**Ideation Phase**

**Defining the Problem Statements**

|  |  |
| --- | --- |
| **Date** | **26-09-2023** |
| **Team ID** | **716** |
| **Project Name** | **Public Health Awareness Campaign Analysis** |

**Public Health Awareness Campaign Analysis**

**Problem Definition and Design Thinking**

**Introduction**

In this modern world, diseases are spread all over the world. We have to create a awareness on the seriousness of the various diseases. Here we have a dataset of various diseases occurred in the past years. We will predict the future possibility of occurrence of diseases using dataset by applying data analytics.

In this document, we will outline the problem statement, the steps involved in solving it, and the design thinking approach that will guide our project.

**Problem Statement**

Objective: Develop a model that can predict the occurrence of various diseases in future to create an Awareness

Data: We have a dataset containing various diseases occurred in the past years(e.g., chicken pox,malaria,small pox etc.) along with their corresponding count. This data will be used to train and evaluate our model.

**Key Challenges:**

1. Data Quality: Ensuring the dataset is clean, complete, and free of errors.

2. Feature Selection: Identifying the most relevant features for accurate count predictions.

3. Model Selection: Choosing the appropriate Tool for the task.

4. Model Evaluation: Evaluating the model's performance using appropriate metrics.

5. Deployment: Creating a user-friendly interface or API for end-users to make predictions.

**Design Thinking Approach**

**Empathize:**

Before diving into solving the problem, it's crucial to empathize with the users and understand their needs. In this case, our primary users are peoples . We need to gather insights into what factors are most important to them when considering their health conditions and how accurate predictions can benefit them.

**Actions:**

- Conduct surveys or interviews with potential users to gather their perspectives.

- Analyse historical data to identify the most occurred disease.

- Seek feedback from domain experts in the hospital industry.

**Define:**

Based on our understanding of the problem and the users' needs, we will define clear objectives and success criteria for our project.

**Objectives:**

- Develop a model that achieves a Mean Absolute Error (MAE) of less than $X on the test data.

- Create a user-friendly web application for users to input their health observations(Health Conditions)

**Ideate:**

Brainstorm potential solutions and approaches to address the problem. This phase involves thinking creatively and considering various algorithms and techniques for diseases occurrence in future.

**Actions:**

- Explore different machine learning algorithms such as linear regression, decision trees, random forests, and neural networks.

- Experiment with feature engineering techniques to enhance model performance.

- Consider incorporating external data sources (e.g., neighbours affected by any disease, already the person was affected by any other diseases) to improve predictions.

**Prototype**

Create a prototype of the model and the user interface for Diseases occurrence prediction.

**Actions:**

- Develop a Cognos script for data pre-processing, model training, and evaluation.

- Create a simple web interface using tools like Flask or Django to allow users to input their health details.

- Test the prototype with a subset of the dataset to ensure it meets performance objectives.

**Test**

Evaluate the model's performance using appropriate metrics and gather feedback from users.

**Actions:**

- Split the dataset into training and testing sets.

- Train the model on the training set and evaluate it on the testing set.

- Use metrics such as MAE, Root Mean Square Error (RMSE), and R-squared to assess model performance.

- Collect user feedback on the web interface for usability and accuracy.

**Implement**

Once the prototype meets the defined objectives and receives positive feedback, proceed with full implementation.

**Actions:**

- Train the final model on the entire dataset.

- Deploy the model as part of a production-ready web application.

- Conduct thorough testing to ensure the application is robust and user-friendly.

**Iterate**

Continuous improvement is essential. Gather user feedback and iterate on the model and interface to enhance accuracy and usability.

**Actions:**

- Monitor the model's performance and retrain it periodically with updated data.

- Address user feedback and make necessary improvements to the web interface.

- Stay informed about advancements in the model for potential enhancements.

**Conclusion**

In this document, we've outlined our approach to solving the problem of Public Health Awareness Campaign. We've defined the problem, identified key challenges, and laid out a design thinking approach that involves empathizing with users, defining objectives, ideating potential solutions, prototyping, testing, implementing, and iterating.

Our ultimate goal is to develop an accurate and user-friendly solution that provides valuable insights for peoples. By following this structured approach, we aim to create a reliable tool that contributes positively to the hospital industry.