

Deteksi Potensi bunuh diri

kelompok 6

Our Great Team



22031554024

NAHDLIYAH ZAHRAH



RESHAR FALDI JULIANDA

22031554025







22031554039

NUR HALIZAH AMRITA

Latar Belakang

Masalah bunuh diri adalah isu serius dalam kesehatan mental yang memerlukan perhatian khusus. Menangani masalah ini dapat membantu mencegah tragedi yang menghancurkan kehidupan individu dan keluarga. Salah satu pendekatan untuk mendeteksi potensi risiko bunuh diri adalah melalui analisis teks, khususnya dengan menggunakan metode Linear Support Vector Machine (SVM) dan model Bag-of-Words (BoW). Dengan merancang dan mengimplementasikan model menggunakan metode ini, diharapkan dapat dikembangkan alat yang efektif dalam mendukung upaya pencegahan bunuh diri dengan menganalisis potensi risiko melalui data teks.

Data Set

 Tweet 	 Suicide 
The "Tweet" column contains text content from Twitter, reflecting users' language, thoughts, emotions, and discussions	"Potentially Suicidal" indicates distress/self-harm concerns, "Not Suicidal" signifies absence of such distress
1778 unique values	Not Suicide post 63% Potential Suicide p... 37%
making some lunch	Not Suicide post
@Alexia You want his money.	Not Suicide post
@dizzyhrvy that crap took me forever to put together. iâ m going to go sleep for DAYS	Potential Suicide post

Preprocessing Data

1. `htm.unescape(text)`:

Mendekode teks HTML entities. Contohnya, mengubah `&` menjadi `&`.

2. `convert_abbrev(text)`:

Mengganti singkatan dalam teks dengan bentuk lengkapnya, jika ada di kamus abbreviations.

3. `lowercase(text)`:

Mengubah seluruh teks menjadi huruf kecil.

4. `remove_unnecessary_char(text)`:

Menghapus karakter yang dianggap tidak perlu, seperti newline (`\n`), username (user), dan URL.

5. `remove_nonalphanumeric(text)`:

Menghapus karakter non-alphanumeric, kecuali spasi.

6. `expand2_contractions(text)`:

Memperluas kontraksi dalam teks. Misalnya, mengubah `"won't"` menjadi `"will not"`.

7. `remove_stopword(text)`:

Menghapus kata-kata penghubung (stop words) berdasarkan kamus stop words bahasa Inggris.

8. `remove_punctuation(text)`:

Menghapus tanda baca dari teks.

9. `english_lemmatization(text)`:

Melakukan lemmatization pada kata-kata dalam teks menggunakan WordNet lemmatizer.

Linear SVM n BoW

lalu memilih model Linear SVM dan Bow dengan akurasi 0.90

[illegible]

METODE



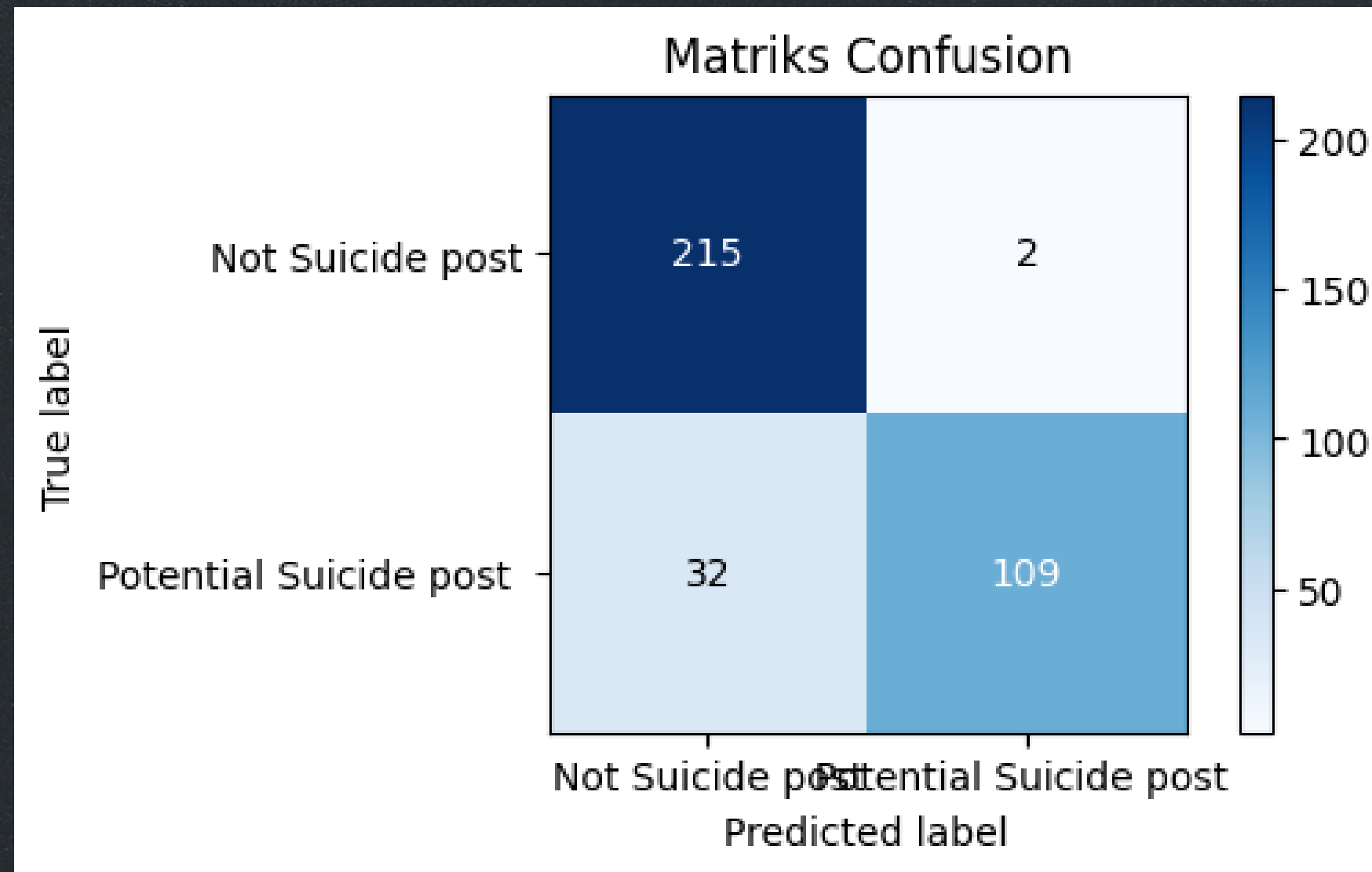
Linear SVM

linear SVM digunakan untuk masalah klasifikasi biner, di mana tujuannya adalah untuk menemukan hyperplane yang memisahkan dua kelas data dengan margin maksimum. Dalam konteks deep learning, linear SVM digunakan sebagai pengganti lapisan softmax dalam model deep learning untuk meningkatkan kinerja dan efek regularisasi.

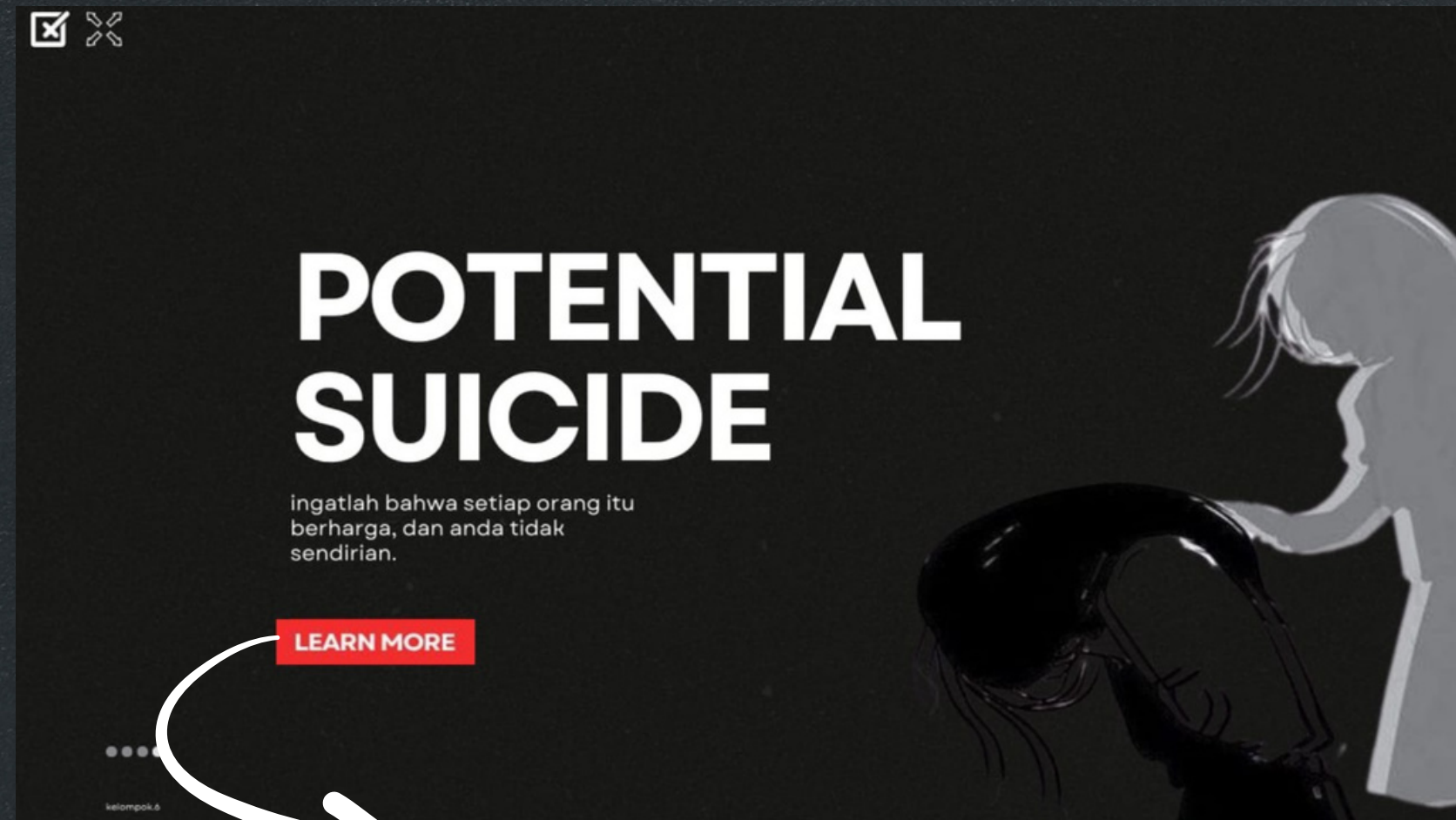
Bag Of Word

Bag of Words (BoW) adalah model yang serbaguna dan dapat digunakan sebagai algoritma pemilihan fitur dan klasifikasi dokumen serta gambar. Metode BoW merekam jumlah kemunculan setiap kata yang dibuat untuk setiap jenis instansi atau kata, tanpa memperhatikan urutan kata atau tata bahasa

Matriks Confusion



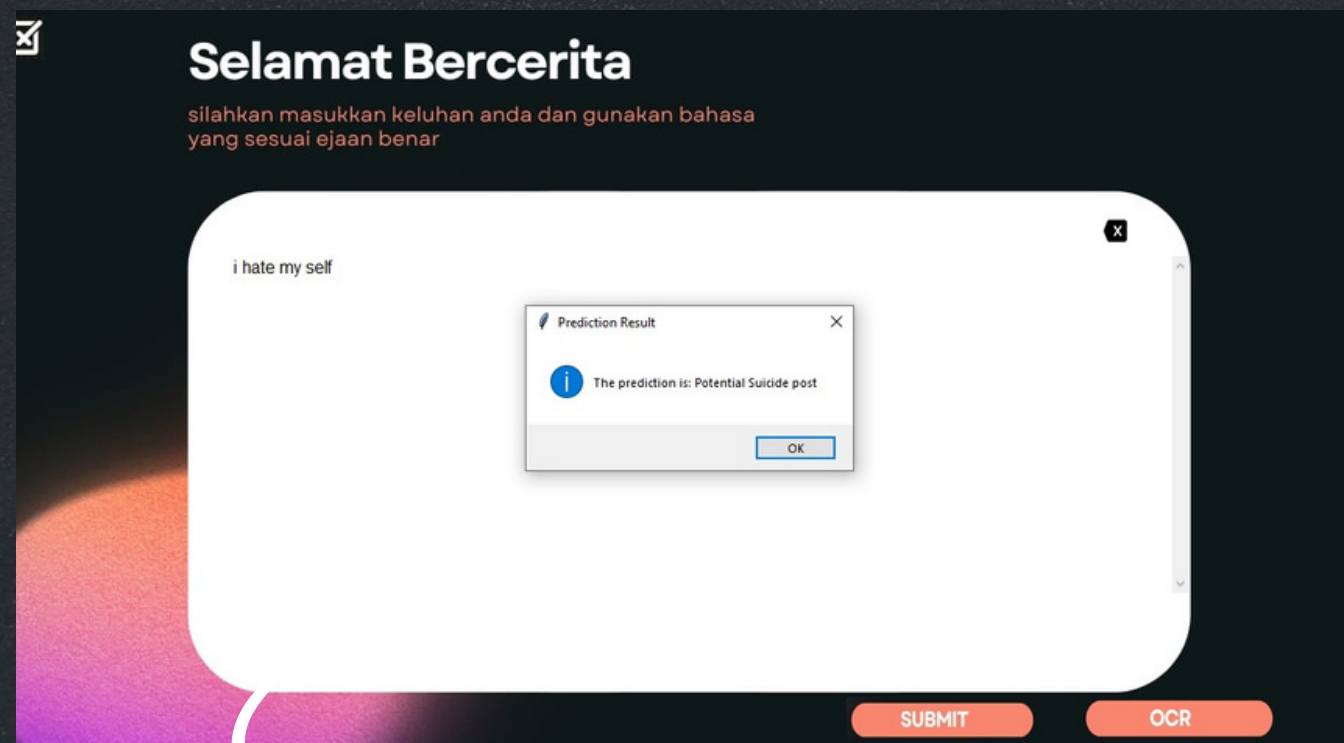
Tampilan GUI



Tampilan Awal

tampilan awal. pengguna dapat melanjutkan ke halaman berikutnya dengan tombol 'Learn More'

Tampilan GUI

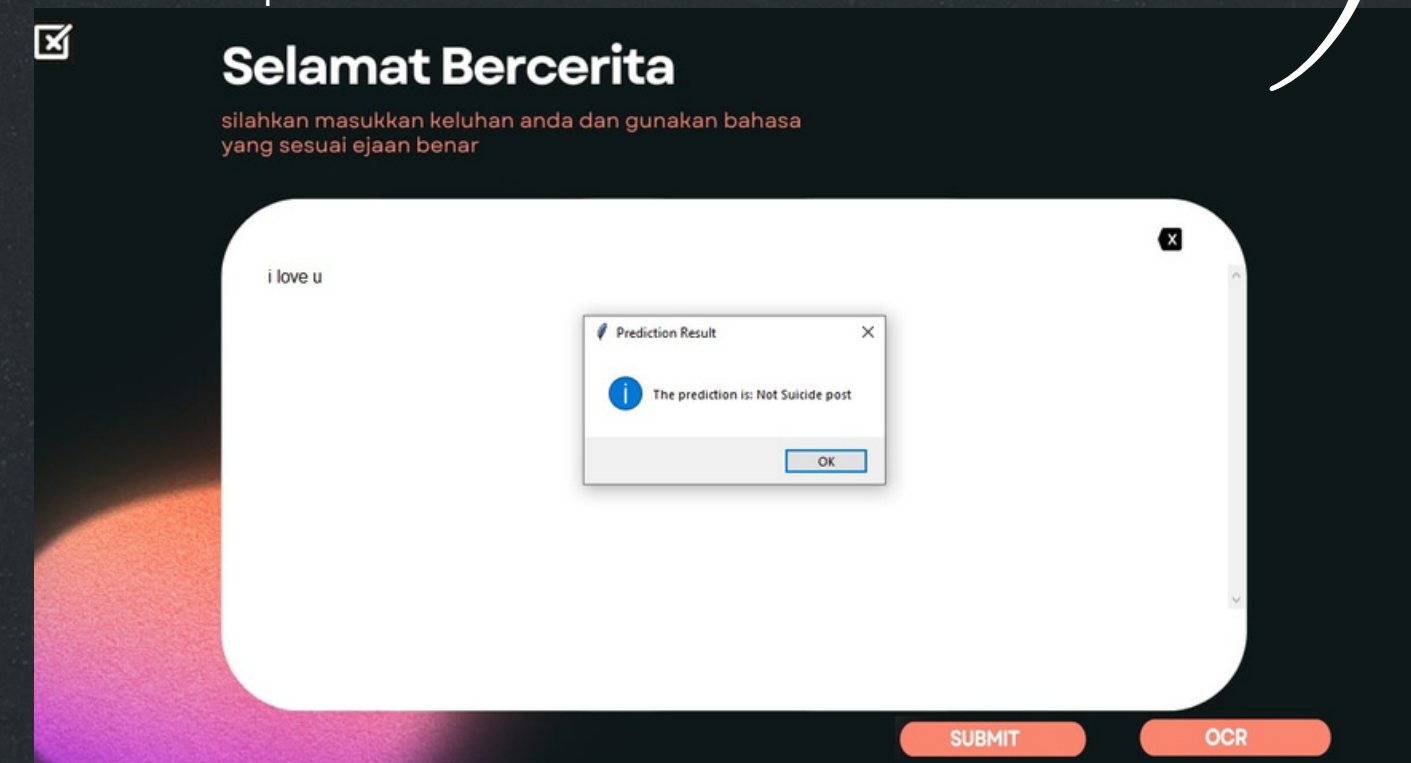


Potential Suicide

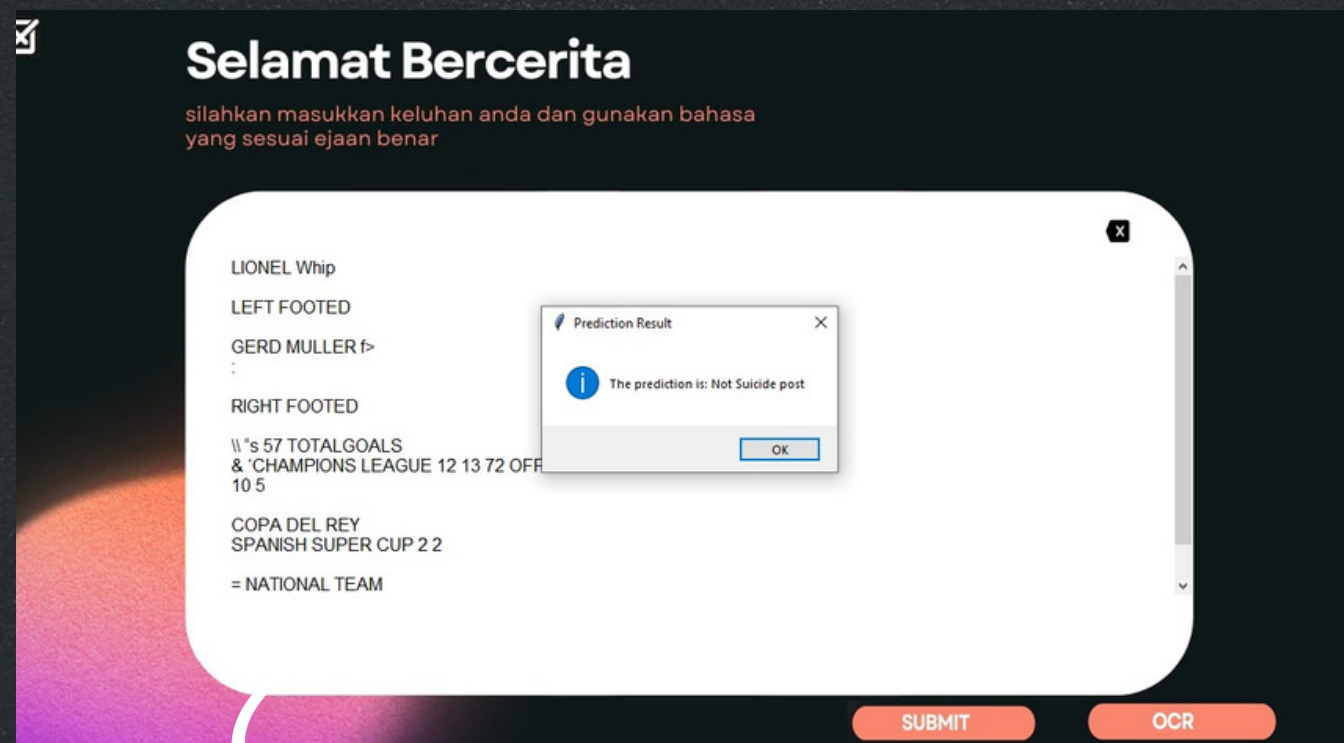
jika memasukkan kata (input manual) ke ranah bunuh diri (negatif) maka prediksinya Berpotensi Bunuh Diri

Not Potential Suicide

jika memasukkan kata (input manual) ke ranah yang jauh dari kata bunuh diri (positif) maka prediksinya Tidak Berpotensi Bunuh Diri



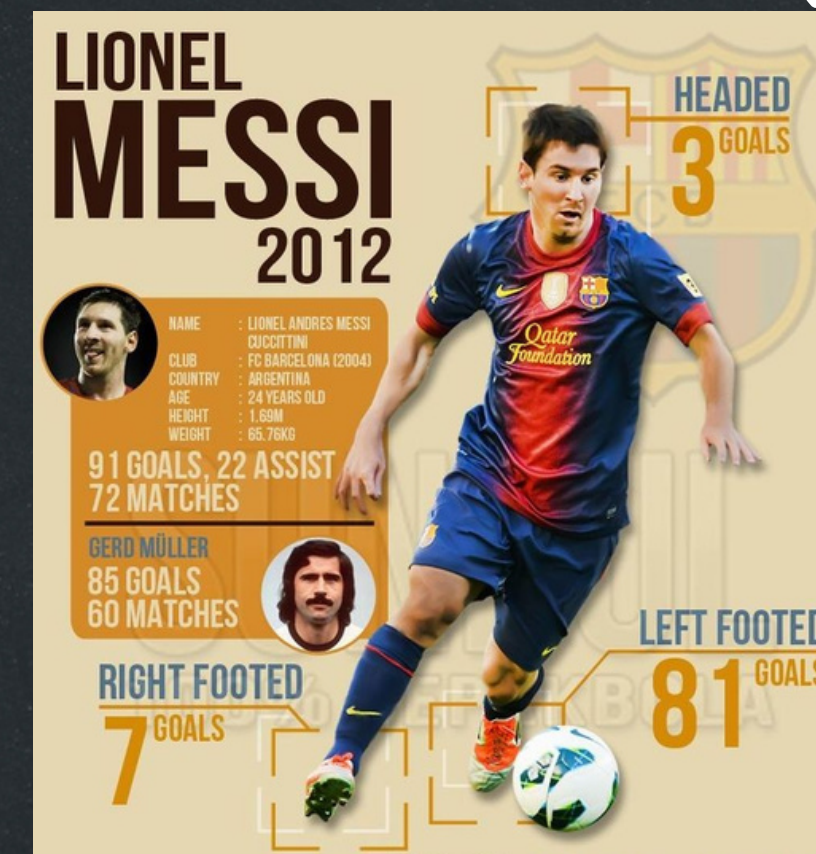
Tampilan GUI



Input Kata tombol 'OCR'
pengguna dapat memasukkan kata kata melalui tombol 'OCR', di mana tombol ini memungkinkan pengguna untuk menginput gambar yang dapat mendeteksi kata kata.

Gambar input 'OCR'

gambar berikut adalah contoh gambar yang di input pada melalui tombol 'OCR'





THANK YOU
FOR ALL ATTENTION

