SPAMSENSE: SISTEM PENGESANAN MESEJ SPAM

Ummi Anas Azhar

Fakulti Teknologi dan Sains Maklumat

43600 Universiti Kebangsaan Malaysia

Abstrak

SpamSense ialah sistem pengesanan mesej spam berasaskan web yang dibangunkan menggunakan pendekatan gabungan Naive Bayes dan Model Bahasa Besar (LLM) bagi meningkatkan ketepatan klasifikasi mesej spam. Dengan menggunakan sistem ini, pengguna dapat mengenal pasti mesej yang mungkin spam melalui analisis yang lebih mendalam dan tepat. SpamSense dibina menggunakan rangka kerja Flask dan mengaplikasikan seni bina *Model-View-Controller (MVC)* untuk memastikan modulariti dan fleksibiliti sistem yang mudah diselenggara dan dikembangkan. Sistem ini menawarkan ciri-ciri seperti klasifikasi mesej secara masa nyata, penyerlahan perkataan penting yang dikenal pasti sebagai spam, penjelasan (*reasoning*) kepada mesej spam, dan kemas kini pangkalan data secara dinamik berdasarkan input pengguna untuk menangani corak spam baharu.

Keistimewaan utama sistem SpamSense terletak pada gabungan antara model Naive Bayes dan Model Bahasa Besar (LLM), di mana Naive Bayes digunakan untuk klasifikasi pantas berdasarkan kebarangkalian, manakala LLM membantu dalam pemahaman konteks mesej yang lebih kompleks. Gabungan ini membolehkan sistem membuat keputusan lebih tepat dan adaptif, mengurangkan isu kesalahan klasifikasi yang sering berlaku dalam pendekatan model tunggal. Selain itu, sistem ini turut menyokong pembelajaran berterusan dengan membenarkan kemas kini perkataan spam baharu dalam pangkalan data, sekali gus menjadikan SpamSense lebih responsif terhadap mesej spam yang semakin berkembang.

Secara keseluruhannya, SpamSense memberikan penyelesaian yang lebih tepat dan mesra pengguna dalam pengesanan spam serta menyediakan pengalaman yang lebih berskala dan sesuai digunakan dalam persekitaran yang pelbagai bahasa, terutamanya dalam Bahasa Malaysia.

Abstract

SpamSense is a web-based spam message detection system developed using a hybrid approach combining Naive Bayes and Large Language Models (LLMs) to enhance the accuracy of spam message classification. With this system, users can identify potentially spam messages through deeper and more precise analysis. SpamSense is built using the Flask framework and adopts the Model-View-Controller (MVC) architecture to ensure modularity and system flexibility, making it easy to maintain and extend. The system offers features such as real-time message classification, highlighting of key words identified as spam, spam reasoning, and dynamic database updates based on user input to adapt to new spam patterns.

The main strength of SpamSense lies in its integration of the Naive Bayes model and LLM, where Naive Bayes is used for fast classification based on probabilities, while LLM aids in understanding more complex message contexts. This combination enables the system to make more accurate and adaptive decisions, reducing misclassification issues that often occur with single-model approaches. Additionally, the system supports continuous learning by allowing new spam words to be updated in the database, making SpamSense more responsive to the evolving nature of spam messages.

Overall, SpamSense provides an accurate and user-friendly solution for spam detection and offers a scalable experience suitable for multilingual environments particularly in Malay language usage.

1.0 PENGENALAN

Dalam era digital yang serba maju ini, mesej spam telah menjadi salah satu ancaman utama dalam komunikasi harian melalui saluran seperti SMS, e-mel, dan platform media sosial. Mesej spam bukan sahaja mengganggu pengguna, tetapi juga digunakan sebagai saluran untuk serangan *phishing*, *malware*, dan penipuan kewangan. Menurut laporan CyberSecurity Malaysia (2023), kes jenayah siber yang melibatkan mesej spam semakin meningkat, menuntut sistem pengesanan spam yang lebih pintar dan berkesan untuk mengenal pasti kandungan yang berniat jahat secara automatik.

Kebanyakan sistem pengesanan spam yang ada pada masa kini masih bergantung kepada pendekatan berasaskan peraturan yang tidak begitu fleksibel, terutamanya apabila berhadapan dengan mesej spam yang lebih kompleks dan tidak konvensional. Oleh itu, projek

ini bertujuan untuk memperkenalkan SpamSense, sebuah sistem yang menggabungkan teknik Naive Bayes dan Model Bahasa Besar (LLM). Naive Bayes digunakan untuk klasifikasi pantas berdasarkan kebarangkalian, manakala LLM berfungsi untuk memahami konteks mesej yang lebih mendalam dan kompleks. Gabungan kedua-dua model ini membolehkan sistem membuat keputusan yang lebih tepat dan adaptif.

Selain daripada kemampuan pengesanan spam yang lebih tepat, SpamSense turut menyediakan fungsi kemas kini pangkalan data dan penyerlahan perkataan spam secara masa nyata. Sistem ini dibangunkan menggunakan Flask dan seni bina *Model-View-Controller* (MVC) untuk memastikan sistem yang modular, mudah diurus, dan senang disesuaikan mengikut keperluan masa depan. Dengan menggunakan pendekatan hibrid ini, SpamSense menawarkan penyelesaian yang lebih moden dan efisien untuk pengesanan mesej spam.

Di samping fungsi teknikal utama, SpamSense turut memperkenalkan elemen gamifikasi melalui permainan kuiz interaktif bertemakan *"Mario Spam Hunter Quiz"*. Ciri ini dibangunkan bagi menyokong usaha pendidikan pengguna dalam mengenal pasti ciri-ciri utama mesej spam melalui pendekatan yang menyeronokkan dan mudah difahami. Kuiz ini direka dengan elemen permainan seperti pengumpulan mata, halangan visual dan animasi, serta soalan bertema spam yang dipaparkan dalam suasana permainan berasaskan platform klasik. Pendekatan gamifikasi ini bukan sahaja menarik minat pengguna, malah dapat meningkatkan kefahaman mereka terhadap teknik pengesanan spam dan risiko sebenar yang dibawa oleh mesej spam. Inisiatif ini menjadikan SpamSense bukan sahaja alat pengesanan, tetapi juga platform pendidikan atas talian yang interaktif.

2.0 KAJIAN LITERATUR

Dalam era digital yang pesat berkembang, ancaman keselamatan siber seperti mesej spam menjadi salah satu cabaran terbesar dalam komunikasi harian. Penggunaan platform komunikasi seperti e-mel, SMS, dan aplikasi sosial semakin meluas, yang secara langsung membuka ruang kepada peningkatan penyebaran mesej spam (Mohd Hamizi, M. A. F., 2023). Mesej spam sering mengandungi pautan berniat jahat, tawaran palsu, atau kandungan penipuan yang berpotensi membahayakan pengguna dari segi keselamatan dan kewangan. Ancaman ini menjadi semakin kompleks dengan mesej spam menjadi ancaman utama dalam komunikasi digital, mengandungi pautan berniat jahat dan kandungan penipuan yang membahayakan keselamatan pengguna (BH Online, 2015).

Projek ini membangunkan sistem pengesanan mesej spam menggunakan gabungan dua pendekatan utama iaitu Naive Bayes dan *Large Language Models (LLM)* (Ahmadi et al., 2025). dalam

sistem yang dinamakan ‘SpamSense’ Sistem ini dilatih dengan dataset SMS Spam Collection untuk mengklasifikasikan mesej sebagai spam atau bukan spam. Naive Bayes akan digunakan untuk membuat keputusan klasifikasi berdasarkan kebarangkalian, manakala LLM digunakan untuk memperkayakan proses pengesanan spam dengan memberikan konteks yang lebih mendalam dalam mengenal pasti pola mesej spam yang lebih kompleks. Sistem ini dibangunkan menggunakan Flask dan menyediakan antaramuka mesra pengguna untuk analisis mesej secara interaktif. Selain itu, sistem ini turut diintegrasikan dengan pangkalan data dinamik bagi menyimpan perkataan spam baharu, menjadikan sisstem lebih adaptif terhadap ancaman spam yang berubah-ubah. Pendekatan ini diharapkan dapat meningkatkan ketepatan pengesanan spam, memastikan keberkesanan jangka panjang, serta relevan untuk pelbagai platform komunikasi yang digunakan oleh pengguna.

Keperluan untuk sistem pengesanan spam yang lebih responsif, efisien, dan adaptif adalah sangat mendesak dalam memastikan komunikasi digital yang lebih selamat (System Design School, 2023). Dengan menggunakan gabungan Naive Bayes dan LLM, SpamSense bukan sahaja dapat mengenal pasti mesej spam dengan lebih tepat, tetapi juga memperbaiki cara sistem menyesuaikan diri dengan ancaman spam yang lebih baharu. Keupayaan untuk terus memperbaharui pangkalan data dengan perkataan spam yang dikumpulkan serta kemampuan LLM untuk memahami konteks mesej menjadikan SpamSense lebih fleksibel berbanding dengan sistem pengesanan spam tradisional. Projek ini bertujuan membangunkan sistem yang bukan hanya membantu pengguna mengenal pasti mesej spam, tetapi juga memastikan sistem ini tetap relevan dalam menghadapi pola spam yang berubah.

Selain daripada pengesanan teknikal, literatur terkini turut menekankan kepentingan pendidikan pengguna dalam meningkatkan kepekaan terhadap mesej spam. Pendekatan gamifikasi semakin diiktiraf sebagai strategi berkesan untuk menyampaikan kandungan pendidikan digital dengan lebih menarik dan efektif. Menurut Baah, Govender dan Subramaniam (2024), penggunaan gamifikasi mampu meningkatkan motivasi, penglibatan serta retensi kognitif dalam kalangan pelajar digital. Kajian oleh Onduto (2021) pula menunjukkan bahawa latihan keselamatan siber berasaskan permainan berjaya meningkatkan kesedaran pengguna terhadap ancaman siber seperti spam sehingga 40%. Sehubungan itu, *SpamSense* turut mengintegrasikan modul kuiz interaktif dinamakan *Mario Spam Hunter Quiz* sebagai komponen gamifikasi yang membolehkan pengguna menguji kefahaman terhadap ciri-ciri mesej spam melalui format permainan platform. Pendekatan ini menjadikan sistem bukan sahaja alat pengesanan, malah berfungsi sebagai medium literasi keselamatan siber yang menyeronokkan dan berimpak.

3.0 METODOLOGI

Pembangunan sistem *SpamSense* dijalankan secara sistematik melalui pendekatan beriterasi yang melibatkan fasa analisis keperluan, reka bentuk sistem, pembangunan aplikasi, pengujian

dan penilaian. Metodologi ini dirancang bagi memastikan setiap fungsi sistem dibangunkan secara modular, responsif dan memenuhi objektif utama pengesanan serta pendidikan mengenai mesej spam.

3.1 ANALISIS KEPERLUAN

Projek *SpamSense* dibangunkan untuk mengesan mesej spam melalui gabungan model statistik (Naive Bayes) dan model bahasa besar (LLM – DeBERTa-v3). Analisis keperluan sistem melibatkan keperluan fungsian seperti pengesanan spam, penyerlahan perkataan mencurigakan, kemaskini pangkalan data, penjanaan skor spam, dan modul kuiz. Keperluan bukan fungsian merangkumi prestasi sistem yang responsif, antara muka mesra pengguna, dan kebolehskalaan sistem untuk pelbagai bahasa serta platform.

Jadual 1.1 Keperluan Fungsian Sistem SpamSense

KEPERLUAN FUNGSIAN SISTEM (KF)	
KF 1	Memasukkan mesej teks
KF 1.1	Sistem menyediakan kotak input untuk menerima mesej teks
KF 2	Menganalisis mesej spam
KF 2.1	Sistem mengklasifikasikan mesej sebagai spam atau bukan spam
KF 3	Penyerlahan perkataan spam
KF 3.1	Sistem menyerlahkan perkataan yang menyebabkan diklasifikasikan sebagai spam
KF 4	Peratusan kebarangkalian spam
KF 4.1	Sistem mengira peratusan berdasarkan perkataan spam dikesan berbanding jumlah keseluruhan perkataan mesej.
KF 4.2	Sistem memaparkan peratusan spam kepada pengguna bersama hasil klasifikasi.
KF 5	Mengemaskini pangkalan data dengan perkataan spam baharu
KF 5.1	Sistem menyimpan perkataan spam yang dikesan ke dalam pangkalan data
KF 6	Hiburan dan Pendidikan (Permainan Kuiz – Mario Spam Hunter)
KF 6.1	Sistem menyediakan permainan kuiz yang bertujuan untuk mendidik pengguna tentang mesej spam.
KF 6.2	Sistem memaparkan soalan dalam bentuk kuiz interaktif semasa permainan.

KF 6.3	Sistem menyimpan skor dan nama pemain dalam papan markah tempatan (localStorage).

Jadual 3.2 Keperluan bukan Fungsian Sistem SpamSense

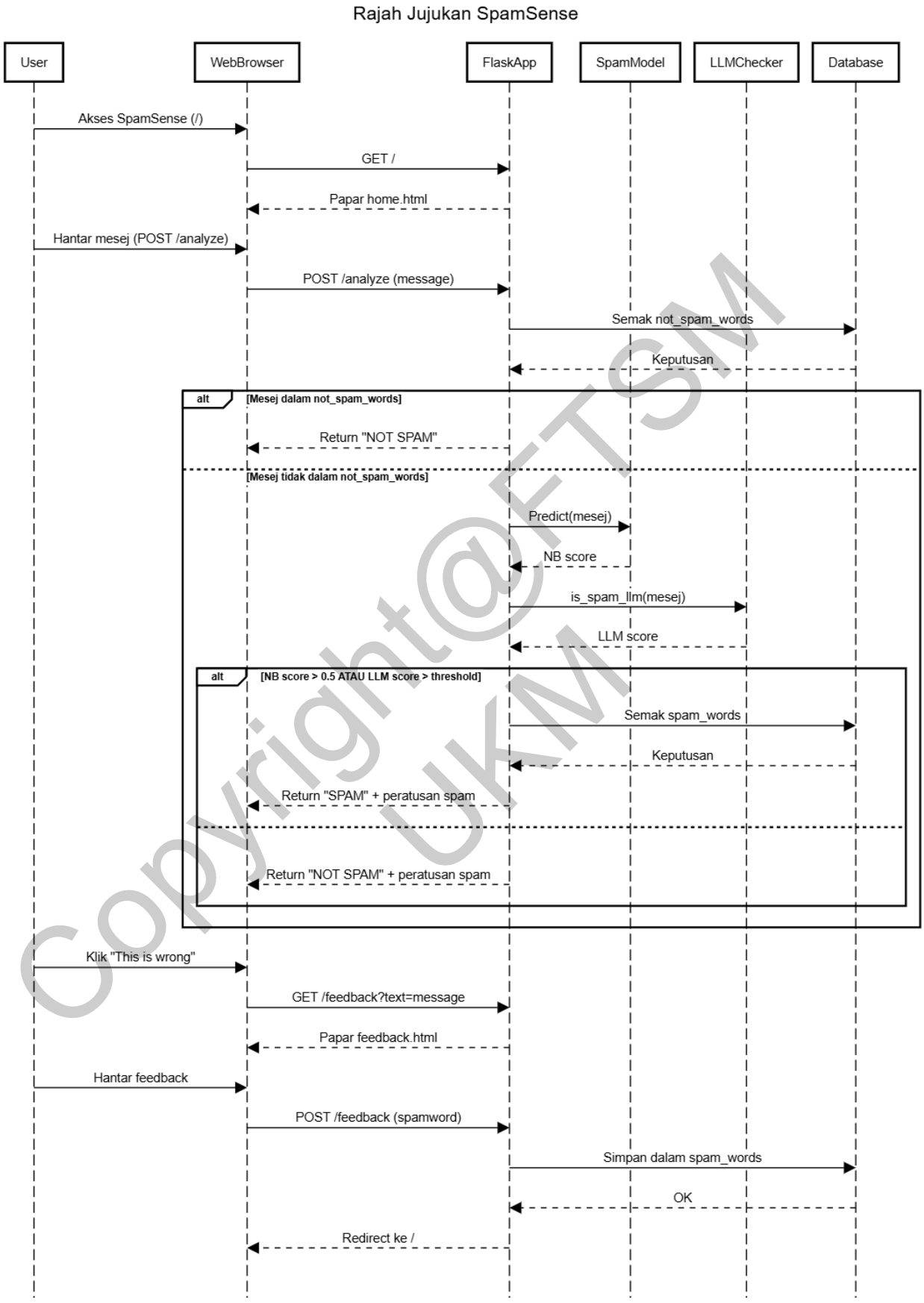
KEPERLUAN BUKAN FUNGSIAN SISTEM (KBF)	
KBF 1	Prestasi
KBF 1.1	Sistem memproses input dan memberikan hasil yang benar dan relevan.
KBF 2	Kebolehgunaan
KBF 2.1	Antaramuka sistem mudah digunakan, mesra pengguna, dan responsif
KBF 3	Pengalaman Pengguna
KBF 3.1	Sistem menyediakan elemen gamifikasi bagi menarik minat pengguna untuk memahami ciri-ciri mesej spam.
KBF 3.2	Permainan kuiz direka bentuk dengan antaramuka yang menarik dan responsif.
KBF 4	Skalabiliti
KBF 4.1	Sistem boleh diintegrasikan dengan pangkalan data dan dapat menyesuaikan diri dengan peningkatan keperluan

Sistem direka menggunakan Flask sebagai backend dan SQLite sebagai pangkalan data, dengan model Naive Bayes disimpan melalui modul pickle. LLM dilaksanakan melalui fungsi *is_spam_llm()* berasaskan transformers dari pustaka *Hugging Face*. Antara muka dibina dengan HTML, CSS (Bootstrap 5) dan JavaScript untuk menyokong interaksi dan visualisasi hasil klasifikasi spam.

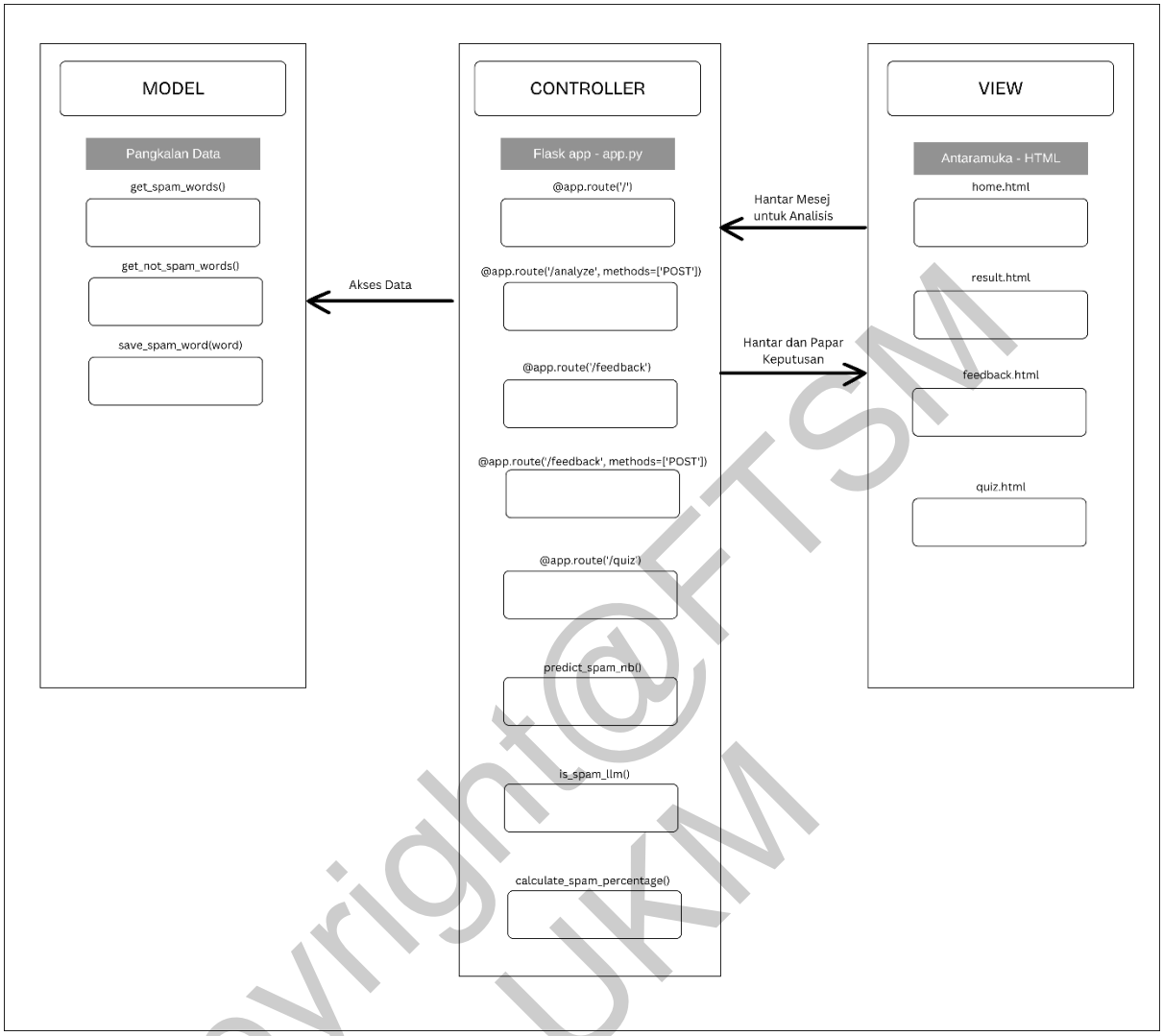
3.2 REKA BENTUK MODEL KONSEPTUAL

Sistem SpamSense direka menggunakan seni bina *Model-View-Controller* (MVC) yang memisahkan antara logik pemprosesan, paparan, dan pengurusan data. Reka bentuk ini membolehkan pembangunan sistem dilakukan secara modular dan memudahkan proses penyelenggaraan serta peluasan fungsi di masa hadapan. Lapisan pengawal dikendalikan melalui fail app.py yang mengurus laluan seperti /, */analyze, /feedback*, dan /*quiz*, manakala lapisan model bertanggungjawab terhadap operasi utama sistem seperti pengesanan spam, pengiraan skor, dan kemaskini pangkalan data. Antaramuka sistem pula dibina dengan prinsip minimalis yang memudahkan pengguna menjalankan analisis mesej atau menyertai kuiz interaktif. Komponen utama sistem terdiri daripada beberapa kelas yang saling berinteraksi,

antaranya kelas *User* yang mewakili pengguna sistem untuk memasukkan mesej dan melihat keputusan, serta kelas SpamSense sebagai komponen pusat yang menjalankan proses analisis melalui fungsi *runAnalysis()* dan *displayResult()*. Untuk tugas klasifikasi, sistem menggunakan kelas NaiveBayes yang melaksanakan fungsi *predict_spam_nb()* bagi menentukan sama ada mesej tersebut spam atau tidak, dan kelas LLM yang menyediakan fungsi *is_spam_llm()* dan *calculate_spam_percentage()* bagi analisis kontekstual dan pengiraan skor spam. Data yang dikendalikan oleh sistem disimpan dan dikemaskini melalui kelas *Database,* termasuk perkataan spam dan bukan spam yang boleh dimuat naik oleh pengguna melalui fungsi *save_spam_word()* dan dipanggil semula menggunakan *get_spam_words()*. Tambahan pula, sistem mengandungi kelas *Quiz* yang merangkumi fungsi permainan seperti *startQuiz(), submitQuiz(),* dan *saveScoreToLocal()* yang memberikan pengalaman pembelajaran interaktif berasaskan kuiz bertemakan permainan Super Mario. Keseluruhan reka bentuk ini menghasilkan satu sistem pengesanan spam yang menyeluruh, adaptif, dan mendidik.



Rajah 3.1 Rajah Jujukan SpamSense



Rajah 3.4 Reka Bentuk Seni Bina MVC SpamSense

4.0 HASIL

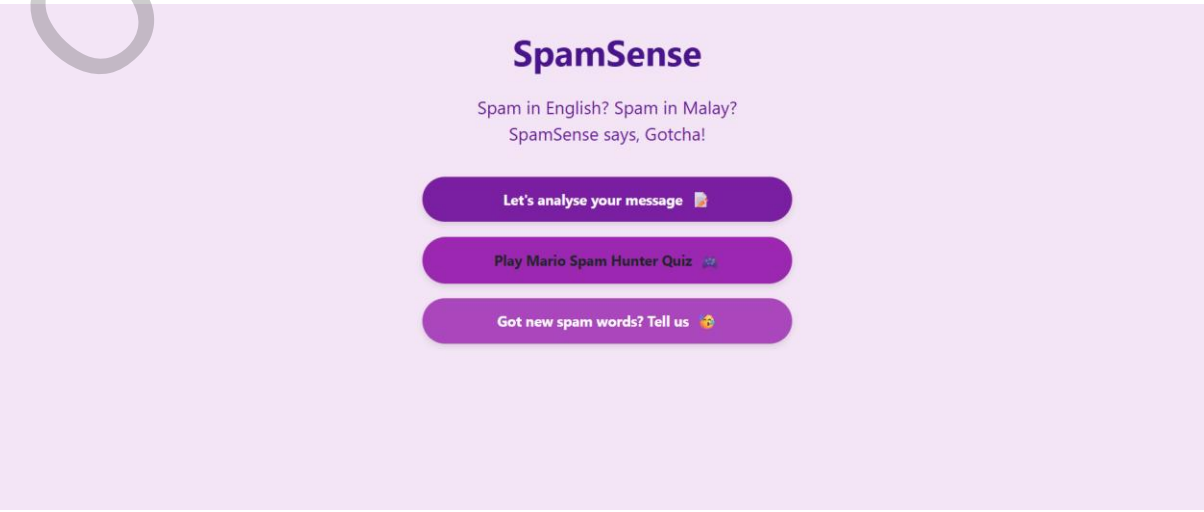
4.1 PEMBANGUNAN APLIKASI

SpamSense dibangunkan secara berfasa bermula daripada peringkat reka bentuk antaramuka sehinggalah kepada pembangunan logik sistem secara menyeluruh. Dalam pembangunan sistem ini, beberapa teknologi dan perisian digunakan untuk menyokong keperluan teknikal dan fungsian sistem. Bagi menghasilkan antara muka pengguna yang mesra dan responsif, teknologi seperti HTML, CSS (Bootstrap 5), dan JavaScript digunakan. Reka bentuk ini menekankan elemen minimalis bagi memudahkan pengguna memasukkan mesej, menerima keputusan klasifikasi spam, dan memberikan maklum balas secara terus melalui pelayar web.

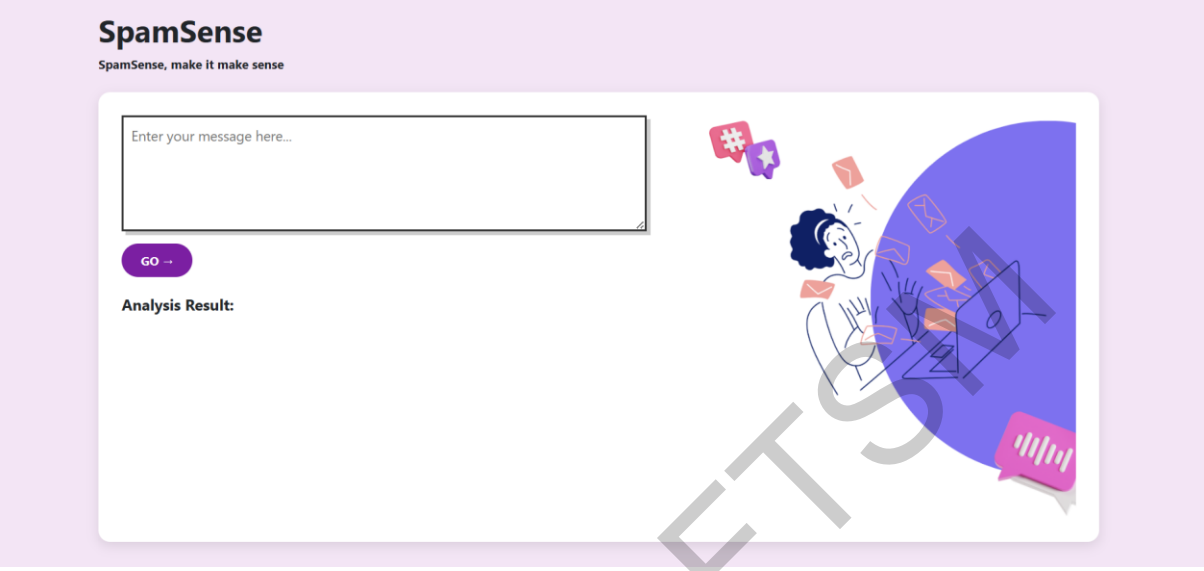
Bahagian backend sistem dibangunkan menggunakan kerangka kerja Flask, yang menyediakan sambungan antara pengguna dan logik pemprosesan model pembelajaran mesin. Untuk fungsi klasifikasi mesej, sistem menggabungkan dua pendekatan iaitu model statistik Naive Bayes dan model bahasa besar DeBERTa-v3. Model Naive Bayes digunakan untuk membuat klasifikasi awal berdasarkan kebarangkalian perkataan, manakala model DeBERTa-v3 dilaksanakan melalui fungsi *is_spam_llm()* yang menggunakan pendekatan zero-shot classification bagi memahami konteks mesej. Fungsi *calculate_spam_percentage()* digunakan untuk mengira skor spam akhir melalui gabungan kedua-dua model tersebut. Bagi pengurusan data, sistem menggunakan SQLite sebagai pangkalan data utama bagi menyimpan perkataan spam, bukan spam dan maklum balas pengguna.

Menariknya, sistem ini turut disokong oleh modul permainan kuiz interaktif *bertemakan Super Mario* sebagai pendekatan gamifikasi dalam pendidikan keselamatan siber. Permainan kuiz ini dibangunkan menggunakan HTML, CSS dan JavaScript sepenuhnya, serta menggunakan *localStorage* untuk menyimpan skor pemain. Pengguna boleh memilih nama dan bahasa (Bahasa Melayu atau Inggeris) sebelum menjawab soalan kuiz berkaitan ciri-ciri mesej spam. Modul ini bukan sahaja meningkatkan penglibatan pengguna, tetapi juga berfungsi sebagai alat literasi digital bagi mendidik pengguna tentang ancaman spam secara lebih menyeronokkan (Baah et al. 2024; Onduto 2021). Seluruh aplikasi dibina secara modular agar mudah untuk dikembangkan dan diintegrasikan dengan komponen tambahan pada masa hadapan.

Antaramuka Spamsense:

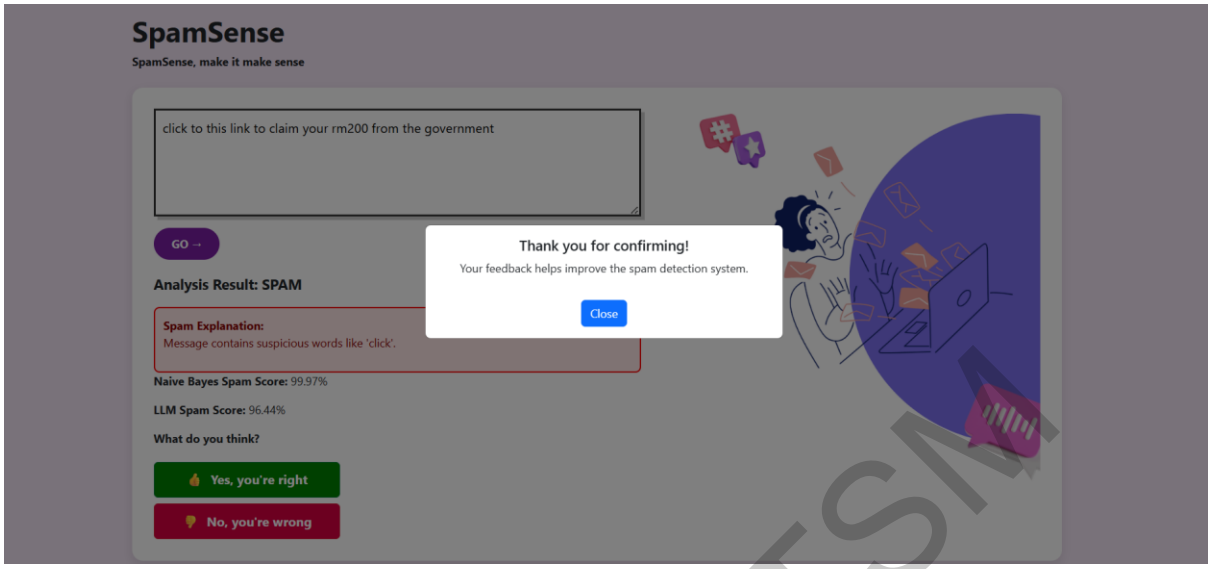


Rajah 4.1 Antaramuka laman utama SpamSense

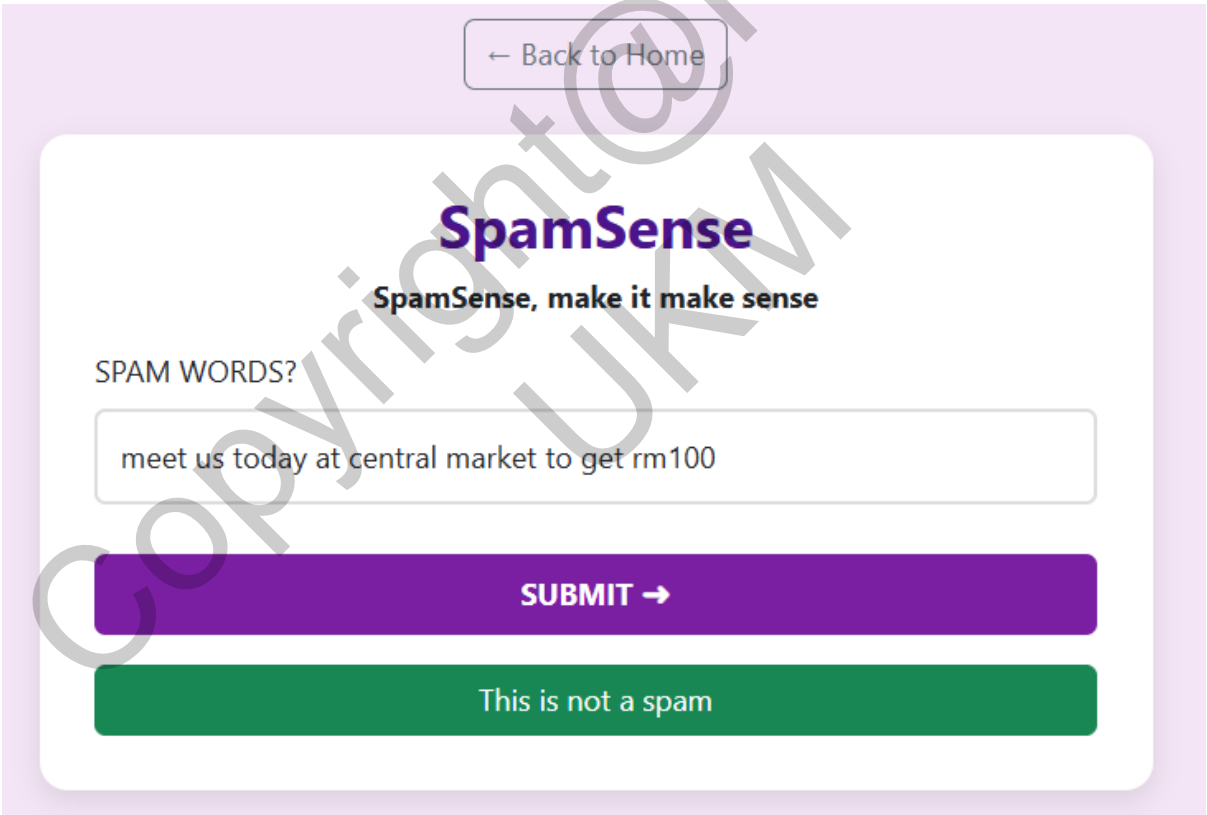


Rajah 4.2 Antaramuka utama

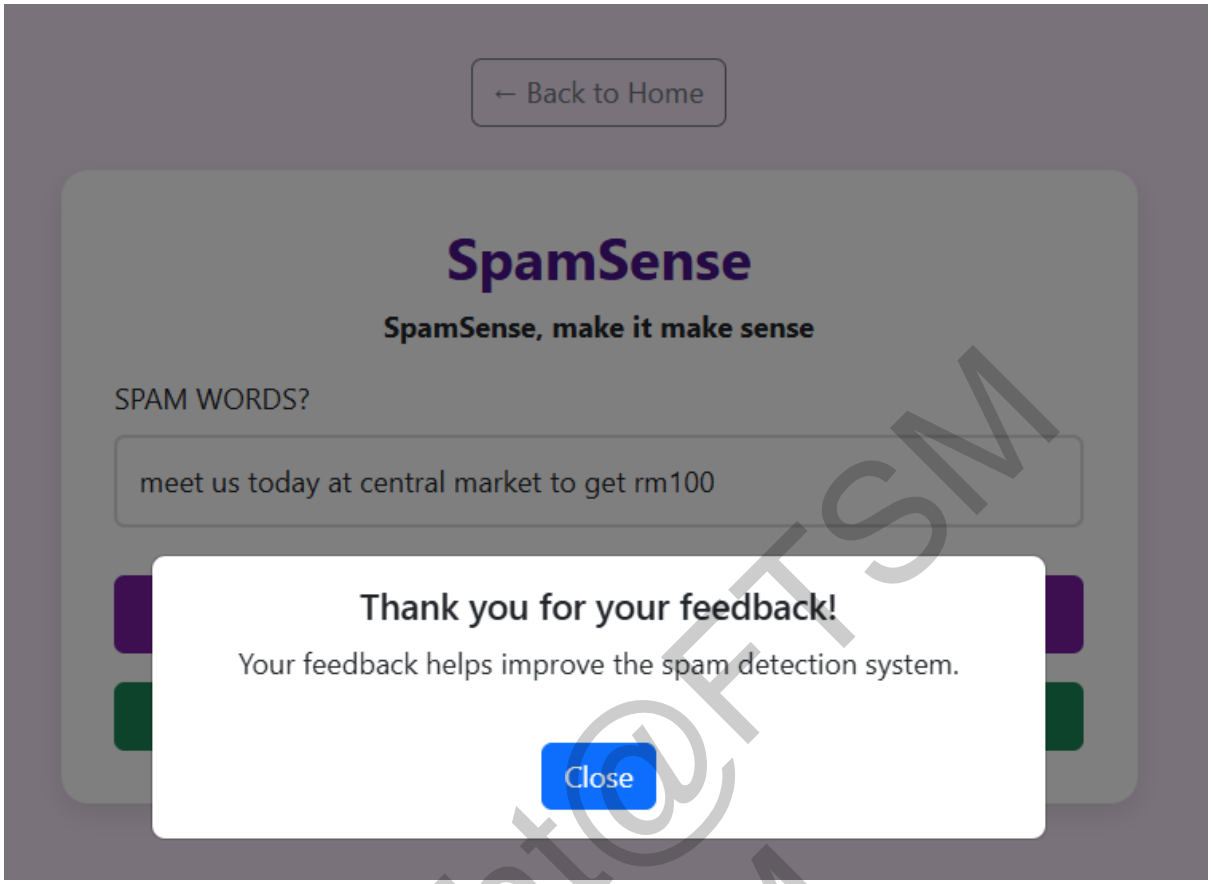
Rajah 4.3 Antaramuka apabila SpamSense telah berjaya menganalisa teks dan mengeluarkan keputusan 'Spam'



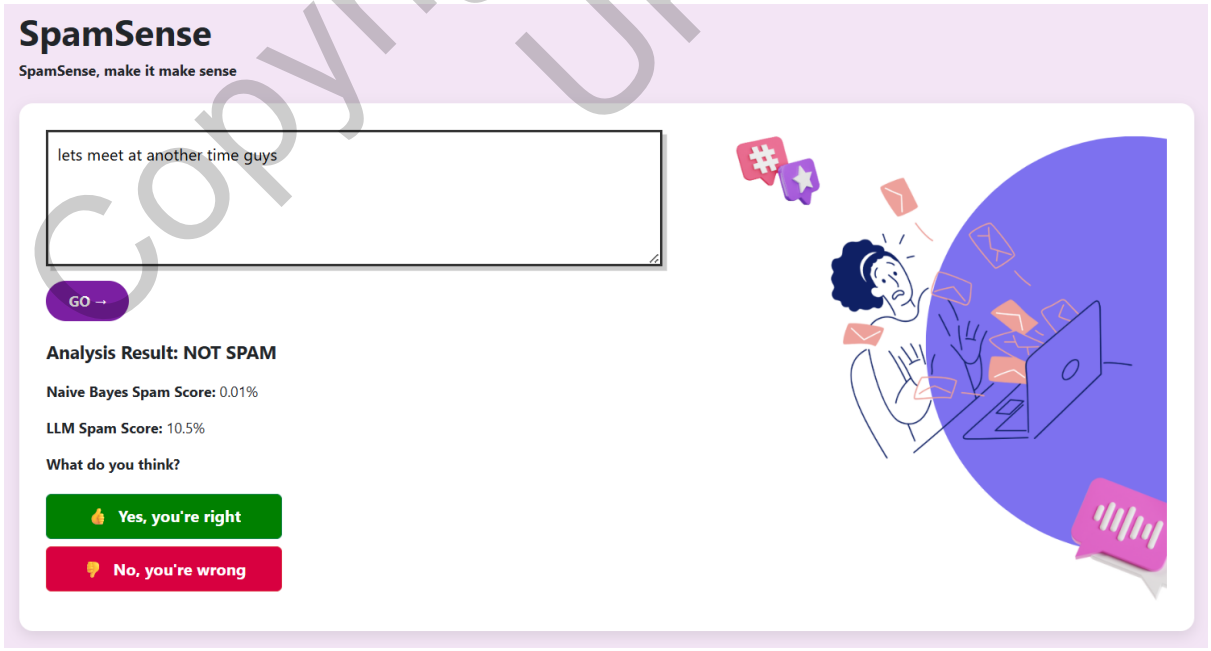
Rajah 4.4 Antaramuka apabila butang "Yes, you're right" ditekan



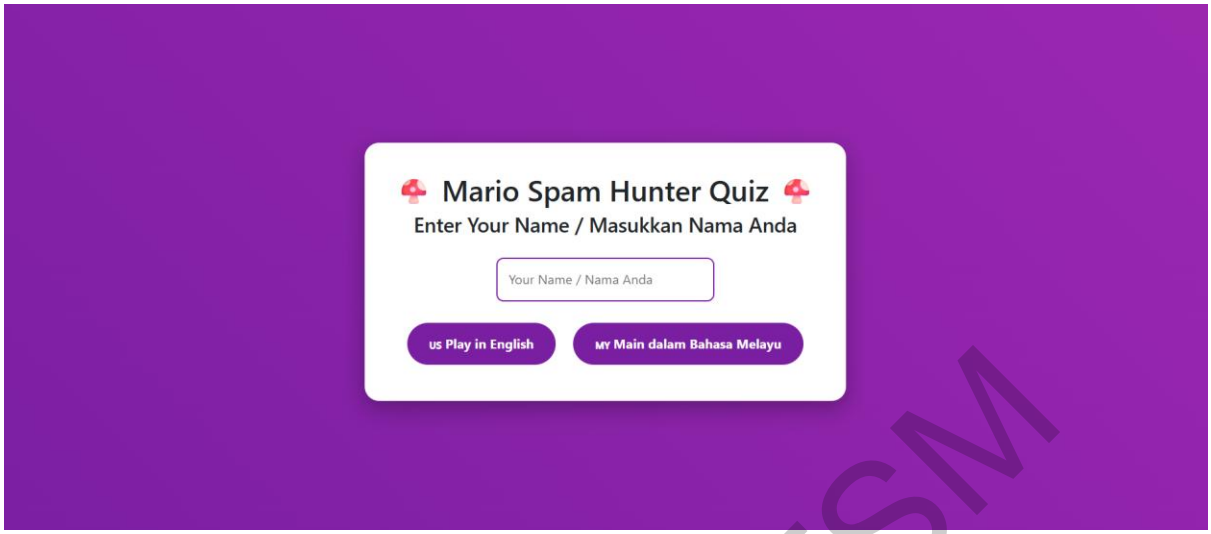
Rajah 4.5 Antaramuka apabila butang "No, you're wrong" ditekan



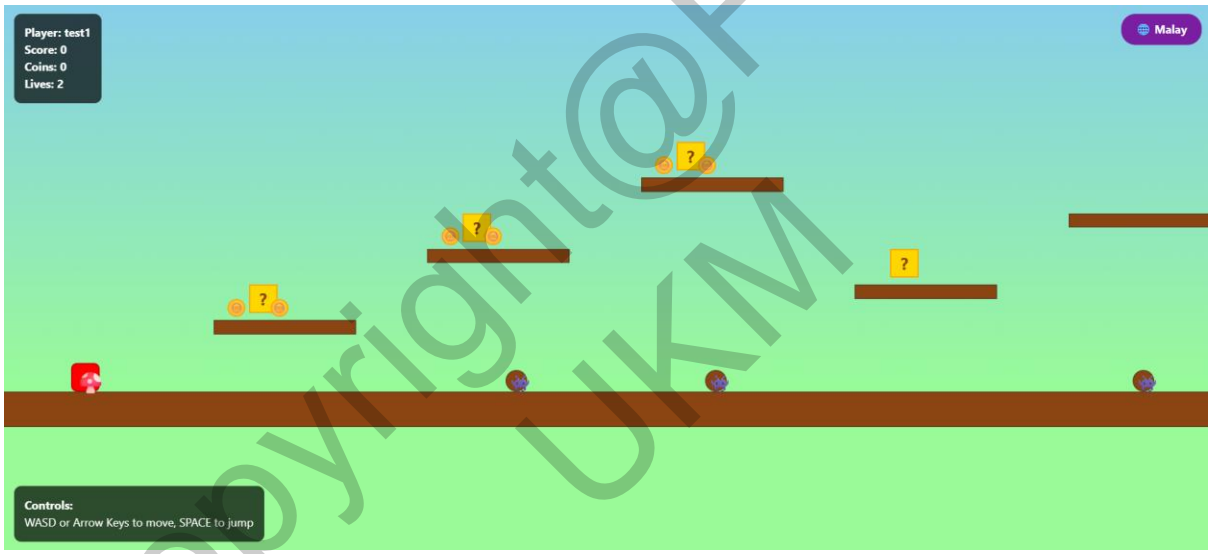
Rajah 4.6 Antaramuka apabila butang "SUBMIT" ditekan. Mesej spam telah dimasukkan ke dalam pangkalan data



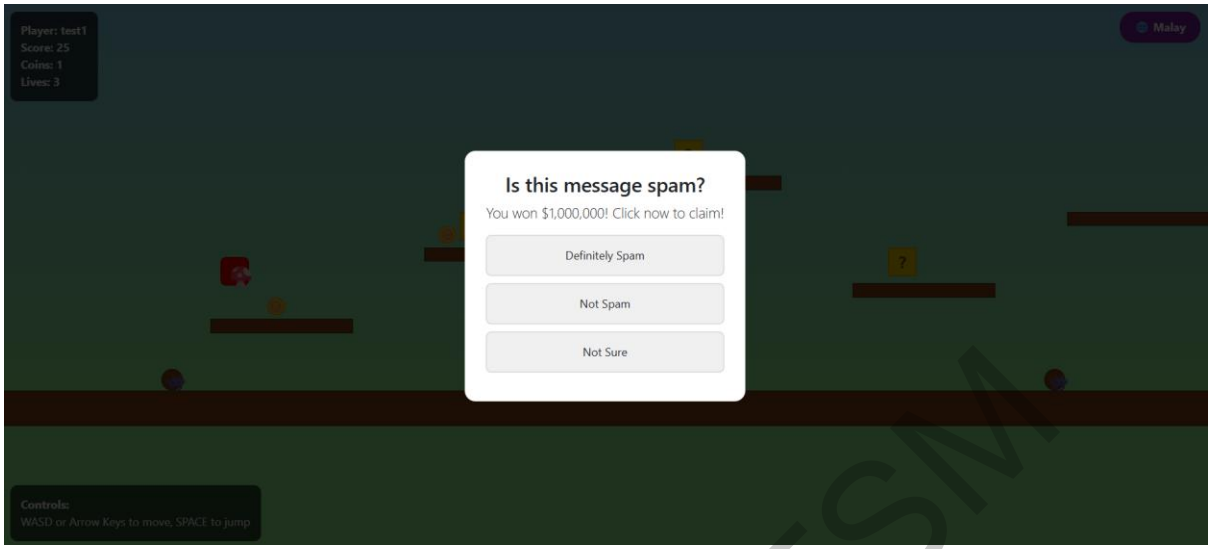
Rajah 4.7 Antaramuka apabila SpamSense telah berjaya menganalisa teks dan mengeluarkan keputusan 'Not Spam'



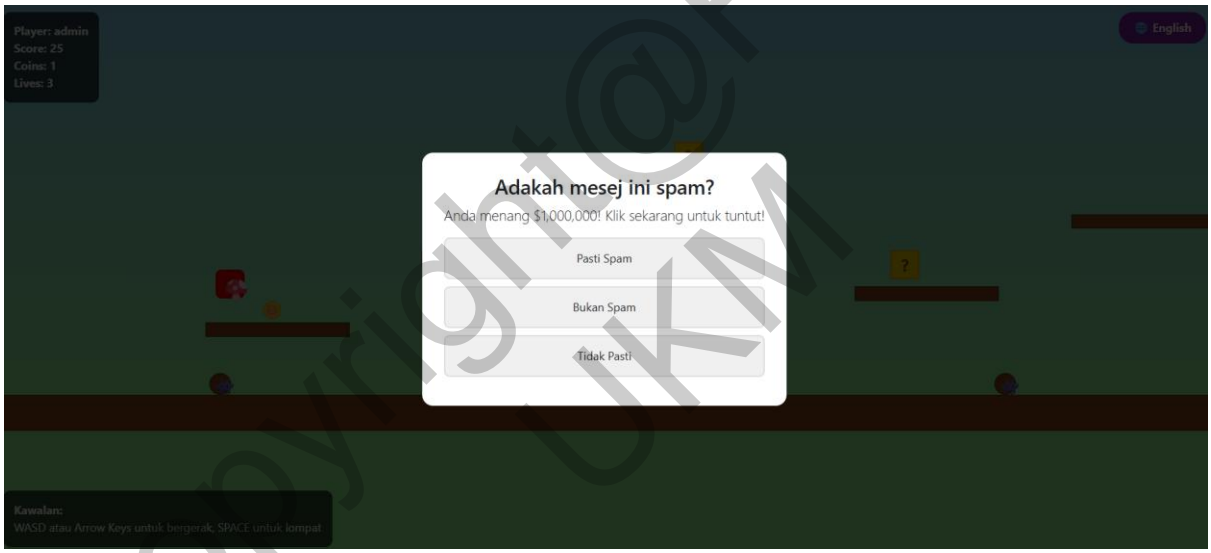
Rajah 4.8 Antaramuka utama permainan kuiz



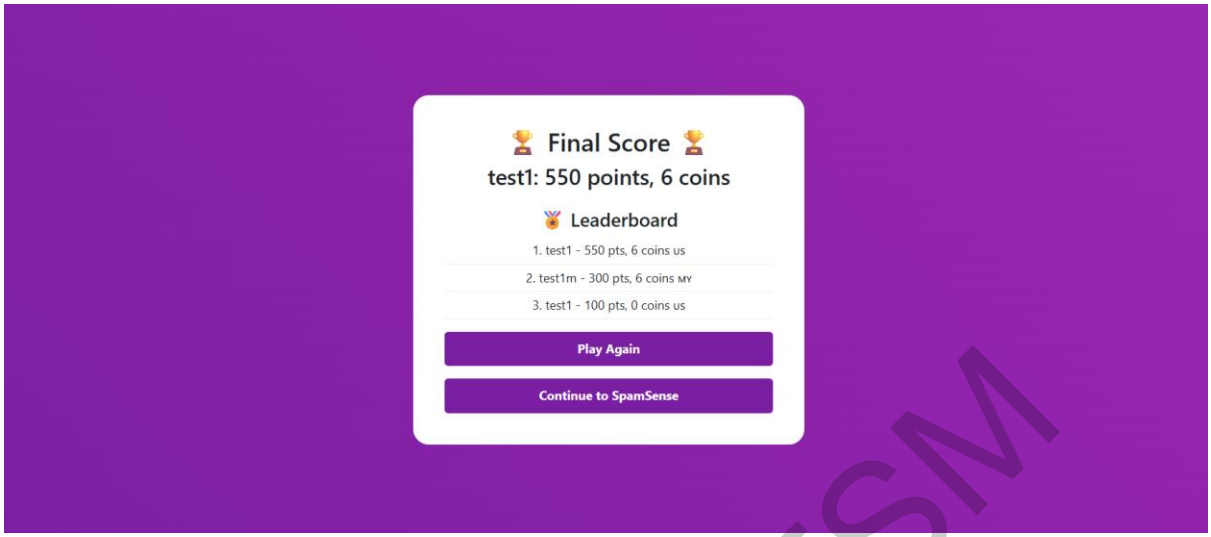
Rajah 4.9 Antaramuka permainan *Mario Spam Hunter Quiz*



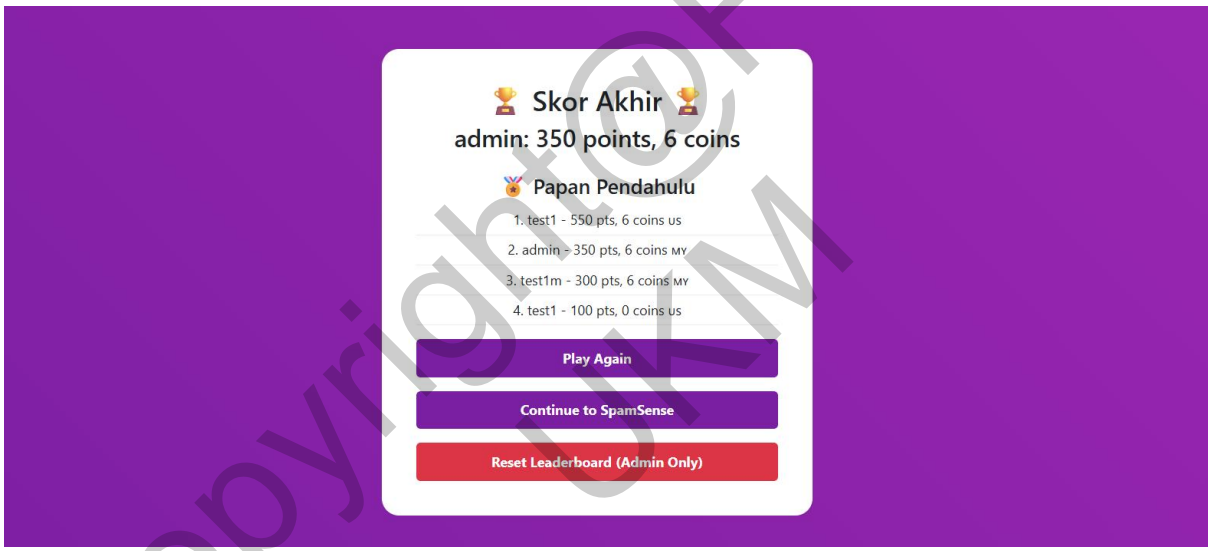
Rajah 4.10 Soalan kuiz didalam permainan (BI)



Rajah 4.11 Soalan kuiz didalam permainan (BM)



Rajah 4.12 Paparan skor markah permainan kuiz (Pengguna)



Rajah 4.13 Paparan skor markah permainan kuiz (Admin)

4.2 PENILAIAN APLIKASI

Proses penilaian dilaksanakan bagi memastikan semua ciri dalam sistem *SpamSense* seperti input mesej, klasifikasi spam, penyerlahan perkataan, pemaparan skor spam, dan permainan kuiz berfungsi dengan baik. Pengujian fungsian dijalankan bagi mengenal pasti sebarang ralat pada proses klasifikasi dan penyimpanan data spam. Selain itu, pengujian kebolehgunaan turut dijalankan melibatkan sekumpulan pengguna akhir yang mewakili pengguna awam bagi menilai kemudahan penggunaan antara muka dan kefahaman terhadap hasil klasifikasi yang dipaparkan. Ujian ini memastikan sistem bukan sahaja berfungsi seperti yang dirancang, malah

mampu memenuhi keperluan dan jangkaan pengguna dari segi pengalaman interaksi, ketepatan keputusan, dan keberkesanan ciri pendidikan melalui gamifikasi kuiz.

4.2.1 PENGUJIAN FUNGSIAN

Jadual ini menyenaraikan tujuh keperluan fungsian utama sistem *SpamSense* berserta tahap risiko dan asas pengujian bagi setiap fungsi. Fungsi KF001 hingga KF002 merangkumi proses asas sistem iaitu input mesej dan klasifikasi spam, di mana KF002 dikategorikan sebagai berisiko tinggi kerana ia merupakan teras sistem pengesanan. KF003 dan KF004 pula berkaitan dengan penyerlahan perkataan spam dan kemas kini pangkalan data; kedua-duanya penting dalam memberi maklumat tambahan kepada pengguna dan memastikan sistem terus belajar daripada input baharu.

Seterusnya, KF005 dan KF006 menyokong kebolehfahaman keputusan sistem, dengan memaparkan skor spam dan memberikan penjelasan klasifikasi. Ini bertujuan meningkatkan ketelusan dan kebolehpercayaan pengguna terhadap sistem. Akhir sekali, KF007 memperkenalkan elemen gamifikasi melalui permainan kuiz, yang memberikan pengalaman pembelajaran interaktif dan menyokong objektif kesedaran terhadap mesej spam. Setiap fungsi diuji berdasarkan *output* yang dijangka sama ada dari segi paparan, penyimpanan data atau tindak balas sistem terhadap input pengguna.

Jadual 4.1 Asas Pengujian

ID Fungsi	**Keperluan Fungsian**	**Tahap Risiko**	**Asas Pengujian**
KF001	Memasukkan teks mesej	Rendah	Input mesej berjaya diproses dan dipaparkan di antaramuka.
KF002	Menganalisis mesej spam	Tinggi	Sistem mengklasifikasikan mesej sebagai SPAM/BUKAN SPAM dengan tepat berdasarkan model ML & LLM.

KF003	Menyerlahkan perkataan spam	Sederhana	Perkataan spam (cth: "*free*", "*win*") diserlahkan dalam penjelasan hasil.
KF004	Kemaskini pangkalan data spam	Tinggi	Perkataan baharu yang disahkan sebagai spam disimpan dalam pangkalan data (*spam_words.db*).
KF005	Memaparkan peratusan spam (NB & LLM)	Sederhana	Sistem menunjukkan skor spam (%) dari Naive Bayes dan LLM dalam *output*
KF006	Memberikan penjelasan (*reasoning*) klasifikasi	Sederhana	Sistem memberikan penjelasan tekstual
KF007	Permainan Kuiz Spam Mario	Sederhana	Memberikan pengalaman interaktif, mesra pengguna dan sebagai pendekatan pembelajaran melalui ciri gamifikasi kuiz berkaitan spam.

4.2.2 PENGUJIAN KEBOLEHGUNAAN

Pengujian kebolehgunaan merupakan proses penting yang melibatkan pengguna akhir untuk menilai sejauh mana sistem SpamSense mudah digunakan, memenuhi keperluan pengguna, dan berfungsi seperti yang diharapkan sebelum dilaksanakan dalam persekitaran sebenar. Pengujian ini bertujuan untuk mengumpul maklum balas pengguna mengenai aspek kemudahan penggunaan, kefahaman, dan kepuasan terhadap sistem yang dibangunkan.

Pengujian kebolehgunaan ini dilaksanakan dengan lima objektif utama. Pertama, menilai tahap kemudahan penggunaan antaramuka SpamSense dari aspek intuitif, navigasi, dan

kebolehgunaan untuk memastikan pengalaman pengguna yang optimum. Kedua, mengukur ketepatan hasil klasifikasi spam melalui penilaian pengguna terhadap kesesuaian keputusan sistem serta menilai kefahaman mereka terhadap justifikasi analisis yang diberikan. Ketiga, mengenal pasti ciri-ciri sistem yang paling bernilai seperti paparan skor spam, penyerlahan kata kunci, atau penjelasan klasifikasi untuk penambahbaikan berfokus. Keempat, menilai keberkesanan penjelasan klasifikasi dalam meningkatkan literasi keselamatan digital pengguna dengan menganalisis sejauh mana maklumat tersebut mendidik dan membina kesedaran. Akhir sekali, mengumpul cadangan penambahbaikan khusus daripada pengguna untuk penyempurnaan sistem, merangkumi aspek fungsian, kandungan, dan pendidikan. Objektif-objektif ini dirangka bagi memastikan SpamSense bukan sahaja memenuhi keperluan teknikal sebagai alat pengesanan spam, tetapi juga berfungsi sebagai platform pembelajaran yang efektif.

Pengujian kebolehgunaan ini dilaksanakan melalui instrumen soal selidik digital yang diedarkan kepada 10 orang responden terpilih menggunakan platform Google Form, dengan reka bentuk yang menggabungkan tiga format penilaian utama iaitu skala Likert 1-5 (1=Sangat Tidak Setuju, 5=Sangat Setuju) untuk mengukur persepsi kualitatif terhadap aspek seperti kemudahan penggunaan antaramuka dan ketepatan klasifikasi. Kedua, soalan dikotomi (Ya/Tidak) bagi menilai kefahaman asas pengguna terhadap output sistem dan yang ketiga, soalan terbuka untuk mengumpul cadangan penambahbaikan spesifik.

Jadual 4.2 *Skala Interpretasi Min*

Skor Min	**Interpretasi**
1.00-2.32	Rendah
2.33-3.65	Sederhana
3.66-5.00	Tinggi

Berdasarkan Jadual 4.3, analisis statistik deskriptif terhadap data yang diperoleh menunjukkan prestasi sistem yang memberangsangkan dengan skor min antara 3.8 hingga 5.0 bagi semua parameter utama iaitu kemudahan antaramuka (4.2), ketepatan klasifikasi (3.8), kefahaman hasil analisis (4.0), kepuasan pengguna (4.6), kemudahan pemahaman sistem (4.5), ketiadaan isu teknikal (5.0), dan kesediaan untuk mencadangkan sistem kepada individu lain (4.1). Min keseluruhan item menunjukkan SpamSense berada pada tahap skor min tinggi iaitu 4.3.

Jadual 4.3 *Kes Guna*

No	Item	Min
1.	Kemudahan penggunaan antaramuka	4.2
2.	Ketepatan klasifikasi spam	3.8
3.	Kefahaman hasil analisis	4.0
4.	Kepuasan pengguna	4.6
5.	Kemudahan pemahaman sistem	4.5
6.	Tiada masalah teknikal	5.0
7.	Cadangan kepada individu lain	4.1
	Min keseluruhan	4.3

Berdasarkan analisis komprehensif terhadap data yang diperoleh, dapat disimpulkan bahawa SpamSense menunjukkan prestasi yang memberangsangkan dari aspek kebolehgunaan dengan purata skor min melebihi 4.0 bagi kebanyakan parameter utama, meskipun ketepatan klasifikasi mencatatkan skor yang sedikit lebih rendah pada 3.8, menunjukkan ruang untuk penambahbaikan dalam algoritma pengesanan spam. Sementara itu, komponen gamifikasi memperoleh maklum balas yang sangat positif dengan skor min konsisten melebihi 4.0 bagi semua aspek penilaian, membuktikan keberkesanannya sebagai alat pendidikan interaktif. Lebih membanggakan, 90% responden melaporkan pemahaman yang jelas terhadap hasil analisis SpamSense, manakala 100% pengguna mengesahkan ketiadaan sebarang masalah teknikal ketika menggunakan permainan kuiz, yang bukan sahaja mencerminkan kestabilan sistem tetapi juga kejayaan dari segi rekabentuk pengalaman pengguna (UX). Data-data ini secara kolektif menunjukkan bahawa pendekatan bersepadu yang menggabungkan sistem pengesanan spam dengan elemen gamifikasi pendidikan telah berjaya mencapai objektif asal projek dalam menyediakan penyelesaian yang bukan sahaja berfungsi dengan efektif tetapi juga mudah diakses dan difahami oleh pengguna akhir.

Berdasarkan maklum balas pengguna, beberapa penambahbaikan kritikal dicadangkan untuk meningkatkan keberkesanan SpamSense dan komponen kuiznya. Penambahbaikan utama termasuk, penyempurnaan algoritma klasifikasi melalui latihan semula dengan dataset yang lebih komprehensif dan pelbagai untuk meningkatkan ketepatan pengesanan spam. Kedua, pengintegrasian fungsi baharu seperti pengesanan nombor *scam* automatik untuk melindungi pengguna daripada penipuan telekomunikasi. Ketiga, penyusunan semula antaramuka pengguna dengan reka bentuk maklumat yang lebih intuitif, termasuk visualisasi

data yang jelas dan hierarki maklumat yang terstruktur. Bagi komponen kuiz, cadangan penambahbaikan merangkumi pengembangan bank soalan dengan lebih banyak soalan interaktif berasaskan senario realistic, dan mengemaskini kandungan berkala dengan contoh spam terkini dan teknik penipuan mutakhir untuk memastikan kerelevanan pendidikan. Penambahbaikan ini diharapkan dapat meningkatkan kedua-dua aspek fungsian sistem dan nilai pendidikannya, sekaligus memperkukuh kedudukannya sebagai penyelesaian komprehensif untuk kesedaran dan pencegahan spam.

5.0 KESIMPULAN

Secara keseluruhan, SpamSense diterima baik oleh pengguna dengan tahap kebolehgunaan yang tinggi, membuktikan bahawa gabungan pendekatan tradisional dan moden berkesan dalam meningkatkan ketepatan pengesanan spam serta kesedaran pengguna. Antaramuka sistem yang mudah digunakan dan penjelasan klasifikasi yang jelas membantu pengguna memahami keputusan sistem, manakala integrasi elemen kuiz interaktif dengan unsur gamifikasi berjaya meningkatkan minat terhadap topik keselamatan digital. Walau bagaimanapun, terdapat ruang untuk penambahbaikan seperti penambahan fungsi pengesanan nombor *scam*, pemantauan berterusan bagi meningkatkan ketepatan klasifikasi, serta pengembangan ciri seperti pengecaman entiti nombor dan e-mel serta pengesanan spam pelbagai bahasa. Sistem ini bukan sahaja berfungsi sebagai alat klasifikasi yang efektif, tetapi juga sebagai platform pendidikan yang menyeluruh dalam memerangi ancaman spam.

6.0 PENGHARGAAN

Projek ini tidak akan berjaya tanpa bimbingan pensyarah penyelia saya, rakan-rakan seperjuangan, serta pengguna yang memberi maklum balas dalam fasa pengujian. Saya juga ingin merakamkan terima kasih kepada semua pihak yang telah membantu secara langsung dan tidak langsung sepanjang pembangunan SpamSense.

Geran Fakulti Teknologi Sains dan Maklumat, FTSM.

7.0 RUJUKAN

Ahmadi, M., et al. 2025. Leveraging large language models for cybersecurity: Enhancing SMS spam detection with robust and context-aware text classification. *arXiv*. https://arxiv.org/abs/2502.11014 [15 Julai 2025]

Baah, C., Govender, I. & Subramaniam, P.R. 2024. Enhancing Learning Engagement: A Study on Gamification's Influence on Motivation and Cognitive Load. *Education Sciences* 14(10): 1115. https://www.mdpi.com/2227-7102/14/10/1115 [26 Julai 2025].

Berita Harian. (2015, Ogos 23). *Serangan siber senjata baharu ancam keselamatan negara*. https://www.bharian.com.my/taxonomy/term/61/2015/08/76416/serangan-siber-senjata-baharu-ancam-keselamatan-negara

Hamizi, M. A. F. (2023). Penggunaan media sosial sebagai media baharu dan impaknya terhadap masyarakat Malaysia. *Perspektif: Jurnal Sains Sosial dan Kemanusiaan, 15*(3), 24–37. https://doi.org/10.37134/perspektif.vol15.sp.3.2023

Onduto, B. 2021. *Gamification of Information Security Awareness and Training*. Universiti Turku. https://www.utupub.fi/bitstream/handle/10024/152929/Onduto_Barack_Thesis_Final.pdf [26 Julai 2025].

System Design School. (2023, Oktober 10). *Understanding and implementing spam detection techniques*. https://systemdesignschool.io/blog/spam-detection