Analisis Sentimen IMDb dengan RoBERTa-base (Fine-tuning)

```
# Install libraries
!pip install transformers datasets evaluate scikit-learn
```

```
Requirement already satisfied: transformers in /usr/local/lib/python3.11/dist-packages (4.54.0)
Requirement already satisfied: datasets in /usr/local/lib/python3.11/dist-packages (4.0.0)
Collecting evaluate
  Downloading evaluate-0.4.5-py3-none-any.whl.metadata (9.5 kB)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.11/dist-packages (1.6.1)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from transformers) (3.18.0)
Requirement already satisfied: huggingface-hub<1.0,>=0.34.0 in /usr/local/lib/python3.11/dist-packages (from transformer
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages (from transformers) (2.0.2)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from transformers) (25.0)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages (from transformers) (6.0.2)
Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.11/dist-packages (from transformers) (2024.11
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from transformers) (2.32.3)
Requirement already satisfied: tokenizers<0.22,>=0.21 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.
Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.5.3)
Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.11/dist-packages (from transformers) (4.67.1)
Requirement already satisfied: pyarrow>=15.0.0 in /usr/local/lib/python3.11/dist-packages (from datasets) (18.1.0)
Requirement already satisfied: dill<0.3.9,>=0.3.0 in /usr/local/lib/python3.11/dist-packages (from datasets) (0.3.8)
Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages (from datasets) (2.2.2)
Requirement already satisfied: xxhash in /usr/local/lib/python3.11/dist-packages (from datasets) (3.5.0)
Requirement already satisfied: multiprocess<0.70.17 in /usr/local/lib/python3.11/dist-packages (from datasets) (0.70.16)
Requirement already satisfied: fsspec<=2025.3.0,>=2023.1.0 in /usr/local/lib/python3.11/dist-packages (from fsspec[http]
Requirement already satisfied: scipy>=1.6.0 in /usr/local/lib/python3.11/dist-packages (from scikit-learn) (1.16.0)
Requirement already satisfied: joblib>=1.2.0 in /usr/local/lib/python3.11/dist-packages (from scikit-learn) (1.5.1)
Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/lib/python3.11/dist-packages (from scikit-learn) (3.6.
Requirement already satisfied: aiohttp!=4.0.0a0,!=4.0.0a1 in /usr/local/lib/python3.11/dist-packages (from fsspec[http]<
Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-h
Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub<1.0
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->trans
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (3.
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->transformer
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->transformer
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas->datasets)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas->datasets) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas->datasets) (2025.2
Requirement already satisfied: aiohappyeyeballs>=2.5.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0
Requirement already satisfied: aiosignal>=1.4.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0.
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0.
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.8.2->pandas-
Downloading evaluate-0.4.5-py3-none-any.whl (84 kB)
  ──────────────────────────────────────  84.1/84.1 kB 3.0 MB/s eta 0:00:00
Installing collected packages: evaluate
Successfully installed evaluate-0.4.5
```

1. Load Dataset IMDb

```
from datasets import load_dataset

# Dataset IMDb
dataset = load_dataset("stanfordnlp/imdb")
dataset
```

Show hidden output

2. Load Tokenizer dan Model RoBERTa-base

```
from transformers import RobertaTokenizerFast, RobertaForSequenceClassification

tokenizer = RobertaTokenizerFast.from_pretrained("roberta-base")
model = RobertaForSequenceClassification.from_pretrained("roberta-base", num_labels=2)
```

Show hidden output

3. Tokenisasi Dataset

```
def tokenize_function(example):
    return tokenizer(example["text"], truncation=True, padding="max_length", max_length=512)
```

```
tokenized_datasets = dataset.map(tokenize_function, batched=True)
```

⇥ Map: 100%   25000/25000 [00:22<00:00, 970.17 examples/s]

    Map: 100%   25000/25000 [00:20<00:00, 1252.76 examples/s]

    Map: 100%   50000/50000 [00:44<00:00, 1085.75 examples/s]

∨ 4. Membuat Data Collator

```
from transformers import DataCollatorWithPadding

data_collator = DataCollatorWithPadding(tokenizer=tokenizer)
```

∨ 5. Evaluasi: Akurasi

```
!pip install evaluate
from evaluate import load
import numpy as np
from sklearn.metrics import precision_recall_fscore_support

accuracy = load("accuracy")
f1 = load("f1")

def compute_metrics(eval_pred):
    logits, labels = eval_pred
    predictions = np.argmax(logits, axis=-1)
    precision, recall, f1, _ = precision_recall_fscore_support(labels, predictions, average='binary')
    acc = accuracy.compute(predictions=predictions, references=labels)
    return {
        'accuracy': acc['accuracy'],
        'f1': f1,
        'precision': precision,
        'recall': recall
    }
```

⇥ Requirement already satisfied: evaluate in /usr/local/lib/python3.11/dist-packages (0.4.5)
    Requirement already satisfied: datasets>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from evaluate) (4.0.0)
    Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages (from evaluate) (2.0.2)
    Requirement already satisfied: dill in /usr/local/lib/python3.11/dist-packages (from evaluate) (0.3.8)
    Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages (from evaluate) (2.2.2)
    Requirement already satisfied: requests>=2.19.0 in /usr/local/lib/python3.11/dist-packages (from evaluate) (2.32.3)
    Requirement already satisfied: tqdm>=4.62.1 in /usr/local/lib/python3.11/dist-packages (from evaluate) (4.67.1)
    Requirement already satisfied: xxhash in /usr/local/lib/python3.11/dist-packages (from evaluate) (3.5.0)
    Requirement already satisfied: multiprocess in /usr/local/lib/python3.11/dist-packages (from evaluate) (0.70.16)
    Requirement already satisfied: fsspec>=2021.05.0 in /usr/local/lib/python3.11/dist-packages (from fsspec[http]>=2021.05.
    Requirement already satisfied: huggingface-hub>=0.7.0 in /usr/local/lib/python3.11/dist-packages (from evaluate) (0.34.1
    Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-packages (from evaluate) (25.0)
    Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from datasets>=2.0.0->evaluate) (3.1
    Requirement already satisfied: pyarrow>=15.0.0 in /usr/local/lib/python3.11/dist-packages (from datasets>=2.0.0->evaluat
    Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages (from datasets>=2.0.0->evaluate) (
    Requirement already satisfied: aiohttp!=4.0.0a0,!=4.0.0a1 in /usr/local/lib/python3.11/dist-packages (from fsspec[http]>
    Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-h
    Requirement already satisfied: hf-xet<2.0.0,>=1.1.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.
    Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests>=2.19.
    Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests>=2.19.0->evaluate)
    Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests>=2.19.0->eva
    Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests>=2.19.0->eva
    Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/dist-packages (from pandas->evaluate)
    Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packages (from pandas->evaluate) (2025.2)
    Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-packages (from pandas->evaluate) (2025.2
    Requirement already satisfied: aiohappyeyeballs>=2.5.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0
    Requirement already satisfied: aiosignal>=1.4.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0.
    Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1
    Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0
    Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4
    Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0.
    Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp!=4.0.0a0,!=4.0
    Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages (from python-dateutil>=2.8.2->pandas-

∨ 6. Setup Trainer

```
from transformers import TrainingArguments, Trainer

training_args = TrainingArguments(
    output_dir="./results",
```

```
        eval_strategy="epoch",
        save_strategy="epoch",
        learning_rate=2e-5,
        per_device_train_batch_size=8,
        per_device_eval_batch_size=8,
        num_train_epochs=4,
        weight_decay=0.01,
        logging_dir="./logs",
        report_to="none",
    )

    trainer = Trainer(
        model=model,
        args=training_args,
        train_dataset=tokenized_datasets["train"].shuffle(seed=42).select(range(500)),
        eval_dataset=tokenized_datasets["test"].select(range(5000)),
        data_collator=data_collator,
        compute_metrics=compute_metrics,
    )
```

✕ 7. Fine-tuning Model

```
trainer.train()
```

⤵ /usr/local/lib/python3.11/dist-packages/torch/nn/modules/module.py:1750: FutureWarning: `encoder_attention_mask` is depr
    return forward_call(*args, **kwargs)
    ▭▭▭▭▭▭▭▭▭▭▭▭▭ [64/252 00:43 < 02:11, 1.43 it/s, Epoch 1/4]

Epoch	Training Loss	Validation Loss

▭▭▭▭▭▭▭▭▭▭▭▭▭ [611/625 02:12 < 00:03, 4.59 it/s]
▭▭▭▭▭▭▭▭▭▭▭▭▭ [252/252 12:52, Epoch 4/4]

Epoch	Training Loss	Validation Loss	Accuracy	F1
1	No log	0.158464	0.931200	0.000000
2	No log	0.935378	0.801600	0.000000
3	No log	0.304767	0.941400	0.000000
4	No log	0.503668	0.900800	0.000000

```
/usr/local/lib/python3.11/dist-packages/torch/nn/modules/module.py:1750: FutureWarning: `encoder_attention_mask` is depr
  return forward_call(*args, **kwargs)
/usr/local/lib/python3.11/dist-packages/torch/nn/modules/module.py:1750: FutureWarning: `encoder_attention_mask` is depr
  return forward_call(*args, **kwargs)
/usr/local/lib/python3.11/dist-packages/torch/nn/modules/module.py:1750: FutureWarning: `encoder_attention_mask` is depr
  return forward_call(*args, **kwargs)
TrainOutput(global_step=252, training_loss=0.28974139501178076, metrics={'train_runtime': 774.4411,
'train_samples_per_second': 2.583, 'train_steps_per_second': 0.325, 'total_flos': 526222110720000.0, 'train_loss':
0.28974139501178076, 'epoch': 4.0})
```

✕ 8. Evaluasi Model

```
trainer.evaluate()
```

⤵ /usr/local/lib/python3.11/dist-packages/torch/nn/modules/module.py:1750: FutureWarning: `encoder_attention_mask` is depr
    return forward_call(*args, **kwargs)
    ▭▭▭▭▭▭▭▭▭▭▭▭▭ [625/625 02:16]
    {'eval_loss': 0.5036680698394775,
     'eval_accuracy': 0.9008,
     'eval_f1': 0.0,
     'eval_runtime': 136.4499,
     'eval_samples_per_second': 36.643,
     'eval_steps_per_second': 4.58,
     'epoch': 4.0}

✕ 9. Simpan Model

```
model.save_pretrained("./roberta-imdb")
tokenizer.save_pretrained("./roberta-imdb")
```

⤵ ('./roberta-imdb/tokenizer_config.json',
   './roberta-imdb/special_tokens_map.json',
   './roberta-imdb/vocab.json',
   './roberta-imdb/merges.txt',
   './roberta-imdb/added_tokens.json',
   './roberta-imdb/tokenizer.json')

```python
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay

# Get predictions on the evaluation set
predictions = trainer.predict(tokenized_datasets["test"].select(range(5000)))
predicted_labels = np.argmax(predictions.predictions, axis=1)
true_labels = predictions.label_ids

# Create a confusion matrix
cm = confusion_matrix(true_labels, predicted_labels)
disp = ConfusionMatrixDisplay(confusion_matrix=cm, display_labels=["Negative", "Positive"])

# Plot the confusion matrix
disp.plot(cmap=plt.cm.Blues)
plt.title("Confusion Matrix")
plt.show()
```
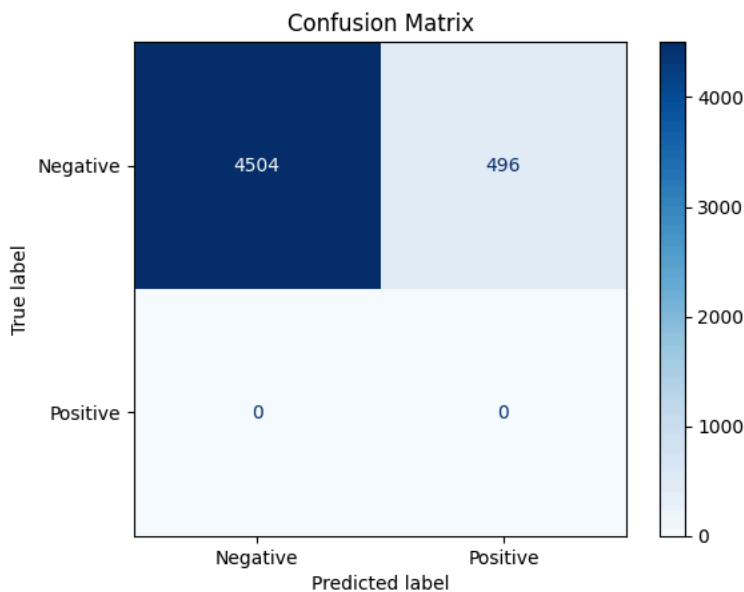
```
/usr/local/lib/python3.11/dist-packages/torch/nn/modules/module.py:1750: FutureWarning: `encoder_attention_mask` is depr
    return forward_call(*args, **kwargs)
```
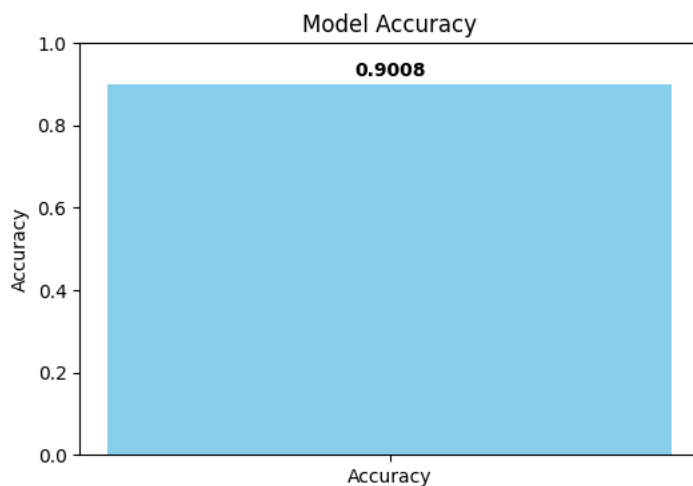


```python
# Get the evaluation accuracy
eval_results = trainer.evaluate()
accuracy_score = eval_results['eval_accuracy']

# Create a bar chart for accuracy
plt.figure(figsize=(6, 4))
plt.bar(['Accuracy'], [accuracy_score], color=['skyblue'])
plt.ylim(0, 1)
plt.ylabel('Accuracy')
plt.title('Model Accuracy')
for i, v in enumerate([accuracy_score]):
    plt.text(i, v + 0.02, f"{v:.4f}", ha='center', fontweight='bold')
plt.show()
```

```
/usr/local/lib/python3.11/dist-packages/torch/nn/modules/module.py:1750: FutureWarning: `encoder_attention_mask` is depr
    return forward_call(*args, **kwargs)
```
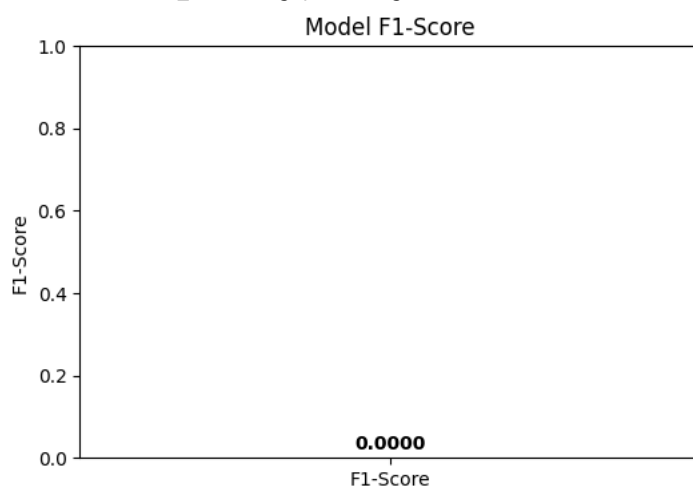[625/625 02:15]


```python
from sklearn.metrics import f1_score

# Get predictions on the evaluation set
predictions = trainer.predict(tokenized_datasets["test"].select(range(5000)))
predicted_labels = np.argmax(predictions.predictions, axis=1)
true_labels = predictions.label_ids

# Calculate F1 score
f1 = f1_score(true_labels, predicted_labels)

# Create a bar chart for F1-score
plt.figure(figsize=(6, 4))
plt.bar(['F1-Score'], [f1], color=['lightgreen'])
plt.ylim(0, 1)
plt.ylabel('F1-Score')
plt.title('Model F1-Score')
for i, v in enumerate([f1]):
    plt.text(i, v + 0.02, f"{v:.4f}", ha='center', fontweight='bold')
plt.show()
```

```
/usr/local/lib/python3.11/dist-packages/torch/nn/modules/module.py:1750: FutureWarning: `encoder_attention_mask` is depr
    return forward_call(*args, **kwargs)
```



```python
import pandas as pd

# IMPORTANT: Please re-run the cell that defines `compute_metrics` and the cell
# that initializes the `Trainer` before running this cell.

# Run evaluation
eval_results = trainer.evaluate()

# Create a DataFrame from the evaluation results
df_results = pd.DataFrame([eval_results])

# Columns to display
```

```
columns_to_display = {
    'eval_accuracy': 'Accuracy',
    'eval_f1': 'F1-Score',
    'eval_precision': 'Precision',
    'eval_recall': 'Recall'
}

# Filter to only columns that exist in the results
existing_columns = {k: v for k, v in columns_to_display.items() if k in df_results.columns}

df_display = df_results[list(existing_columns.keys())]
df_display.rename(columns=existing_columns, inplace=True)


display(df_display)
```

[625/625 10:33]

```
/tmp/ipython-input-1436541997.py:24: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-
  df_display.rename(columns=existing_columns, inplace=True)

	Accuracy	F1-Score
0	0.9008	0.0

Time series