



# ANALISA SENTIMENT PADA TWEET

Challenge Platinum level  
Dsc binar wave 7

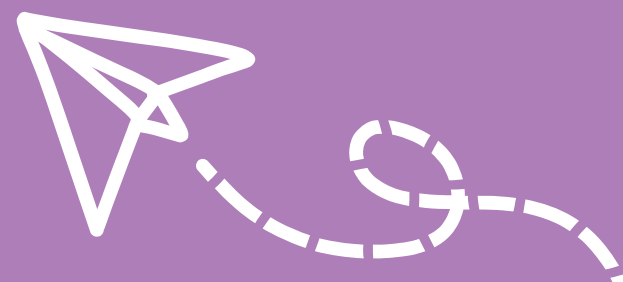
**Hermawan**  
**Rahmatina Ari A**  
**Rona Wenas**

# Latar belakang

Menurut KBBI, sentimen berarti pendapat yang didasarkan pada perasaan yang berlebihan terhadap sesuatu.

Analisis sentimen sendiri merupakan proses menganalisis teks untuk menentukan apakah emosi pada teks tersebut positif, negative, atau netral.

Analisis sentimen tidak hanya digunakan di media social seperti Twitter saja. Banyak perusahaan melakukan analisis sentiment untuk meningkatkan produk dan layanan, seperti analisis berdasarkan ulasan dari pelanggan

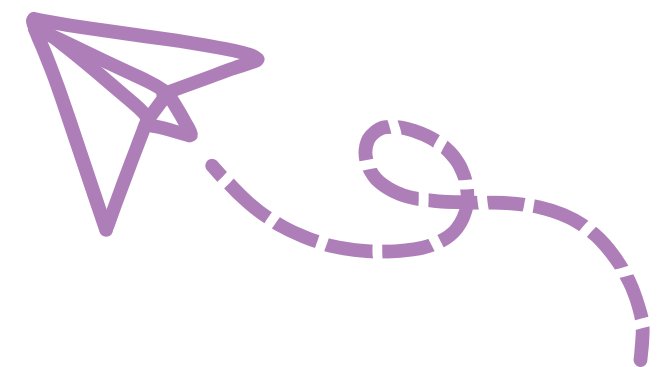


# Tujuan

Mengidentifikasi sebaran data tweet positif, negatif, dan netral

Membuat model yang optimal untuk memprediksi sentimen dari sebuah teks

Membuat API yang dapat mengklasifikasikan sentimen yang dihasilkan



# Data Preparation

11.000

Baris data

2

Kolom.

Tweet Bahasa Indonesia

Label positif, negatif, netral

0

Missing Values

67

Data duplikat

```
#Cek Data  
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 11000 entries, 0 to 10999  
Data columns (total 2 columns):  
#   Column      Non-Null Count  Dtype  
---  -  
0   Kalimat    11000 non-null   object  
1   Label      11000 non-null   object  
dtypes: object(2)  
memory usage: 172.0+ KB
```

```
# Cek Missing Values
```

```
df.isna().sum()
```

```
Kalimat    0  
Label      0  
dtype: int64
```

```
# Cek duplikasi data
```

```
print('{} data duplikat'.format(df.duplicated().sum()))
```

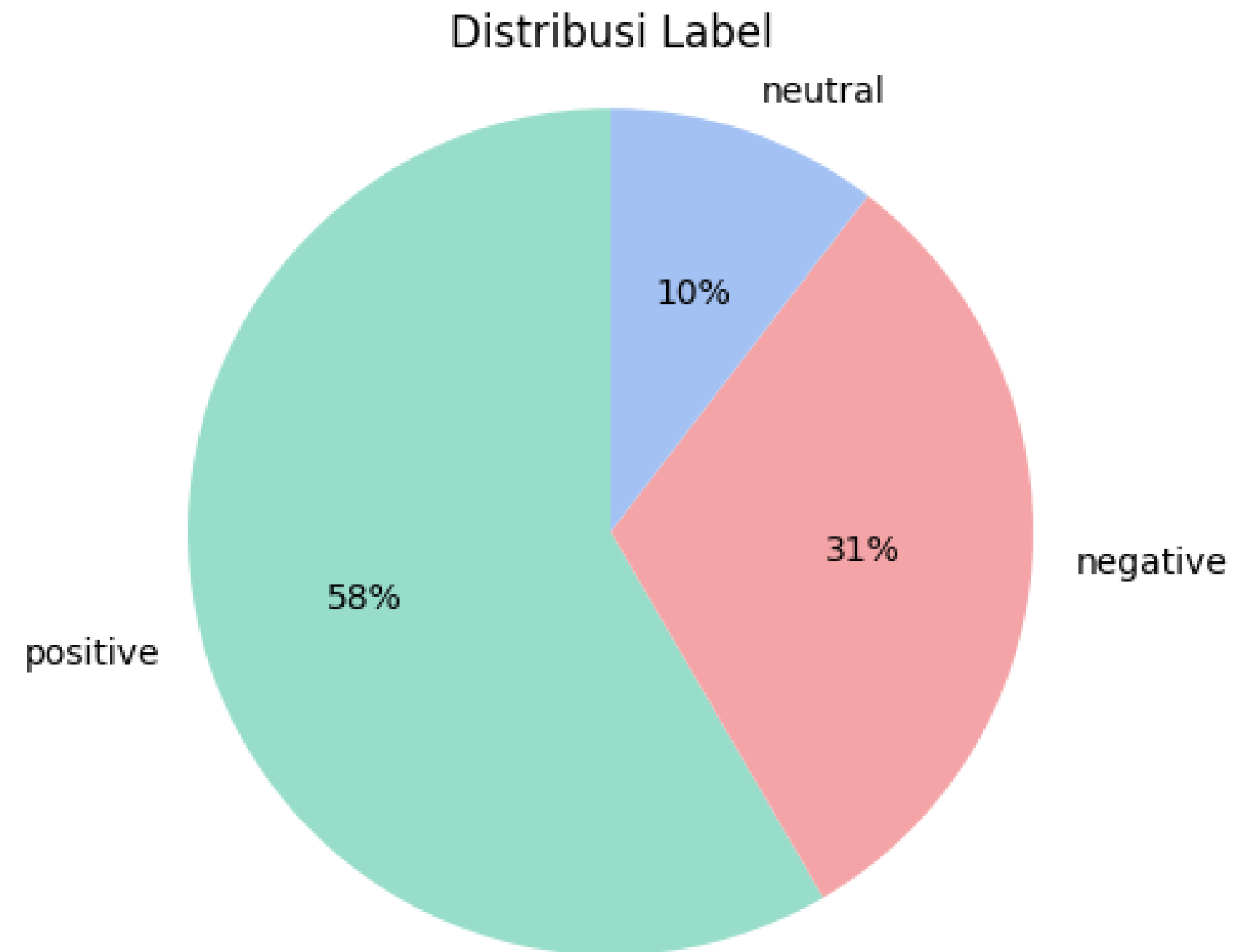
```
67 data duplikat
```

# Exploratory Data Analysis

Sentiment Positif: 6416 Tweet

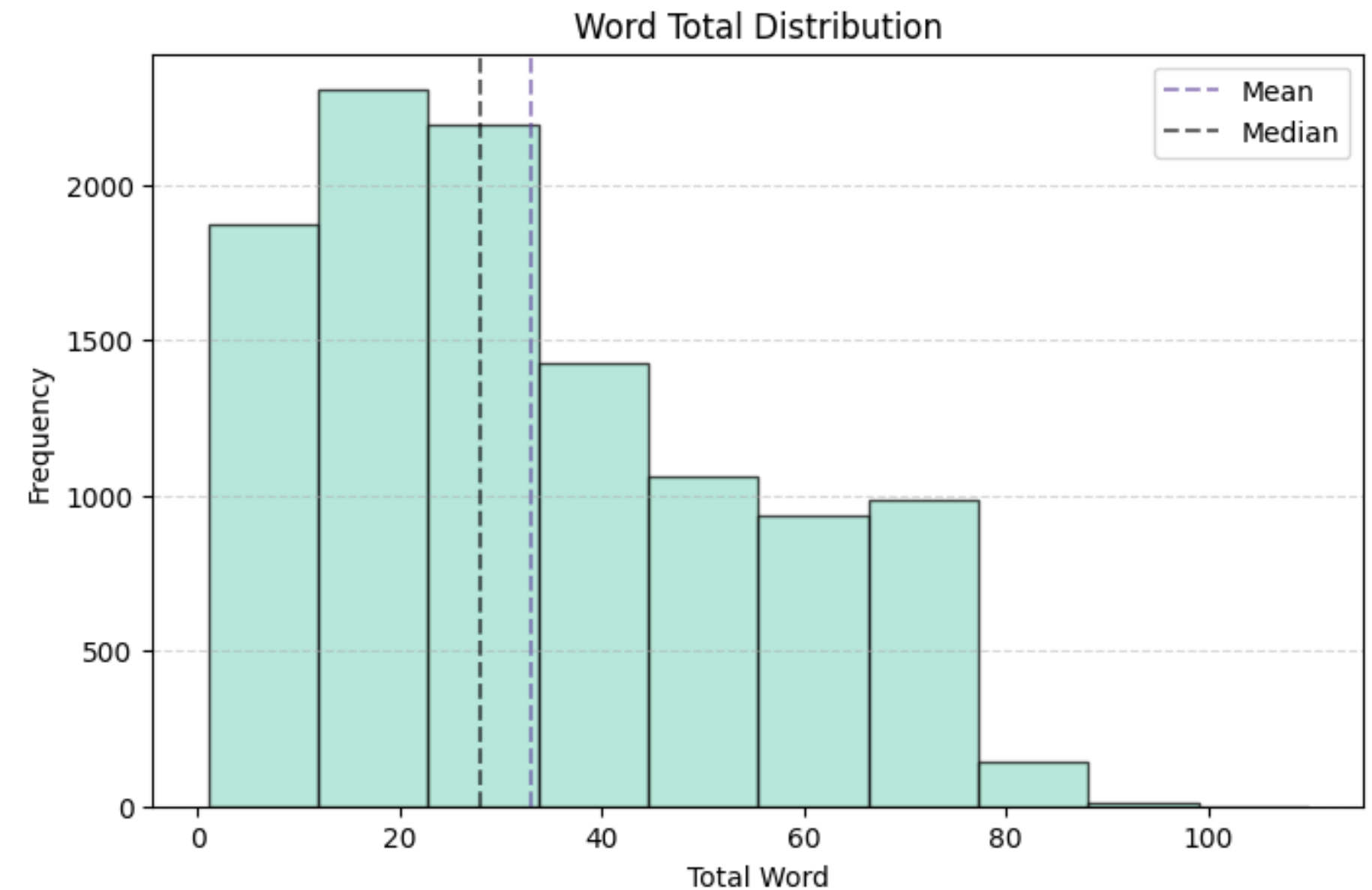
Sentiment Negatif: 3436 Tweet

Sentiment Netral: 1148 Tweet

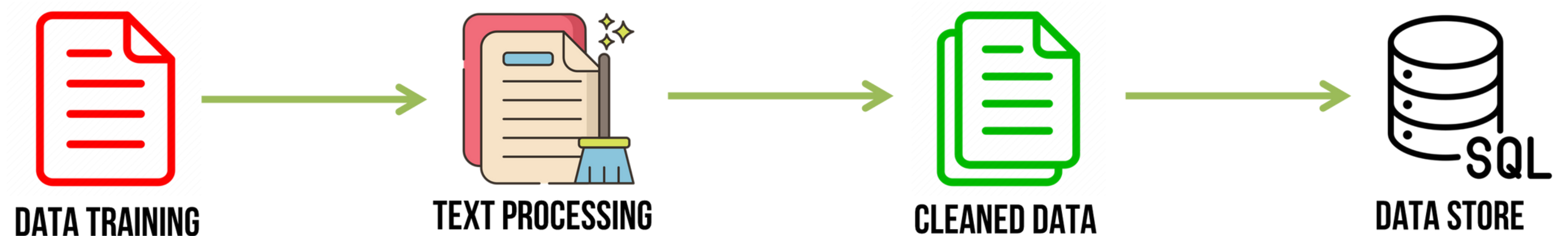


# Exploratory Data Analysis

- Modus: 20 – 30 kata
- Median: 28 kata
- Mean: 33 kata



# Text Normalization



Sebelum: nkri bersyariah , fpi , revolusi akhlaq , suara keadilan , \_\_laugh\_\_ semua pada tersandung kasus , fpi , ulama nya tersandung obrolan mesum tidak benar

Sesudah: negara kesatuan republik indonesia bersyariah front pembela islam revolusi akhlaq suara keadilan laugh semua pada tersandung kasus front pembela islam

# Machine Learning Preparation

## KLASIFIKASI LABEL FEATURE

Memisahkan fitur (X) dan label (Y) pada dataset



## FEATURE EXTRACTION

Mengekstraksi fitur dan mengubah data teks menjadi bilangan vector  
(Tokenizer dan Pad Sequence)

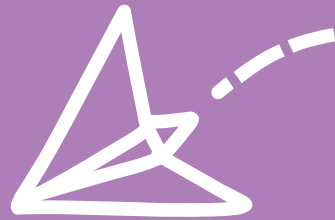


## TRAIN – TEST DATA SPLIT

Membagi data menjadi data train dan data test dengan perbandingan 80:20

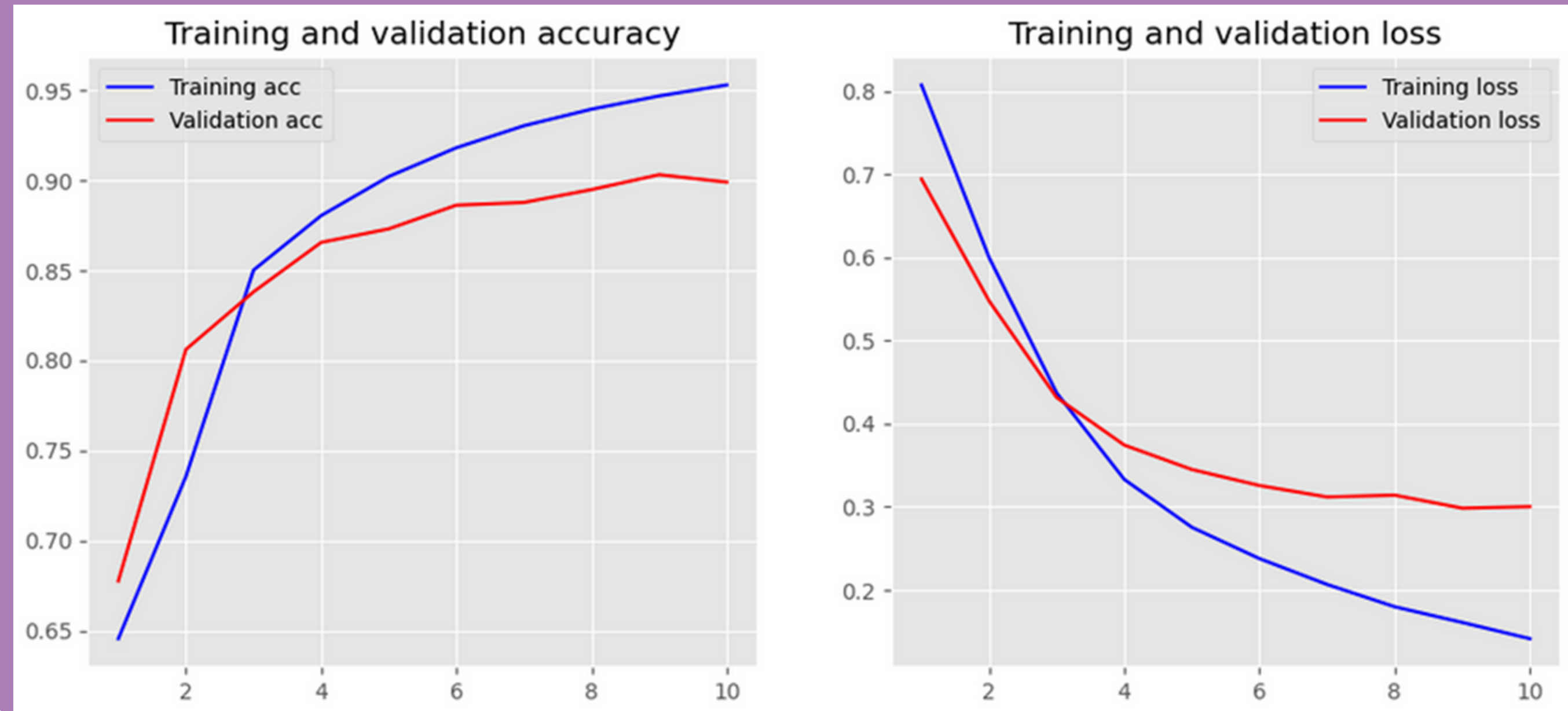


# Hasil Analisis



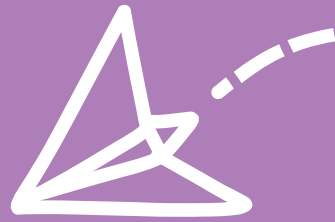
## MODEL TRAINING DAN EVALUATION (LSTM)

Parameter	
Activation	Softmax
Learning Rate	0.00005
Epoch	10
Batch size	10
Embedding dimension	200
Cross validation	5
Regularization	
Dropout	0.2
Early Stopping	
• Monitor	val_loss
• Mode	min
• Verbose	1
• Patience	2



Dengan menggunakan model LSTM dihasilkan rata-rata akurasi sebesar 0.94 dan capaian loss pada iterasi ke-5 Sebesar 0.14 dan val\_loss 0.33

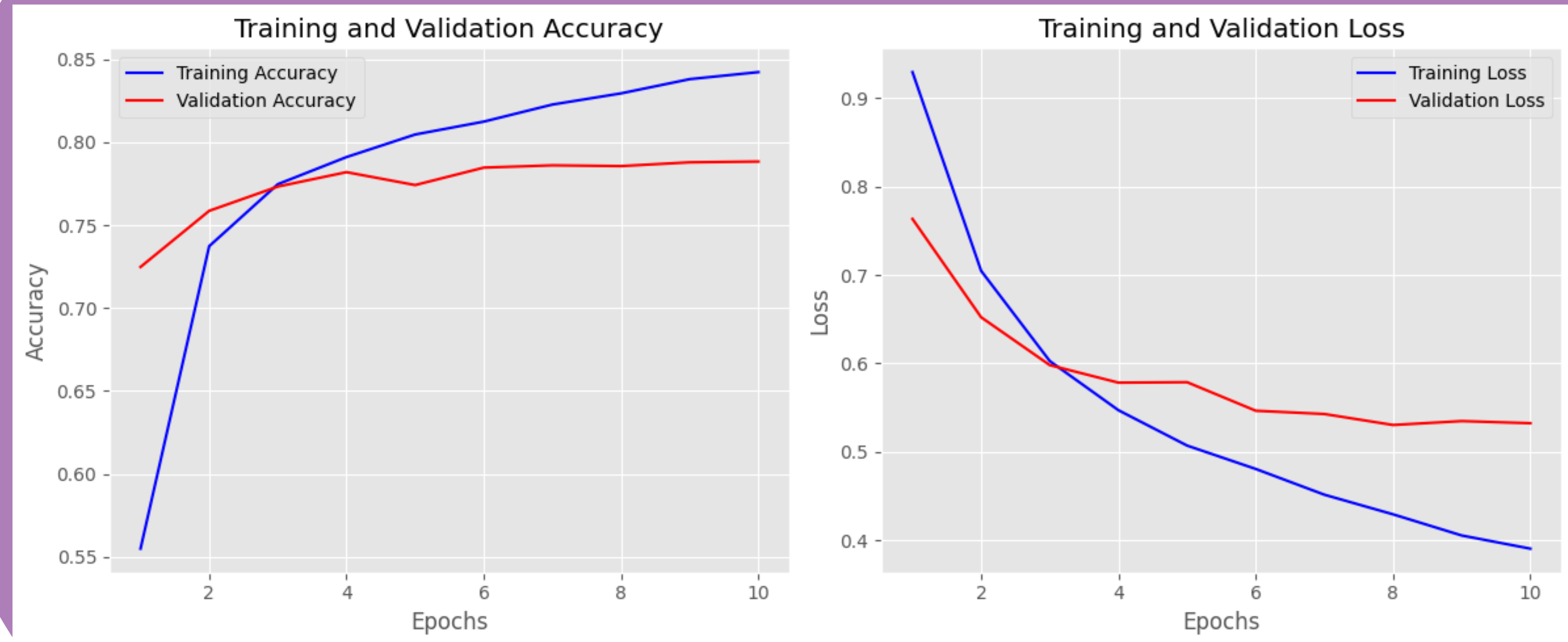
# Hasil Analisis



## MODEL TRAINING DAN EVALUATION (RNN)

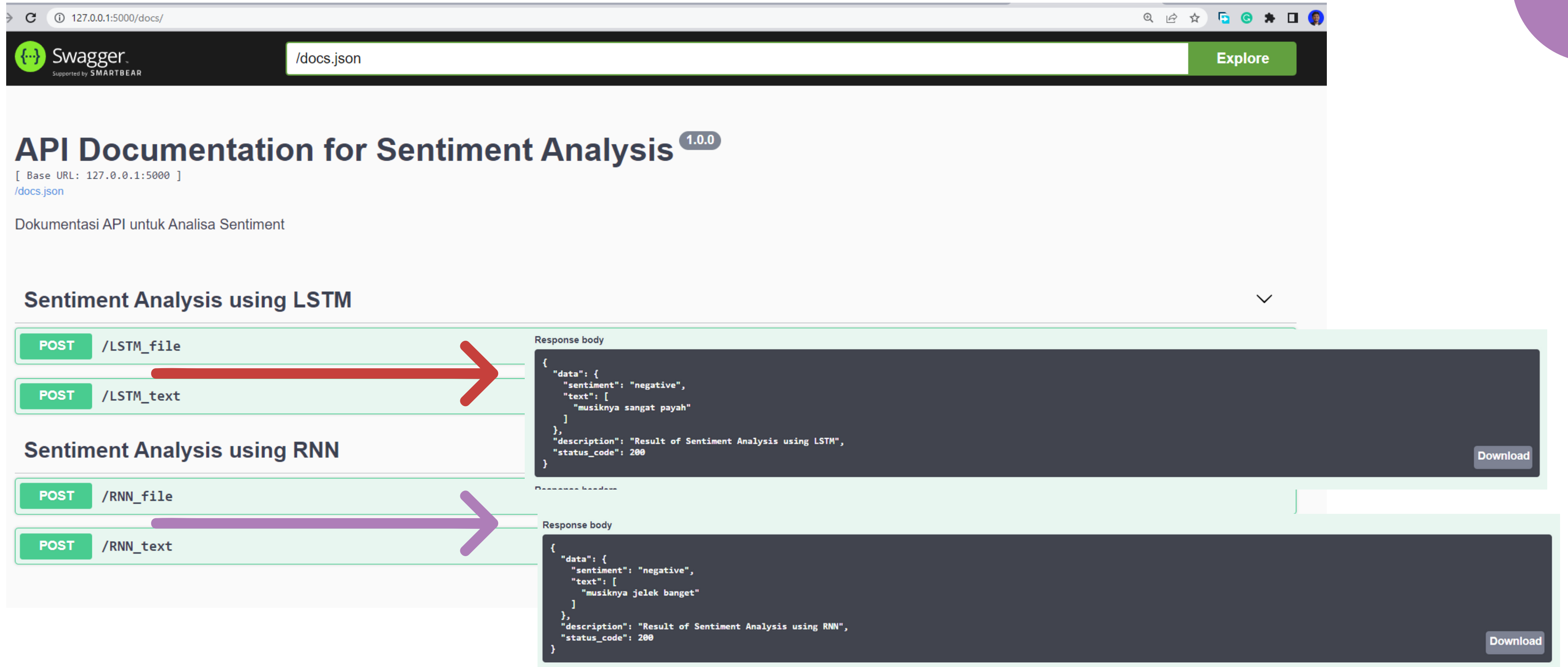
Parameter	
Activation	Softmax
Learning Rate	0.00005
Epoch	10
Batch size	10
Embedng dimension	200
Cross validation	5

Regularization	
Dropout	0.2
Early Stopping	
• Monitor	val_loss
• Mode	min
• Verbose	1
• Patience	2



Dengan menggunakan model LSTM dihasilkan rata-rata akurasi sebesar 0.94 dan capaian loss pada iterasi ke-5 Sebesar 0.14 dan val\_loss 0.33

# Prediksi dan Deployment



The image shows a Swagger API documentation page for a Sentiment Analysis service. The page is titled "API Documentation for Sentiment Analysis 1.0.0" and includes a base URL of "127.0.0.1:5000/docs/". The documentation is organized into two main sections: "Sentiment Analysis using LSTM" and "Sentiment Analysis using RNN". Each section lists two endpoints: a POST endpoint for file uploads and a POST endpoint for text input. Red and purple arrows point from the endpoints to their respective response body examples. The response bodies are JSON objects containing sentiment analysis results.

Swagger  
Supported by SMARTBEAR

/docs.json Explore

## API Documentation for Sentiment Analysis 1.0.0

[ Base URL: 127.0.0.1:5000 ]  
[/docs.json](#)

Dokumentasi API untuk Analisa Sentiment

### Sentiment Analysis using LSTM

**POST** /LSTM\_file

**POST** /LSTM\_text

**Response body**

```
{
  "data": {
    "sentiment": "negative",
    "text": [
      "musiknya sangat payah"
    ]
  },
  "description": "Result of Sentiment Analysis using LSTM",
  "status_code": 200
}
```

[Download](#)

### Sentiment Analysis using RNN

**POST** /RNN\_file

**POST** /RNN\_text

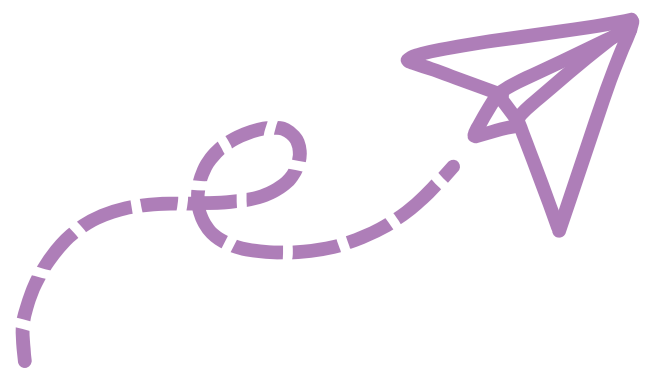
**Response body**

```
{
  "data": {
    "sentiment": "negative",
    "text": [
      "musiknya jelek banget"
    ]
  },
  "description": "Result of Sentiment Analysis using RNN",
  "status_code": 200
}
```

[Download](#)



## Hasil

1. Dataset terdiri dari 58% data sentiment positif (6416 data), 31% data sentiment negative (3436 data), dan 10% data sentiment netral (1148 data)
  2. Model LSTM memiliki performa yang lebih baik dari model RNN maupun ANN, dibuktikan dengan hasil capaian akurasi yang lebih tinggi yakni 0.94 dan loss yang lebih rendah yaitu 0.16
- 



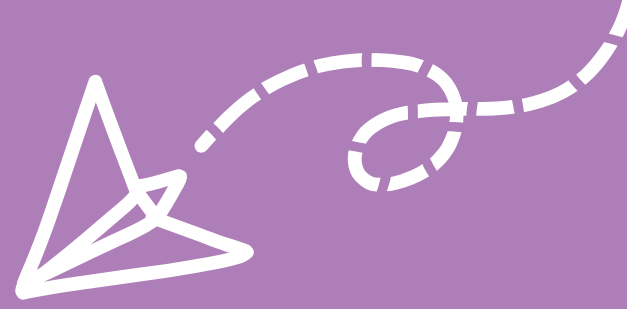
# Kesimpulan

Secara kesimpulan, deep learning merupakan teknik yang sangat kuat dalam mengenali pola-pola rumit dalam data dan menghasilkan hasil yang lebih akurat dibandingkan dengan pendekatan machine learning tradisional.

Kelebihan utama deep learning adalah **fleksibilitasnya** yang memungkinkan aplikasinya pada berbagai jenis data, termasuk gambar, suara, sinyal, dan sebagainya. Namun, deep learning juga memiliki beberapa kendala, yaitu kebutuhan akan data yang cukup besar, yang dapat menjadi tantangan saat bekerja dengan dataset terbatas dan berpotensi menyebabkan overfitting.



Selain itu, deep learning juga memerlukan perangkat keras yang kuat dan komputasi yang intensif untuk melatih model dengan efisien.



# Terima kasih

For Documentation

