**Ideation Phase**

**Defining the Problem Statements**

|  |  |
| --- | --- |
| **Date** | **26-09-2023** |
| **Team ID** | **3921** |
| **Project Name** | **Spam prediction using powered AI** |

**AI-powered spam classifier using natural language processing(NLP) and machine learning**

**Problem Definition and Design Thinking**

**Introduction:**

The task at hand is to develop a machine learning model that can accurately predict whether an incoming message is spam or not using Natural Language Processing (NLP) techniques. Spam prediction is a critical problem in communication platforms, aiming to protect users from unwanted and potentially harmful messages. 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 machine learning model that can predict whether an incoming message is spam or not with a high level of accuracy using NLP and ML techniques.

Data: We have a dataset containing various features of messages (e.g., text content, sender information, message type, etc.) along with labels indicating whether they are spam or not. This data will be used to train and evaluate our machine learning model.

**Key Challenges:**

1.Data Quality: Ensuring the dataset is clean, well-labeled, and free of errors.

2.Feature Extraction: Extracting meaningful text features for accurate spam prediction using NLP.

3.Model Selection: Choosing the appropriate NLP and ML algorithms for the task.

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

5.Deployment: Creating a user-friendly interface or API for users to check messages for spam.

**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 individuals who receive messages and want to filter out spam. We need to gather insights into what types of messages they consider spam and how accurate predictions can benefit them.

**Actions:**

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

- Analyze historical spam patterns and common characteristics of spam messages in text.

- Seek feedback from experts in email filtering, NLP, and cybersecurity.

**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 machine learning model that achieves a high accuracy rate (e.g., 95% or higher) in classifying messages as spam or not using NLP and ML.

- Create a user-friendly web or mobile app for users to check messages for spam in real-time.

**Ideate:**

Brainstorm potential solutions and approaches to address the problem. This phase involves thinking creatively and considering various NLP and ML algorithms and techniques for spam prediction.

**Actions:**

- Explore different NLP techniques such as TF-IDF, Word Embeddings (e.g., Word2Vec, GloVe), and neural networks.

- Experiment with feature engineering techniques specific to NLP to enhance model performance.

- Consider incorporating sender reputation and message metadata as features for prediction.

**Prototype:**

Create a prototype of the NLP and ML model and the user interface for spam prediction.

**Actions:**

- Develop a Python script for data preprocessing, NLP feature extraction, model training, and evaluation.

- Create a web or mobile app prototype with a simple and intuitive user interface for real-time spam checking.

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

**Test:**

Evaluate the model's performance using appropriate NLP and ML 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 accuracy, precision, recall, and F1-score to assess model performance.

- Collect user feedback on the app for usability and accuracy.

**Implement:**

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

**Actions:**

- Train the final NLP and ML model on the entire dataset.

- Deploy the model as part of the web or mobile app.

- Conduct thorough testing to ensure the app 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 update it periodically with new data.

- Address user feedback and make necessary improvements to the app's interface and NLP model.

- Stay informed about advancements in NLP and ML for potential enhancements.

**Conclusion:**

In this document, we've outlined our approach to solving the problem of spam prediction using NLP and ML. 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 helps individuals filter out spam messages effectively using advanced NLP and ML techniques. By following this structured approach, we aim to create a reliable tool that enhances communication and cybersecurity in digital messaging platforms