**NAANMUDHALVAN-IBM SKILL**

**ARTIFICIAL INTELLIGENCE**

**GROUP PROJECT**

**Project Title\* Build a smarter AI-powered spam classifier**

**Phase I. Submission**

|  |  |  |  |
| --- | --- | --- | --- |
| S.No | Group Members  Name | **Naan Mudhalvan** ID | Email ID |
| 1 | AROORAN B | au820321106007 |  |
| 2 | SAKTHI GANESH P | au820321106032 |  |
| 3 | YOGESHWARAN | au820321106041 | yogeshmaddy692@gmail.com |
| 4 | MANO BALAN B | au820321106025 |  |

**Problem statement and Design thinking**

**Problem statement**

In machine learning, spam filtering protocols use instance-based or memory-based learning methods to identify and classify incoming spam emails based on their resemblance to stored training examples of spam emails.

**Design thinking**

* We’ve all been the recipient of spam emails before. Spam mail, or junk mail, is a type of email that is sent to a massive number of users at one time, frequently containing cryptic messages, scams, or most dangerously, phishing content.
* While spam emails are sometimes sent manually by a human, most often, they are sent using a bot. Most popular email platforms, like Gmail and Microsoft Outlook, automatically filter spam emails by screening for recognizable phrases and patterns.
* A few common spam emails include fake advertisements, chain emails, and impersonation attempts. While these built-in spam detectors are usually pretty effective, sometimes, a particularly well-disguised spam email may fall through the cracks, landing in your inbox instead of your spam folder.
* Clicking on a spam email can be dangerous, exposing your computer and personal information to different types of malware. Therefore, it’s important to implement additional safety measures to protect your device, especially when it handles sensitive information like user data.

### Prerequisites

* First, we’ll import the necessary dependencies. Pandas is a library used mostly used by data scientists for data cleaning and analysis.
* [Scikit-learn](https://scikit-learn.org/), also called Sklearn, is a robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling, including classification, regression, clustering, and dimensionality reduction via a consistent interface.

Conclusions:

By using this process we are build a smarter AI-powered spam classifier