarily develop a system that can detect an outbreak through multiple factors using the multivariate linear regression algorithm, collect data through satellite machines and process these data into easily readable information, then for this information to be presented in a web application.

The consequent objectives of the study is to analyze the relationships of the different factors that may contribute to a possible outbreak of infection, compare these factors and show which factors are the primary causes of the outbreak.

Chapter 3 RESEARCH DESIGN AND METHODOLOGY

Description of the Proposed Study

The study aims to primarily develop a system that can detect an outbreak through multiple factors using the multivariate linear regression algorithm, collect data through deployable machines, process these data into easily readable information, then display this information into an easily readable format in a web application.

An Agile development method was used for this study to ensure the timeliness and accuracy of the results. This development cycle allows for constant updates for the system and allows frequent changes to fix any problem that may arise. This study is meant to facilitate the proper dissemination of data to authorities, as this is the case, with the situation still developing and better information being released daily, an Agile development was necessary. The agile methodology was conceived primarily in the collecting, testing, and authentication of various dynamic variables which may have worked with this study's ML model in various stages of development.

Methods and Proposed Enhancements

The multivariate linear regression algorithm was chosen as this study's ML model because it is a statistical technique that can be shaped through the use of different variables to show a distinct result that can guide authorities in decision making. The researchers finalized on four variables namely the population, number of vaccinated, high risk individuals, using these variables this study can ascertain or predict the number of infected individuals.

The dataset used for this study is a dataset based on the different factors that could contribute to the likelihood that an outbreak could happen within a certain area. A study by Phipps (2020) on the robust estimates of a global infection rate was used as the base infection rate for our study.

The first factor considered for the dataset is the population of an area. According to a study by Kadi (2020), their findings suggest that COVID-19 infection rates are moderately correlated with higher populations, and the death cases are higher with larger populations.

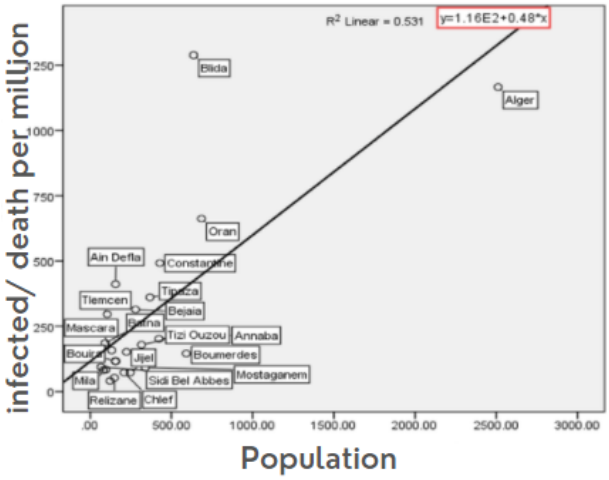


Figure 1 : Relationship between Infection Rate and Population. Source: Kadi (2020)

The second factor determined was the number of high-risk occupations in an area. Healthcare workers(HCWs) at the frontline provide a substantial risk of infection during the COVID-19 outbreak. According to a study by Zheng (2020), HCWs infection rates are significantly higher than that of non-HCWs.

Parameters	Total, N	Confirmed Cases, n (%)	CIR, %
Sex			
Female	84 078	1776 (72.28)	2.11
Male	33 022	681 (27.72)	2.06
Types of occupation			
Nurse	57 700	1279 (52.06)	2.22
Doctor	43 100	826 (33.62)	1.92
Medical staff	16 300	352 (14.33)	2.16

Figure 2 : Relationship between Infection Rate and High-risk Occupations. Source: Zheng (2020)

The third factor determined was the number of vaccinated people in an area. Vaccinated People reduces the risk of an infection and accelerates viral clearance. According to a study by Levine-Tiefenbrun (2021), vaccinated people tend to be less infectious than unvaccinated people.

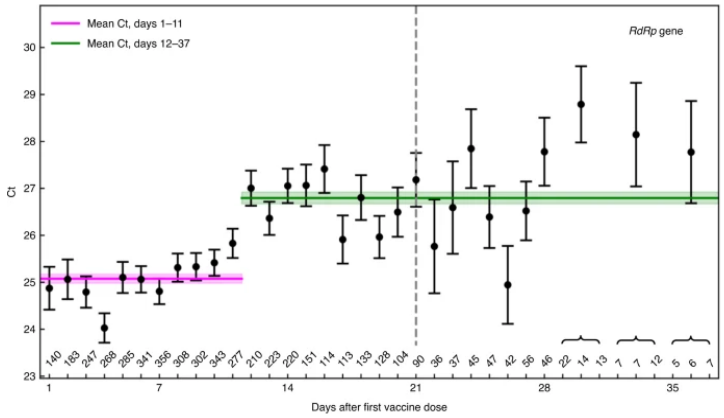


Figure 3 : Relationship between Infection Rate and Vaccinated Individuals. Source: Levine-Tiefenbrun (2021)

And the last factor determined was the location. Due to the COVID-19's spread among the world, the virus has affected each location differently. A study conducted by Chang (2021) states that COVID-19 fosters better transmission dynamics due to some underlying factors in a location like wind speed and area density.

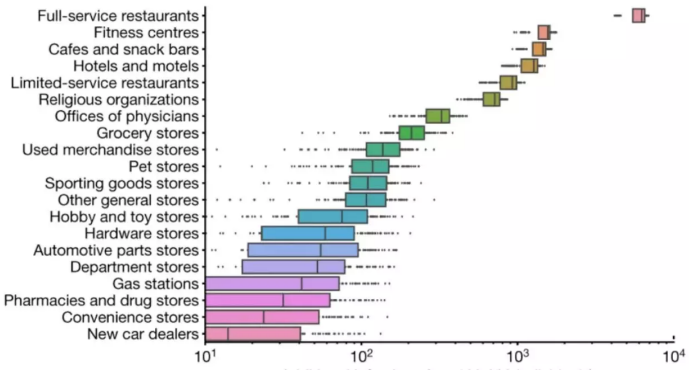


Figure 4 : Relationship between Infection Rate and Location. Source: Chang (2021)

The dataset suits the purpose of showing how a multivariate linear regression algorithm could detect outbreaks and also show the relationship between the different accompanying variables and how they instigate an outbreak. Missing values in particular the "No. of Infect" have been subjected to a mean value interpolated using past values.

This study is a set of software technologies integrated to detect and limit possible outbreaks of the COVID-19 infection by performing contact tracing. As such, the system is intended to support:

- the citizens, by tracing potentially high risk individuals that enter an area;

- health authorities, in managing the population and their status to rapidly check and get in touch with infected, possibly infected, and quarantined patients;
- researchers and experts, by collecting anonymized data if needed, and offering a data source to precisely monitor the trends in the movement of the population.

To achieve such goals, this study is composed of an RFID, to anonymously trace all the contacts that enter a particular space, checking for their temperature and using a machine-learning algorithm to detect if there are abnormal patterns developing in a certain area.

This study's infrastructure serves as a data backend. The authorized personnel can access such data if the study's users are pre-registered to the platform, the authorized staff can update the status of users. Thus, through a web frontend, the authorities can use the data on the server to actively supervise the population, update their status (Vaccine, High-Risk Occupations), and monitor the potential spread of the infection outbreak.

Components and Design

System Architecture

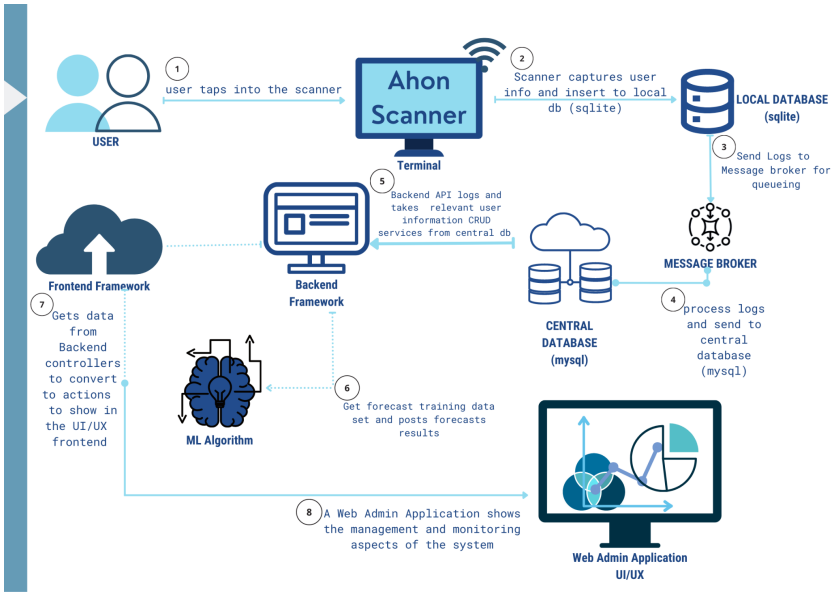


Figure 5 : System Architecture of the Proposed System

The outbreak prediction system includes a web application where the users can check for a complete report on the data collected and visualizations on the movement of transmissions, recorded cases in an area, and overall well-being in a certain location. The following figures show the general system overview of the proposed work and the components and functionalities are described further in the succeeding sections. With the given system architecture, the LGU or relevant authorities will be in

control of the data that the system collects and will have the monitoring ability over these locations.

Figure 6 shows the user dashboard of the web-based application for the system. Where the user can check for the users that have pinged in that area currently, the number of infected people in relation to the general people, and the percentage of infected people.

Figure 7 shows the dashboard people will see when tapping their details into the system.

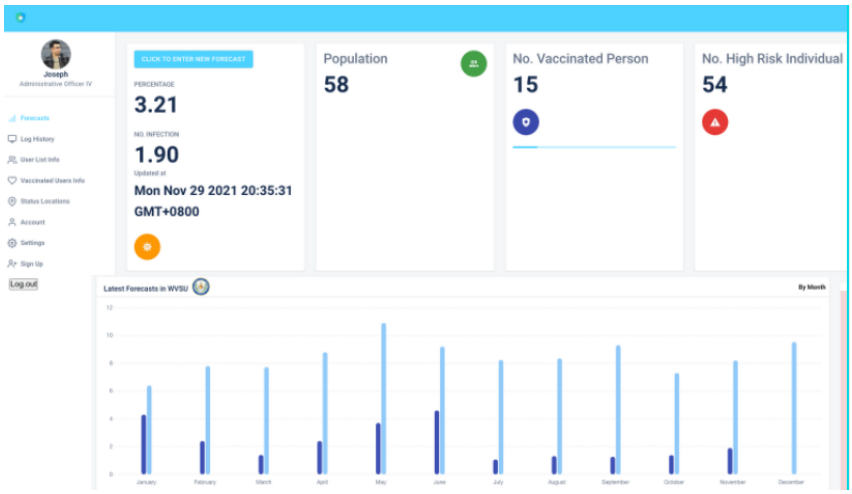


Figure 6 : Web-based dashboards for clients

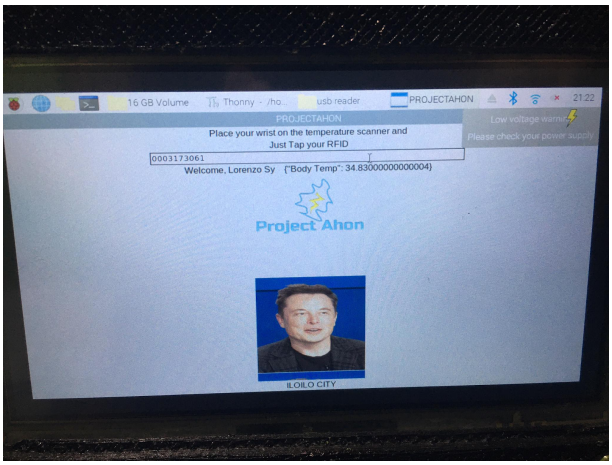


Figure 7 : Dashboard for the Deployable Device

Database Design

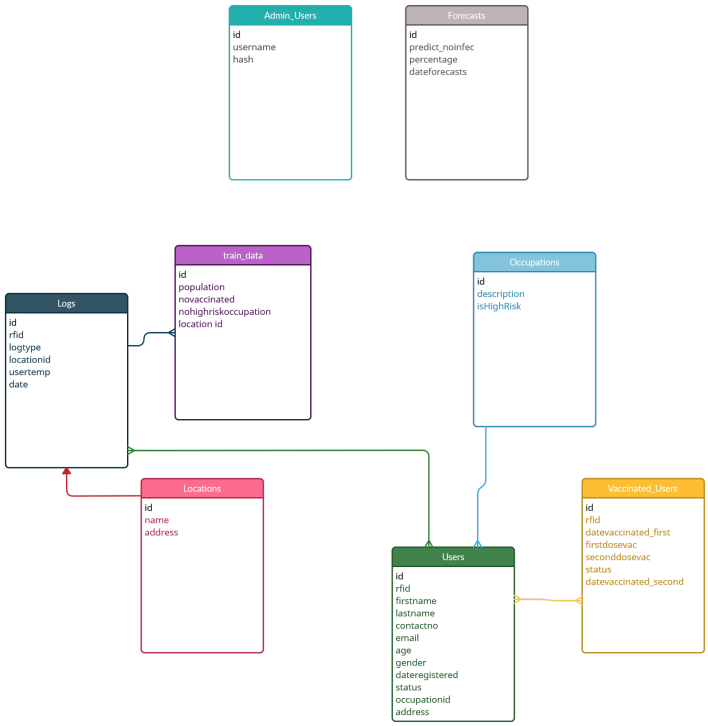


Figure 8 : Database Design for the Proposed System

Procedural and Object-Oriented Design

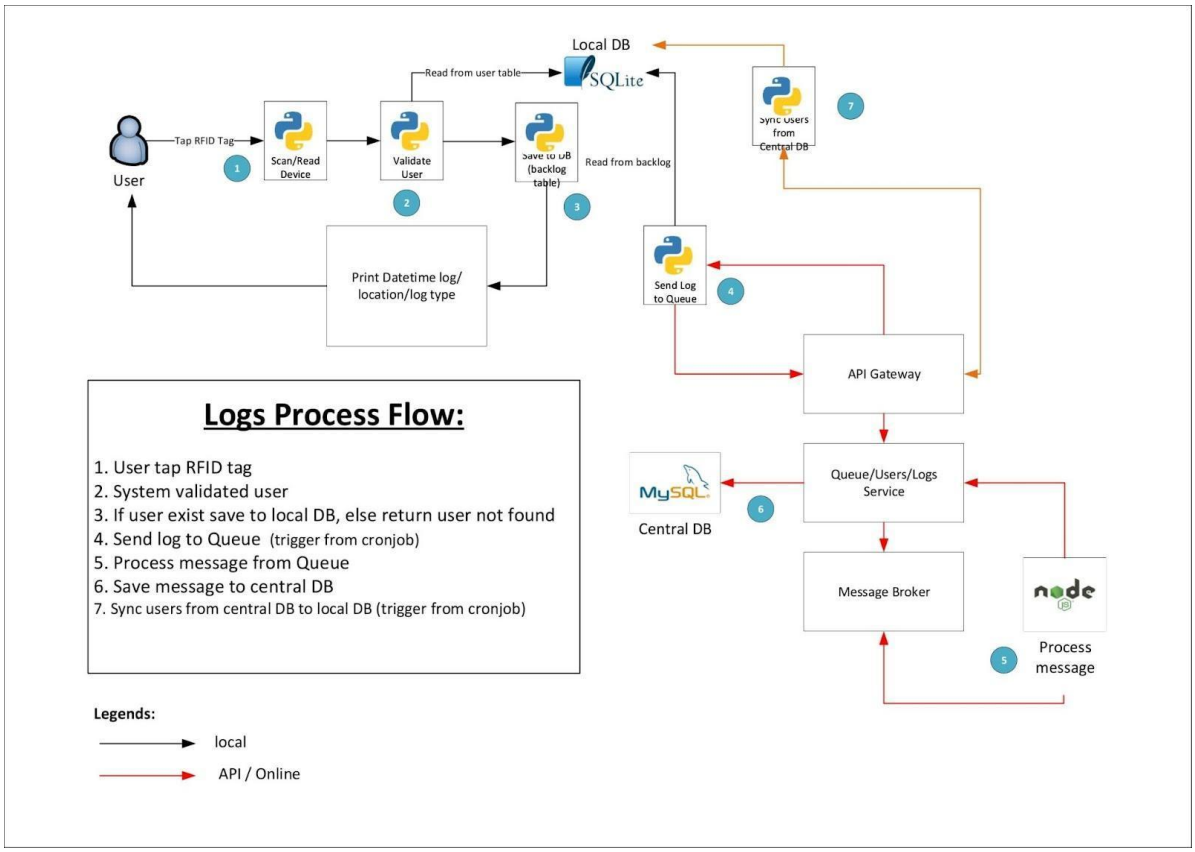


Figure 9 : Procedural Flow for the Proposed System

Algorithm Data Flow Diagram

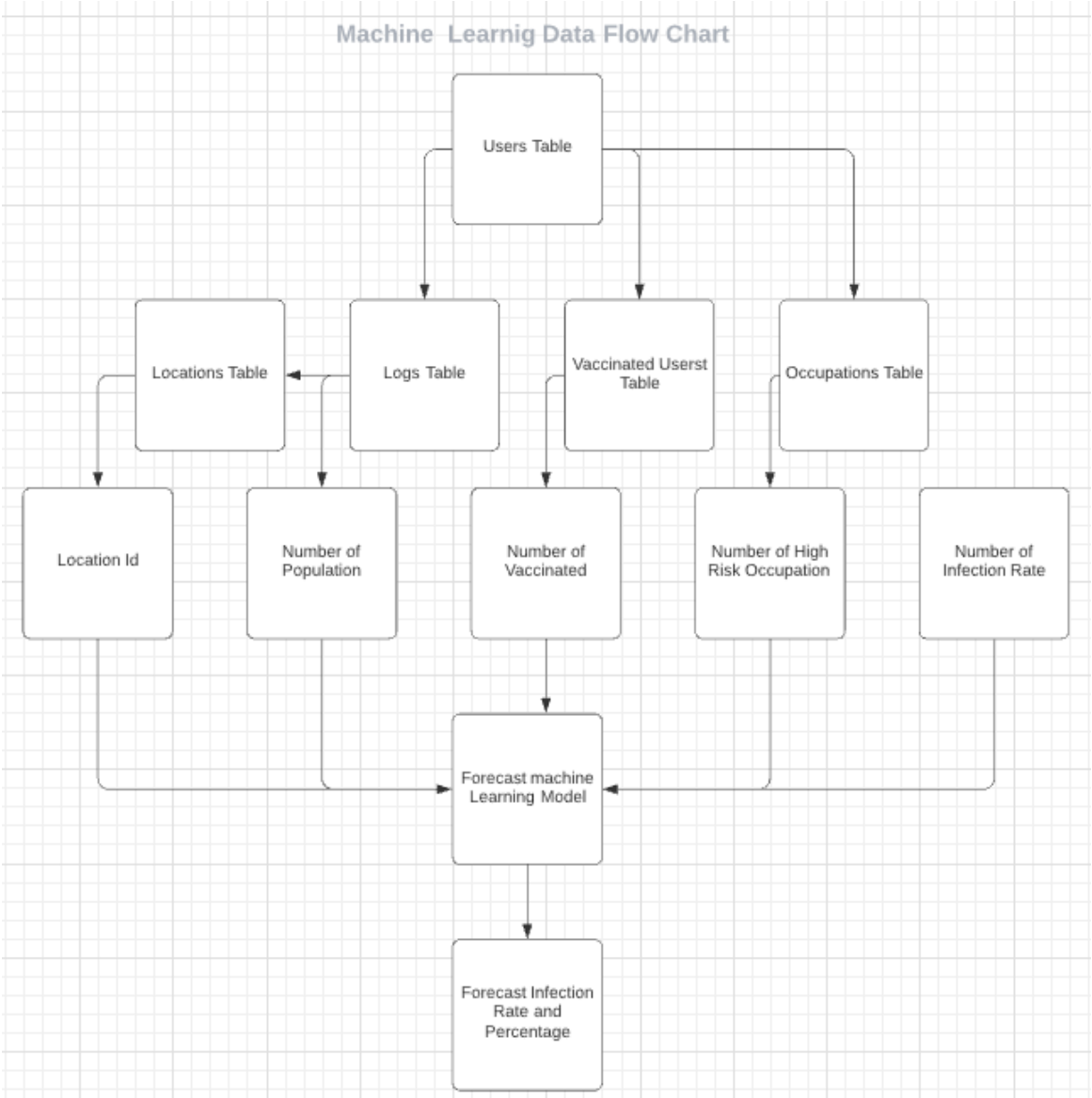


Figure 10: Algorithm Data Flow Diagram of the Proposed Study

Application Data Flow Diagram

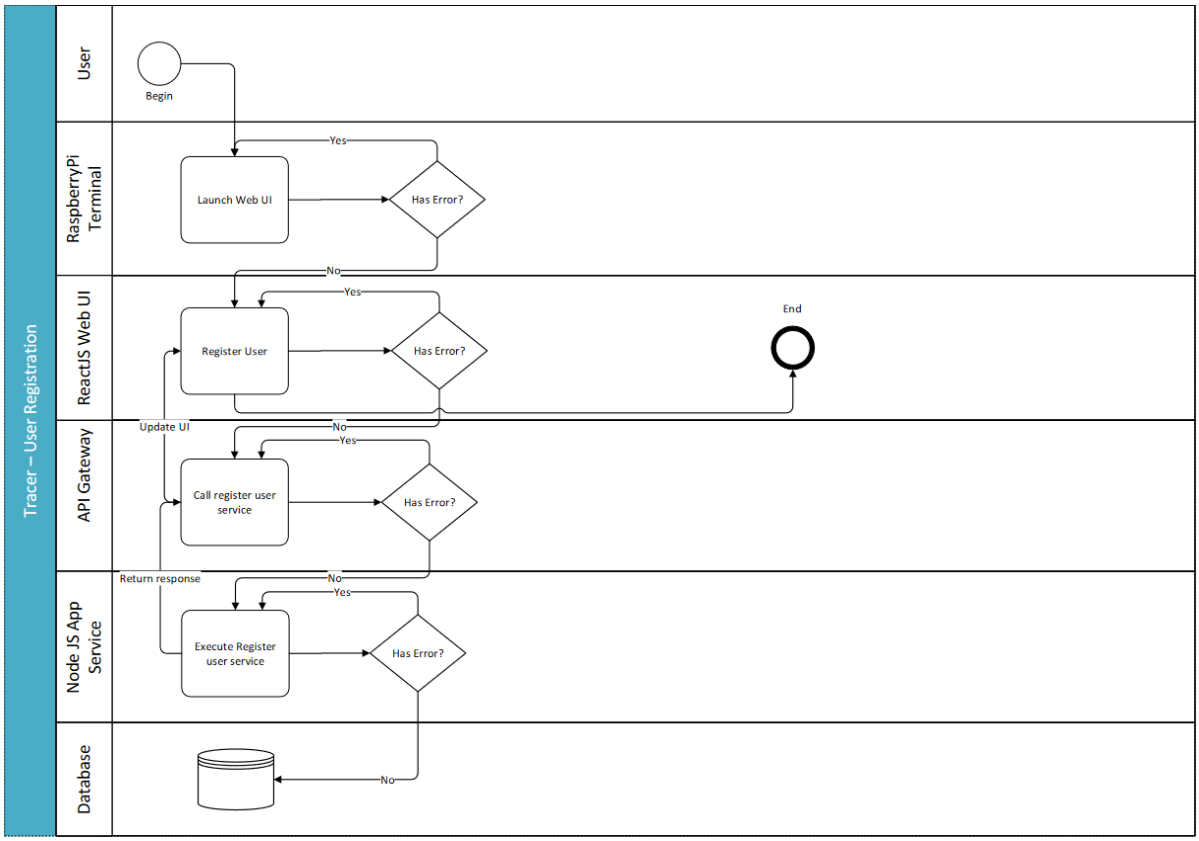


Figure 11: Application Data Flow Diagram of the Proposed Study

Methodology

System Development Life Cycle

The Systems development life cycle used in this research included analysis, planning, system design, system development, testing, and evaluation. By following this development life cycle, the study is able to test the capabilities of the system while still allowing for major developments in how the system behaves.

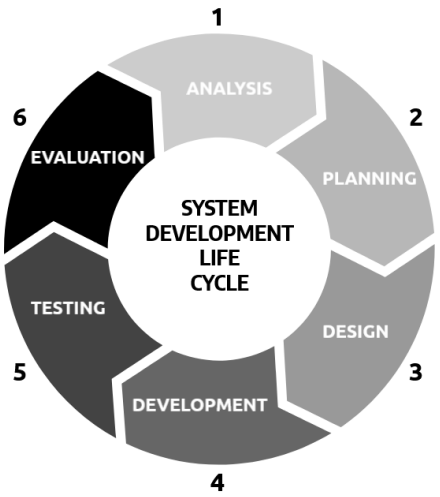


Figure 12: System Development Cycle of the Proposed Study

CHAPTER 4 RESULTS AND DISCUSSION

Implementation

This study aimed to propose an outbreak prediction system and an accompanying web application. The three main components of the system are the remote monitoring device, the database, and the web application. The monitoring device is composed of a raspberry pi, with a temperature sensor and an rfid scanner, all enclosed in a 3d printed casing. The RFID scanner is responsible for scanning the RFID's that are registered for tracking, the temperature scanner to scan for real time temperatures of every entry, and the wireless raspberry pi system to communicate with the server.

The population was based on the general population of Iloilo and on various conditions which may appear based on time and date.

The number of vaccinated people in a certain area indicates how population density could affect the risk of a certain area. This is based on the general percentage of vaccinated people which was then calculated to fit Iloilo's population.

The number of high risk populations was also considered, this was based on a study that showed how many

health workers or frontline workers there are in iloilo which was then scaled down to fit our location.

The location determines where the algorithm's predictions will focus, separating every prediction score based on its population.

This data was then processed and sent into a web application that could display the processed data.

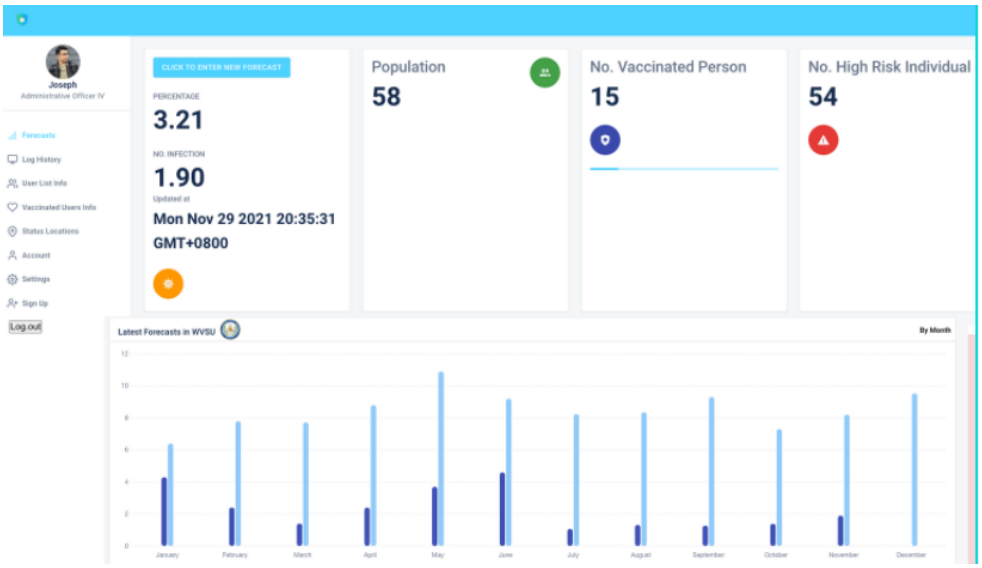


Figure 13 : Web-based Dashboard Application for the System

The application will allow the user to check all basic data collected and also check the users logged and filter by location.

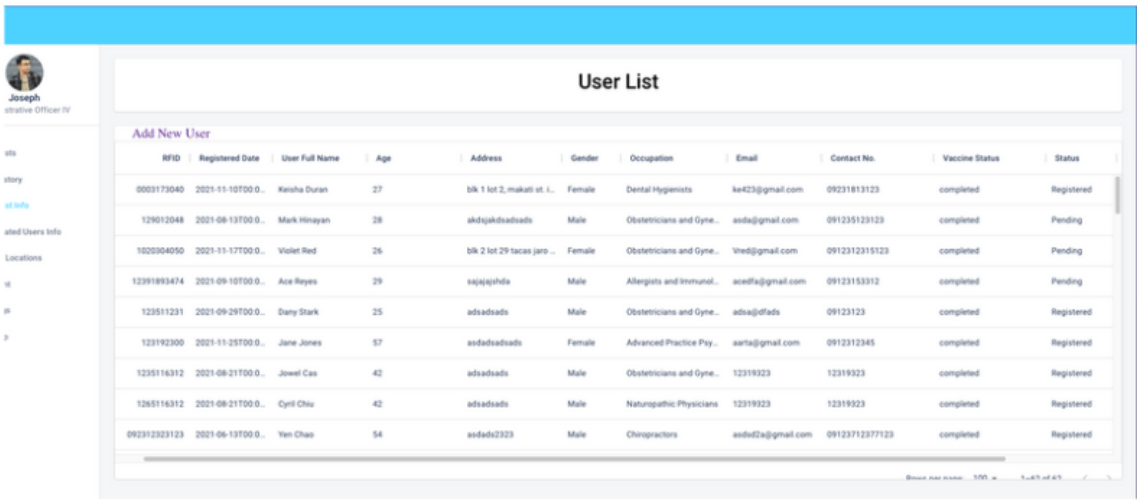


Figure 14 : User List dashboard where users can filter by different characteristics

To evaluate the quality, performance, and usability characteristics of the system, a 5-point rating scale based on the ISO/IEC 25010 standards was used. Evaluation was done remotely via Zoom and in person in a cafe.

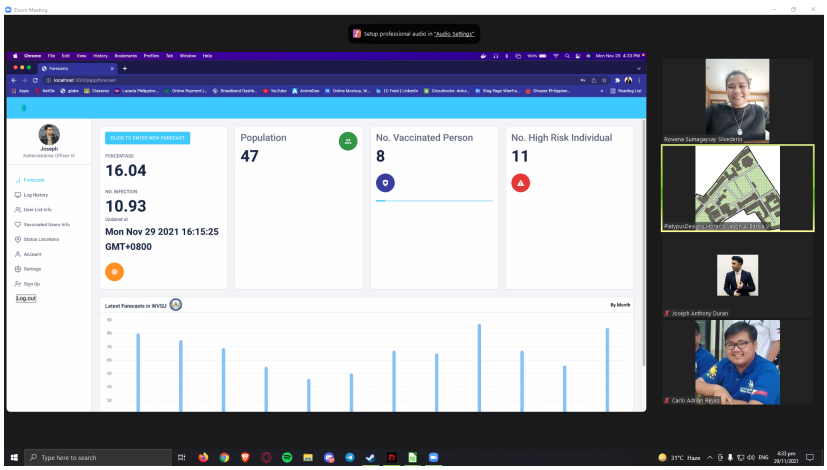


Figure 15 : Remote System Evaluation via Zoom

Results Interpretation and Analysis

The proposed system upon conducting a comparative analysis between other contact tracing applications or methods like Manual contact tracing system, Staysafe.Ph and WeTrace were different in their characteristics.

For the purposes of this paper, Staysafe.ph was used as the main differential as it is widely used within Iloilo.

The proposed system differed primarily in how they gathered information. The study utilizes an entire system for its identification while Staysafe.ph uses a QR code for it and relies on the population itself to log themselves in.

The proposed system also differs in how data is authenticated, the proposed system’s data is first validated by relevant authorities before being entered into the database. Staysafe.ph’s information on the other hand is entered by its own users at their discretion without any proper authentication.

The final and primary difference between the proposed system and the Staysafe.ph is that the proposed system

utilizes a machine learning algorithm as a byproduct of our contact tracing.

For this study, the algorithm quickly became an integral part of the system, to test its reliability.

Figure 16 shows the predicted linear regression relationship between the population and predicted number of infected people.

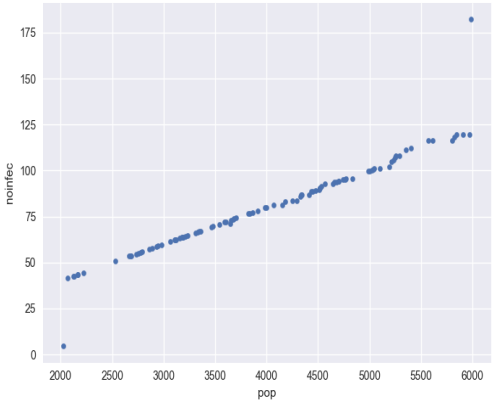


Figure 16 : Relationship between Infected Cases and Population

Figure 16 shows the relationship between the number of infected people, and the population, with the number of predicted infected people rising the larger population.

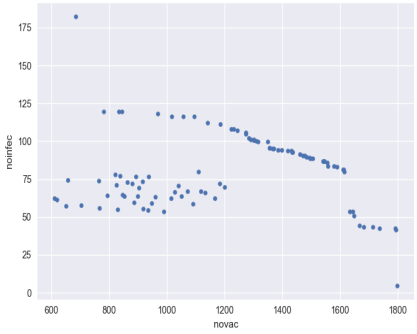


Figure 17 : Relationship between Infected Cases and Number of Vaccinated People

Figure 17 shows the relationship between the number of infected people, and the number of vaccinated people, with the number of predicted infected people decreasing the more vaccinated people there are.

Figure 18 : Relationship between Infected Cases and Number of High-risk Individuals

Figure 18 shows the relationship between the number of infected people, and the number of high risk individuals, with the number of predicted infected people rising with the number of high risk individuals. Figures 16, 17, and 18 show the summary of the data collected in the trials. These relationships found will be essential to the factors for the prediction model. The same model will be used to compute the regression equation to predict forthcoming risk.

Prediction model based on Regression Equation

The prediction model using the training data for the trials will be the input to the equation given below.

$$Y = aX + b$$

$$Y = a + b1X1 + b2X2 + b3X3 + b4X4$$

Equation 1 is a regression equation for the risk score. The variable Y is a predicted model, while A is the dependent variable coefficient and b1,b2,b3, and b4 are the model factors computed based on the input values. While x1 is the population parameter, x2 is the number of vaccinated people, x3 is the number of high risk individuals, and x4 is the location. This will forecast the probability of an imminent risk in the rise of infected people. With the input training data for the trials, the following model was computed.

Prediction Model

The prediction model for one dependent and four independent variables is presented in this Equation

$$Y = a + b1X1 + b2X2 + b3X3 + b4X4$$

With this equation, the results could be used by the users or clients to help the decision making when to classify a location as high risk or not, which also applies to the ongoing data that the system receives every update.

System Evaluation Results

Evaluation Metrics

To evaluate the accuracy of the regression model, the Mean Absolute Error(MAE) is calculated from the input and output data of the model. The MAE is calculated by calculating the absolute difference between the actual values and the predicted values, therefore giving the study a clear comparison between the actual values and how close the predicted values are.

Table 1 : First 10 Actual and Predicted Readings

Computation of Mean Absolute Error

The mean absolute error used the first 10 trials of both the actual data and the predicted data using this formula. Where Yi are the predicted values and Xi are the actual values, and n total number of data points in a sample.

$$\mathbf{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

First the mean for the predicted values are calculated

$$\overline{Yi} = \frac{(\sum y1 + y2 ... + y10)}{n}$$

$$\overline{Yi} = \frac{(81 + 81 + 79 + 69 + 78 + 89 + 88 + 54 + 58 + 59)}{10}$$

$$\overline{Yi} = 73.6$$

then the mean for the actual values are then calculated

$$\overline{Xi} = \frac{(\sum x1 + x2 + ... x10)}{n}$$

$$\overline{Xi} = \frac{(77 + 78 + 77 + 66 + 81 + 88 + 88 + 57 + 58 + 57)}{10}$$

$$\overline{Xi} = 72.7$$

the differences are then calculated between the absolute mean of the predicted values and the absolute mean of the actual values to solve for the Mean Absolute Error.

$$MAE = \left| \sum \overline{Yi} - \sum \overline{Xi} \right|$$

$$MAE = |73.6 - 72.7|$$

$$MAE = 0.9$$

This result implies that with a Mean Absolute Error of 0.9, the predicted values are on average ±0.9 away from the actual values.

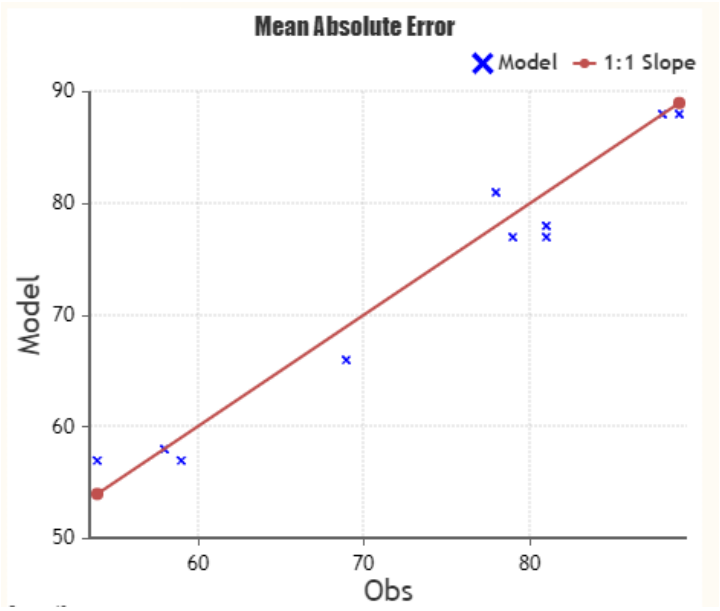


Figure 19 : Graph of Mean Absolute Error

To evaluate the system in terms of performance and usability, the system was evaluated by respondents rating the system in terms of functional suitability, performance efficiency, compatibility, usability, reliability, maintainability, and portability in accordance with ISO/IEC:25010 standards.

Figure 20 : Characteristics tested by ISO/IEC25010

Based on the evaluations done by respondents, the system garnered these averages:

Table 2 : Averages Scores of the Evaluation

Based on the results, this implies that the system performed with up to standards in terms of maintainability, portability, and performance efficiency, with the lowest

score on compatibility implying that the system may not be compatible with some setups.

CHAPTER 5 Summary, Conclusions, and Recommendations

Summary of the Proposed Study Design and Implementation

The study aimed to develop an outbreak detection system that could gather, analyze, interpet, and then display the data in a web application.

The outbreak prediction was done using a multivariate linear regression algorithm to predict the number of infections that may arise in a given day, this was calculated using a dataset made up of 4 independent variables and 1 dependent variable, with the independent variables being the number of population in an area, the number of vaccinated people within the area, the number of high-risk occupations in the area, and the location or area itself, with the dependent variable being the number of infections given that day.

Data Gathering is done through a deployable satellite machine that could operate wirelessly from the central database, this machine was made using a raspberry pi as its main computer, with an RFID scanner for identification and

temperature scanner for introductory data, the machine also comes with a 5.5-inch display for easy viewing.

Information gathered by the satellite machine was then transferred to the central database where it is managed, and stored.

The data is then processed and displayed in a web application for easy data dissemination.

Summary of Findings

Through analysis of the data and the graphs presented, the most significant variables that emerged from the algorithm are namely the number of high-risk individuals, and the location of an area all follow the same general trend; the number of infections increase as this number increases as well. Looking at the other variable number of vaccinated people, the trend of the regression outputs that the algorithm has made is trending downward, with the number of infected people decreasing as the number of vaccinated people in the area increases. This proves the relationships between location, population, and infected people as directly proportional, while the relationship between the number of vaccinated people to the number of infected people are inversely proportional.

Conclusions

With the covid 19 pandemic and the increasing number of cases in the country, outbreak prediction has been the top priority in the philippines, manual contact tracing for possible cases have always been slow and unreliable to be the main combating force against the pandemic and has always needed a change.

This paper, specifically the Algorithm and its forecasts due to uncontrollable circumstances stemming from the ongoing pandemic uses simulated data, and is therefore fully theoretical.

This paper developed an outbreak detection based on a regression algorithm that bases its forecasts from the population, the number of high-risk individuals, and the location. These factors were considered as triggers for an outbreak to occur, thus becoming the main factors of the regression algorithm. The regression model was computed with the assumption that high densities in population with a number of high-risk individuals in a certain location is the leading cause for an outbreak. This model is then used

to forecast likelihood of an outbreak using the data that is submitted to the server. A web application is then used for data dissemination. Machine learning has proven to be a very promising and reliable tool to help in crises like this. Scientists all over the world are experimenting with different machine learning based models to help end this pandemic.

Recommendations

Based on the finds and conclusions presented in this study, the following recommendations are suggested:

1. The researchers recommend that the algorithm be tested with data given by the government. Since testing done in this study was only done with publicly accessible data.
2. The researchers recommend adding a feature in the machine learning algorithm that allows the changing of the weight of the vaccine status, to allow for testing in scenarios where the vaccine is not as effective as it was.
3. The four factors determined by the researchers are population size, number of vaccinated people, number of high-risk individuals, and location. The researchers recommend to discover more variables to further the accuracy of the model.

References

Ahmar, A. S., & del Val, E. B. (2020). SutteARIMA: Short-term forecasting method, a case: Covid-19 and stock market in Spain. *Science of The Total Environment*, *729*, 138883. https://doi.org/10.1016/j.scitotenv.2020.138883

Ali, W. (2020). Online and Remote Learning in Higher Education Institutes: A Necessity in light of COVID-19 Pandemic. *Higher Education Studies*, *10*(3), 16. https://doi.org/10.5539/hes.v10n3p16

Altmann, S., Milsom, L., Zillessen, H., Blasone, R., Gerdon, F., Bach, R., Kreuter, F., Nosenzo, D., Toussaert, S., & Abeler, J. (2020). Acceptability of App-Based Contact Tracing for COVID-19: Cross-Country Survey Study. *JMIR MHealth and UHealth*, *8*(8), e19857. https://doi.org/10.2196/19857

Amit, A. M., Pepito, V. C., & Dayrit, M. (2021). Early response to COVID-19 in the Philippines. *Western Pacific Surveillance and Response Journal*, *12*(1), 56–60. https://doi.org/10.5365/wpsar.2020.11.1.014

Bengio, Y., Janda, R., Yu, Y. W., Ippolito, D., Jarvie, M., Pilat, D., Struck, B., Krastev, S., & Sharma, A. (2020). The need for privacy with public digital contact tracing during the COVID-19 pandemic. *The Lancet Digital Health*, *2*(7), e342–e344. https://doi.org/10.1016/s2589-7500(20)30133-3

Bhaskar, A., Ponnuraja, C., Srinivasan, R., & Padmanaban, S. (2020). Distribution and growth rate of COVID-19 outbreak in Tamil Nadu: A log-linear regression approach. *Indian Journal of Public Health*, *64*(6), 188. https://doi.org/10.4103/ijph.ijph_502_20 de Lusignan, S., Lopez Bernal, J., Zambon, M.,

Akinyemi, O., Amirthalingam, G., Andrews, N., Borrow, R., Byford, R., Charlett, A., Dabrera, G., Ellis, J., Elliot, A. J., Feher, M., Ferreira, F., Krajenbrink, E., Leach, J., Linley, E., Liyanage, H., Okusi, C., . . . Hobbs, R. (2020). Emergence of a Novel Coronavirus (COVID-19): Protocol for Extending Surveillance Used by the Royal College of General Practitioners Research

and Surveillance Centre and Public Health England. *JMIR Public Health and Surveillance*, *6*(2), e18606. https://doi.org/10.2196/18606

Education: From disruption to recovery. (2021, October 5). UNESCO. https://en.unesco.org/covid19/educationresponse

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D., & Fraser, C. (2020). Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, *368*(6491). https://doi.org/10.1126/science.abb6936

Franco, M., Rodrigues, B., Killer, C., Scheid, E. J., de Carli, A., Gassmann, A., Schonbachler, D., & Stiller, B. (2021). WeTrace: A privacy-preserving tracing approach. *Journal of Communications and Networks*, 1–16. https://doi.org/10.23919/jcn.2021.000021

Gautam Jamdade, P., & Gautamrao Jamdade, S. (2021). Modeling and prediction of COVID-19 spread in the Philippines by October 13, 2020, by using the VARMAX time series method with preventive measures. *Results in Physics*, *20*, 103694. https://doi.org/10.1016/j.rinp.2020.103694

Hapal, K. (2021). The Philippines' COVID-19 Response. *Journal of Current Southeast Asian Affairs*, *40*(2), 224-244. https://doi.org/10.1177/1868103421994261

James, A., Plank, M., Hendy, S., Binny, R., Lustig, A., Steyn, N., Nesdale, A., & Verrall, A. (2020). Successful contact tracing systems for COVID-19 rely on effective quarantine and isolation. *Successful Contact Tracing Systems for COVID-19 Rely on Effective Quarantine and Isolation*. Published. https://doi.org/10.1101/2020.06.10.20125013

Juneau, C. E., Briand, A. S., Pueyo, T., Collazzo, P., & Potvin, L. (2020). Effective Contact Tracing for COVID-19: A Systematic Review. *Effective Contact*

Tracing for COVID-19: A Systematic Review. Published. https://doi.org/10.1101/2020.07.23.20160234

Kadi, N., & Khelfaoui, M. (2020). Population density, a factor in the spread of COVID-19 in Algeria: statistic study. *Bulletin of the National Research Centre*, *44*(1). https://doi.org/10.1186/s42269-020-00393-x

Lacsa, J. E. M. (2021). StaySafe.ph App: public health tool or creeping surveillance app? *Journal of Public Health*. Published. https://doi.org/10.1093/pubmed/fdab267

McKibbin, W., & Fernando, R. (2021). The Global Macroeconomic Impacts of COVID-19: Seven Scenarios. *Asian Economic Papers*, *20*(2), 1–30. https://doi.org/10.1162/asep_a_00796

Phillips, S., & Williams, M. A. (2021). Confronting Our Next National Health Disaster — Long-Haul Covid.

New England Journal of Medicine, *385*(7), 577–579. https://doi.org/10.1056/nejmp2109285

Phipps, S. J., Grafton, R. Q., & Kompas, T. (2020). Robust estimates of the true (population) infection rate for COVID-19: a backcasting approach. *Royal Society Open Science*, 7(11), 200909. https://doi.org/10.1098/rsos.200909

Radil, S. M., Castan Pinos, J., & Ptak, T. (2020). Borders resurgent: towards a post-Covid-19 global border regime? *Space and Polity*, *25*(1), 132–140. https://doi.org/10.1080/13562576.2020.1773254

Rashid, H., Haworth, E., Shafi, S., Memish, Z. A., & Booy, R. (2008). Pandemic influenza: mass gatherings and mass infection. *The Lancet Infectious Diseases*, *8*(9), 526–527. https://doi.org/10.1016/s1473-3099(08)70186-5

Russell, E. S., Zheteyeva, Y., Gao, H., Shi, J., Rainey, J. J., Thoroughman, D., & Uzicanin, A. (2016).

Reactive School Closure During Increased Influenza-Like Illness (ILI) Activity in Western Kentucky, 2013: A Field Evaluation of Effect on ILI Incidence and Economic and Social Consequences for Families. *Open Forum Infectious Diseases*, *3*(3). https://doi.org/10.1093/ofid/ofw113

Sharon, T. (2020). Blind-sided by privacy? Digital contact tracing, the Apple/Google API and big tech's newfound role as global health policy makers. *Ethics and Information Technology*. Published. https://doi.org/10.1007/s10676-020-09547-x

Statista. (2021a, September 29). *Number of nurses per 10,000 population in the Philippines 2020, by region*. https://www.statista.com/statistics/1122083/philippines-number-of-nurses/

Statista. (2021b, November 22). *Number of nurses per 10,000 population in the Philippines 2020, by region*. https://www.statista.com/statistics/1122083/philippines-number-of-nurses/

Technology for COVID-19 Surveillance. (2020, May 12). Https://Www.Lguvscovid.Ph/Content/Technology-for-Covid-19-Surveillance. Retrieved October 12, 2021, from https://www.lguvscovid.ph/content/technology-for-covid-19-surveillance

Tee, M. L., Tee, C. A., Anlacan, J. P., Aligam, K. J. G., Reyes, P. W. C., Kuruchittham, V., & Ho, R. C. (2020). Psychological impact of COVID-19 pandemic in the Philippines. *Journal of Affective Disorders*, *277*, 379-391. https://doi.org/10.1016/j.jad.2020.08.043

Vlasov, V., & Barkova, O. (2021). The impact of the COVID-19 pandemic on economic, socialand political processes in global, regional and national aspects. *Научный Компонент*, *1*, 60-66. https://doi.org/10.51980/2686-939x_2021_1_60

Appendices

Appendix A

Letter to the Adviser

Attachment 3

	INVITATION LETTER FOR ADVISER	Document No.	WVSU-ICT-SOI-03-F03
		Issue No.	1
		Revision No.	0
	WEST VISAYAS STATE UNIVERSITY	Date of Effectivity:	April 27, 2018
		Issued by:	CICT
		Page No.	Page 1 of 2

March 08, 2021

Evans Sansolis

CICT-West Visayas State University
Luna St. La Paz Iloilo City, 5000

Dear Evans Sansolis ,

The undersigned are/BS in Computer Science Research 1/Thesis 1 students of CICT, this university. Our thesis/capstone project title is *"Project Ahon: An Agile Tracking System Using Mobile Embedded Systems for COVID-19 Contact Tracing and Data Collection"*.

Knowing of your expertise in research and on the subject matter, we would like to request you to be our **ADVISER**.

We are positively hoping for your acceptance. Kindly check the corresponding box and affix your signature in the space provided. Thank you very much.

Respectfully yours,

1. Horacio Jason A. Barba V
2. Carlo Reyes
3. Jovan Earl Hinayan
4. Joseph Anthony Duran

PS:

Advisers, are task to work with the students in providing direction and assistance as needed in their thesis/capstone project. They shall meet with the students weekly or as needed to provide direction, check on progress and assist in resolving problems until such a time that the students passed their defenses and submit their final requirements, as well as, preparing their evaluations and grades.

Action Taken:
● I Accept.
☐ Sorry. I don't accept.

EVANS SANSOLIS
Signature over printed name of the Adviser

Appendix B

Letter to the Editor

	TECHNICAL EDITOR'S ENDORSEMENT FORM (For Thesis Manuscript)	Document No.	WVSU-ICT-SOI-03-F11
		Issue No.	1
		Revision No.	0
	WEST VISAYAS STATE UNIVERSITY	Date of Effectivity:	April 27, 2018
		Issued by:	CICT
		Page No.	Page 1 of 1

Respectfully endorsed to the **English Editor**, the attached manuscript of the thesis entitled:

Project Ahon: An Agile Tracking System using Mobile Embedded Systems for COVID-19 Tracing

Said manuscript was presented to me and was reviewed and edited in terms of technical specifications, correctness of diagrams and other technical matters. The corrections and suggestions was carried and implemented by the proponents whose names are listed hereunder.

Now therefore, I hereby **ENDORSE** the said thesis manuscript to the English Editor/Grammarian for **English Grammar Editing.**

_____
Technical Editor's Name & Signature

Date: _____

Group Members:

1. Horacio Jason A. Barba V
2. Carlo Reyes
3. Jovan Earl Hinayan
4. Joseph Anthony Duran

1.

	ENGLISH EDITOR/GRAMMARIAN'S ENDORSEMENT FORM (For Thesis Manuscript)	Document No.	WVSU-ICT-SOI-03-F12
		Issue No.	1
		Revision No.	0
	WEST VISAYAS STATE UNIVERSITY	Date of Effectivity:	April 27, 2018
		Issued by:	CICT
		Page No.	Page 1 of 1

Respectfully endorsed to the **Thesis Format Editor**, the attached manuscript of the thesis entitled:

Project Ahon: An Agile Tracking System using Mobile Embedded Systems for COVID-19 Tracing

Said manuscript was presented to me for English grammar editing, corrections has been made and the proponents whose names are listed hereunder implemented said corrections and changes in the revised manuscript.

Now therefore, I hereby **ENDORSE** the said thesis manuscript for **Thesis Format Editing**.

_____
English Editor/Grammarian's Name and Signature

Date: _____

Group Members:

1. Horacio Jason A. Barba V
2. Carlo Reyes V
3. Jovan Earl Hinayan
4. Joseph Anthony Duran

	THESIS FORMAT EDITOR'S ENDORSEMENT FORM (For Thesis Manuscript)	Document No.	WVSU-ICT-SOI-03-F13
		Issue No.	1
		Revision No.	0
	WEST VISAYAS STATE UNIVERSITY	Date of Effectivity:	April 27, 2018
		Issued by:	CICT
		Page No.	Page 1 of 1

Respectfully endorsed to the **Thesis Coordinator**, the attached manuscript of the thesis entitled:

Project Ahon: An Agile Tracking System using Mobile Embedded Systems for COVID-19 Tracing

Said manuscript was presented to me and has checked the preliminaries, thesis document convention and end matters, made some corrections which was implemented by the proponents whose names are listed hereunder.

Now therefore, I hereby **ENDORSE** said manuscript to the Thesis Coordinator for appropriate action.

_____
Thesis Format Editor's Name and Signature

Date: _____

Group Members:

1. Horacio Jason A. Barba V
2. Carlo Reyes
3. Jovan Earl Hinayan
4. Joseph Anthony Duran

Appendix C

Disclaimer

This software project and its corresponding documentation entitled "Project Ahon: An Agile Tracking System using Mobile Embedded Systems for COVID-19 Tracing" is submitted to the College of Information and Communications Technology, West Visayas State University, in partial fulfillment of the requirements for the degree, Bachelor of Science in Computer Science. It is the product of our own work, except where indicated text.

We hereby grant the College of Information and Communications Technology permission to freely use, publish in local or international journal/conferences, reproduce, or distribute publicly the paper and electronic copies of this software project and its corresponding documentation in whole or in part, provided that we are acknowledged.

Horacio Jason A. Barba V

Joseph Anthony M. Duran

Jovan Earl P. Hinayan

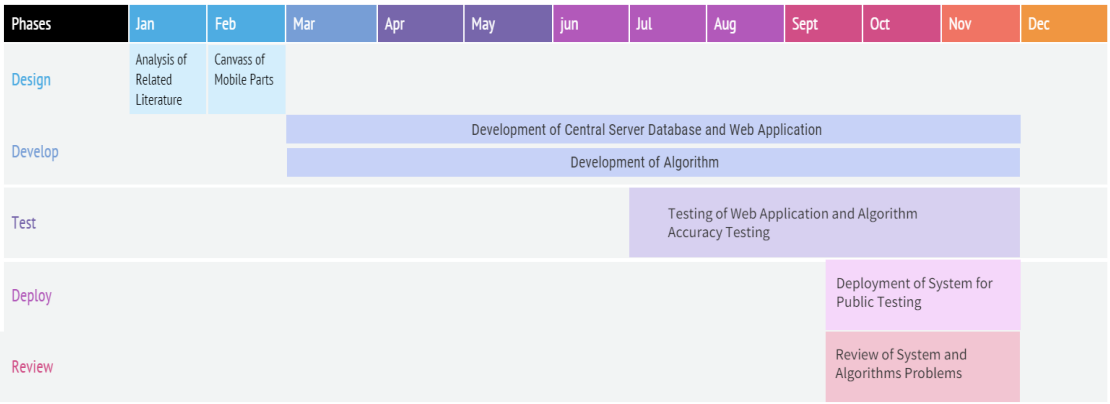
Carlo Adrian R. Reyes

June 2022

Appendix D

Gantt Chart

Project Ahon Development Gantt Chart



Appendix E

Data Dictionary

Appendix F

Sample Program Codes

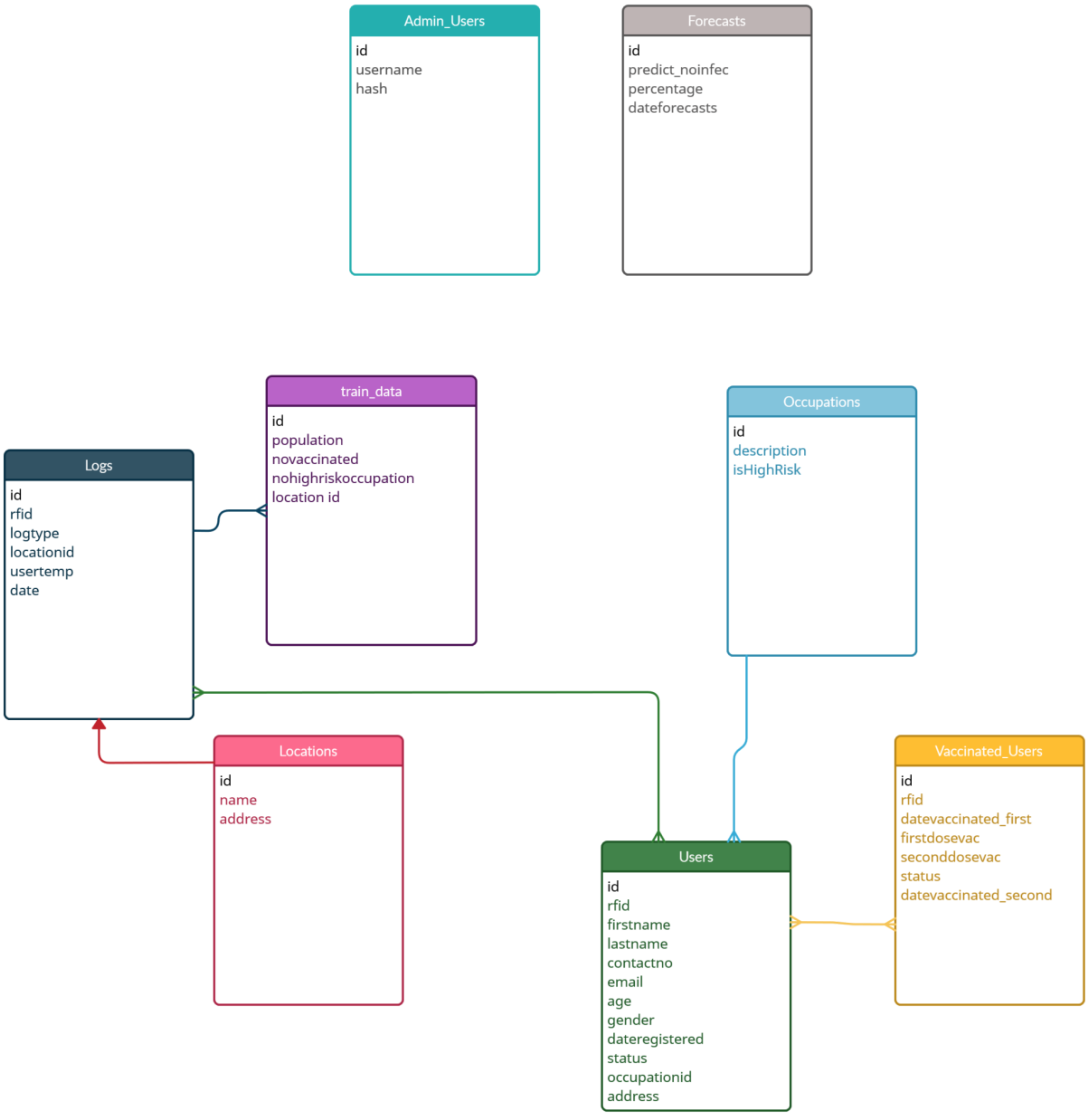
```
def predict_noinfec():

    global reg
    df['noinfec'] =
df['noinfec'].fillna(df['noinfec'].interpolate(method ='linear',
limit_direction='forward'))
    reg = linear_model.LinearRegression()
    reg.fit(df[['pop', 'novac', 'occup', 'location']],df.noinfec)
    reg.coef_
    reg.intercept_
    global last_line
    last_line = pd.DataFrame(df.iloc[-1:,:4].values)
    print(reg.predict(last_line))
```

```
def percentage():
    population = pd.DataFrame(df.iloc[-1:,:1].values)
    a = population
    b = reg.predict(last_line)
    percent = b / a * 100
    print(float(percent.iloc[0]))
```

Appendix G

Entity-Relationship Diagram



Appendix H

System Evaluation Tool

Please rate the following software characteristics.
The objective of this survey is to evaluate the quality of the developed software.

ALL INDIVIDUAL RESPONSES WILL REMAIN CONFIDENTIAL.

Name (Optional):_____

Gender: [✓] Male [ ] Female

Scale	***Description***
5	Excellent
4	Very Good
3	Good
2	Fair
1	Poor

Characteristic	Sub-characteristics	Rating				
		1	2	3	4	5
Functional Suitability	Functional completeness					✓
	Functional correctness					✓
	Functional appropriateness					✓
Performance Efficiency	Time behaviour					✓
	Resource utilization					✓
	Capacity					✓
Compatibility	Co-existence				✓	
	Interoperability				✓	
Usability	Interoperability					✓
	Appropriateness recognisability					✓
	Learnability					✓
	Operability					✓
	User error protection				✓	

	User interface aesthetics				✓	
	Accessibility					✓
Characteristic	Sub-characteristics	Rating				
		1	2	3	4	5
Reliability	Availability				✓	
	Fault tolerance					✓
	Recoverability					✓
	Confidentiality					✓
	Integrity					✓
	Non-repudiation					✓
	Accountability					✓
	Authenticity					✓
	Modularity					✓
	Reusability					✓
Maintainability	Analysability					✓
	Modifiability					✓
	Modifiability					✓
Portability	Adaptability					✓
	Installability					✓
	Replaceability					✓