NLP KLASIFIKASI SENTIMENT PADA COMMENT TWITTER MENGGUNAKAN DECISION TREE

KELOMPOK 3A :

ANTON WIJAYA - UNTIDAR

FATIH - UTY

NATAL

LINK GOOGLE COLABB

- https://colab.research.google.com/drive/1SCAQ3ipHM_6dO2HqyHblu3cRhXAaAwXt?usp=sharing

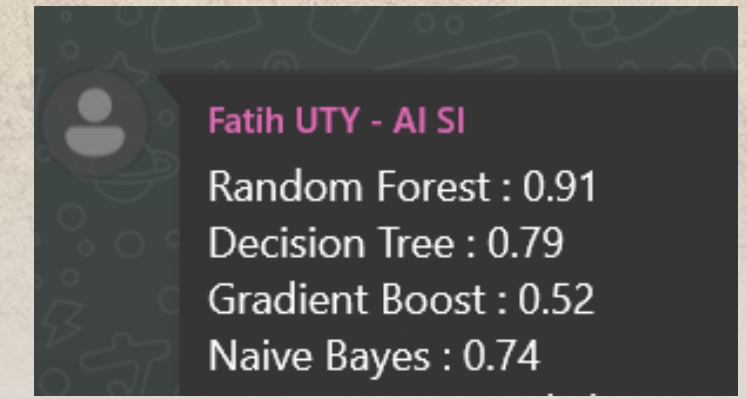
DESKRIPSI DATASET

- Dataset yang digunakan adalah dataset comment twitter yang didaptkan dari situs Kaggle. Kolom data yang digunakan adalah sentiment, branch (sumber), dan comment.

	Branch	Sentiment	Tweet
0	Borderlands	Positive	I am coming to the borders and I will kill you...
1	Borderlands	Positive	im getting on borderlands and i will kill you ...
2	Borderlands	Positive	im coming on borderlands and i will murder you...
3	Borderlands	Positive	im getting on borderlands 2 and i will murder ...
4	Borderlands	Positive	im getting into borderlands and i can murder y...
5	Borderlands	Positive	So I spent a few hours making something for fu...
6	Borderlands	Positive	So I spent a couple of hours doing something f...
7	Borderlands	Positive	So I spent a few hours doing something for fun...
8	Borderlands	Positive	So I spent a few hours making something for fu...
9	Borderlands	Positive	2010 So I spent a few hours making something f...

DESKRIPSI MODEL AI

- Model NLP yang dibangun adalah model AI yang dapat mengklasifikasikan sentiment dari data comment. Alogaritma yang digunakan adalah alogaritma decision tree.
- Alogaritma ini dipilh karena alogaritma ini memiliki akurasi tertinggi selain dari alogaritma random forest (model asli).



HASIL EVALUASI MODEL

Predicted Labels

Classification Report:

	precision	recall	f1-score	support
Irrelevant	0.83	0.61	0.70	2511
Negative	0.70	0.85	0.77	4296
Neutral	0.82	0.63	0.72	3498
Positive	0.70	0.80	0.75	3887
accuracy			0.74	14192
macro avg	0.76	0.72	0.73	14192
weighted avg	0.75	0.74	0.74	14192

```
Actual vs Predicted Labels with Tweet Text:
Tweet: women please work stores like circuit home hardware get bored condescending men thinking ’ know another screw
Actual Label: Positive
Predicted Label: Negative
----------
Tweet: zai absolute beast reddit fe
Actual Label: Neutral
Predicted Label: Positive
----------
Tweet: hard fought indeed …
Actual Label: Neutral
Predicted Label: Neutral
----------
Tweet: rhandlerr ’ superstar solo wins criminal deserved pictwittercomhilloxffi
Actual Label: Irrelevant
Predicted Label: Irrelevant
----------
```

EVALUASI MODEL

Model yang dibangun menunjukkan kinerja yang cukup baik dengan akurasi keseluruhan sebesar 74%, menunjukkan **kemampuan yang cukup untuk mengklasifikasikan data dengan benar.** Meskipun presisi, recall, dan skor F1 bervariasi di antara kelas-kelas yang berbeda, model ini secara konsisten menunjukkan tingkat presisi dan recall yang cukup tinggi untuk kelas "Negatif" dan "Positif", sementara kelas "Netral" dan "Tidak Relevan" memiliki performa yang sedikit lebih rendah. Meskipun demikian, rata-rata makro dan rata-rata terbobot dari metrik evaluasi menunjukkan distribusi kelas yang seimbang, menandakan bahwa **model ini mampu menangani berbagai jenis kelas dengan baik.**

MODEL AI

```python
Tweet_Sentiment_Ai = joblib.load('rf_model_decisontree.pkl') #LOAD model MLnya

# definisi kan emoji sesuai dengan sentimentnya
sentiment_emojis = {
    'Positive': '😄',
    'Negative': '😒',
    'Neutral': '😐',
    'Irrelevant': '🤖'
}

#  masukan tweets disini !!
tweets = [
    "I just finished playing Borderlands and it was absolutely amazing! Can't wait for the next one!",
    "I'm really disappointed with the latest Borderlands update. It ruined the game for me.",
    "Haven't played Borderlands in a while. Need to catch up on the latest updates.",
    "Help me superman"
]


# Transform example tweets into TF-IDF features
tweets_tfidf = vectorizer.transform(tweets)

# Prediksi tweets
predictions = Tweet_Sentiment_Ai.predict(tweets_tfidf)

# Print the predicted sentiment and corresponding emoji for each example tweet
for tweet, prediction in zip(tweets, predictions):
    sentiment = prediction
    emoji = sentiment_emojis[sentiment]  # Get the emoji directly from the mapping without a default value
    print("Tweet:", tweet)
    print("Sentiment:", sentiment)
    print("Emoji:", emoji)
    print()
```

```
Tweet: I just finished playing Borderlands and it was absolutely amazing! Can't wait for the next one!
Sentiment: Positive
Emoji: 😄

Tweet: I'm really disappointed with the latest Borderlands update. It ruined the game for me.
Sentiment: Negative
Emoji: 😒

Tweet: Haven't played Borderlands in a while. Need to catch up on the latest updates.
Sentiment: Neutral
Emoji: 😐

Tweet: Help me superman
Sentiment: Irrelevant
Emoji: 🤖
```

TERIMA KASIH