Progress Evaluation: Milestone 2

Project: *The Phisherman - AI-Powered Phishing Email Detector*

Team Members:
Curtis Jones - *cjones2022@my.fit.edu*
Elton Batista - *ebatista2022@my.fit.edu*
Jordan Chesley - *jchesley2022@my.fit.edu*

Faculty Advisor: Khaled Slhoub - *kslhoub@fit.edu*

Client: Khaled Slhoub - *College of Engineering and Science: Department of Electrical Engineering and Computer Science*

Milestone 2 Task Matrix

Task	*Completion*	*Jordan*	*Elton*	*Curtis*	*To Do*
Detection Demo	*100%*	*Web App*	*AI Model Training*	*AI Model Testing*	*Completed (note: integrated model will have to be swapped out after testing*
Jira Setup	*80%*	*Create epics and add user stories in a singular document*			*Integrate epics and user stories into jira, develop sprints for each milestone, and assign story points*
AI model testing	*50%*	*Format datasets for model testing*	*Create scripts for collecting data (Accuracy and performance)*	*Test three selected models / collect and compare data / select a model for project use*	*Retrain and retest datasets on better hardware, select an artificial intelligence model*
AI model training	*50%*	*Select and format datasets for model training*	*Integrate the model into a web browser*	*Train the model and analyze possible degradation in performance over the browser*	*Model has to be integrated via design documentation into a microservice, additional training to fine tune the model must be completed*

Research Gmail integration	*90%*	*Test implementation of Oauth with different email domains*	*Research gmail API tools in relation to project requirements*	*Research possible alternatives to direct gmail API integration*	*Completed*

Milestone 2 Task Discussion:

Detection Demo

Completed a detection demo using the selected artificial intelligence model. This simple demo demonstrates the main functionality of the model. The demo takes input from the user, consisting of the sender, subject, and body of an email. The demo will display a confidence score back to the user, indicating whether the model believes the email is phishing or legitimate.

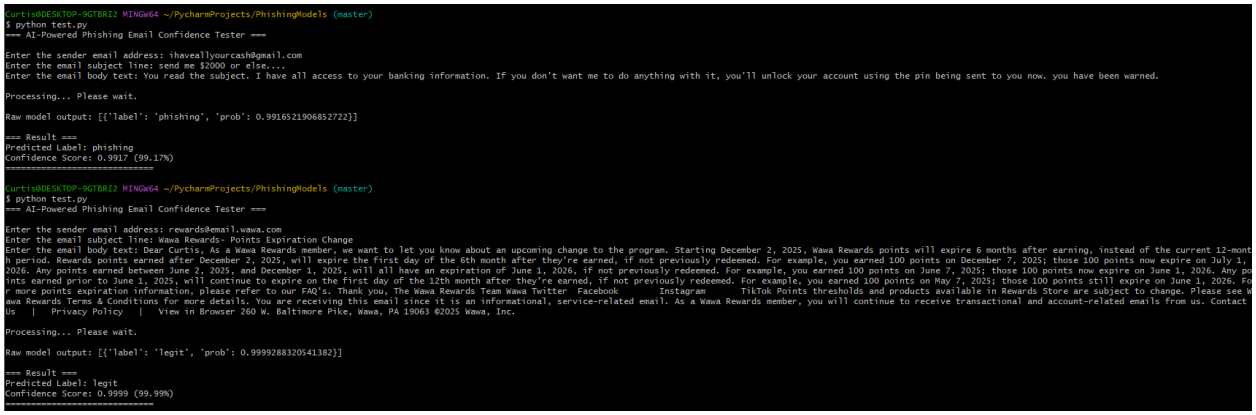
Jira Setup

A backlog for all future tasks has been compiled into a Google Doc.

AI model testing

We formatted all the datasets to be used for testing and training. We also developed a python script to automatically clean the datasets to work with the models. We used the langformers library in python to train 3 different artificial intelligence models (RoBERTa, DistilBERT, and Llama 3). We also separately tested the GPT LLM using their API. After training the models individually on the same dataset, we used three different datasets to individually test the models for performance and accuracy. Unfortunately, the models did not test well after training. This could of been caused by a variety of factors, but can be easily fixed through retraining on larger datasets using better hardware. The scripts for training and testing are completed.

Example of model functioning (RoBERTa)



AI model training

We have yet to use the model in a web interface outside of the demo. This includes testing it on large datasets through the web interface. We also plan on fine tuning the model through rigorous testing. The langformers library trains the model based on various configurations we give it. We plan on researching google codespace to give us better hardware for testing that should result in more accurate models. In addition to this, we may research developing a model without the langformer library if we find it will give us better performance and accuracy.

Research Gmail integration

We were able to make OAuth work using Google and Apple. OAuth requires payment for integration with different email providers, so we will not be integrating Microsoft into our project. We also have to test whether we will be able to integrate third party email providers, such as @fit.edu. However, we are ready to integrate OAuth as outlined in milestone 3.

Discussion of contribution of each team member to the current Milestone

Curtis Jones

- Researched Non LLM models (distilBERT and roBERTa)
- I trained and tested the non LLM models by developing python scripts.
- The scripts I developed:
    - Train models using the langformers library
    - Test models on performance and accuracy
    - Parse CSV files into a format the model can read
- Actively researching google codespace to train models on more models with better performance (Hardware directly affects the quality of trained models)

Elton Batista

- I mainly worked on the GPT-5 and LLaMA 4 parts of the project. My job was to get them working with our phishing email dataset and see how well they could tell the difference between real and fake emails. I spent a good chunk of time figuring out how to format the inputs properly and testing different ways to get better results from the models.
- Once things were up and running, I ran a bunch of tests to compare how each model performed. I wasn't just looking at accuracy — I also paid attention to how they handled tricky emails that looked legit but were actually phishing. It was interesting to see how each model responded to stuff like urgent language or weird links.
- I also helped explain how the models worked to the rest of the team and made sure our final report broke things down in a way that made sense. I didn't want it to just be a bunch of numbers — I wanted us to actually understand what was going on and be able to talk about it clearly.

Jordan Chesley

- Formatted the CSV files of datasets to match each other
- Researched Google OAuth and Apple ID
- Created and configured Google Cloud Console app for Google OAuth and Gmail APIs
- Created and configured an Apple Bundle ID and Apple Service IDs for OAuth integration and the iCloud API.
- Got OAuth login working in a local environment through Google and Apple

Milestone 3 Task Matrix

Task	*Jordan*	*Elton*	*Curtis*
Jira Setup	*Continue adding user Stories as needed*	*Implement epics and user stories into Jira*	*Choose methodology for assigning story points and construct sprints in JIRA*
AI model testing	*Parse and evaluate collected data to choose model*	*Extensively test LLM models with datasets*	*Retrain Models using google codespace for more accurate testing, extensively test non LLM models with datasets*
AI model training	*Collect more quality datasets and combine into one large dataset for testing*	*Research different ways to train the selected model (either from scratch or using a different library) and do performance/accuracy testing*	*Train and test selected model on larger dataset*
Web Interface Construction	*Implement OAuth API to feed user's emails into web application*	*Format user's emails into web application*	*Develop AI Model Processing service on backend*

Milestone 3 Task Discussion:

Jira Setup

We will hold a meeting to discuss the assignment of story points for each user story in order to develop sprints for each milestone. In addition to this, we plan to assign a manager for each specific sprint in order to divide management responsibilities between the three of us.

AI model testing

As described before, the models did not perform well, so we will have to retrain the models on better hardware using google codespace and finetune the parameters of the training to properly

test the models. We will then have a datasheet to choose a model to integrate into our web application.

AI model training

Once we have a model chosen, we can train it on a large dataset and integrate it into the web application. We also will look at different training methods to construct the most efficient model possible for the final product.

Web Interface Construction

We will start the construction of the web interface. The web interface for milestone 3 will include:

- OAuth integration
    - OAuth login
    - Feeding user emails into the web application
- AI Model integration
    - Returning the confidence score for emails
        - Giving a list of 100 emails that have the highest confidence score

1. Date(s) of meeting(s) with Client during the current milestone:

   10/6

   10/20

2. Client feedback on the current milestone
    - ... (if Client and Faculty Advisor are the same, write "see Faculty Advisor Feedback below")
    - ...
    - ...
3. Date(s) of meeting(s) with Faculty Advisor during the current milestone: …

   10/6

   10/20

4. Faculty Advisor feedback on each task for the current Milestone
    - Task 1: Satisfied
    - Task 2: Satisfied
    - Task 3: Satisfied
    - Task 4: Satisfied
5. Faculty Advisor Signature: ___Dr Slhoub_____ Date: ___10/27/25_____