Capstone Project Proposal

<Abhijna Yaji>

Business Goals

Project Overview and Goal What is the industry problem you are trying to solve? Why use ML/AI in solving this task? Be as specific as you can when describing how ML/AI can provide value. For example, if you're labeling images, how will this help the business?	Communication online has become ever so important in the current age and much of the important information is conveyed through mails. On a day an average of 121 mails are received by a person. Spam emails are also a part of this huge pile of incoming mails. These emails pose as a security threat and also cause wastage of time going through them. This repetitive task of checking if a mail is spam is tedious and time consuming. This repetitive task is better suited for ML/AI to classify the mails and keep the inbox safe and organized containing only relevant and important mails. Thus, the project proposed is a spam classifier which can be added to our email application to sort out the spam mails.
Business Case Why is this an important problem to solve? Make a case for building this product in terms of its impact on recurring revenue, market share, customer happiness and/or other drivers of business success.	Electronic mails play an important role of communication both in the professional as well as unprofessional world even though there are several other parallel applications with similar features. Emails are used in several places such as signing up for newsletter, applying for jobs, sending documents to colleagues and more. Thus, a spam classifier plays a vital role to safeguard the inbox from unwanted mails thus can be used by many on a day-to-day basis, which increases the share to market value with higher rate of usage. It also provides customer happiness and relieves them from the stress of scanning through a pile of spam emails.
Application of ML/AI What precise task will you use ML/AI to accomplish? What	ML/AI can be utilized to identify the span mail present in the list of emails and can stored in a single category. A logistic regression model can be utilized to classify the emails and natural language processing along with

business outcome or objective will you achieve?	neural networks can help detect a spam email. The experience of going through mails containing information becomes easy and less time is spent on insignificant malicious mails. Users can be protected against hackers utilizing mails to phish information or infect the system with virus or bugs.

Success Metrics

Success Metrics What business metrics will you apply to determine the success of your product? Good metrics are clearly defined and easily measurable. Specify how you will establish a baseline value to provide a point of comparison.	The rate of usage of the spam classifier, the percentage of mails that are deleted present in the spam bin, percentage of mails additionally added to the spam bin, customer retention rate metrics help determine the success of the product. The baseline values to determine the product success are as follows: Rate of customer usage: 80% and above Customer Retention rate: 90% and above Percentage of mails which are marked as spam is deleted: 95% and above Percentage of mails additionally added to the spam bin: 10% and below The customer experience regarding the product through feedbacks and the cost spent per customer are two other valid metrics which help determine the growth and success of the product.

Data

Data Acquisition	The data will be obtained from the email application consisting of the various emails that the person had

Where will you source your data from? What is the cost to acquire these data? Are there any personally identifying information (PII) or data sensitivity issues you will need to overcome? Will data become available on an ongoing basis, or will you acquire a large batch of data that will need to be refreshed?	received. A database of malicious email addresses as well as links is collected from various organizations are required which can be found on open-source intelligent feeds. The cost to acquire this data would be around $10000 (depends on the organizations). Consent of the email owner to access the meta data of the mail and in some cases the content in case of ambiguity is required to perform this task. A brief on the way the data will be processed to establish trust will be provided while taking their consent. The data will be available on an ongoing basis as per the frequency of the incoming mails.
Data Source Consider the size and source of your data; what biases are built into the data and how might the data be improved?	Considering the size of the data the percentage of emails which are spam is usually less than 10% which causes a bias of predicting the mail as not spam to obtain a higher accuracy. The data can be improved by providing a balanced data of spam and non-spam emails while training. In case of ambiguity the user can be involved in classifying the mail as spam or not and this can be recorded for future references. Smart selection of data can be done by using data with low confidence, uncertainty, novelty and important for spam class.
Choice of Data Labels What labels did you decide to add to your data? And why did you decide on these labels versus any other option?	The data labels that will be included is “spam” and “not spam” this is adopted mainly because identifying the spam emails and isolating them from the important emails is the main objective of this project.

Model

Model Building How will you resource building the model that you need? Will you outsource model training and/or hosting to an external platform, or will you build the model using an in-house team,	The model will be built using an in-house team as the email application is a product of the company and will have quicker access to the user database. A cloud system will be used to store vast amounts of data needed for analysis and classification. Flutter and firebase can be used to build the app and for the website development python and Django can be utilized. The in-house team will consist of cyber security experts, data scientists, data and system engineers and AI experts.

and why?	Co-ordination between these teams can help build the model in an efficient and cost- effective manner.
Evaluating Results Which model performance metrics are appropriate to measure the success of your model? What level of performance is required?	A higher recall is required to ensure we accurately determine the spam mails which pose a security threat to the user. A balance high of precision and recall can be targeted by achieving a higher F1 score. Log loss can be utilized to predict the probability of the mail being classified correctly. A high F1 score of 0.8 and above is required. A log loss of 0.1 and below can highlight that the model is classifying correctly.

Minimum Viable Product (MVP)

Design What does your minimum viable product look like? Include sketches of your product.	Minimum viable product consists of the email application interface consisting of spam filter enable option, main folders such as inbox, spam, bin; information on when the filtering was last performed and markers highlighting safe mails and mails which maybe spam but are not put in the spam folder.
Use Cases What persona are you designing for? Can you describe the major epic-level use cases your product addresses? How will users access this product?	The product is designed for user persona consisting of individuals and professionals of any age group. The major use cases would be filter out hazardous mails which steal personal information, increasing user productivity, securing important details present in the email. The main challenge would be to classify all the fraudulent mails as spam and not cause additional efforts for the user to classify more mails as spam or correct the mails classified as spam to not spam. The epic use case would be to protect the user information and mails from being hacked through phishing mails and click baits. This would prevent a huge damage that would cause to the user if private data regarding finances or personal life fall into the hands of a hacker.
Roll-out How will this be adopted? What	This will be adopted as a secure email spam classifier product. The go-to market plan would be to find investors to support and develop this product. This can

does the go-to-market plan look like?	be released both as a website as well as an app on the app store or play store. The pricing strategy would be to avail this feature for free for the existing email application users and then charge them a small fee with increase in customer interaction and satisfaction of the product. The new email applications users can avail this product for free but for a shorter time compared to existing users after which a fee would be charged. The distribution strategy would be to market the product globally highlighting the security, transparency and user-friendliness of the product.

Post-MVP-Deployment

Designing for Longevity How might you improve your product in the long-term? How might real-world data be different from the training data? How will your product learn from new data? How might you employ A/B testing to improve your product?	The product can be improved by clustering the user base into several groups based on their age category, profession, behaviour and mail usage purpose. Common features of the spam mails received by them can be utilized to perform better classification. The training data for this product will include real world data of people under various age-groups and professions to have a diversified data. On receiving new data, the model will predict the probability of the mail being a spam if the percentage is close to 50% then the user will be indicated that the mail maybe a spam mail, and receives user feedback to extract the necessary features to indicate the mail is spam or not. The A/B testing can be employed and done until satisfying performance metrics are met on a significant sample size having a diversified user group.
Monitor Bias How do you plan to monitor or mitigate unwanted bias in your model?	The user can be involved in classifying the mail as spam or not and key features of this mail can be recorded for future references. Feedback obtained from the user can be collected and incorporated to enhance the performance of the model and mitigate unwanted bias which may pose as a security threat to the user.