

PENGESANAN EMOSI MULTIMODAL DALAM PERKONGSIAN MUZIK DI MEDIA SOSIAL MENGGUNAKAN PEMBELAJARAN MESIN

¹Iliana Hanin binti Rusli, ²Dr. Fadhilah binti Rosdi

*Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia, 43600 UKM Bangi,
Selangor Darul Ehsan, Malaysia*

Abstrak

Platform di media sosial mempunyai satu jumlah data yang besar yang dihasilkan setiap hari oleh pengguna-pengguna platform tersebut, termasuk data seperti teks, gambar, video, dan muzik. Data-data ini memegang nilai sentimen yang boleh memberikan maklumat tentang emosi, perasaan dan pendapat pengguna terhadap sesuatu perkara. Justeru tujuan projek ini dibangunkan adalah untuk mengklasifikasikan emosi pengguna dengan lebih tepat dengan melakukan analisis gabungan data muzik dan teks melalui perkongsian pengguna di media sosial berdasarkan penggunaan multimodal pendekatan pembelajaran mesin. Walaupun terdapat banyak kaedah pembelajaran mesin yang dihasilkan untuk mengesan emosi berdasarkan data teks pendek dan sebagainya, rata-rata kaedah tersebut tidak mengambil kira gabungan diantara data yang memainkan peranan terhadap emosi pengguna. Model bagi pengesanan emosi pengguna melalui gabungan data muzik dan teks masih kurang mendapat perhatian. Bagi mengatasi kekurangan ini, projek ini menggunakan model Pembelajaran Mesin Multimodal yang menggabungkan ciri-ciri muzik dan kandungan teks daripada perkongsian media sosial bagi tujuan pengesanan emosi secara menyeluruh. Metodologi yang digunakan adalah pengumpulan data muzik dan teks, pemprosesan data, latihan model menggunakan algoritma Pemprosesan Bahasa Tabii, Bidirectional Encoder Representations from Transformers (BERT), dan ujian model untuk menilai ketepatan dalam mengklasifikasikan emosi tersebut. Projek ini dijangka dapat meningkatkan ketepatan pengesanan emosi melalui kaedah multimodal, yang boleh digunakan dalam pelbagai aplikasi pengesanan emosi.

Kata kunci: Multimodal, Pengesanan Emosi, Pembelajaran mesin

Abstract

Social media platforms generate a large amount of data every day by the users of the platform, including data such as text, images, videos, and music. These data hold sentimental value that can provide information about the user's emotions, feelings, and opinions on a matter. Therefore, this project aimed to classify user emotions more accurately by performing a combined analysis of music and text data through user sharing on social media based on the use of various machine learning approach models. Although there are many machine learning methods produced to detect emotions based on short text data and so on, the average method does not take into account the combination of data that plays a role in the user's emotions. Therefore, the model for user emotion detection through the combination of music and text data still lacks attention. To overcome this problem, this project will use a multimodal machine learning model, which combines music and text analysis from social media sharing. The methodology used is music and text data collection, data processing, model training using natural language processing algorithms, Bidirectional Encoder Representations from Transformers (BERT), and model testing to assess accuracy in classifying the emotion. This project is expected to improve the accuracy of emotion detection through multimodal methods, which can be used in various emotion detection applications.

Keywords: Multimodal, Emotion Detection, Machine Learning

1.0 PENGENALAN

Dalam era digital kini, dapat dilihat bahawa transformasi teknologi dan digital yang lebih maju pada masa ini telah mencipta sejarah dalam kehidupan manusia, yang memberikan banyak faedah dan kelebihan kepada semua pengguna. Industri muzik dalam era digital juga tidak ketinggalan. Muzik tidak dapat dipisahkan dalam kehidupan harian kerana ia telah menjadi satu medium yang sangat penting untuk mengekspresikan emosi dan juga perasaan pengguna. Kini sudah menjadi satu trend kepada rata-rata pengguna platform media sosial seperti di Instagram, Twitter dan Spotify untuk mereka berkongsi muzik yang mengungkapkan atau menggambarkan keadaan emosi, perasaan atau memori yang tersimpan.

Keperluan untuk mengenal pasti emosi melalui perkongsian muzik ini timbul kerana emosi yang dikongsi ini dapat memberikan pandangan yang mendalam terhadap perasaan pengguna dan berpotensi untuk digunakan dalam pelbagai bidang seperti pemasaran digital, psikologi, dan sokongan kesihatan mental. Pengesanan emosi ini dapat mengetahui sama ada

seseorang mungkin mempunyai tekanan perasaan, tekanan mental, yang mungkin dalam proses percubaan membunuh diri yang memerlukan bantuan daripada pakar psikiatri. Berdasarkan perkongsian daripada portal rasmi Kementerian Kesihatan Mental Malaysia, Pertubuhan Kesihatan Sedunia (WHO) melaporkan bahawa bunuh diri merupakan di antara sepuluh penyebab utama kematian di dunia ini.

Emosi turut memainkan peranan yang penting dalam bidang pembelajaran mesin (*Machine Learning*), khususnya dalam sistem pengesanan emosi. Terdapat pelbagai kaedah yang telah digunakan berdasarkan penyelidikan terdahulu untuk mengesan emosi. Antaranya, melalui data satu modaliti seperti data teks, audio, video, atau imej sahaja. Namun, pendekatan yang lebih terkini yang melibatkan penggunaan data multimodal, iaitu gabungan dua atau lebih jenis data yang berbeza, untuk memperkayakan maklumat yang dapat dianalisis oleh model. Sistem Pengesanan Emosi Multimodal Melalui Perkongsian Muzik di Media Sosial merupakan sistem yang dibangunkan menggunakan model pembelajaran mesin seperti *Support Vector Machine* (SVM), *Decision Tree*, dan *Random Forest*. Pembelajaran model ini adalah berdasarkan gabungan set data multimodal yang terdiri daripada data muzik dan teks untuk mengklasifikasikan emosi pengguna kepada empat kategori utama iaitu *happy*, *sad*, *angry* dan *calm*. Dengan pendekatan ini, sistem dapat memberikan pemahaman emosi yang lebih menyeluruh dan tepat berbanding penggunaan data daripada satu modaliti sahaja.

2.0 KAJIAN LITERATUR

Kajian yang berkait rapat dalam bidang pengesanan emosi yang berasaskan daripada data ini berkembang dengan maju mengikut peredaran masa. Beberapa kajian terdahulu yang telah menunjukkan dan membuktikan kesan yang besar dalam bidang pengesanan emosi manusia melalui perkongsian di media sosial ini.

Kajian yang dijalankan oleh Krommyda dan rakan-rakan pada tahun 2021 ini memberi tumpuan dalam pengumpulan data daripada *Twitter*, dengan menggunakan kaedah penyesuaian (*harmonization*) dan anotasi berasaskan peraturan untuk menyediakan set data berkualiti tinggi bagi tujuan klasifikasi emosi. Proses anotasi data menggunakan *Emoji Python Library* untuk

mengklasifikasikan penggunaan emoji ke dalam kategori emosi, NRC *Emotion Lexicon*, dan *WordNet* digunakan untuk mengekstrak data ke dalam set sinonim kognitif. Model utama yang digunakan adalah *Long Short-Term Memory* (LSTM) dan model *Baseline Classifiers* digunakan untuk membanding prestasi yang ditunjukkan antara model. Pendekatan ini menunjukkan LSTM sebagai model yang utama untuk menangkap kerumitan emosi daripada data teks, manakala *Baseline Classifiers* berfungsi sebagai pembandingan untuk menilai kaedah pembelajaran mendalam. Hasil kajian menunjukkan bahawa penggunaan model LSTM dalam kajian ini mampu memberikan purata ketepatan yang tinggi untuk menganalisis sentimen dalam data yang dikumpul.

Seterusnya, kajian oleh Bharat Gaiind dan rakan-rakan pada tahun 2019 menumpukan perhatian menggunakan pendekatan pertama iaitu NLP dan *Emotion-Words Set* (EWS) untuk mencari persamaan daripada enam kategori emosi asas, dan penggunaan *Degree-Words Set* yang memainkan peranan kepada bagaimana kadar mempengaruhi intensiti emosi dalam ayat. Pendekatan kedua iaitu menggunakan pembelajaran mesin iaitu *Support Vector Machine* (*Sequential minimal optimization*, SMO) dan *Decision Tree* untuk mengesan emosi pengguna. Pendekatan ini digunakan untuk memastikan bahawa data latihan disediakan dengan baik. Gabungan kekuatan untuk analisis bahasa menggunakan NLP dan pembelajaran mesin untuk menyediakan kaedah pengesanan emosi menunjukkan keputusan peratusan yang lebih tepat dalam menganalisis data teks pendek berbanding dengan hanya menggunakan pendekatan tunggal. Algoritma yang dibangunkan dalam kajian ini menunjukkan pendekatan pengesanan emosi yang digunakan sangat berkesan dan boleh digunakan untuk aplikasi dunia nyata lebih-lebih lagi untuk mengklasifikasikan data teks pendek.

Kajian oleh Malhotra dan Jindal (2020) menggunakan satu bentuk pembelajaran mendalam multimodal untuk mengesan kemurungan dan tingkah laku bunuh diri melalui analisis daripada data media sosial. Algoritma dalam model ini menggunakan pra-pemprosesan tokenisasi teks, *Word2Vec* digunakan untuk mengekstrakan ciri yang merujuk kepada perwakilan ciri sesuatu perkataan dalam bentuk vektor, dan turut menggunakan teknik pembelajaran mendalam multimodal berdasarkan analisis pelbagai jenis data termasuk teks, imej, video dan emoji, di mana mengambil kira data emoji sebagai satu maklumat penting terhadap emosi

pengguna dan mempengaruhi ketepatan pengesanan emosi dalam data teks. Menggabungkan data teks dan imej ke dalam *Neural Network* lapisan *Fully Connected*, bagi mengesan pos yang menunjukkan kemurungan atau tingkah laku bunuh diri. Bagi pemprosesan data jenis video, *Faster-RCNN* digunakan untuk mengenal pasti orang dalam setiap rangkaian gambar tersebut. Hasil daripada gabungan multimodal ini memberikan skor purata yang tinggi yang menunjukkan ketepatan analisis dan menyediakan model yang lebih kukuh dalam pengesanan tingkah laku pengguna.

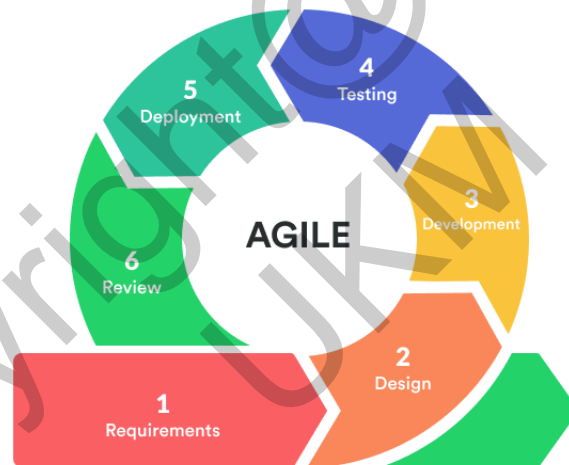
Berikut adalah perbandingan terhadap kajian lepas berkaitan dengan pengesanan emosi pengguna di media sosial:

Jadual 1.0 Ringkasan perbandingan terhadap kajian lepas berkaitan pengesanan emosi pengguna di media sosial

Artikel	Algoritma	Rumusan	Ketepatan Model (%)
Krommyda et al. 2021	<i>LSTM</i>	Pembangunan sistem	91.90
	<i>SVM-SGD</i>	model untuk	86.86
	<i>XGBoost</i>	mengklasifikasikan emosi	84.45
	<i>Naive Bayes</i>	berdasarkan teks pendek	77.01
	<i>Decision Tree</i>	di media sosial.	84.69
	<i>Random forest</i>		80.35
Gaing et al. 2019	<i>SMO</i>	Pembangunan sistem	91.70
	<i>J48</i>	pengesanan dan analisis emosi dalam teks pendek di media sosial, terutamanya <i>tweet</i> .	85.40
Malhotra & Jindal 2020	<i>Word2Vec</i>	Pembelajaran mendalam	Tidak dinyatakan
	<i>VGG-16</i>	multimodal untuk	
	<i>Faster R-CNN</i>	mengesan kemurungan	
	<i>3D CNN (Conv3D)</i>	dan tingkah laku bunuh	
	<i>Weighted Score Fusion</i>	diri daripada analisis pos media sosial.	

3.0 METODOLOGI

Metodologi yang digunakan dalam pembangunan keseluruhan projek Pengesanan Emosi Multimodal dalam Perkongsian Muzik di Media Sosial adalah menggunakan model Agile. Model proses pembangunan Agile ini merupakan satu model yang menekankan fleksibiliti dan reponsif terhadap sesuatu perubahan. Metodologi ini dipilih kerana berdasarkan data projek pengesanan emosi ini, keperluan dan juga teknik yang akan digunakan boleh mempunyai perubahan dengan cepat berdasarkan penemuan baharu. Model ini membenarkan perubahan yang sesuai tanpa mengganggu keseluruhan proses pembangunan sistem dan memastikan sistem memenuhi keperluan sebenar. Rajah 1.0 menunjukkan gambaran model Agile yang terdiri daripada beberapa fasa utama.



Rajah 1.0 Model Agile

Sumber: Chathmini Jayathilaka 2020

Fasa Analisis Keperluan

Dalam fasa yang pertama ini keperluan sesuatu sistem dikaji dengan lebih mendalam untuk memastikan setiap aspek sistem mempunyai matlamat yang jelas dan boleh dicapai. Dalam projek ini, analisis yang perlu dilakukan adalah analisis yang melibatkan pemilihan platform

yang diperlukan, analisis terhadap keperluan data, dan teknik multimodal untuk pengesanan emosi.

Fasa Reka Bentuk

Fasa berkaitan dengan struktur sistem dan seni bina dibangunkan. Antara aspek utama reka bentuk adalah menentukan pilihan seni bina sistem, struktur pembinaan pembelajaran mesin, serta pilihan penggunaan papan pemuka.

Fasa Pembangunan

Pembangunan sistem dijalankan secara berulang berdasarkan keperluan yang telah ditetapkan. Setiap peringkat akan dibangunkan dalam perulangan kecil, diuji dan dikemaskini sebelum beralih kepada peringkat yang seterusnya. Bahagian yang terlibat dalam fasa ini adalah proses pengumpulan data muzik dan teks, prapemprosesan data, membina model pembelajaran mesin dan juga membina antara muka.

Fasa Pengujian

Selepas proses pembangunan dijalankan, ia perlu diuji untuk memastikan model pembelajaran mesin mencapai nilai ketepatan yang diinginkan. Metrik penilaian model boleh digunakan untuk mengesan prestasi model dalam sistem pengesanan emosi pengguna.

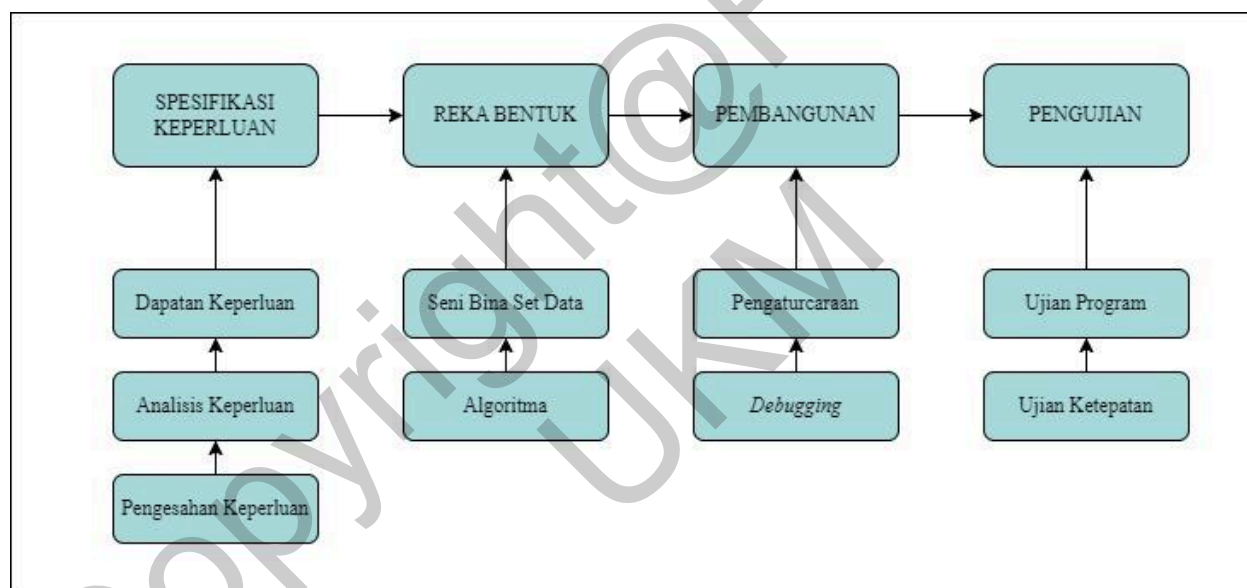
Fasa Penempatan dan Pemantauan

Fasa ini akan melancarkan dan menguji sistem dalam persekitaran yang sebenar. Kesilapan yang berlaku atau keperluan yang diperlukan untuk penambahbaikan akan diatasi pada fasa yang seterusnya.

Fasa Penambahbaikan

Dalam fasa ini dapat dilihat bahawa model Agile memberi kelebihan dalam setiap sistem yang ingin dibina. Kelebihan yang diberikan adalah projek yang dijalankan berupaya untuk menyesuaikan keadaan dengan perubahan dan penambahbaikan yang dilakukan secara berterusan. Dalam projek ini, model pembelajaran mesin boleh dipertingkatkan dengan penggunaan set data baharu atau teknik Pemprosesan Bahasa Tabii yang lebih baik untuk meningkatkan ketepatan sistem pengesanan.

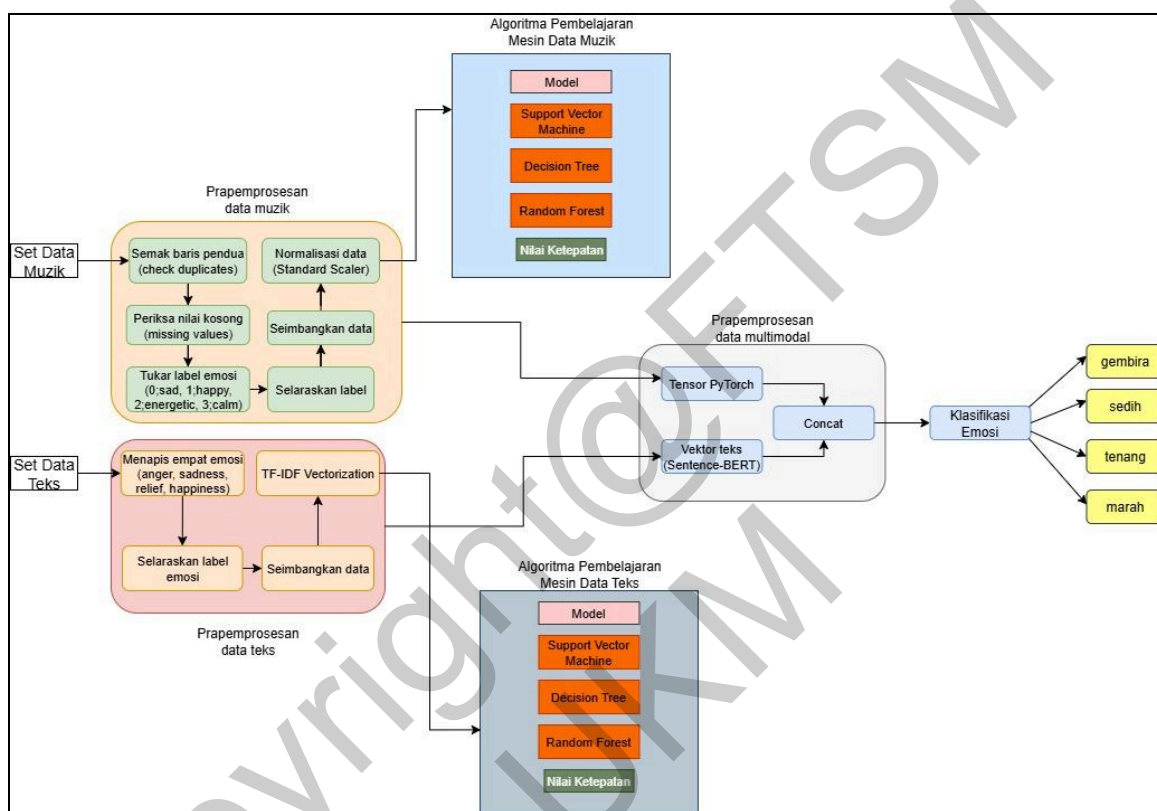
Rajah 2.0 menunjukkan ringkasan terhadap struktur pembangunan projek.



Rajah 2.0 Struktur Pembangunan Projek

Antara aliran proses yang akan digunakan dalam projek ini adalah proses spesifikasi keperluan, reka bentuk projek, pembangunan, dan pengujian terhadap projek tersebut. Dalam proses pembangunan projek yang pertama, spesifikasi keperluan yang digunakan oleh pengkaji adalah untuk mendapatkan data yang diperlukan, menganalisis keperluan data tersebut dan mengesahkan keperluan yang dapat dikenal pasti. Seterusnya, proses pembangunan projek ini diikuti dengan proses reka bentuk. Proses ini bertujuan untuk perancangan kepada teknik dan seni bina sistem yang sesuai untuk digunakan dalam kajian. Proses pembangunan melibatkan

pelaksanaan teknik kepada pengaturcaraan, menguji dan memperbaiki pembangunan sistem. Akhir sekali adalah proses pengujian untuk memastikan sistem dapat berfungsi dengan baik untuk mencapai nilai ketepatan dan objektif kajian. Rajah 3.0 menggambarkan ringkasan terhadap reka bentuk algoritma yang digunakan dalam projek ini.



Rajah 3.0 Reka bentuk algoritma Sistem Pengesanan Emosi Multimodal dalam Perkongsian Muzik di Media Sosial Menggunakan Pembelajaran Mesin

Proses pengumpulan data dilakukan berdasarkan keperluan kajian yang memerlukan jenis data muzik dan juga data teks. Kajian ini menggunakan pendekatan pembelajaran mesin multimodal, yang menunjukkan bahawa pentingnya penyediaan kepada beberapa jenis data termasuklah data teks dan juga muzik. Set data yang pertama, iaitu data muzik yang dimuat turun daripada sumber *Kaggle*, yang bertajuk *278k Emotion Labeled Spotify Songs* disediakan oleh Abdullah Orzan. Berdasarkan set data yang disediakan yang menunjukkan pelbagai ciri audio muzik dalam bentuk nilai angka seperti *danceability*, *energy*, *loudness*, *speechiness*, *acousticness*, *instrumentalness*, *liveness*, *valence*, serta tempo lagu. Nilai ini mewakili dimensi

teknikal yang mencirikan emosi yang terkandung dalam lagu tersebut. Misalnya, atribut *valence* menerangkan tahap positif atau negatif lagu, iaitu nilai yang lebih tinggi akan menunjukkan emosi positif. Set data teks diperoleh daripada set data *Emotion Analysis Based on Text* disediakan oleh Sima Anjali yang mengandungi ayat-ayat pendek yang telah dilabelkan dengan emosi tertentu. Set data ini mengandungi atribut seperti *text* dan *emotion*.

Berdasarkan data muzik, prapemprosesan yang pertama adalah dengan memilih dan menentukan ciri atribut yang mempengaruhi pembelajaran mesin model ini. Sebagai contoh, dengan membuang lajur seperti *Unnamed: 0.1*, *Unnamed: 0*, dan *uri* yang tidak diperlukan untuk menganalisis emosi. Seterusnya, menyemak dan mengendalikan jika data mempunyai data yang hilang dan mengendalikan data tersebut dengan menggantikan menggunakan min, median ataupun padam baris tersebut. Ciri-ciri muzik seperti tempo, *loudness*, *energy*, akustik mempunyai nilai julat yang berbeza. Oleh itu, normalisasi dengan menggunakan Penskalaan *Min-Max* atau Standardisasi boleh dilakukan bagi memastikan model tidak dipengaruhi dengan nilai skala yang besar.

Set data teks mengandungi pelbagai jenis emosi seperti *neutral*, *love*, *happiness*, *sadness*, *relief*, *anger* dan lain-lain. Namun, untuk tujuan pemodelan yang lebih fokus dan seimbang, hanya beberapa kategori emosi utama sahaja yang dipilih. Data yang tidak termasuk dalam kategori emosi yang dipilih akan ditapis keluar untuk meningkatkan keberkesanan model. Bagi menukarkan kandungan teks ke dalam bentuk berangka untuk digunakan oleh algoritma pembelajaran mesin, proses pengekstrakan ciri dijalankan menggunakan kaedah *Term Frequency-Inverse Document Frequency* (TF-IDF). Kaedah ini adalah untuk menilai kepentingan sesuatu perkataan dalam dokumen berbanding keseluruhan korpus, serta membantu mengenal pasti perkataan yang lebih bermakna dalam konteks tanggapan emosi.

Bagi pemprosesan data multimodal, data muzik yang dinormalisasikan dan seimbang ditukar kepada bentuk *tensor* menggunakan *PyTorch* untuk menjadikannya sesuai dimasukkan ke dalam model pembelajaran mesin. Manakala kandungan teks pula ditukarkan kepada vektor semantik menggunakan *Sentence-BERT* yang menghasilkan representasi teks berdimensi tinggi berdasarkan maksud ayat tersebut. Kedua-dua bentuk perwakilan ini kemudiannya digabungkan

dalam satu vektor gabungan yang mewakili ciri multimodal. Gabungan ini membolehkan model memahami konteks emosi dengan lebih teliti sebelum menjalankan klasifikasi ke dalam kategori seperti gembira (*happy*), sedih (*sad*), tenang (*calm*) dan marah (*angry*).

Model pembelajaran mesin yang dipilih dengan mengikut kesesuaian data muzik dan teks ini adalah seperti *Support Vector Machine*, *Decision Tree* dan *Random Forest*, yang berfungsi dengan pendekatan yang berbeza dalam mengenal pasti pola emosi. Setelah model dilatih, ia akan diuji pada data ujian untuk menilai tahap ketepatan dan keupayaan model untuk mengenal pasti emosi tersebut. Pemerhatian dibuat berdasarkan nilai ketepatan (*accuracy*), *precision*, *recall* dan skor-F1 untuk mengukur sejauh mana model berjaya untuk mengklasifikasikan emosi dengan betul. Model dengan prestasi terbaik akan dipilih dan seterusnya diintegrasikan ke dalam antara muka pengguna menggunakan platform *Streamlit* bagi membolehkan interaksi secara langsung dengan sistem.

4.0 HASIL

4.1 Pembangunan Model Multimodal

Sebelum model pembelajaran mesin dilatih, proses pencarian hiperparameter terbaik telah dijalankan bagi memastikan setiap model dapat memberikan prestasi yang optimum. Jadual 2.0 dan Jadual 3.0 masing-masing menunjukkan hiperparameter yang digunakan bagi model data muzik dan data teks.

Jadual 2.0 Senarai hiperparameter terbaik bagi model data muzik

Model	Hiperparameter	Nilai
SVM	<i>C</i>	10
	<i>gamma</i>	<i>auto</i>
	<i>kernel</i>	<i>rbf</i>
Decision Tree	<i>criterion</i>	<i>entropy</i>
	<i>max_depth</i>	10

<i>Random Forest</i>	<i>min_samples_leaf</i>	4
	<i>min_samples_split</i>	10
	<i>class_weight</i>	<i>None</i>
	<i>n_estimators</i>	200
	<i>max_depth</i>	20
	<i>min_samples_leaf</i>	4
	<i>min_samples_split</i>	10
	<i>max_features</i>	<i>sqrt</i>
	<i>class_weight</i>	<i>balanced</i>

Jadual 3.0 Senarai hiperparameter terbaik bagi model data teks

Model	Hiperparameter	Nilai
SVM	<i>C</i>	10
	<i>class_weight</i>	<i>balanced</i>
	<i>kernel</i>	<i>linear</i>
<i>Decision Tree</i>	<i>criterion</i>	<i>gini</i>
	<i>splitter</i>	<i>random</i>
	<i>max_depth</i>	<i>None</i>
	<i>min_samples_split</i>	2
	<i>min_samples_leaf</i>	1
	<i>class_weight</i>	<i>balanced</i>
<i>Random Forest</i>	<i>n_estimators</i>	200
	<i>max_depth</i>	<i>None</i>
	<i>min_samples_split</i>	2
	<i>min_samples_leaf</i>	1
	<i>max_features</i>	<i>sqrt</i>
	<i>class_weight</i>	<i>balanced</i>

Selepas penentuan parameter terbaik, model dilatih menggunakan set data yang telah dipraproses. Bagi data muzik, set ciri dan label dibahagikan kepada data latihan dan data pengujian menggunakan fungsi *train_test_split* dengan 80% data untuk latihan dan 20% untuk pengujian, serta diseragamkan menggunakan kaedah *StandardScaler* bagi memastikan semua ciri berada dalam skala yang sama. Manakala bagi data teks, pembahagian dilakukan dengan cara yang sama menggunakan *X_tfidf* sebagai ciri teks hasil pengekstrakan TF-IDF, bagi memastikan pengagihan kelas yang seimbang dalam set latihan dan penguji melalui parameter *stratify=y*. Bagi pembangunan model multimodal, ciri-ciri daripada data muzik dan teks digabungkan secara mendatar untuk membentuk satu vektor representasi gabungan. Data multimodal ini turut dibahagikan menggunakan kaedah yang sama dan digunakan untuk melatih model pembelajaran mesin bagi mengenal pasti emosi berdasarkan dua jenis modaliti serentak.

4.2 Perbandingan Antara Model Pengesanan Emosi

Bahagian ini membandingkan prestasi model klasifikasi emosi berdasarkan tiga pendekatan utama iaitu model berasaskan muzik, model berasaskan teks, dan model multimodal. Jadual 2.0 menunjukkan ringkasan terhadap perbandingan model.

Jadual 4.0 Perbandingan mengikut jenis data

Model	Data Muzik		Data Teks		Data Multimodal	
	Ketepatan	Skor F1	Ketepatan	Skor F1	Ketepatan	Skor F1
<i>SVM</i>	84%	0.84	99%	0.99	96%	0.96
<i>Decision Tree</i>	85%	0.85	98%	0.98	82%	0.82
<i>Random Forest</i>	87%	0.87	97%	0.97	94%	0.94

Model *Random Forest* bagi data muzik menunjukkan prestasi terbaik dengan ketepatan 87%. Model ini mampu untuk menangkap corak kompleks dalam ciri-ciri akustik seperti *valence*, tempo, dan *danceability*. Sementara itu, *Decision Tree* mencatatkan ketepatan 85% yang menunjukkan kemampuan sederhana untuk membuat keputusan berdasarkan struktur data ini.

Model SVM pula berada pada kedudukan paling rendah dengan 84% ketepatan untuk dibandingkan antara ketiga-tiga model ini. Walaupun perbezaan yang ditunjukkan ini kecil, ia menunjukkan bahawa algoritma *ensemble* seperti *Random Forest* lebih efektif untuk pengelasan emosi berdasarkan data muzik.

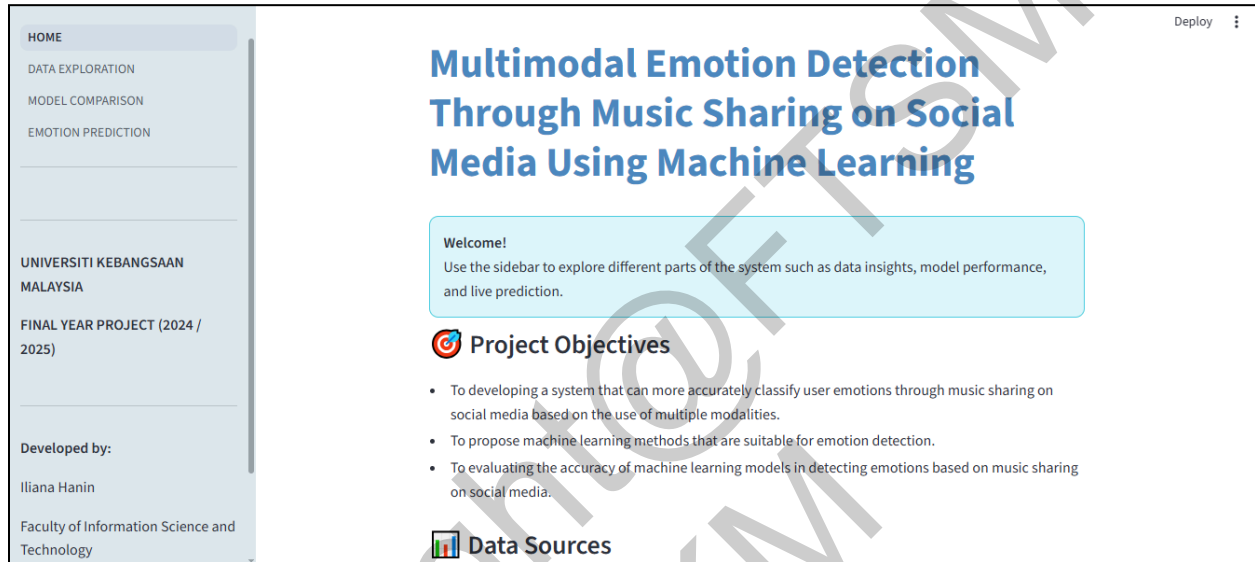
Perbandingan model terhadap data teks menunjukkan bahawa model SVM memberikan prestasi tertinggi dengan ketepatan 99% menunjukkan pengujian yang sangat baik dalam mengenal pasti emosi berdasarkan ciri linguistik. Hal ini berlaku disebabkan oleh keupayaan SVM untuk mengawal data berdimensi tinggi seperti TF-IDF. *Decision Tree* dan *Random Forest* juga mencapai prestasi tinggi, dengan ketepatan 98% dan 97%. Ketiga-tiga model ini menunjukkan bahawa data teks lebih kaya dengan maklumat emosi yang boleh diekstrak dengan tepat oleh pelbagai jenis algoritma.

Dalam gabungan data teks dan muzik (data multimodal), model SVM sekali lagi muncul sebagai model terbaik dengan ketepatan 96%, membuktikan keupayaannya untuk mengintegrasikan dua jenis modaliti dan menghasilkan keputusan yang stabil. *Random Forest* juga menunjukkan prestasi kukuh (94%), manakala *Decision Tree* mencatatkan penurunan ketara dengan ketepatan 82%. Pengurangan ini menunjukkan bahawa model padat seperti *Decision Tree* berkemungkinan tidak cukup kompleks untuk mengendalikan interaksi antara dua modaliti data yang berbeza.

4.3 Pembangunan Antara Muka Streamlit

Antara muka sistem ini dibangunkan menggunakan *Visual Studio Code* (VS Code) sebagai persekitaran pembangunan dan *Streamlit* sebagai rangka kerja utama untuk membina aplikasi web interaktif. Papan pemuka ini dibina supaya pengguna boleh melakukan ramalan emosi secara langsung (*real-time*) dengan memasukkan teks kapsyen media sosial dan ciri akustik muzik. Prose ramalan emosi ini dibangunkan menggunakan model SVM yang telah dilatih dan diuji pada data multimodal. Pemilihan model SVM sebagai model utama dalam halaman ini adalah berdasarkan dapatan pengujian yang menunjukkan bahawa model SVM memberikan prestasi terbaik. Rajah 4.0 hingga Rajah 9.0 menunjukkan paparan antara muka yang telah

dibangunkan. Reka bentuk terhadap antara muka ini memfokuskan kepada kebolegunaan dan kebolehcapaian dengan membahagikan paparan sistem kepada beberapa bahagian utama yang boleh diakses melalui bar sisi yang terletak di bahagian sebelah kiri antara muka tersebut. Struktur utama antara muka ini terdiri daripada empat bahagian iaitu *Home*, *Data Exploration*, *Model Comparison*, serta *Emotion Prediction*.



Rajah 4.0 Antara Muka Sistem Pengesanan Emosi Multimodal

DATA EXPLORATION

MODEL COMPARISON

EMOTION PREDICTION

UNIVERSITI KEBANGSAAN MALAYSIA

FINAL YEAR PROJECT (2024 / 2025)

Developed by:

Iliana Hanin

© iliana2025

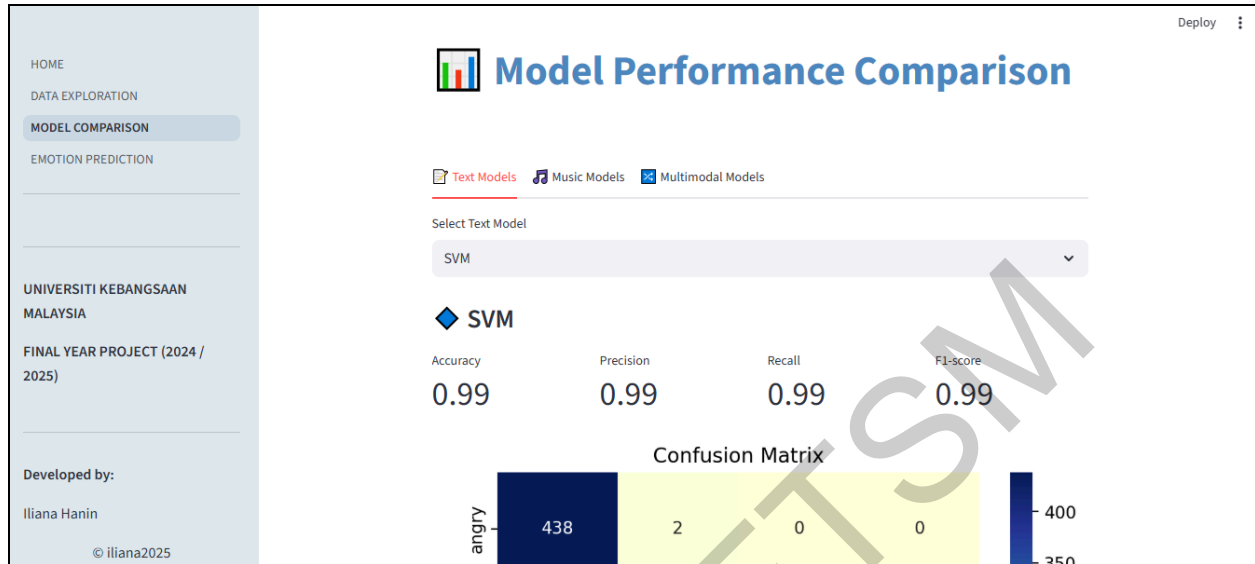
Data Exploration

Music Data Overview

Raw Dataset Preview (Music):

	Unnamed: 0.1	Unnamed: 0	duration (ms)	danceability	energy	loudness	speechiness	acousticness
0	0	0	195000	0.611	0.614	-8.815	0.0672	0.0
1	1	1	194641	0.638	0.781	-6.848	0.0285	0.0
2	2	2	217573	0.56	0.81	-8.029	0.0872	0.0
3	3	3	443478	0.525	0.699	-4.571	0.0353	0.0
4	4	4	225862	0.367	0.771	-5.863	0.106	0.0
5	5	5	166920	0.572	0.837	-7.876	0.0367	0.0
6	6	6	193133	0.725	0.687	-6.465	0.0596	0.0
7	7	7	253000	0.675	0.547	-4.999	0.0481	0.0
8	8	8	216187	0.516	0.692	-4.842	0.0279	0.0

Rajah 5.0 Antara Muka Halaman Penerokaan Data



Rajah 6.0

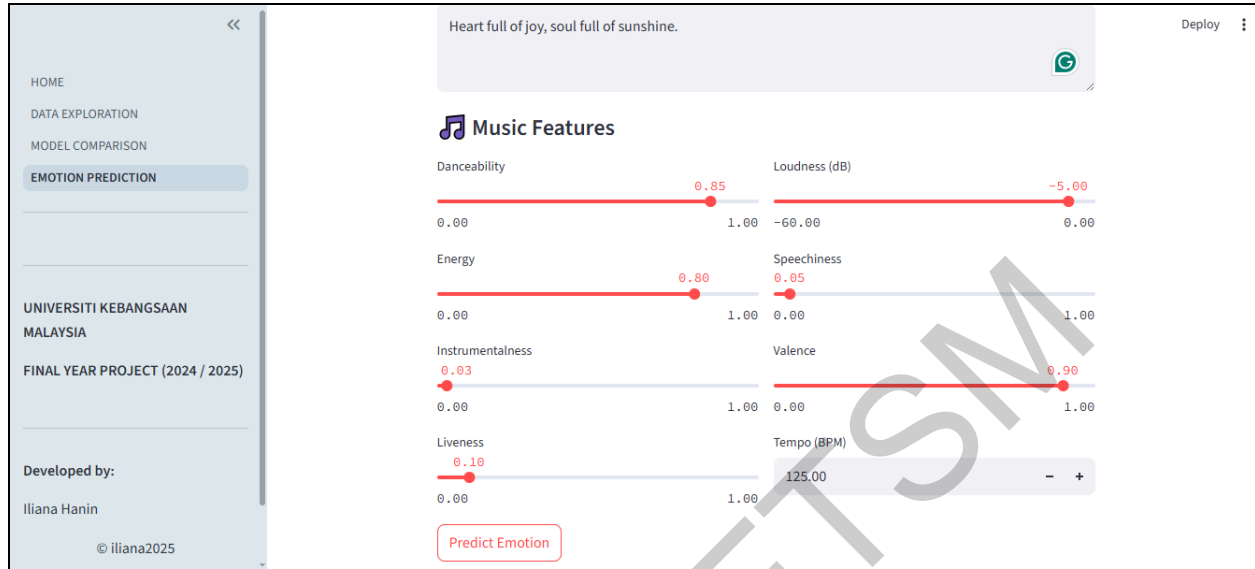
Antara Muka Halaman Perbandingan Model



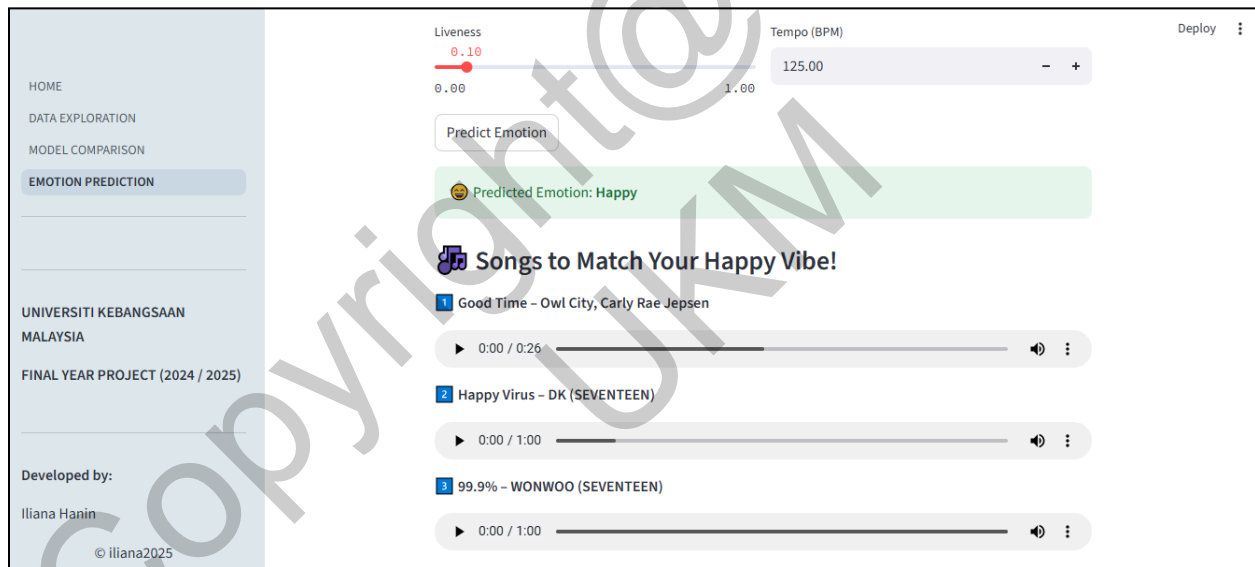
Rajah 7.0

Antara Muka Halaman Ramalan Emosi

Pengguna boleh menaip kapsyen dan melaraskan pelbagai ciri audio seperti *danceability*, *energy*, *instrumentalness*, *liveness*, *loudness*, *speechiness*, *valence*, dan tempo melalui peluncur interaktif. Input ini akan diproses oleh model untuk mengklasifikasikan emosi pengguna kepada salah satu daripada empat kategori utama iaitu *happy*, *sad*, *angry* atau *calm*.



Rajah 8.0 Contoh input pengguna



Rajah 9.0 Keputusan klasifikasi emosi berdasarkan input pengguna

Melalui contoh input pengguna seperti yang ditunjukkan dalam Rajah 5.0, pengguna telah menaip kapsyen “*Heart full of joy, soul full of sunshine*” serta melaraskan pelbagai ciri audio seperti *danceability* (0.85), *energy* (0.80), *valence* (0.90) dan tempo (125 BPM) yang menunjukkan nilai ciri yang cenderung kepada emosi positif. Input ini kemudiannya diproses oleh sistem, dan hasil klasifikasi yang ditunjukkan dalam Rajah 6.0 meramalkan emosi pengguna sebagai *Happy*. Selain itu, sistem turut mencadangkan senarai lagu yang sepadan dengan emosi

tersebut, sekali gus memperlihatkan keberkesanan integrasi antara mengenal pasti emosi dan cadangan muzik dalam aplikasi ini.

5.0 KESIMPULAN

Secara keseluruhannya, projek ini telah berjaya memenuhi objektif utama, iaitu membangunkan model pembelajaran mesin yang dapat mengklasifikasikan emosi pengguna dengan lebih tepat melalui perkongsian muzik di media sosial berdasarkan penggunaan multimodal yang menggabungkan data teks dan muzik. Melalui penggunaan pelbagai model pembelajaran mesin seperti *Support Vector Machine*, *Decision Tree*, dan *Random Forest*, sistem ini telah menunjukkan prestasi yang baik dalam klasifikasi emosi dengan model SVM multimodal mencatatkan ketepatan yang tertinggi. Kejayaan ini membuktikan bahawa gabungan daripada pelbagai modal data dapat meningkatkan keberkesanan pengesanan emosi berbanding penggunaan data tunggal sahaja.

Sepanjang proses pembangunan sistem ini, beberapa cabaran yang berlaku diluar jangkaan dapat dikenal pasti dan dapat ditangani dengan baik secara berperingkat, khususnya dalam proses penyelarasan data daripada dua sumber berbeza serta pemilihan algoritma yang sesuai. Selain itu, antara muka pengguna yang dibina menggunakan *Streamlit* juga telah berjaya menyediakan paparan yang interaktif dan mesra pengguna untuk menggambarkan hasil pengesanan emosi. Projek ini secara tidak langsung telah memberi sumbangan kepada bidang pemrosesan bahasa tabii dan pembelajaran mesin dalam konteks aplikasi emosi di media sosial.

Melalui projek ini juga, ia berpotensi untuk memberi sumbangan bermakna kepada komuniti, khususnya dalam aspek kesihatan mental. Dengan keupayaan sistem ini mengesan emosi pengguna berdasarkan perkongsian muzik dan teks di media sosial, ia boleh digunakan sebagai alat sokongan awal untuk mengenal pasti individu yang mungkin mengalami tekanan emosi atau masalah kesihatan mental. Sistem ini boleh membantu pihak yang memainkan peranan penting seperti ahli psikologi, kaunselor atau organisasi kesihatan untuk mengenal pasti corak emosi yang tidak stabil dan mengambil langkah intervensi awal. Justeru, projek ini bukan sahaja memberi nilai dalam bidang teknologi dan akademik, malah turut membuka ruang kepada

pembangunan aplikasi sosial yang lebih peka terhadap kesejahteraan mental pengguna pada era digital ini.

6.0 PENGHARGAAN

Saya ingin merakamkan setinggi-tinggi penghargaan dan jutaan terima kasih kepada penyelia saya, Dr. Fadhilah binti Rosdi, yang tidak putus memberikan banyak bantuan, bimbingan, dorongan, dan tunjuk ajar yang amat berharga sepanjang tempoh tugas ini dijalankan. Beliau bukan sahaja membimbing dari segi teknikal, malah turut memberi semangat yang membantu saya untuk kekal fokus dan berdisiplin dalam melaksanakan setiap fasa projek.

Kepada keluarga tercinta, khususnya Encik Rusli bin Ismail, Puan Norliza binti Ismail, kakak saya Izyan Hanis binti Rusli, serta abang saya Ikhmal Hakim bin Rusli, terima kasih yang tidak terhingga kerana telah menjadi sumber kekuatan dan inspirasi saya. Doa, sokongan moral, dan kasih sayang yang tidak pernah putus amatlah saya hargai. Saya tujukan juga penghargaan buat sahabat-sahabat seperjuangan, Eiza Athira, Intan Humaira, Puteri Alissa, Ain Batrisyia serta rakan-rakan yang lain yang sentiasa memberi sokongan antara satu sama lain. Kehadiran anda semua telah menceriakan perjalanan akademik ini dan menyumbang kepada kejayaan projek ini.

7.0 RUJUKAN

Associate, A.S.F.C.O. 2021. Understanding TF-IDF for Machine Learning | Capital One. <https://www.capitalone.com/tech/machine-learning/understanding-tf-idf/>.

Aris-Ai. (n.d.). GitHub - aris-ai/Audio-and-text-based-emotion-recognition: A multimodal approach on emotion recognition using audio and text. <https://github.com/aris-ai/Audio-and-text-based-emotion-recognition>.

Emotion analysis based on text. 2024. <https://www.kaggle.com/datasets/simaanjali/emotion-analysis-based-on-text>.

Gaind, B., Syal, V. & Padgalwar, S. 2019. Emotion Detection and Analysis on Social Media. <https://arxiv.org/abs/1901.08458>.

- Krommyda, M., Rigos, A., Bouklas, K. & Amditis, A. 2021b. An experimental analysis of data annotation Methodologies for emotion Detection in short text posted on social media. *Informatics* 8(1): 19.
- M, B. 2024. Decision Trees: Split Methods & Hyperparameter Tuning. <https://www.analyticsvidhya.com/blog/2024/03/decision-trees-split-methods-hyperparameter-tuning/>.
- Malhotra, A. & Jindal, R. 2020. Multimodal Deep Learning based Framework for Detecting Depression and Suicidal Behaviour by Affective Analysis of Social Media Posts. *EAI Endorsed Transactions on Pervasive Health and Technology* 6(21): e1.
- Singh, H. 2025. 10 Techniques to solve Imbalanced Classes in Machine Learning (Updated 2025). <https://www.analyticsvidhya.com/blog/2020/07/10-techniques-to-deal-with-class-imbalance-in-machine-learning/>.
- Yen, S.-J. & Lee, Y.-S. 2006. Under-Sampling approaches for improving prediction of the minority class in an imbalanced dataset. *Intelligent Control and Automation*, hlm. 731–740.
- 278k Emotion labeled Spotify songs. 2023. https://www.kaggle.com/datasets/abdullahorzan/moodify-dataset?select=278k_labelled_uri.csv.

Iliana Hanin binti Rusli (A193127)
Dr. Fadhilah binti Rosdi
 Fakulti Teknologi & Sains Maklumat,
 University Kebangsaan Malaysia