UNIVERSITI TEKNIKAL MALAYSIA MELAKA

FAKULTI TEKNOLOGI MAKLUMAT DAN KOMUNIKASI

SEMESTER 1 2022/2023

BITI3533 ARTIFICIAL INTELLIGENCE PROJECT MANAGEMENT

PROJECT TITLE:

SPAM EMAIL FILTERING SYSTEM

GROUP NAME: C

PREPARED BY:

NAME	MATRIC NO
1. KAVINRAJ A/L SIVAKUMAR	B032010230
2. MUHAMMAD FARHAN BIN MOHD NOOR	B032010270
3. TONG YING FAN	B032120061

PREPARED FOR:

PROFESOR TS. DR. BURHANUDDIN BIN MOHD ABOOBAIDER

1.0 Forum on GitHub

1. What you understand about GitHub? Why do you think it is useful to manage the project?

Github is an online hosting service for the development of software and every version of the software development using Git. Github is very crucial for project management because it allows all the individual team members to develop, integrate or copy any prior version of the project software without jeopardizing the main branch of the project software.

2. How GitHub can be used to manage project related to artificial intelligence and software development?

Github can be used to manage to track and record any version of the project development through local server and the global server. This means that each individual can add or develop their part of the module and integrate it later on to the main branch.

3. List down other alternative tools can be used for AI project management like GitHub.

1.Proofhub

2.Azure boards

3.Aws CodeCommit

4. Rank the alternative tools given in Question 3 and which is the best?

1.Azure Boards

2.Proofhub

3.Aws Codecommit

Other than github,Azure boards is the best

5. Justify your answer given in Question 4.

Azure boards provide plenty of resources and tools to jumpstart any projects such as analytical tools,AI machine learning, databases and so much more. Not only that Azure DevOps provides tools for a group of developers to collaborate together on any kinds of software development.

2.0 GitHub

SPAM EMAIL DETECTION IN EMAIL USING MACHINE LEARNING

This is a project that detects and recognizes the spam in an email. This project uses Artificial Intelligence(AI) techniques such as Machine Learning and Neural Network(NN).

The project is developed by following Project Management Plan(PMP) to better organize the lifecycle of the project. Several charting techniques such as Gantt Chart, Milestone Chart, Work Breakdown Structure(WBS) are used to manage the project.

The PMP-Plan developed to simulate collaboration between third party organizations.

PROJECT MANAGEMENT PLAN

PROJECT SUMMARY

Project Summary gives a brief overview of the project objectives and team members details.

PROJECT PLANNING

Planning processes include devising and maintaining a workable scheme to ensure that the project addresses the organization's needs. Gantt Chart, Work Breakdown Structure(WBS) is documented in this phase.

PROJECT IMPLEMENTATION

Project Implementation define the project deliverables which included software requirement, Intelligent system architecture and etc.

PROJECT EXECUTION

Project Execution is the execution phase where the result, coding blocks and technical implementation is defined and documented.

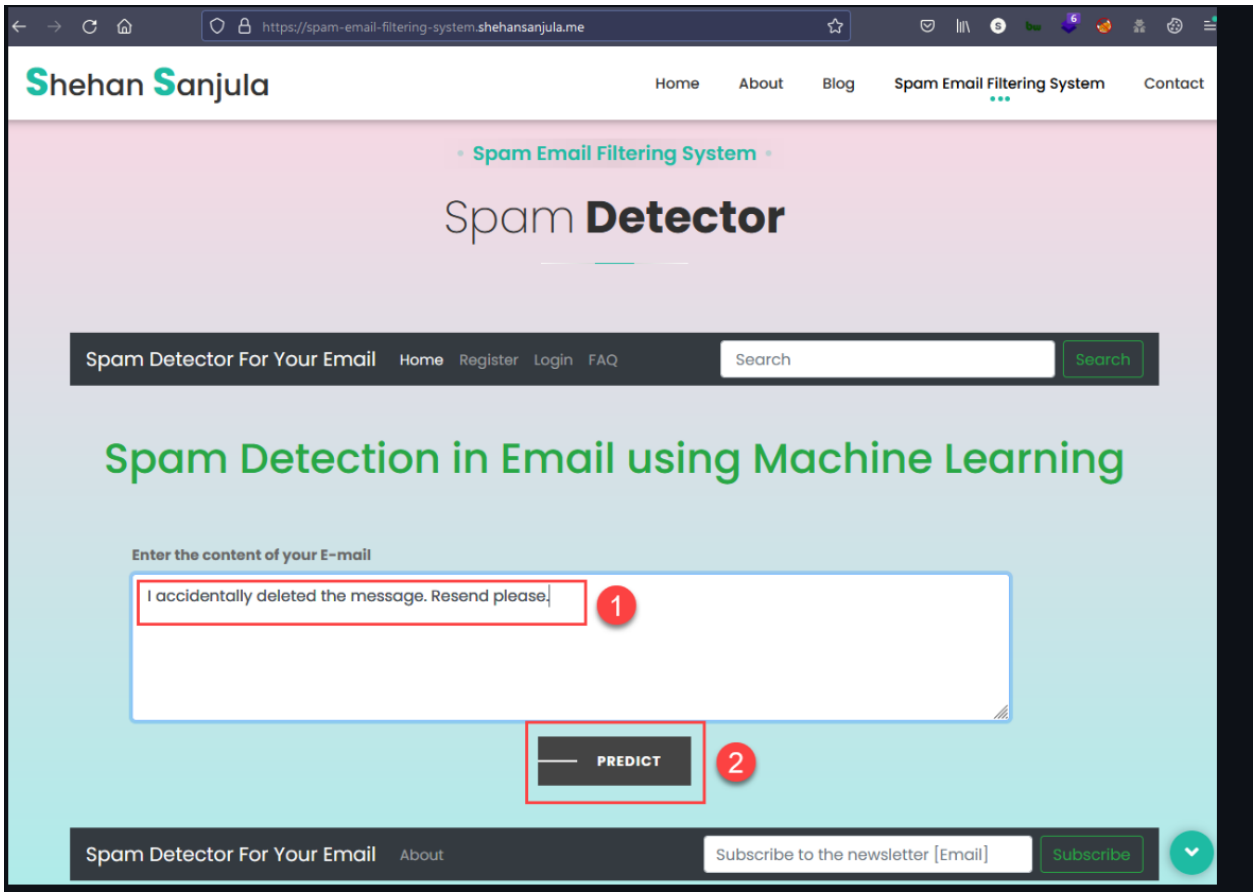
PROJECT CLOSING

Project Closing indicate the end of the project lifecycle where the project need to document closing document and lesson learnt report.

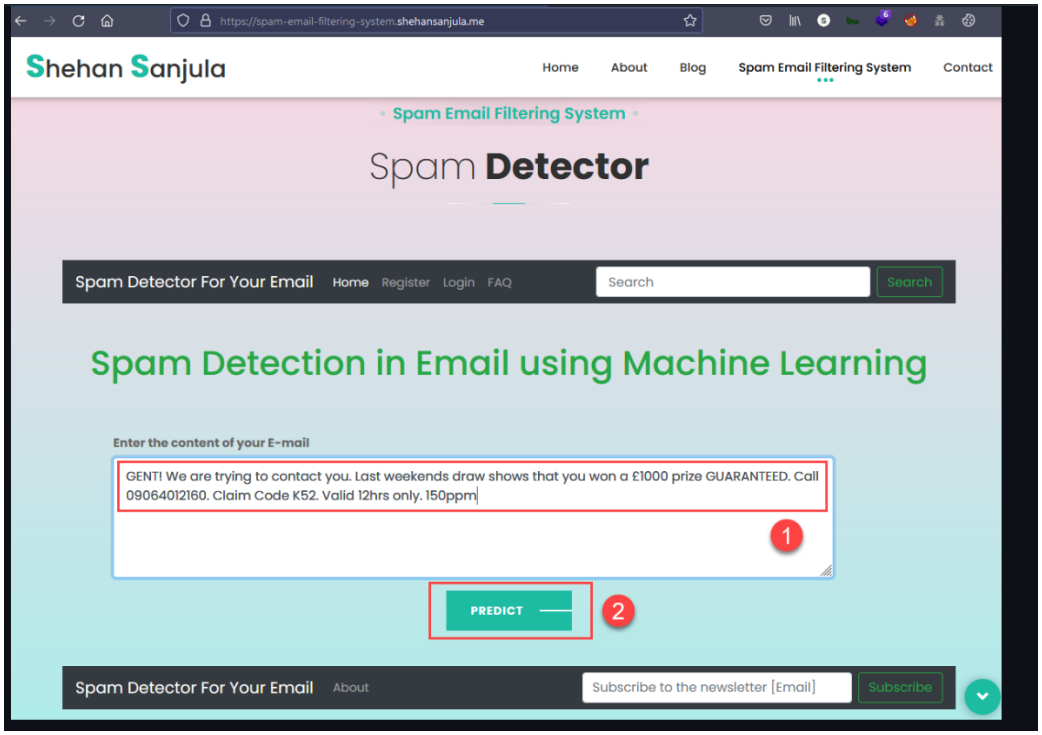
SAMPLE RESULT

Output 1: The output shows the model predicting an email text that is supposed to be ham

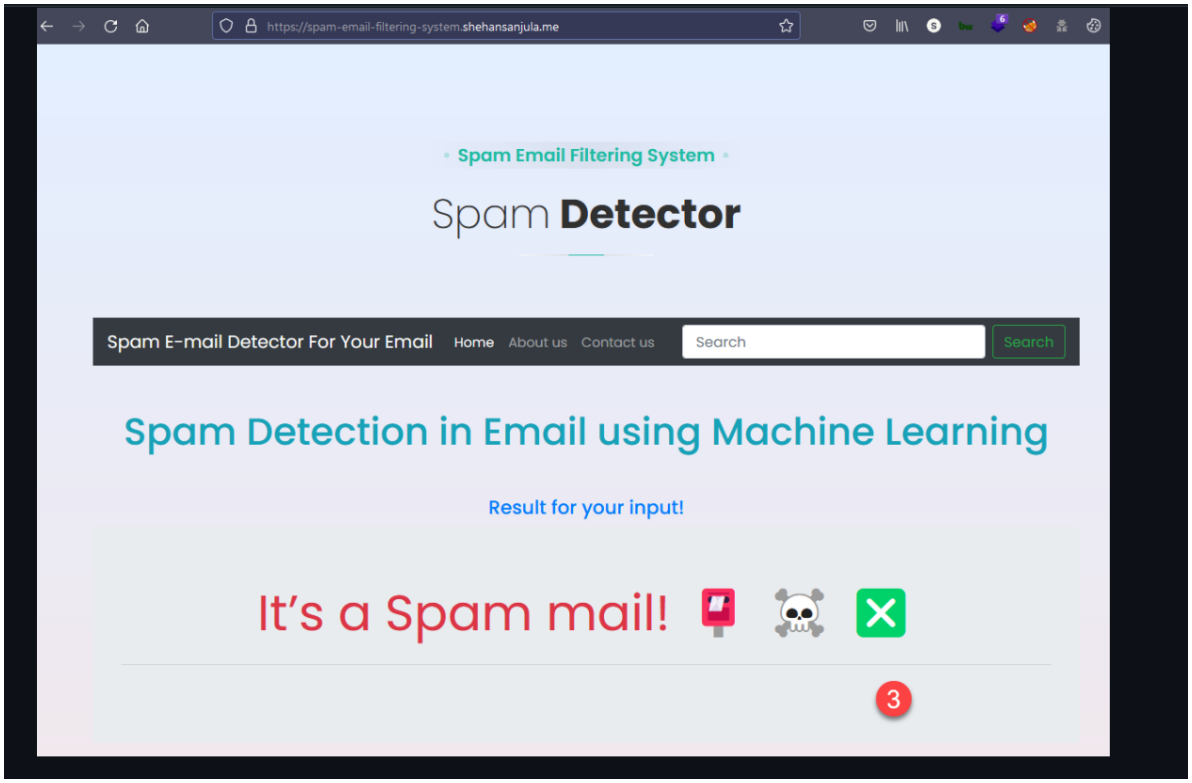
The image 1 shows the user entering an email text and clicking predict button



Output 2: The output shows the model predicting an email text that is supposed to be spam



The image 3 shows the user entering an email text and clicking predict button



3.0 PROJECT SUMMARY

Nowadays, spam has become a serious internet issue, statistics show that spam already accounted for 55% for all messages in 2017 and same with previous year. Spam is also known as unsolicited bulk email has led to the increasing use of email as email provides the perfect ways to send unwanted advertisements or junk messages by irresponsible users and resulting in clutter in the mailbox of people all around the world.

Employees in Robert Bosch would require top level security as they will be receiving constant emails from multiple organizations regarding project enquiries, thus organizing incoming email and removal of spam that may affect workspace performance is important

A specific algorithm is used to learn to filter and classify these messages. These algorithms are used for classifying objects into different classes which provide input and output to have a self-learning program to complete the task given. The message will be classified as spam or ham. This algorithm is used to predict the probability and classification of data outcome.

Client: Robert Bosch Sdn Bhd

Project title: Spam Email Filtering System

Team Members:

1. KAVINRAJ A/L SIVAKUMAR B032010230

2. MUHAMMAD FARHAN BIN MOHD NOOR B032010270

3. TONG YING FAN B032120061

Objective:

1. To develop an email spam filtering system based on machine learning algorithm

2. To develop a system that can classify messages in an email using Natural Language Processing(NLP)

3. To develop a system that can identify threats from emails

4.0 Project Planning

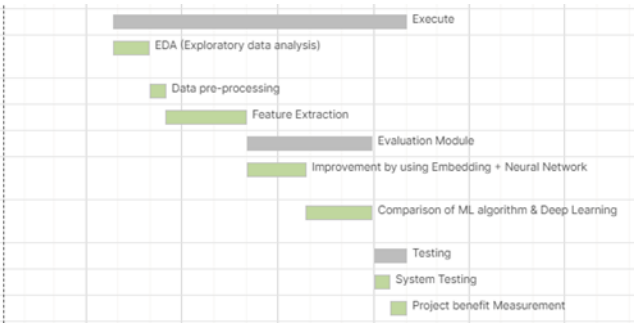
Project Management Life Cycle

The project management life cycle represents and is documented in Gantt Chart form to help staff follow the guidelines and achieve the milestone from time to time to increase the chance of achieving the goals. The WBS breakdown all modules and tasks into several modules so that every member can play their roles in every stage of the life cycle.

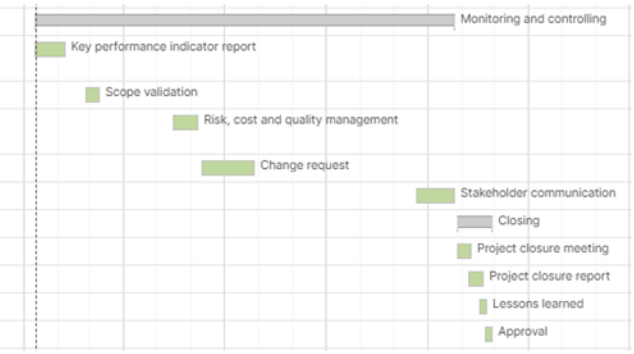
Gantt Chart

Initiating	12/01/22	01/31/23	44d	Kobu Paar	Initiating
Construct Business Case Modu	12/01/22	12/28/22	20d	Kobu Paar	Construct Business Case Module
Pitch Initiatives Towards Stakel	12/29/22	01/16/23	13d	Kobu Paar	Pitch Initiatives Towards Stakeholder
Develop Project Charter	01/21/23	01/26/23	5d	Kobu Paar	Develop Project Charter
Communication With Stakeholder	01/21/23	01/30/23	7d	Kobu Paar	Communication With Stakeholder
Planning	02/01/23	04/25/23	60d	Kobu Paar	Planning
Hold team planning meeting	02/01/23	02/03/23	3d	Kobu Paar	Hold team planning meeting
Prepare team contract	02/04/23	02/07/23	3d	Kobu Paar	Prepare team contract
prepare scope statement	02/08/23	02/09/23	2d	Kobu Paar	prepare scope statement
Prepare WBS	02/10/23	02/15/23	4d	Kobu Paar	Prepare WBS
Prepare Schedule and Cost Baseline	02/16/23	02/25/23	7d	Kobu Paar	Prepare Schedule and Cost Baseline
Identify,discuss, and prioritize risk	02/26/23	03/02/23	5d	Kobu Paar	Identify,discuss, and prioritize risk
Design Spam Detection System	03/03/23	03/09/23	5d	Kobu Paar	Design Spam Detection System
Software Design	03/10/23	03/16/23	5d	Kobu Paar	Software Design
Interface Design				Kobu Paar	
Input Page					
Output Page					
API Management					
Create Design Specification	03/24/23	03/30/23	5d	Kobu Paar	Create Design Specification
Deliver Final Design	03/31/23	04/05/23	4d	Kobu Paar	Deliver Final Design

Execute	04/26/23	01/31/24	201d	y	yingfan te
EDA (Exploratory data analysis)	04/26/23	05/30/23	25d	y	yingfan te
Data pre-processing	05/31/23	06/15/23	12d	y	yingfan te
Feature Extraction	06/15/23	08/31/23	56d	y	yingfan te
Evaluation Module	09/01/23	12/29/23	86d	y	yingfan te
Improvement by using Embedding + Neural Network	09/01/23	10/27/23	41d	y	yingfan te
Comparison of ML algorithm & Deep Learning	10/27/23	12/29/23	46d	y	yingfan te
Testing	01/01/24	01/31/24	23d	y	yingfan te
System Testing	01/01/24	01/15/24	11d	y	yingfan te
Project benefit Measurement	01/16/24	01/31/24	12d	y	yingfan te



Monitoring and controlling	01/10/23	01/24/24	272d	Kavinraj sival
Key performance indicator report	01/10/23	02/06/23	20d	Kavinraj sival
Scope validation	02/25/23	03/09/23	10d	Kavinraj sival
Risk, cost and quality management	05/15/23	06/06/23	17d	Kavinraj sival
Change request	06/10/23	07/27/23	35d	Kavinraj sival
Stakeholder communication	12/21/23	01/24/24	25d	Kavinraj sival
Closing	01/27/24	**02/27/24**	23d	Kavinraj sival
Project closure meeting	01/27/24	02/08/24	10d	Kavinraj sival
Project closure report	02/06/24	02/19/24	10d	Kavinraj sival
Lessons learned	02/16/24	02/22/24	5d	Kavinraj sival
Approval	02/21/24	02/27/24	5d	Kavinraj sival



Work Breakdown Structure

	Initiating	12/01/22	01/31/23	44d	Kobu Pase
⊞	Construct Business Case Mod.	12/01/22	12/28/22	20d	Kobu Pase
⊞	Pitch Initiatives Towards Stakel	12/29/22	01/16/23	13d	Kobu Pase
⊞	Develop Project Charter	01/21/23	01/26/23	5d	Kobu Pase
⊞	Communication With Stakeholder	01/21/23	01/30/23	7d	Kobu Pase
	Planning	02/01/23	04/25/23	60d	Kobu Pase
	Hold team planning meeting	02/01/23	02/03/23	3d	Kobu Pase
	Prepare team contract	02/04/23	02/07/23	3d	Kobu Pase
	prepare scope statement	02/08/23	02/09/23	2d	Kobu Pase
	Prepare WBS	02/10/23	02/15/23	4d	Kobu Pase
⊞	Prepare Schedule and Cost Baseline	02/16/23	02/25/23	7d	Kobu Pase
	Identify,discuss, and prioritize risk	02/26/23	03/02/23	5d	Kobu Pase
	Design database	03/03/23	03/09/23	5d	Kobu Pase
	Software Design	03/10/23	03/16/23	5d	Kobu Pase
	Interface Design	03/17/23	03/23/23	5d	Kobu Pase
	Create Design Specification	03/24/23	03/30/23	5d	Kobu Pase
	Deliver Final Design	03/31/23	04/05/23	4d	Kobu Pase
	Execute	04/26/23	01/31/24	201c	yingtan t
	EDA (Exploratory data analysis)	04/26/23	05/30/23	25d	yingtan t
	Data pre-processing	05/31/23	06/15/23	12d	yingtan t
	Feature Extraction	06/15/23	08/31/23	56d	yingtan t
	Evaluation Module	09/01/23	12/29/23	86d	yingtan t
	Improvement by using Embedding + Neural Network	09/01/23	10/27/23	41d	yingtan t
	Comparison of ML algorithm & Deep Learning	10/27/23	12/29/23	46d	yingtan t
	Testing	01/01/24	01/31/24	23d	yingtan t
	System Testing	01/01/24	01/15/24	11d	yingtan t
	Project benefit Measurement	01/16/24	01/31/24	12d	yingtan t
⊞	**Monitoring and controlling**	01/10/23	01/24/24	272c	Kavinraj silvi
	Key performance indicator report	01/10/23	02/06/23	20d	Kavinraj silvi
	Scope validation	02/25/23	03/09/23	10d	Kavinraj silvi
	Risk, cost and quality management	05/15/23	06/06/23	17d	Kavinraj silvi
	Change request	06/10/23	07/27/23	35d	Kavinraj silvi
	Stakeholder	12/21/23	01/24/24	25d	Kavinraj silvi
⊞	**Closing**	01/27/24	**02/27/24**	23d	Kavinraj silvi
	Project closure meeting	01/27/24	02/08/24	10d	Kavinraj silvi
	Project closure report	02/06/24	02/19/24	10d	Kavinraj silvi
	Lessons learned	02/16/24	02/22/24	5d	Kavinraj silvi
	Approval	02/21/24	02/27/24	5d	Kavinraj silvi

Scope

A work breakdown structure (WBS) is created to map all the necessary tasks and develop the process in which the WBS will be maintained and approved. In this project, this system will detect email in the form of text and image to classify it into spam or ham.

Roles and Responsibilities

Roles	Person In Charge	Responsibilities
Project Manager	Kavinraj A/L Sivakumar	The project manager is the individual responsible for delivering the project. The individual leads and manages the project team, with authority and responsibility from the project board, to run the project on a day-to-day basis. Responsibilities include: · managing the production of the required deliverables · planning and monitoring the project · preparing and maintaining project, stage and exception plans as required · managing project risks, including the development of contingency plans

		· monitoring overall progress and use of resources, initiating corrective action where necessary · applying change control and configuration management processes · reporting through agreed lines on project progress through highlight reports and end-stage assessments · liaison with appointed project assurance representatives to assure the overall direction and integrity of the project · maintaining an awareness of potential interdependencies with other projects and their impact · adopting and applying appropriate technical and quality strategies and standards · preparing any follow-on action recommendations · managing project administration · liaison with appointed project assurance representatives to assure the overall direction and integrity of the project

Procurement Manager	Muhamad Farhan Bin Mohd Noor	A procurement manager is the person responsible for sourcing and buying products and services for their company. The role is sometimes also called a purchasing manager. Responsibilities include: · Solution and product research · Inventory management · Procurement lifecycle optimization · Spend management · RFP management · Vendor ROI analysis · Supplier selection · Risk and compliance management · Supplier relationship management · Vendor negotiation · Procurement data analytics

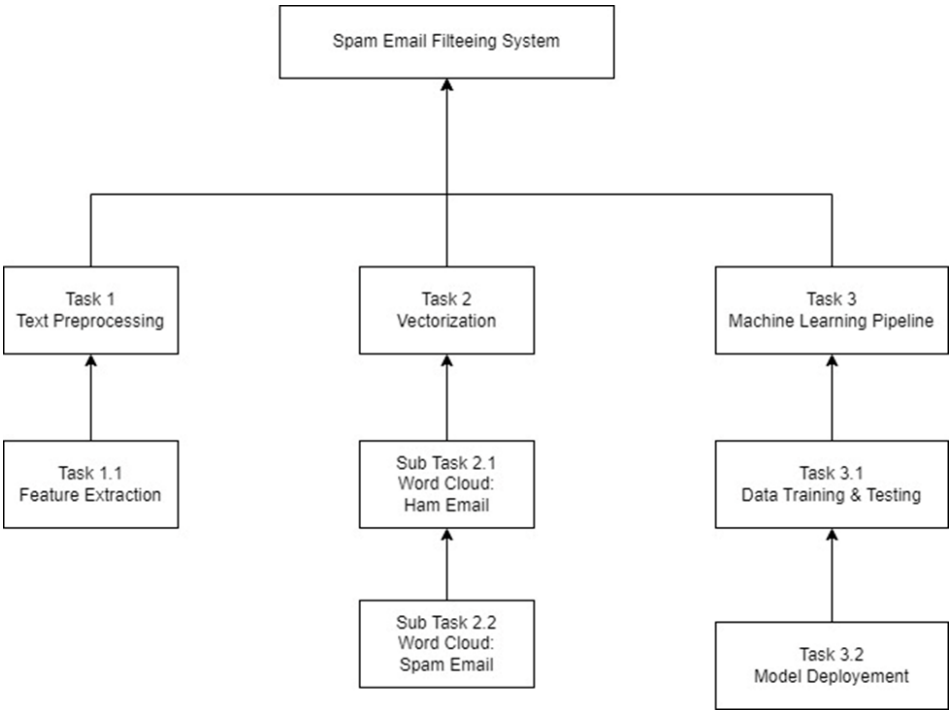
Risk Manager	Tong Ying Fan	A risk manager's overarching role is to help their organization prepare for uncertainties, minimize threats to its interests, and mitigate losses. Responsibilities include: · Provide a methodology to identify and analyze the financial impact of loss to the organization, employees, the public, and the environment. · Examine the use of realistic and cost-effective opportunities to balance retention programs with commercial insurance. · Prepare risk management and insurance budgets and allocate claim costs and premiums to departments and divisions.

Administrative Manager	Kavinraj A/L Sivakumar	Administrative Managers play a role in many parts of a business's operations. This is a top-level role, meaning Administrative Managers work with other leaders and team members in each of the business' departments. Responsibilities include: · Guiding the organization's activities · Ensuring a company is operating securely and effectively · Preparing and reviewing operational reports · Leading and/or participating in meetings · Assisting managers in compiling annual budget information and reports · Hiring and training administrative staff

Financial Analyst	Muhamad Farhan Bin Mohd Noor	A financial analyst is responsible for a wide range of activities including gathering data, organizing information, analyzing historical results, making forecasts and projections, making recommendations, and generating Excel models, presentations, and reports. Responsibilities include: · analyzing the financial needs and reports · making forecasts and projections · manage and monitoring cash flow
Project Scheduler	Tong Ying Fan	The Project Scheduler is responsible for creating, updating and coordinating project schedules. Creates resource usage model and predictions. Responsibilities include: · Work with the project team to develop, maintain and manage the project schedule · Participate in regular project team meetings across multiple projects

		· Regularly review critical path, cost loading, schedule content, use of relationships and lag, constraints and milestones
Quality Manager	Tong Ying Fan	The primary responsibility of a Project quality manager is to ensure the delivery of products or services according to specifications to satisfy the customers or stakeholders. Responsibilities include: · implementing and executing inspection, testing, and evaluation methods to ensure the quality of Spam Email Filtering System · prepare reports to analyze system

WBS (Technical Part)



Responsibility Assignment Matrices (RAM)

R- Compulsory to do the task

A- Only one who responsible

C- Consulted (Two way communication is needed)

I- The person who will be inform upon completion

	Project Manage	Procurement	Risk Manag	Administrative Manag	Financial Analys	Project Schedule	Quality Manager
Initiating							
Construct Business Case Module	R	A	A	A	A	A	I
Pitch Initiatives Towards Stakehold	R	A	A	A	C	C	I
Develop Project Charter	R	C	C	C	C	C	I
Communication With Stakeholder	R	C	C	C	C	C	I
Planning							
Hold team planning meeting	I	R	A	A	I	A	A
Prepare team contract	I	R	C	A	I	A	A
prepare scope statement	I	R	C	A	I	A	A
Prepare WBS	C	R	A	A	I	C	A
Prepare Schedule and Cost Baselir	I	C	C	A	I	C	A
Identify,discuss, and prioritize risk	I	A	R	A	I	C	A
Design Spam Detection System	I	C	C	A	I	R	A
Software Design	I	C	A	A	I	R	A
Interface Design	I	C	A	A	I	R	C
Input Page	I	C	A	A	I	R	C
Output Page	I	C	A	A	I	R	C
API Management	I	C	A	A	I	R	C
Create Design Specification	I	C	A	A	I	R	C
Deliver Final Design	I	C	A	A	I	R	C
Execute							
EDA (Exploratory data analysis)	I	R	A	C	A	C	A
Data pre-processing	A	I	C	C	A	C	R
Feature Extraction	A	C	I	C	A	C	R
Evaluation Module	A	C	C	I	A	A	R
Improvement by using Embedding +	A	C	C	S	I	A	R
Comparison of ML algorithm & Dee	A	A	C	C	I	C	R
Testing	A	A	A	A	I	C	R
System Testing	A	A	A	C		I	R
Project benefit Measurement	A	I	C	C	R	A	I
Monitoring and controlling							
Key performance indicator report	R	C	C	C	A	A	I
Scope validation	A	C	A	A	C	R	I
Risk, cost and quality management	A	A	R	A	A	A	I
Change request	A	R	A	C	C	C	I
Stakeholder communication	A	A	A	C	C	R	I
Closing							
Project closure meeting	I	I	I	I	I	I	I
Project closure report	R	I	I	I	I	I	I
Lessons learned	I	R	I	I	I	I	I
Approval	R	I	I	I	I	I	I

Risk Identification Chart (Quality, Cost, Time)

Control Element	What is likely to go wrong?	How will we know?	What will we do about it?
Quality	The project outcome doesn't reach expectation	During testing phase	Quality manager must monitor and control the project flow
Cost	Actual cost exceeds budget/estimated cost	When report of cost has been generated	Financial analyst should track and manage the cash flow
Time	Project cannot be done on time	Project cannot follow the milestone, next activities cannot start as well	All member should help each other (free member give some help on cannot done project part)

5.0 PROJECT IMPLEMENTATION

IMPLEMENTING THE PROJECT PLAN

Deliverables:

Hardware Requirement: For the hardware, our company requires the usage of a standard computer that is able to consistently run the spam filter therefore the computer part will consist of 8 Gb ram and intel core i5.

Software Requirement:

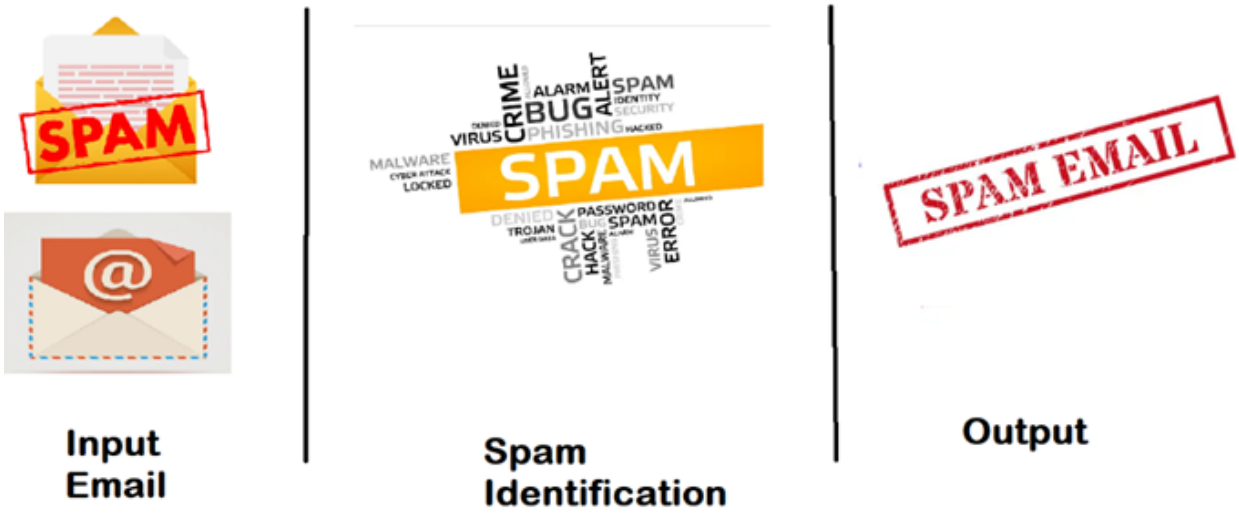
Software Requirement: Our company mainly utilized python programming language as the main language to develop intelligent modules in filtering spam and content. Other than that, our main software is a graphical operating system to develop and create all the documents and algorithms related to this project.

Intelligent System Architecture

The proposed neural network is Logistic Regression. This is because the model is dealing with binary output which is either spam or non-spam. Therefore logistic regression is the most suitable and simple for the project to attune and transform.

Outcome Of the System

2 Main Steps will be processed throughout the system, which is content and spam identifying through the neural network training phase, spam criteria and recognition performance through the testing phase. Provided with the dataset composed of email data which label as spam and non-spam(ham), the model utilized logistic regression to attune the graph transformation to fit the best line on the model that can accurately predict the spam email. After that, the model will be tested to determine the model performance against testing dataset and displaying the information such as Confusion Matrix and recall and accuracy score.

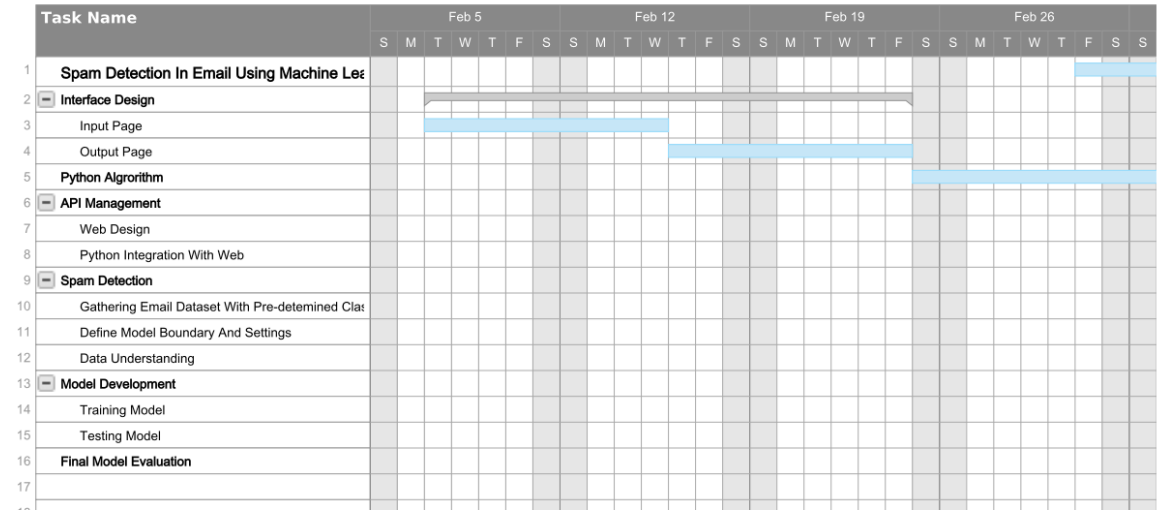


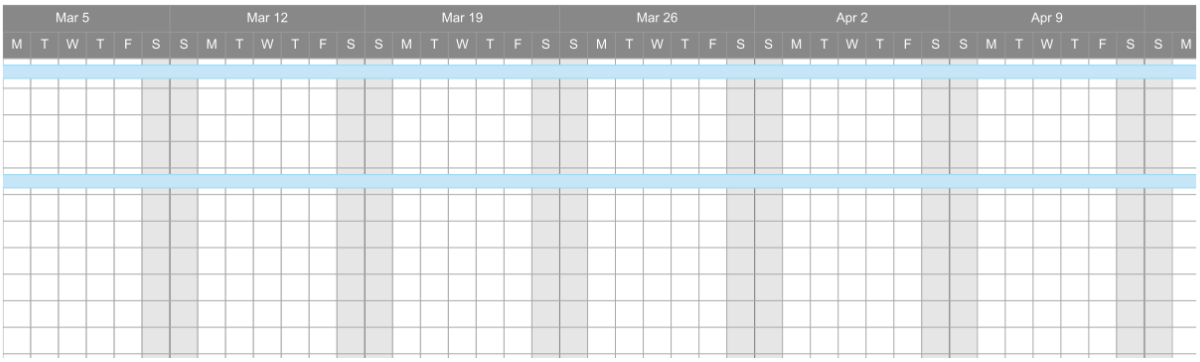
Task and Estimated Costs

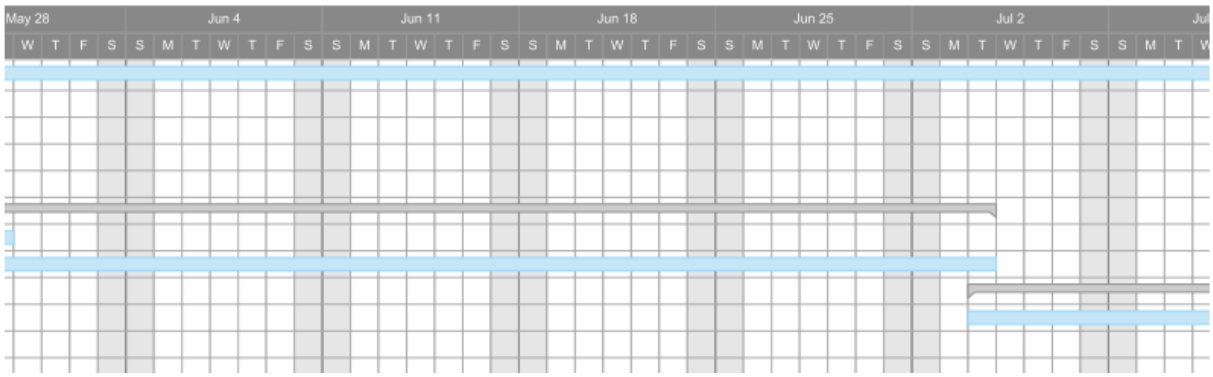
Task	Estimated Costs	Notes
LABOR	RM 13,4500	*Worker and Team members annual salary*
HARDWARE	RM 38,500	*Electronic Devices and System*
SOFTWARE	RM 3,000	*I.e Microsoft Office, Visual Studio*
TRAINING AND SUPPORT	RM 4,900	*Other Quality Of Life Expenses*
RESERVES	RM 75,000	*Backup Budget Incase of incidental cost*
Server Security	RM 3,500	Security Architecture Implementation
Electronic Material	RM 2,060	Other Unaccounted Electronic Material
Server Maintenance	RM 3,100	Server Upkeep Cost

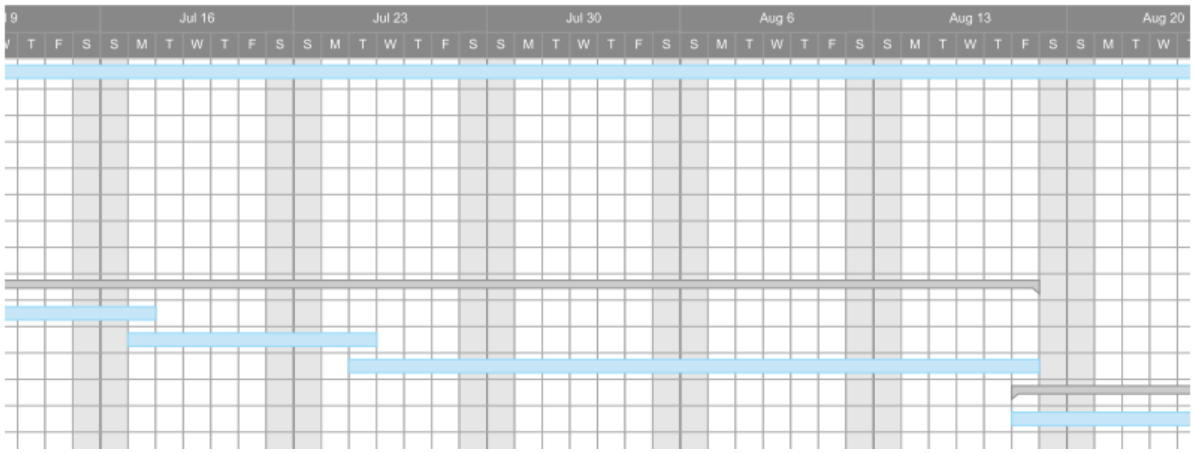
MILESTONE CHART

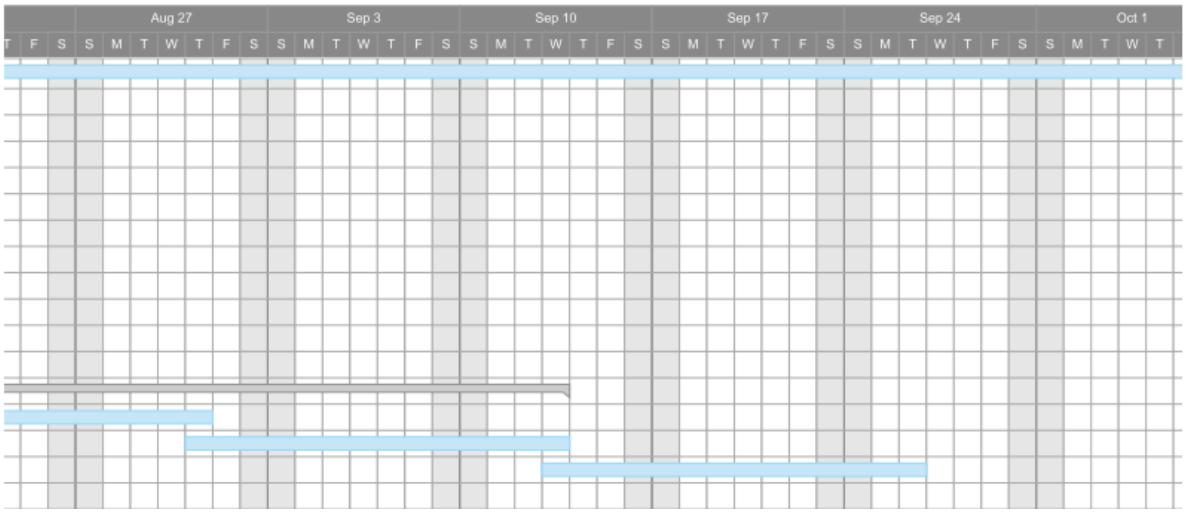
smartsheet

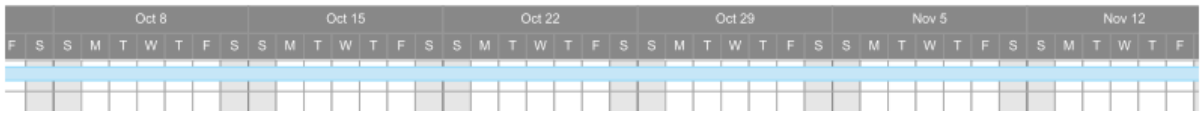


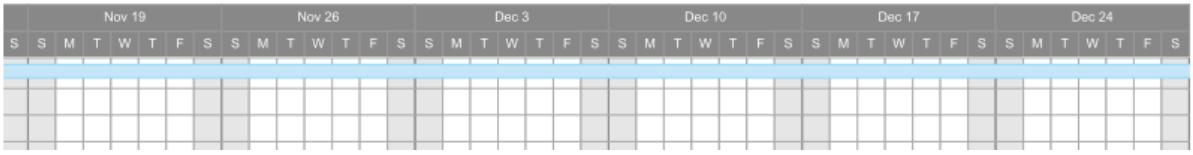


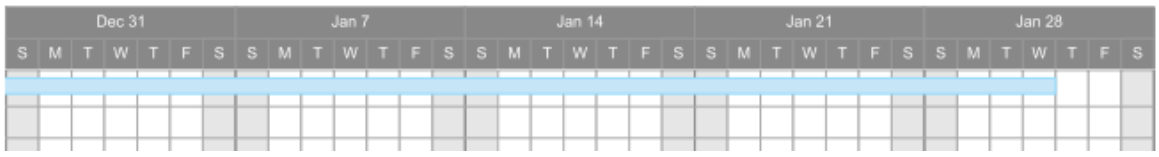



Jul 16
Jul 23
Jul 30
Aug 6
Aug 13
Aug 20


Aug 27
Sep 3
Sep 10
Sep 17
Sep 24
Oct 1


Oct 8
Oct 15
Oct 22
Oct 29
Nov 5
Nov 12


Nov 19
Nov 26
Dec 3
Dec 10
Dec 17
Dec 24


Dec 31
Jan 7
Jan 14
Jan 21
Jan 28

6.0 Project Execution

Project coding

Requirements and dependencies :

```
click==8.1.3
flask==2.1.2
importlib-metadata==4.11.4
itsdangerous==2.1.2
Jinja2==3.1.2
joblib==1.1.0
MarkupSafe==2.1.1
numpy==1.22.4
pandas==1.4.2
python-dateutil==2.8.2
pytz==2022.1
scikit-learn==1.1.1
scipy==1.8.1
six==1.16.0
threadpoolctl==3.1.0
Werkzeug==2.1.2
zipp==3.8.0
gunicorn
matplotlib==3.5.2
streamlit==1.10.0
streamlit-option-menu==0.3.2
```

```
from flask import Flask, render_template, url_for, request
import pandas as pd
import pickle
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score
```

Python source code:

1. This block of code creates an object app which is accessed remotely through localhost that executes a python script and displays it into the html application. The predict function is created with loading data that contains ham and spam classified emails. It labels the spam mail as '0' and ham email as '1'. The data is separated into texts and labels before splitting into training and test data to train the model.

```
app=Flask(__name__)

@app.route('/')
def home():
    return render_template('home.html')

@app.route('/predict', methods=['POST'])
def predict():


    raw_mail_data = pd.read_csv('mail_data.csv')


    mail_data = raw_mail_data.where((pd.notnull(raw_mail_data)),'')


    mail_data.loc[mail_data['Category'] == 'spam', 'Category',] = 0
    mail_data.loc[mail_data['Category'] == 'ham', 'Category',] = 1

    X = mail_data['Message']
    Y = mail_data['Category']

    X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=3)
```

2. The next step is to perform feature extraction using TF-IDf vectorizer to get values that will be used as the inputs for the logistic regression model. The messages from emails are in string

format and need to be fit into feature vectors by converting them to integer. The labels must also be converted into integers which are the Y_train and Y-test variables. The regression based model must be trained using training data before being tested using both training and test data to obtain the accuracy.

```
feature_extraction = TfidfVectorizer(min_df = 1, stop_words='english', lowercase='True')

X_train_features = feature_extraction.fit_transform(X_train)
X_test_features = feature_extraction.transform(X_test)

Y_train = Y_train.astype('int')
Y_test = Y_test.astype('int')

model = LogisticRegression()

model.fit(X_train_features, Y_train)

prediction_on_training_data = model.predict(X_train_features)
accuracy_on_training_data = accuracy_score(Y_train, prediction_on_training_data)

prediction_on_test_data = model.predict(X_test_features)
accuracy_on_test_data = accuracy_score(Y_test, prediction_on_test_data)
```

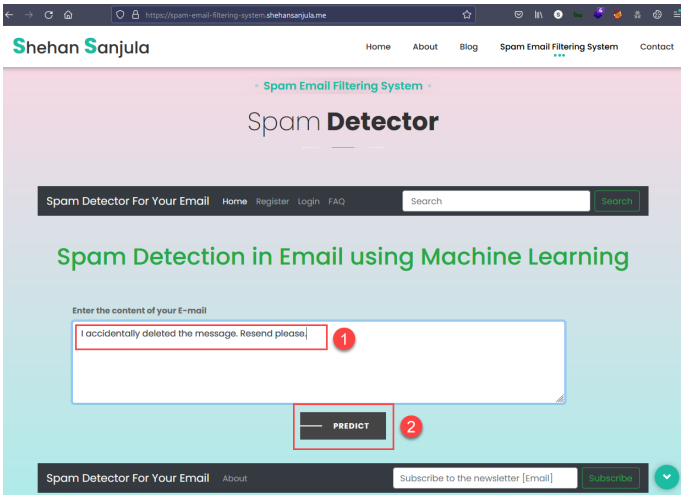
3. The last step is to link the prediction function when the user presses the button predict in the html application. The text are converted into feature vectors before being used as an object by the prediction function. The app will execute the python blocks when the main is called from the html script.

```
    if request.method=='POST':
        comment=request.form['comment']
        data=[comment]


        input_data_features = feature_extraction.transform(data).toarray()


        my_prediction = model.predict(input_data_features)

    return render_template('result.html', prediction=my_prediction)

if __name__ == '__main__':
    app.run(debug=True)
```

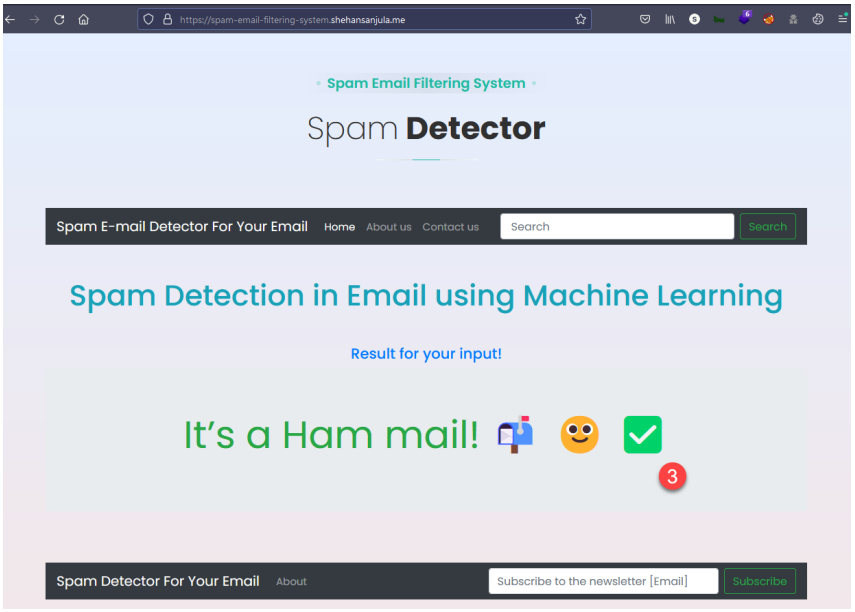
Project output

Output 1: The output shows the model predicting an email text that is supposed to be ham

The image 1 shows the user entering an email text and clicking predict button

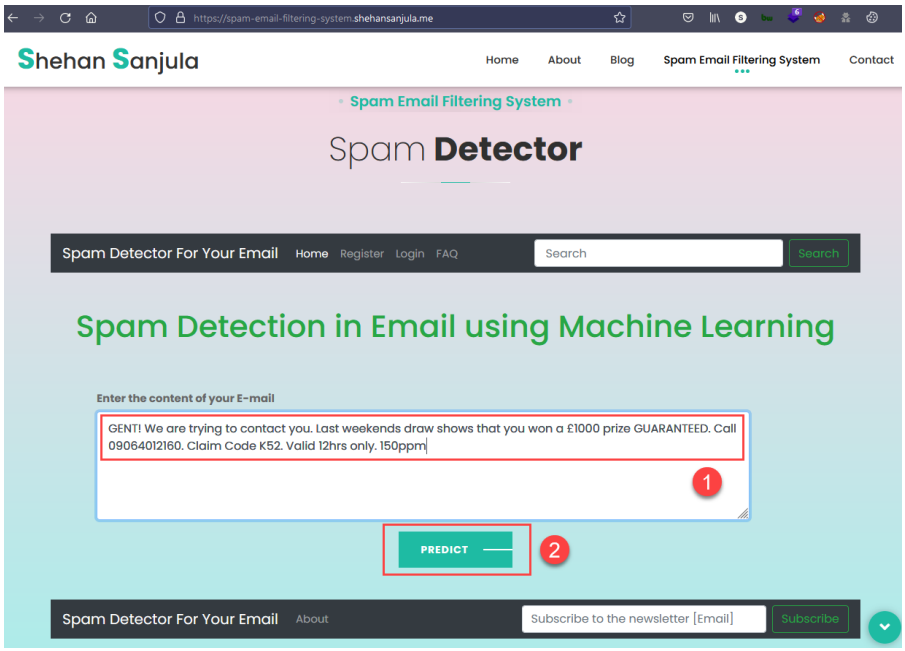


The image 2 shows the model accurately classifying that the email is ham

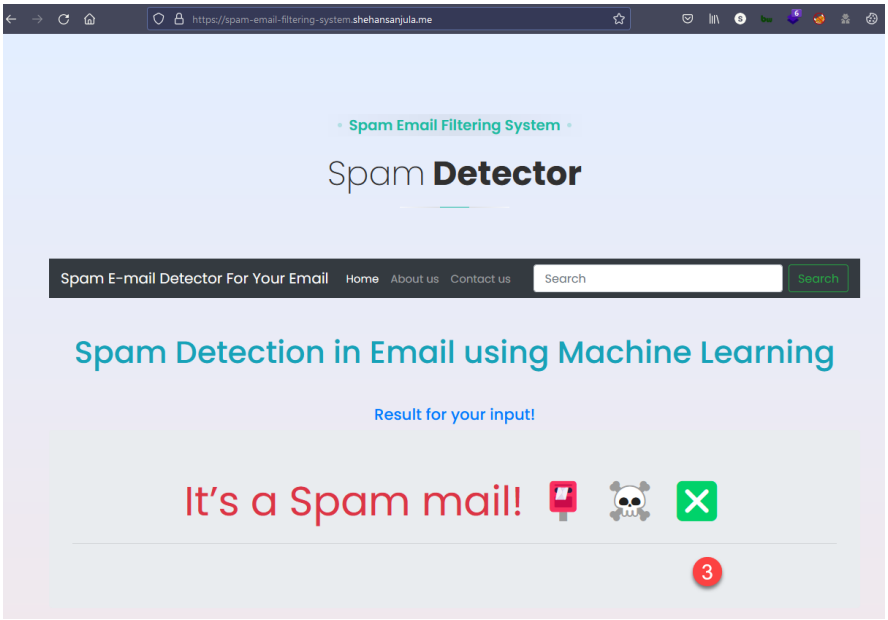


Output 2: The output shows the model predicting an email text that is supposed to be spam

The image 3 shows the user entering an email text and clicking predict button



The image 4 shows the model accurately classifying that the email is spam



7.0 Project Closing

A) Customer Acceptance Form

Customer Acceptance/Project Completion Form

February 20, 2023

Project Name: **Spam Detection in Email Using Machine Learning**

Project Leader: **KAVINRAJ A/L SIVAKUMAR**

I (We), the undersigned, acknowledge and accept delivery of the work completed for this project on behalf of our organization. My (Our) signature(s) attest(s) to my (our) agreement that this project has been completed. No further work should be done on this project.

Name	Title	Signature	Date
Ali Bin Abdul Bakar	Training Director	Ali	February 20, 2023

1. Was this project completed to your satisfaction? ___X____ Yes _____ No

2. Please provide the main reason for your satisfaction or dissatisfaction with this project.

The project met and exceeded my expectations. In my 15 years with this company. I have never seen workers so interested in developing software projects. Kavinraj effectively coordinated all of the people who worked on this project. We worked with several new suppliers, and many things went smoothly.

3. Please provide suggestion on how our organization could improve its project delivery capability in the future.

One suggestion would be to try to improve our estimating and forecasting abilities. The project cost was slightly over budget, even with some reserve built in. Not only that, but the schedule was also slightly late than the expected schedule. The schedule buffer prevented the project from finishing late. We also need to improve the way we forecast the number of people who undertake the projects. The demand for the Spam and filter were higher than expected therefore server maintenance was much higher than expected. Even though that was a pleasant surprise, it was still poor forecasting and causing extra work and budget expenses for the project and support staff.

Thank you for your inputs.

Lesson-Learned Documents

Lesson-Learned Report

February 20, 2023

Project Name: Spam Detection in Email Using Machine Learning

Project Sponsor: Ali bin Abdul Bakar

Project Leader: KAVINRAJ A/L SIVAKUMAR

Project Dates: 1 December 2021 - 27 February 2023

Final Budget: RM 264,560

---

1. Did the project meet scope, time and cost goals?

We did meet scope and time goals, but we had to request an additional RM, which the sponsor approve. We exceeded scope goals by having more people undertake the software projects than planned, primarily in the Spam filtering model.

2. What was the success criteria listed in the project scope statement?

The following statement outlined the project scope and success criteria:

"Our sponsor has stated that the project will be a success if the project are available within one year, if the average project evaluations are at least 3.0 on a 1-5 scale, and if the company recoups the cost of the project in reduced training cost within two years after project completion".

3. Reflect on whether or not you met the project success criteria.

The product was offered well within a year, and the course evaluation averaged 3.1 on a 5.0 scale. We do not know if the project cost will be recouped within two years after completion, but the number of initial customers utilizing the product slightly exceed our expectations. We have slight confidence that the costs will be recouped in two years' time.

4. What were the main lessons your team learned from this project?

The main lessons we learned include the following:

Having good communication was instrumental to project success. We had a separate item in the WBS for stakeholder communications, which was very important. Moving from traditional code-based system to web-based system was a big change for the project, so strong communication played a crucial role on determining the success of the projects. The intranet-site information was great, thanks to support from the IT department. It was also very effective to have different departments create project description posters to hang in their works areas. They showed creativity and team spirit causing increase in the productivity in the teams.

3.0 Final Project Report

Spam Detection in Email Using Machine Learning

1.0 PROJECT OBJECTIVES

The Spam Detection in Email Using Machine Learning done by the KFT Inc. was intended to develop an artificial intelligence system that is able to detect spam from the email. Besides, one of the aims of Company is to develop a system that is able to recognize the characteristics of key spam words. Not only that, but we also intend to build a web-based application programming interface for future ease of use as the integration module.

2.0 SUMMARY OF PROJECTS RESULTS

After the system has been completed, we tested the accuracy of the system. The system has an accuracy of 65%

3.0 ORIGINAL AND ACTUAL SCHEDULE

Actual Schedule

Create Design Specification	03/24/23	03/30/23	5d
Deliver Final Design	03/31/23	04/05/23	4d
Execute	05/01/23	02/05/24	201d
EDA (Exploratory data analysis)	05/01/23	06/02/23	25d

Original

Execute	04/26/23	01/31/24	201c
EDA (Exploratory data analysis)	04/26/23	05/30/23	25d
Data pre-processing	05/31/23	06/15/23	12d
Feature Extraction	06/15/23	08/31/23	56d
Evaluation Module	09/01/23	12/29/23	86d
Improvement by using Embedding + Neural Network	09/01/23	10/27/23	41d
Comparison of ML algorithm & Deep Learning	10/27/23	12/29/23	46d
Testing	01/01/24	01/31/24	23d
System Testing	01/01/24	01/15/24	11d
Project benefit Measurement	01/16/24	01/31/24	12d

The project's members mostly stick to the Gantt Chart schedule so the actual progress fits the original schedule with exception in the designing phases where we are slightly late than expected by a day.

4.0 ORIGINAL AND ACTUAL BUDGET

	OVERALL PROJECT Cost	GRANT REQUEST Amount	PRISM MATCH Amount	MATCH NOT IN PRISM Amount
Sheet #1 Acquisition				
Property Costs	$ 20,570	$ 22,000	$ -	$ -
Incidental Costs	$ 24,057	$ 24,000	$ -	$ -
Administrative Costs	$ 50,300	$ 50,000	$ -	$ -
Indirect Costs	$ 5,000	$ 4,500	$ -	$ -
STotal	$ 99,927	$ 100,500	$ -	$ -
Sheet #2 Design				
Design Costs	$ 181,300	$ 214,000	$ (7,150)	$ 18,750
Indirect Costs	$ 8,660	$ 5,000	$ -	$ -
STotal	$ 189,960	$ 219,000	$ (7,150)	$ 18,750
Sheet #3 Restoration				
Construction Costs	$ 67,850	$ 80,000	$ -	$ -
AA&E	$ 9,750	$ 10,000	$ -	$ -
Indirect Costs	$ 1,920	$ 2,000	$ -	$ -
STotal	$ 79,520	$ 92,000	$ -	$ -
GTOTAL	**$ 369,407**	**$ 411,500**	**$ (7,150)**	**$ 18,750**

Total PRISM Project Budget	**$ 404,350**

The actual budget mostly matches the original budget, which means that the team has followed the budget planned at the beginning of the project and does not use any excessive money with the exception being in the indirect cost mainly in the server maintenance.

5.0 PROJECT ASSESSMENT

5.1 SCOPE

The project scope was initiated from the beginning of the project. For instance, it isn't publicly easy for any user to utilize any third-party spam filtering. Most of the spam filtering services come mostly within other services restricting its usage outside of the environment. So, the system will build upon the web-services so any user can utilize it. Not only that, but the spam

filtering also only works on email that uses English language as the main mode of communication. Therefore, the projects limit down to scope to identify English alphabets and number characters. Therefore, this system cannot be used when another language is embedded in the email.

5.2 OBJECTIVES

KFT Inc. managed to achieve the objectives initialized since the beginning of the projects. A system that is artificial intelligence powered, which utilizes machine learning techniques has been made. Machine learning is used on this project to identify which words contributed to the email being spam and non-spam. Besides, the algorithm is also embedded into the web-services allowing one of the criteria of the project to be met.

5.3 IMPACTS

By creating this system, the accessibility of the user to determine whether their selected email is spam or not have been increased twofold. This accessibility allows the user to independently filter their email and any written mail easily without necessarily relying on pre-built services such as Gmail or outlook. Not only that, any company that wants to utilize the system for their own purpose to build it into their own services can also be done and plucked straight from our system. This allows independent parties or companies to use our services without any difficulties since they are dealing in the same vein as the third-party company.

6.0 TRANSITION PLAN

KFT Inc. staff will provide support when require. The main work of support provided by KFT Inc. staff on the Spam Filter recognition system projects are as shown below:

i) Maintain the system.

ii) Handling errors and updates of the system.

iii) Provide user guidelines for ASD staff to use the system.

iv) Report information to senior management on monthly basis.

Close Contract

CLOSE CONTRACT

KFT Inc. Contract Closure Notice

2/1/2023

This letter provides formal notice that the work contracted with ASD Co. (M) BHD has been completed. KFT Inc. has developed a The Spam Detection in Email Using Machine Learning that can detect and filter any spam email from any email repository given by the user.

Kavinraj, the project manager, has provided the following performance assessment for the cooperation provided:

"We were very pleased with the cooperation provided by the ASD Co. (M) BHD. The staff from ASD have given full cooperation and willing to have a good and stable communication with KFT Inc. to ensure that the entire project can be executed in a smooth manner and pace. We were extremely happy to work together with ASD Co. (M) BHD and looking forward for the next collaboration.

By: KAVINRAJ A/L SIVAKUMAR, CEO, KFT Inc.

Date: 2/1/2023