

**ANALISIS SENTIMEN SISTEM SMART SURIPRENEUR  
BERDASARKAN ULASAN PEMBELI KE ATAS PRODUK JUALAN E-  
DAGANG (SHOPEE)**

MUHAMMAD ADAM SHUKRI BIN RAYMOND AZMAN

SABRINA BINTI TIUN

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

**ABSTRAK**

Bilangan keluarga B40 di Malaysia semakin meningkat dan sudah menjadi masalah yang ketara. Antara sebab masalah ini berlaku ialah kerana hanya satu daripada ahli keluarga tersebut mendapat gaji yang mencukupi dalam keluarga. Kebiasaannya hanya suami atau bapa keluarga tersebut yang mendapat gaji untuk menampung keluarga manakala ibu atau isteri tinggal di rumah untuk tolong menjaga dan membesarkan anak. Untuk menyelesaikan masalah ini adalah untuk mengajar ibu dan isteri keluarga ini untuk membuat perniagaan dalam talian. Untuk membantu ibu dan isteri keluarga ini ialah membina sebuah aplikasi bernama *Smart SuriPreneur*, di mana sistem ini akan membantu mereka dengan mengajar, menolong dan membuat analisis setiap produk mereka. Bahagian analisis sentimen akan berada di dalam bahagian "Sentiment" di dalam sistem *Smart SuriPreneur* ini. Projek ini bertujuan untuk membina model analisis sentimen untuk menganalisis ulasan pembeli produk-produk tersebut dan memaparkan keputusannya dalam bentuk graf di aplikasi *Smart SuriPreneur* di bahagian Sentimen. Analisis sentimen ini merupakan sebahagian daripada pelbagai sistem di dalam laman web *Smart SuriPreneur* yang dibangunkan menggunakan *Laravel*. Sentimen analisis ini menggunakan kaedah pembelajaran mesin. Kaedah pembelajaran mesin yang akan digunakan ialah pembelajaran terselia yang menggunakan algoritma *naive bayes*, regresi logistik dan SVM. Prestasi model akan diuji dengan mengambil kira ketepatan model pembelajaran mesin ini dalam menganalisis setiap ulasan pembeli. Analisis sentimen ini akan berguna apabila

mereka ingin membangunkan perniagaan dalam talian. Mereka boleh melihat ulasan pembeli untuk produk-produk yang dijual. Bahagian sentimen analisis ini akan menggunakan antaramuka dimana pengguna boleh memilih untuk melihat sentimen terhadap produk yang diingini mereka. Selepas dipilih produk, graf bar yang menunjukkan sentimen analisis setiap produk yang mereka cari akan dipaparkan di antaramuka yang tersedia.

Kata kunci: Sentimen, *naive bayes*, regresi logistik, SVM

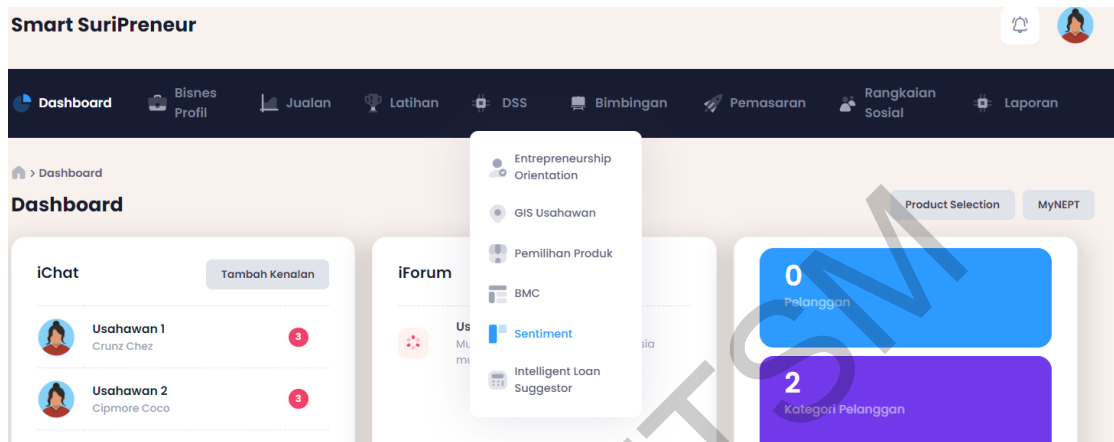
## PENGENALAN

Keluarga B40 di Malaysia merupakan golongan isi rumah yang mendapat pendapatan kurang daripada RM 4850 sebulan (Department of Statistics). Pendapatan isi rumah di Malaysia dibahagi kepada tiga oleh kerajaan iaitu B40, M40 dan T20. Takrifan ini digunakan agar lebih mudah untuk kerajaan menyalurkan bantuan dan subsidi kepada yang perlu. B40 disini bermaksud “Bottom 40 percent” atau “Bawah 40 peratus”. Terdapat pelbagai sebab keluarga B40 ini mendapat pendapatan yang kurang dari RM 4850 (Mysumber.com). Penting bagi negara Malaysia untuk mengatasi masalah B40 ini kerana jika keluarga tersebut tidak mampu menampung hidup dengan pendapatan tersebut, maka akan jatuh ke bawah miskin tegar (ECentral.my).

Antara sebab keluarga B40 ini berada di lingkungan tersebut ialah mereka tidak mendapat gaji yang mencukupi, hanya salah seorang ahli keluarga bekerja dan peluang pekerjaan yang tidak banyak. Jadi ramai keluarga B40 akan mencari jalan lain untuk mendapat pendapatan yang lebih tinggi atau lebih dikenali sebagai pendapatan sampingan. Sebagai contoh, menjadi penghantar makanan, pemandu e-hailing dan berniaga dalam talian. Pendapatan sampingan ini akan membantu mereka untuk keluar daripada golongan B40 dan hidup lebih selesa. Kerajaan juga membantu golongan ini dengan memberikan beberapa jenis subsidi dan potongan cukai agar mampu menampung hidup.

Perniagaan dalam talian akan dapat membantu keluarga b40 ini mendapat pendapatan sampingan. Tetapi, masih ramai yang tidak faham dan pandai untuk memulakan perniagaan dalam talian. Jadi untuk membantu golongan B40 ini, sebuah aplikasi yang bernama *Smart SuriPreneur* di mana akan diadakan sesi pembelajaran bagi mereka. Aplikasi ini dikhususkan kepada ibu dan suri rumah yang tidak bekerja untuk belajar cara memulakan perniagaan dalam talian. Mereka dapat belajar cara memulakan, menguruskan dan meraih keuntungan melalui

aplikasi ini. Sesi pembelajaran akan diadakan dan diajar oleh pelatih-pelatih yang terlatih. Mereka juga dapat melihat sendiri perkembangan produk-produk yang mereka jual.



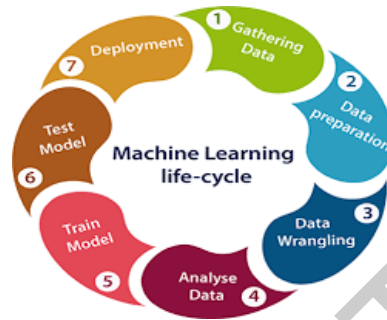
Rajah **Error! No text of specified style in document.** Antara muka Smart SuriPreneur untuk navigasi ke bahagian Sentiment

Projek tahun akhir ini merupakan salah satu modul yang dibangunkan dalam aplikasi *Smart SuriPreneur*. Bahagian projek ini ialah untuk membuat analisis sentimen ke atas ulasan pelanggan di Shopee pada produk-produk yang dijual oleh ibu dan suri rumah ini. Berdasarkan Rajah 1 di atas, bahagian projek ini ialah pada "Sentiment". Analisis sentimen ini dapat membantu mereka melihat bagaimana produk mereka jual itu dilihat oleh pembeli. Mereka dapat membuat keputusan sama ada untuk terus menjual produk itu atau mencari produk lain yang juga dijual di Shopee yang mempunyai analisis sentimen yang lebih positif.

## METODOLOGI KAJIAN

Metodologi membangunkan aplikasi pembelajaran mesin mempunyai banyak peringkat. Aplikasi pembelajaran mesin ini menggunakan model kitaran hayat pembelajaran mesin. Metodologi ini digunakan kerana memberikan lebih pemerhatian terhadap pengumpulan data dan meyakinkan pembangunan kerana setiap fasa adalah dilakukan dengan teliti. Setiap peringkat dalam pembangunan amatlah penting agar aplikasi tersebut berjaya dibangunkan dan berjalan lancar. Rajah 1 menunjukkan kitaran hayat pembelajaran mesin yang diambil dari laman web *JavaTpoint*. Setiap langkah dalam kitaran hayat ini perlu diikuti sehingga akhir projek. Kaedah ini akan mengambil data dan proses data sebelum diuji dalam sebuah

model analisis sentimen dan mencari model terbaik untuk digunakan sebetulnya dan diintegrasikan ke dalam sebuah aplikasi. Kaedah ini dipilih kerana setiap data yang diambil akan di pra-proses dahulu sebelum di masukkan ke dalam model agar lebih tepat dalam membuat analisis apabila model dilatih dan diuji.



Rajah 1 Kitaran Hayat Pembelajaran Mesin

Sumber: (JavaTPoint)

### **Pra Pemrosesan Data (Pengumpulan Data, Penyediaan Data)**

Data yang diperolehi merupakan data daripada aplikasi Shopee. Data yang diambil ini merupakan data yang tidak bersih jadi data yang diambil ini perlu dibersihkan sebelum di proses. Tanpa membersihkan data ini, kemungkinan besar analisis sentimen yang dilakukan tidak tepat dan lari daripada sentimen anggaran jadi amat pentinglah langkah ini dilakukan selepas memperolehi data.

### **Melatih Dan Menguji Model**

Sebelum menguji model, model tersebut perlu dilatih dahulu. Apabila model telah dilatih, prestasi model akan menjadi lebih baik apabila membuat analisis sentimen. Model akan diuji selepas dilatih untuk menguji ketepatan model. Ketepatan model yang lebih tinggi memastikan keyakinan untuk digunakan semasa fasa integrasi dengan sistem *Smart SuriPreneur*.

### **Mengintegrasikan Model**

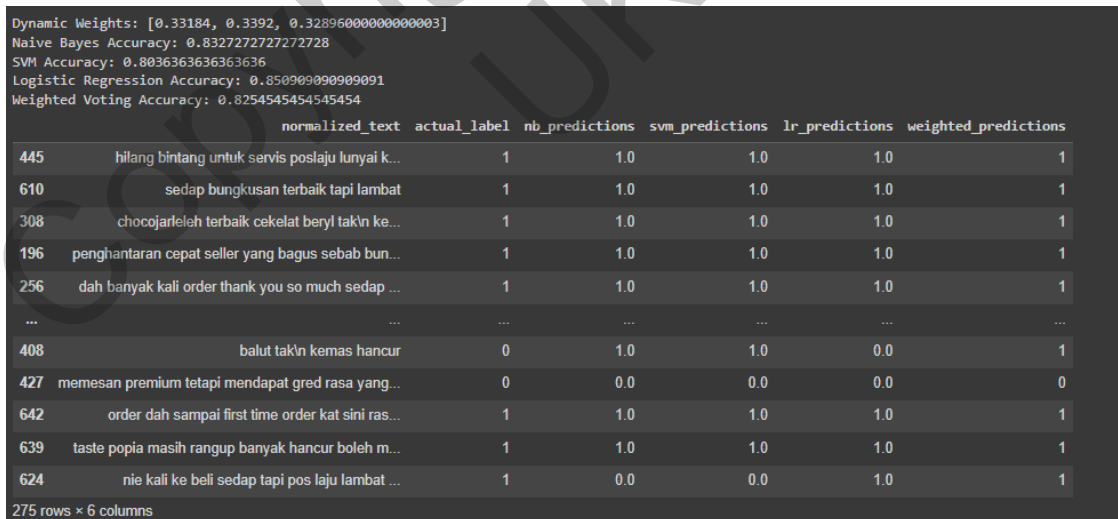
Selepas fasa pengujian, model akan diintegrasikan ke dalam aplikasi *Smart SuriPreneur* dimana setiap suri rumah akan dapat melihat analisis sentimen terhadap produk-produk yang dijual di dalam Shopee. Pengujian juga akan di lakukan pada fasa integrasi untuk melihat

sama ada model berjaya di integrasi ke dalam aplikasi *Smart SuriPreneur*. Fasa integrasi model ini adalah penting kerana ia adalah fasa yang terakhir sebelum pengguna menggunakan sistem ini.

## KEPUTUSAN DAN PERBINCANGAN

Model Analisis Sentimen bagi sistem Smart SuriPreneur berdasarkan ulasan pembeli ke atas produk jualan e-dagang (Shopee) berjaya dibangunkan dan semua dokumentasinya telah dilengkapi. Pembangunan model ini telah menggunakan bahasa pengaturcaraan Python di IDE Google Colaboratory. Kemudian, antara muka di sistem Smart SuriPreneur menggunakan bahasa pengaturcaraan PHP iaitu dengan menggunakan rangka kerja Laravel/Livewire. Seterusnya, proses integrasi model analisis sentimen ke dalam sistem Smart SuriPreneur menggunakan sebuah API yang bernama FastAPI di mana API ini menggunakan bahasa Python.

Model analisis sentimen ini telah memperoleh ketepatan yang tinggi dimana semua model memperoleh ketepatan melebihi 0.8 atau 80%. Undian pemberat dilakukan bagi mengambil kira relevan setiap model untuk mengeluarkan sentimen akhir. Setiap berat model juga ditunjukkan dalam Rajah 3.



Dynamic Weights: [0.33184, 0.3392, 0.32896000000000003]  
 Naive Bayes Accuracy: 0.8327272727272728  
 SVM Accuracy: 0.8036363636363636  
 Logistic Regression Accuracy: 0.850909090909091  
 Weighted Voting Accuracy: 0.8254545454545454

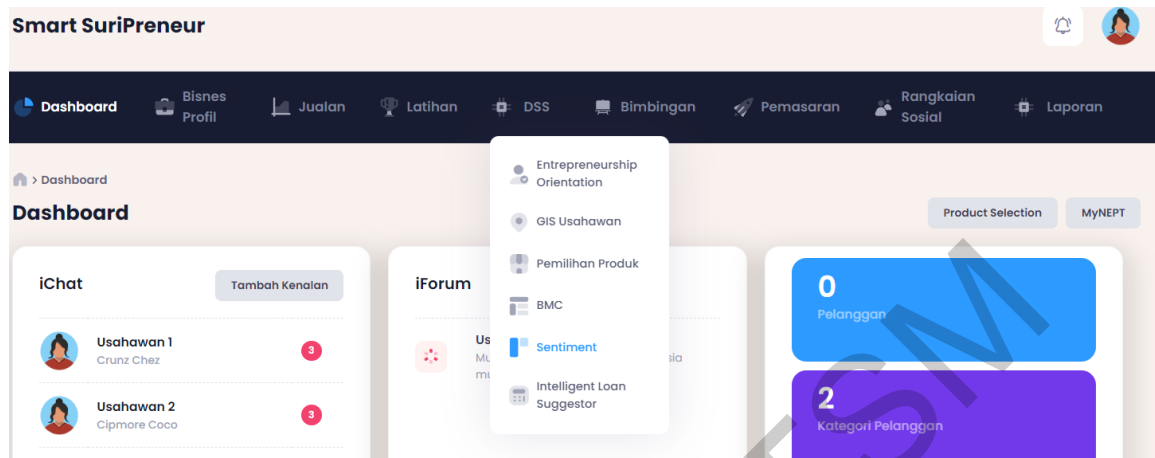
	normalized_text	actual_label	nb_predictions	svm_predictions	lr_predictions	weighted_predictions
445	hilang bintang untuk servis poslaju lunyai k...	1	1.0	1.0	1.0	1
610	sedap bungkus terbaik tapi lambat	1	1.0	1.0	1.0	1
308	chocojarleleh terbaik ceketat beryl tak'n ke...	1	1.0	1.0	1.0	1
196	pengantaran cepat seller yang bagus sebab bun...	1	1.0	1.0	1.0	1
256	dah banyak kali order thank you so much sedap ...	1	1.0	1.0	1.0	1
...	...	...	...	...	...	...
408	balut tak'n kemas hancur	0	1.0	1.0	0.0	1
427	memesan premium tetapi mendapat gred rasa yang...	0	0.0	0.0	0.0	0
642	order dah sampai first time order kat sini ras...	1	1.0	1.0	1.0	1
639	taste popia masih rangup banyak hancur boleh m...	1	1.0	1.0	1.0	1
624	nie kali ke beli sedap tapi pos laju lambat ...	1	0.0	0.0	1.0	1

275 rows x 6 columns

Rajah 3 Ketepatan setiap model dan berat setiap model

Selepas pengguna telah membuat akaun baharu ataupun Log Masuk untuk memasuki sistem Smart SuriPreneur, mereka perlu pergi ke bahagian analisis sentimen. Untuk pergi ke bahagian analisis sentimen, pengguna perlu meletakkan penunjuk tetikus mereka di atas

pilhan 'DSS' dan kemudian pilih 'Sentiment' seperti ditunjukkan dalam Rajah 4.



Rajah 4 Antara Muka Smart SuriPreneur

Apabila pengguna memilih bahagian 'Sentiment', terdapat dua bahagian iaitu bahagian mengisi komen baharu ataupun bahagian paparan graf. Seperti ditunjukkan dalam Rajah 5, pengguna boleh memilih sama ada mereka ingin menambah komen baharu ke dalam sistem ataupun memaparkan sebuah graf sentimen bagi melihat sentimen produk-produk yang telah wujud di dalam pangkalan data. Untuk memaparkan graf, pengguna perlu memasukkan nama produk di bahagian "Carian" dahulu sebelum menekan "Papar".

 The image shows two sections of the application. The top section is titled 'Tambah Komen Baharu' and contains several input fields: 'Nama \*' with a subtext 'Masukkan nama produk', 'Komen \*' with a subtext 'Masukkan review/komen mengenai produk', 'Jualan' with a subtext 'Masukkan jumlah jualan (tidak wajib)', 'Harga' with a subtext 'Masukkan harga (tidak wajib)', and 'Stok' with a subtext 'Masukkan stok (tidak wajib)'. There is a blue 'Hantar' button at the bottom of this section. The bottom section is titled 'Sentimen Analisis Produk' and contains the text 'Masukkan Kata Kunci di Bahagian Carian untuk menyemak analisis sentimen bagi produk tersebut'. It features two columns: 'Positif' with a green bar and 'Negatif' with a red bar. Below each column is a short paragraph explaining the meaning of the bars. At the bottom, there is a search input field labeled 'Carian' and a blue 'Papar' button.

Rajah 5 Antara Muka Bahagian 'Sentiment'

Jika pengguna ingin memasukkan komen baharu ke dalam pangkalan data, mereka boleh mengisi bahagian Tambah Komen Baharu untuk buat sedemikian. Perkara yang wajib di isi dalam borang ini ialah Nama dan Komen produk tersebut sahaja. Jualan, Harga dan Stok pula adalah pilihan sama ada ingin di isi atau tidak. Rajah 6 menunjukkan contoh di mana pengguna telah memasukkan sebuah komen baharu di mana komen baharu itu akan masuk ke pangkalan data dan model analisis sentimen akan meramalkan apakah sentimen komen tersebut.

**Tambah Komen Baharu**

Nama \*

Ayam goreng

Komen \*

SEDAP GILA LA WEH

Jualan

Masukkan jumlah jualan (tidak wajib)

Harga

Masukkan harga (tidak wajib)

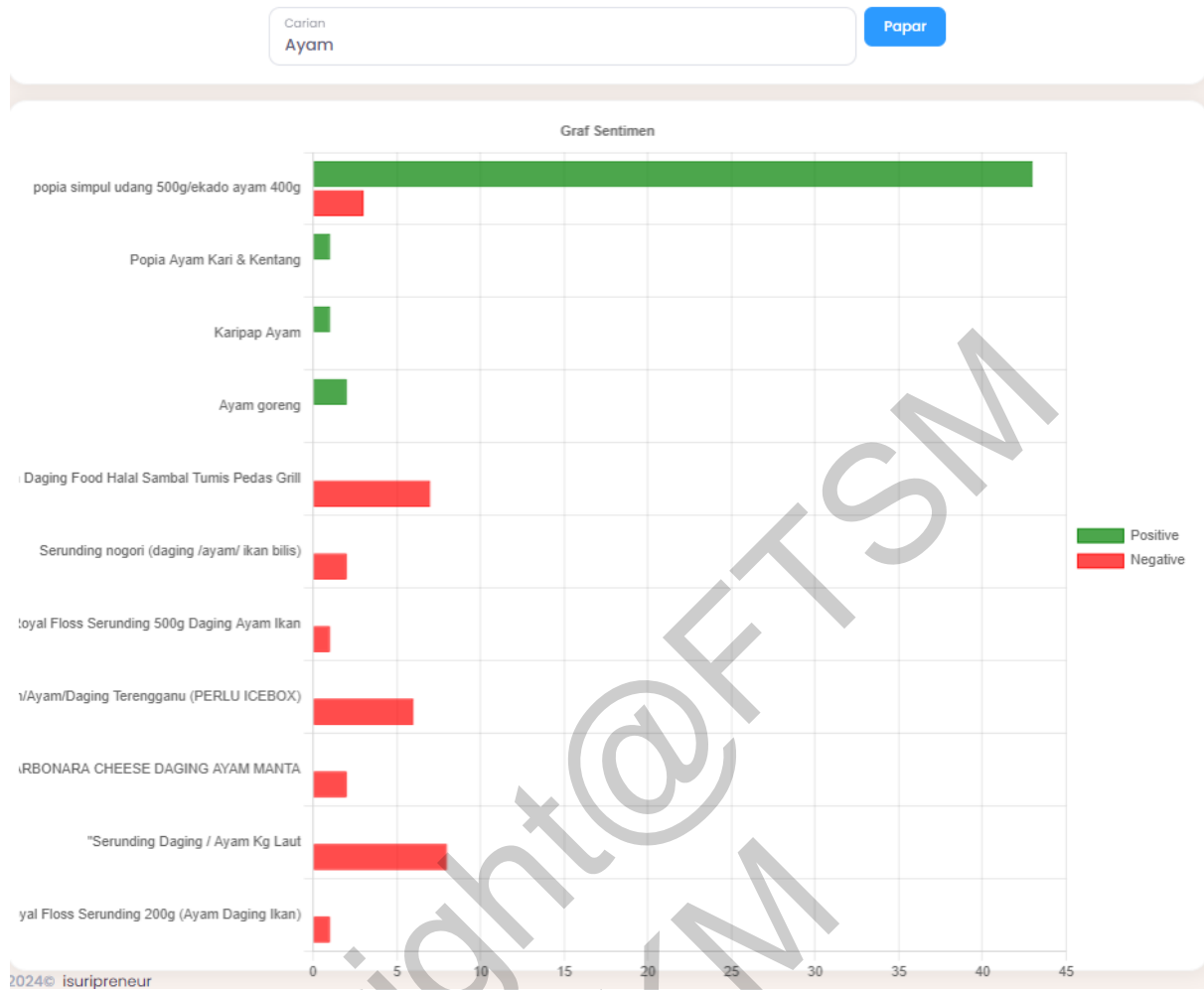
Stok

Masukkan stok (tidak wajib)

Hantar

Rajah 6 Contoh Tambah Komen Baharu

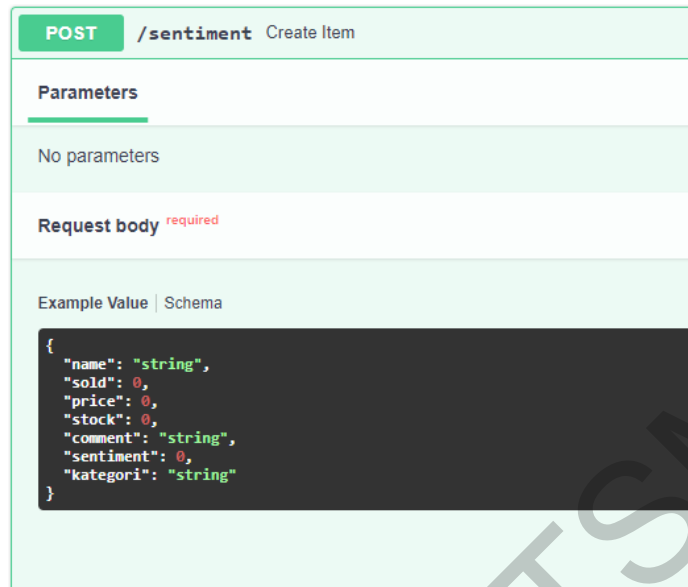
Kemudian, komen yang telah di hantar tersebut akan melalui proses peramalan sentimen dengan menggunakan model analisis sentimen yang telah dibina. Nama, komen dan sentimen produk ini akan disimpan di dalam pangkalan data. Seperti di Rajah 7, jika pengguna ingin melihat sentimen produk ini, mereka boleh memasukkan nama produk ke dalam bahagian carian untuk mengeluarkan sebuah graf sentimen.



Rajah 7 Graf sentimen

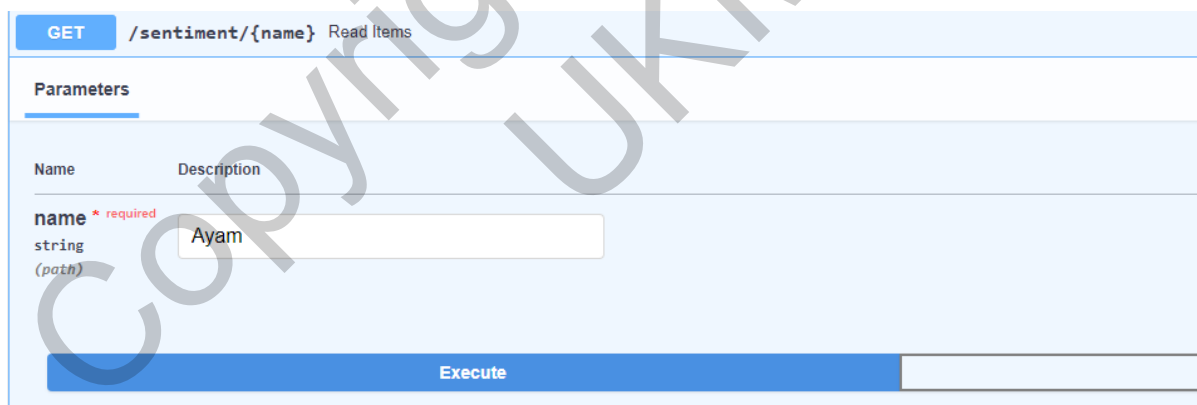
Kesemua proses ini adalah menggunakan FastAPI iaitu dengan kaedah 'POST' dan juga 'GET'. Proses 'POST' adalah untuk mengambil input daripada pengguna selepas pengguna mengisi nama dan komen di antara muka tersebut. Input ini akan di simpan di pangkalan data beserta dengan sentimen dan juga kategori produk. Kategori produk adalah ditetapkan kepada 'Makanan' kerana produk makanan sahaja diterima kerana model telah dilatih dan diuji dalam domain makanan. Rajah 8 menunjukkan struktur kaedah 'POST' di antara muka FastAPI.





Rajah 8 Antara muka kaedah 'POST' di FastAPI

Seterusnya, kaedah 'GET' adalah untuk mendapatkan data daripada pangkalan data untuk memaparkan graf sentimen di antara muka sistem Smart SuriPreneur. Rajah 9 menunjukkan titik akhir untuk mendapatkan data daripada pangkalan data dengan menggunakan kaedah 'GET' beserta dengan carian Ayam seperti di Rajah 7 sebelum ini.



Rajah 9 Antara muka kaedah 'GET' FastAPI

Setelah pengguna memasukkan carian dan menekan butang 'Papar', sebuah blok data yang disusun dalam bentuk JSON akan dihantar ke kod antara muka graf tersebut. Rajah 10 menunjukkan struktur data JSON serta kod 'response' 200 iaitu menunjukkan data berjaya dicapai dan diambil dari pangkalan data.

200

Response body

```
[
  {
    "name": "\npopia simpul udang 500g/ekado ayam 400g",
    "sentiment": 1
  },
  {
    "name": "\npopia simpul udang 500g/ekado ayam 400g",
    "sentiment": 1
  },
  {
    "name": "\npopia simpul udang 500g/ekado ayam 400g",
    "sentiment": 1
  },
  {
    "name": "\npopia simpul udang 500g/ekado ayam 400g",
    "sentiment": 1
  },
  {
    "name": "\npopia simpul udang 500g/ekado ayam 400g",
    "sentiment": 1
  },
  {
    "name": "\npopia simpul udang 500g/ekado ayam 400g",
    "sentiment": 1
  },
  {
    "name": "\npopia simpul udang 500g/ekado ayam 400g",
    "sentiment": 1
  },
  {
    "name": "\npopia simpul udang 500g/ekado ayam 400g",
    "sentiment": 1
  }
]
```

Response headers

Rajah 10 Struktur data JSON

### Cadangan Penambahbaikan

Selepas menjalankan kajian yang menyeluruh, cadangan untuk menambahbaik model analisis sentimen ini ialah dengan melatih dan menguji model ini dengan lebih banyak jenis data. Model ini telah dilatih dalam domain makanan sahaja jadi ia hanya boleh membuat peramalan sentimen hanya terhadap produk dari domain makanan. Untuk menambah baik selepas ini, model boleh dibina dengan lebih banyak domain seperti domain pakaian, servis dan perabot. Jika model dilatih dengan berbagai domain, maka pengguna boleh menggunakan fungsi ‘Sentiment’ di sistem Smart SuriPreneur ini dengan lebih banyak jenis produk yang di jual di Shopee.

Seterusnya, antara muka pada bahagian “Sentiment” juga boleh di tambah baik. Cadangan adalah apabila model dilatih dengan berbagai lagi jenis data dari pelbagai domain maka antara muka boleh di tambah dengan sebuah pilihan kategori. Pilihan ini ialah di mana pengguna dapat memilih kategori produk yang mereka ingin cari. Jadi mereka tidak perlu melihat graf yang terlalu besar yang mempunyai produk yang tidak relevan kepada kategori produk yang mereka ingin cari.

## KESIMPULAN

Projek ini menggunakan pembelajaran mesin untuk melakukan analisis sentimen. Sebuah model analisis sentimen di bina dengan menggunakan tiga model pembelajaran mesin iaitu *naive bayes*, regresi logistik dan mesin support vektor (SVM). Model analisis sentimen ini dibina bagi melakukan analisis sentimen ke atas produk makanan yang dijual di aplikasi Shopee. Model telah dibangunkan menggunakan aplikasi *Google Colaboratory*. Model yang dibina menggunakan tiga model pembelajaran mesin ini akan di lakukan undian pemberat bagi mengetahui model terbaik untuk setiap komen pengguna/pelanggan. Berat merupakan kepentingan atau relevan setiap model tersebut dalam membuat ramalan sentimen terhadap setiap komen. Integrasi ke dalam sistem Smart SuriPreneur pula menggunakan FastAPI dan juga *Visual Studio Code*. Kemudian, Antara muka di Smart SuriPreneur telah di sesuaikan dengan penggunaan model ini dimana ia menggunakan PHP dalam rangka kerja *Laravel/Livewire*.

### **Kekuatan Sistem**

Kekuatan model analisis sentimen ini ialah dapat menganalisis sentimen setiap komen pelanggan dengan ketepatan yang tinggi. Setiap ramalan ini akan disimpan ke dalam pangkalan data yang boleh di akses dengan keselamatan yang tinggi. Ramalan dapat dicapai untuk mengeluarkan graf. Graf sentimen adalah dengan penuh perincian dapat dipaparkan pada antara muka di sistem Smart SuriPreneur ini. Pengguna juga dapat memasukkan komen baharu ke dalam pangkalan data beserta ramalan akan di buat dan juga di simpan dalam pangkalan data. Pengguna memasukkan komen baharu di sistem Smart SuriPreneur juga.

### **Kelemahan Sistem**

Kekangan model analisis sentimen ini pula ialah ia hanya dilatih dan diuji dengan menggunakan domain makanan sahaja. Jika pengguna memasukkan komen dalam domain lain, kemungkinan terdapat kesalahan dalam ramalan analisis sentimen. Kemudian, terdapat kekangan di mana pengguna hanya boleh memasukkan carian nama produk untuk memaparkan graf sentimen, dan bukan memilih kategori produk, ini kerana kategori bagi produk ini adalah ditetapkan kepada makanan.

## PENGHARGAAN

Pertama sekali, penulis kajian ingin memberikan ucapan syukur ke hadrat ilahi kerana penulis kajian dapat menyiapkan projek ini dengan segera dan keadaan sihat.

Seterusnya, sekalung penghargaan dan terima kasih kepada penyelia penulis kajian, Prof Madya Dr. Sabrina Tiun. Beliau telah meluangkan masa yang banyak dalam menunjuk ajar kepada penulis kajian, memberi bantuan yang banyak dan meluangkan masa untuk melakukan perjumpaan untuk membincangkan tentang projek ini.

Selain itu, ucapan terima kasih kepada semua ahli keluarga dan rakan-rakan penulis kajian ini kerana memberikan pelbagai jenis sokongan sepanjang menyiapkan projek ini.

## RUJUKAN

- Department of Statistics, Malaysia. Poverty in Malaysia. Poverty in Malaysia | [https://www.dosm.gov.my/portal\\_main/release-content/poverty-in-malaysia-](https://www.dosm.gov.my/portal_main/release-content/poverty-in-malaysia-) [6 Disember 2023]
- Ecentral.My. Maksud miskin tegar B40 M40 T20. | <https://ecentral.my/maksud-miskin-tegar-b40-m40-t20/#:~:text=Ringkasannya%2C%20dalam%20Laporan%20Survei%20Pendapatan,T20%20%3A%20RM11%2C820%20ke%20atas> [6 Disember 2023].
- Javatpoint. Machine Learning Life Cycle. | <https://www.javatpoint.com/machine-learning-life-cycle> [3 November 2023].
- Mysumber.Com. Apa itu B40, M40 dan T20 Maksud Golongan Isi Rumah Mengikut Pendapatan. | <https://www.mysumber.com/b40-m40-t20.html> [1 November 2023].

*Muhammad Adam Shukri Bin Raymond Azman (A189107)*

*Prof. Madya Dr. Sabrina Tiun*

Fakulti Teknologi & Sains Maklumat

Universiti Kebangsaan Malaysia