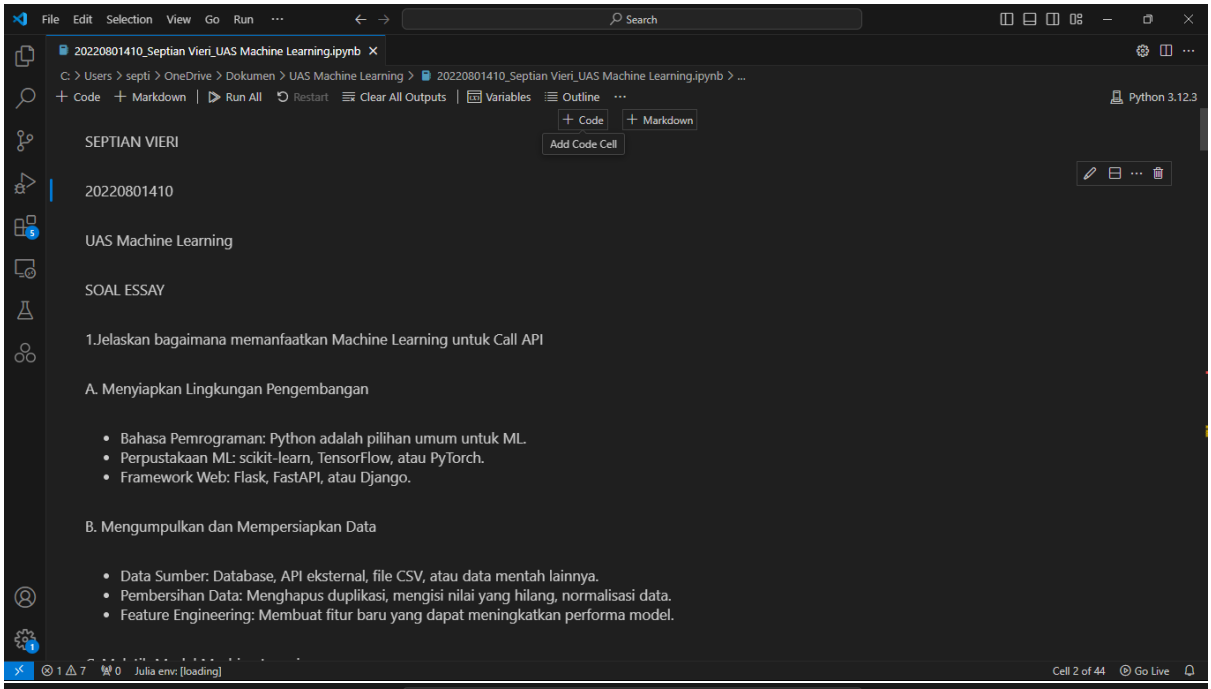
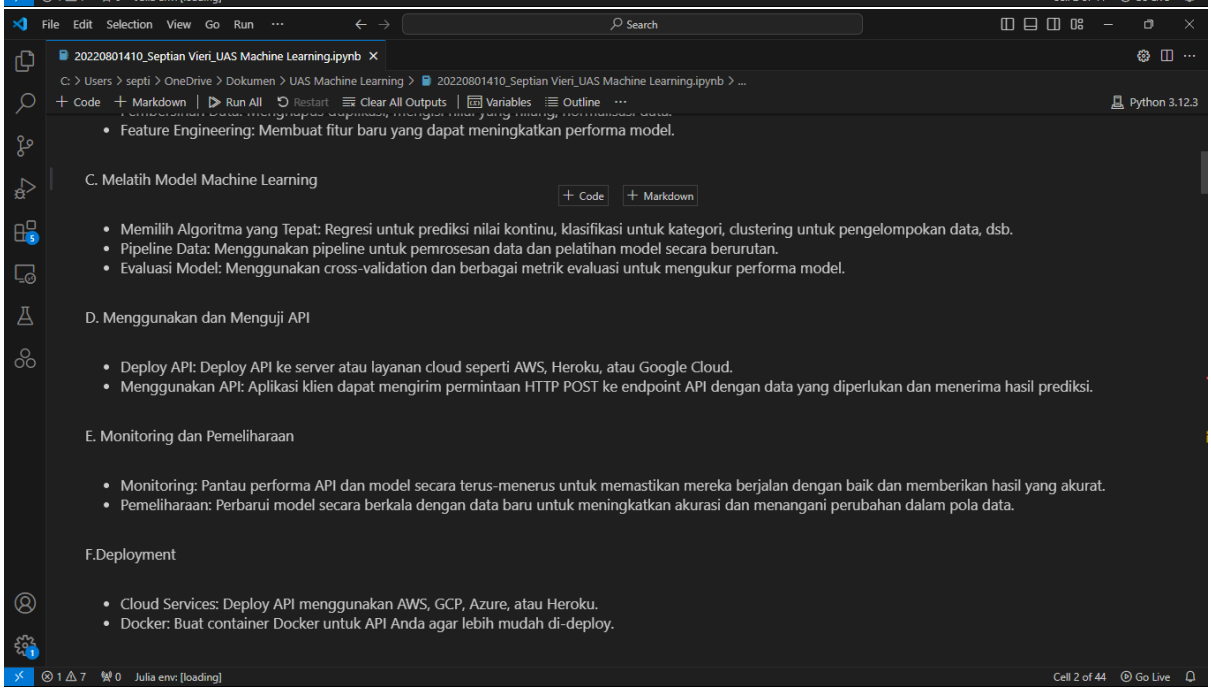
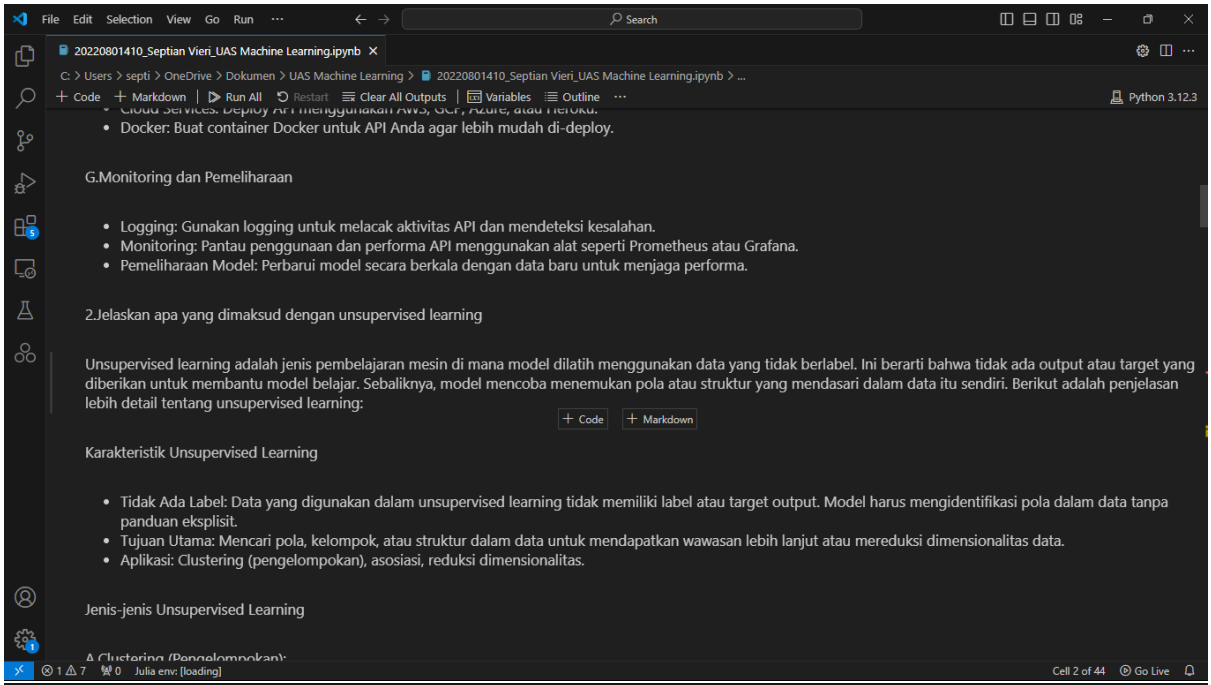
Nama: Septian Vieri

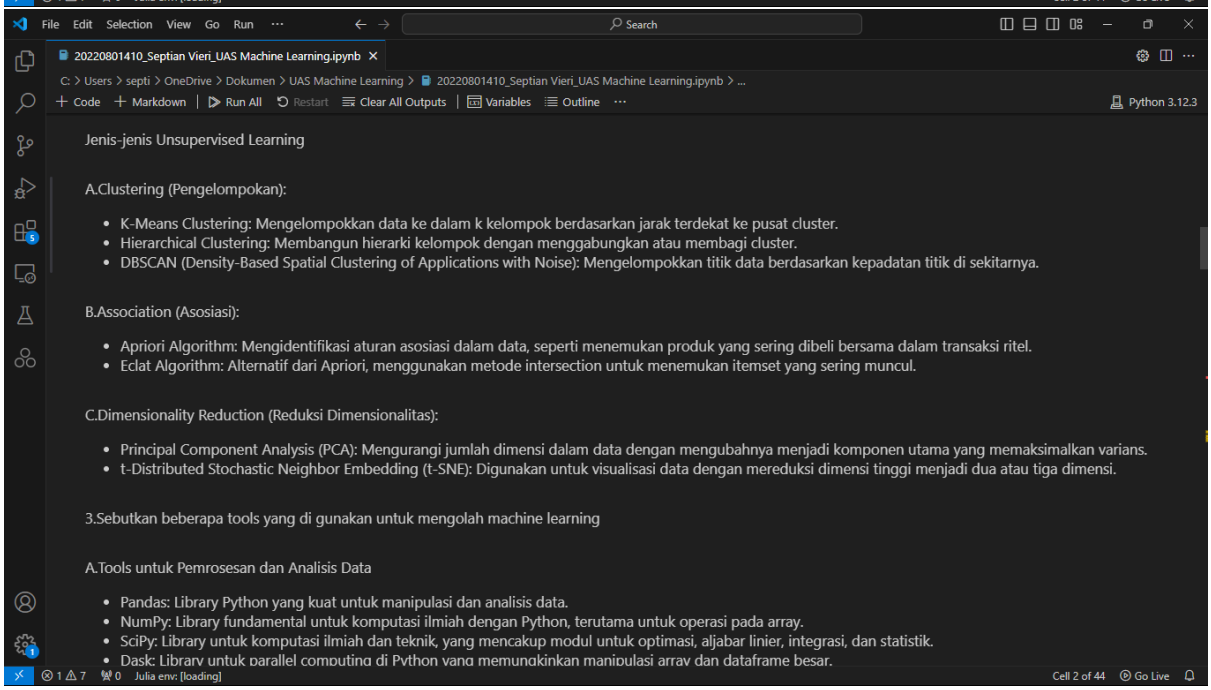
NIM: 20220801410

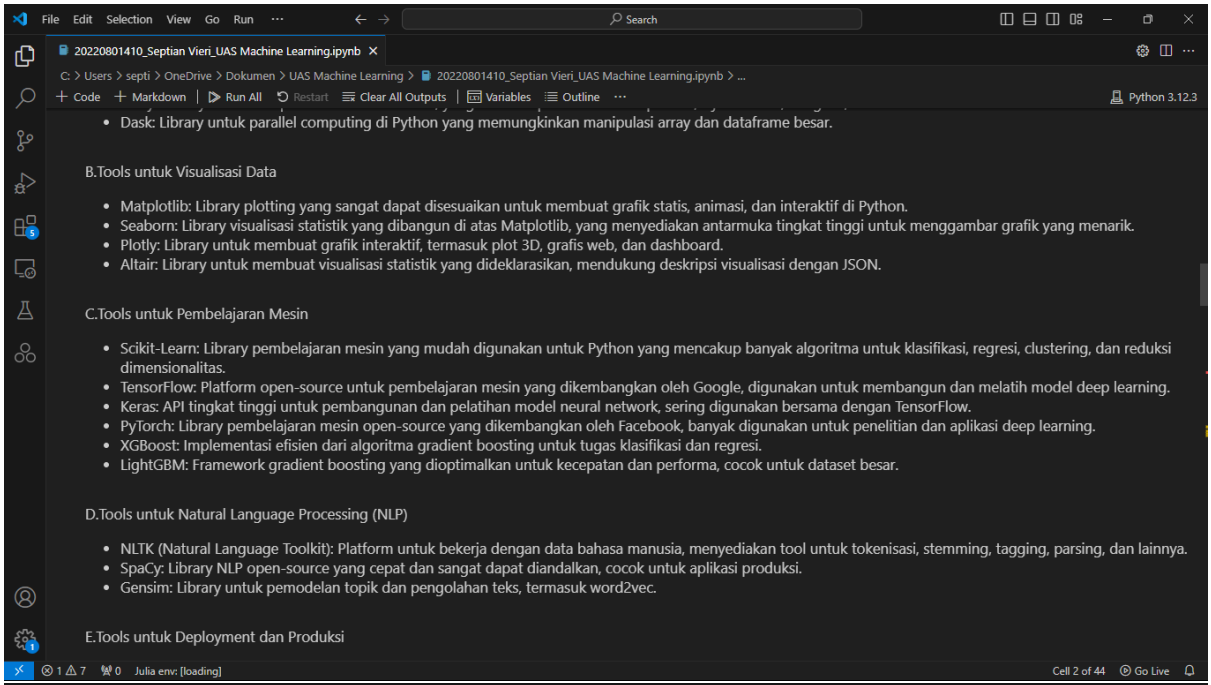
UAS Machine Learning

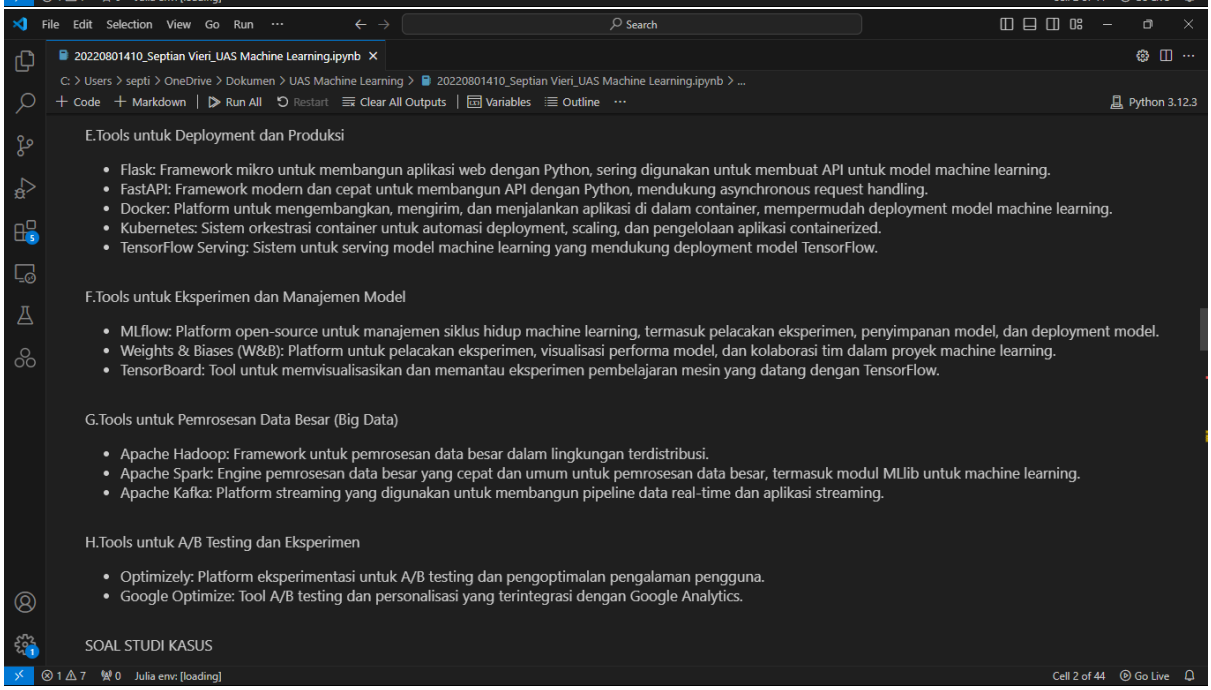


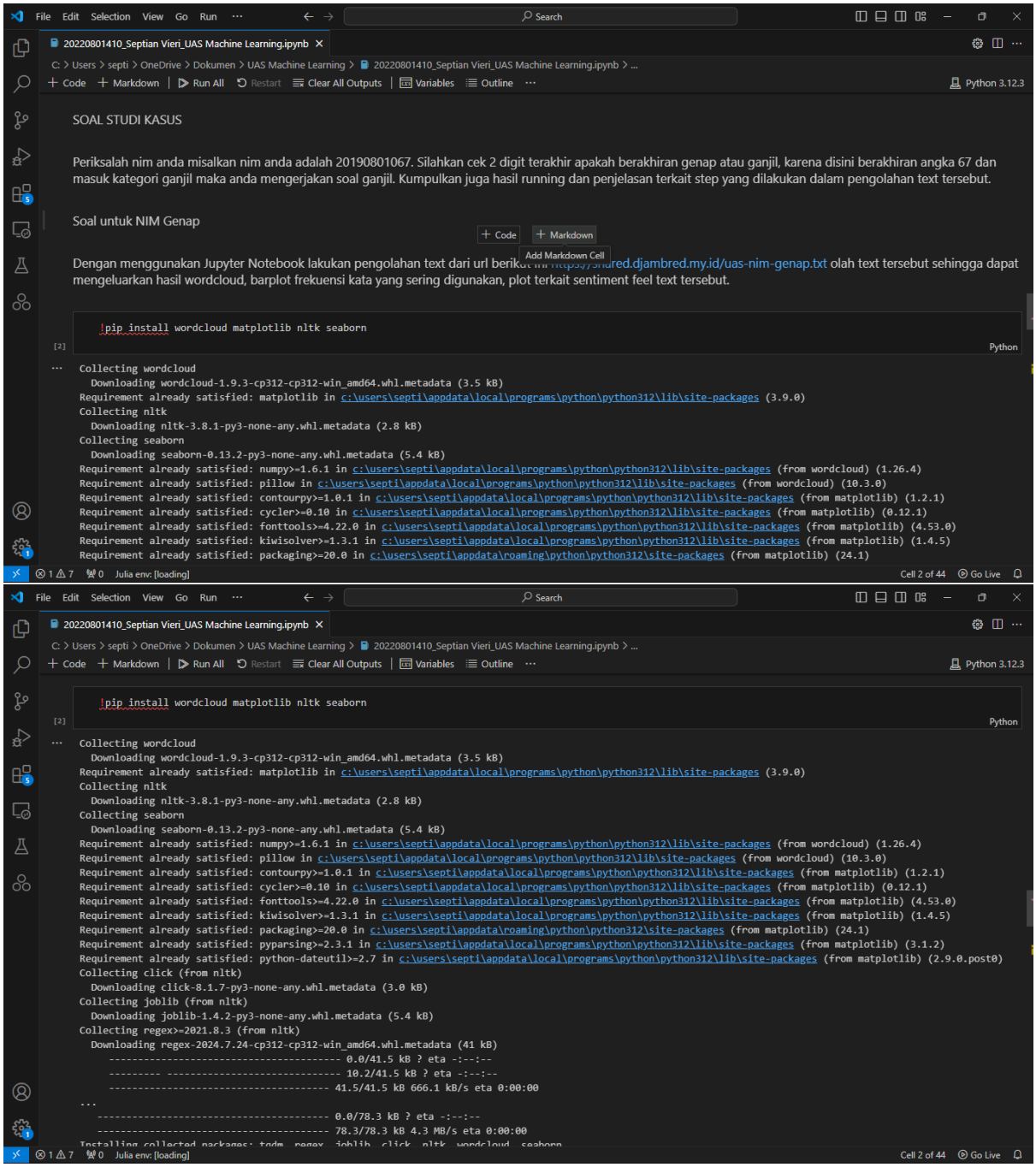



File Edit Selection View Go Run
Search
20220801410_Septian Vieri_UAS Machine Learning.ipynb
C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...
Code Markdown Run All Restart Clear All Outputs Variables Outline
Python 3.12.3
Docker: Buat container Docker untuk API Anda agar lebih mudah di-deploy.
G.Monitoring dan Pemeliharaan
Logging: Gunakan logging untuk melacak aktivitas API dan mendeteksi kesalahan.
Monitoring: Pantau penggunaan dan performa API menggunakan alat seperti Prometheus atau Grafana.
Pemeliharaan Model: Perbarui model secara berkala dengan data baru untuk menjaga performa.
2.Jelaskan apa yang dimaksud dengan unsupervised learning
Unsupervised learning adalah jenis pembelajaran mesin di mana model dilatih menggunakan data yang tidak berlabel. Ini berarti bahwa tidak ada output atau target yang diberikan untuk membantu model belajar. Sebaliknya, model mencoba menemukan pola atau struktur yang mendasari dalam data itu sendiri. Berikut adalah penjelasan lebih detail tentang unsupervised learning:
Code Markdown
Karakteristik Unsupervised Learning
Tidak Ada Label: Data yang digunakan dalam unsupervised learning tidak memiliki label atau target output. Model harus mengidentifikasi pola dalam data tanpa panduan eksplisit.
Tujuan Utama: Mencari pola, kelompok, atau struktur dalam data untuk mendapatkan wawasan lebih lanjut atau mereduksi dimensionalitas data.
Aplikasi: Clustering (pengelompokan), asosiasi, reduksi dimensionalitas.
Jenis-jenis Unsupervised Learning
Julia env: [loading]
Cell 2 of 44
Go Live

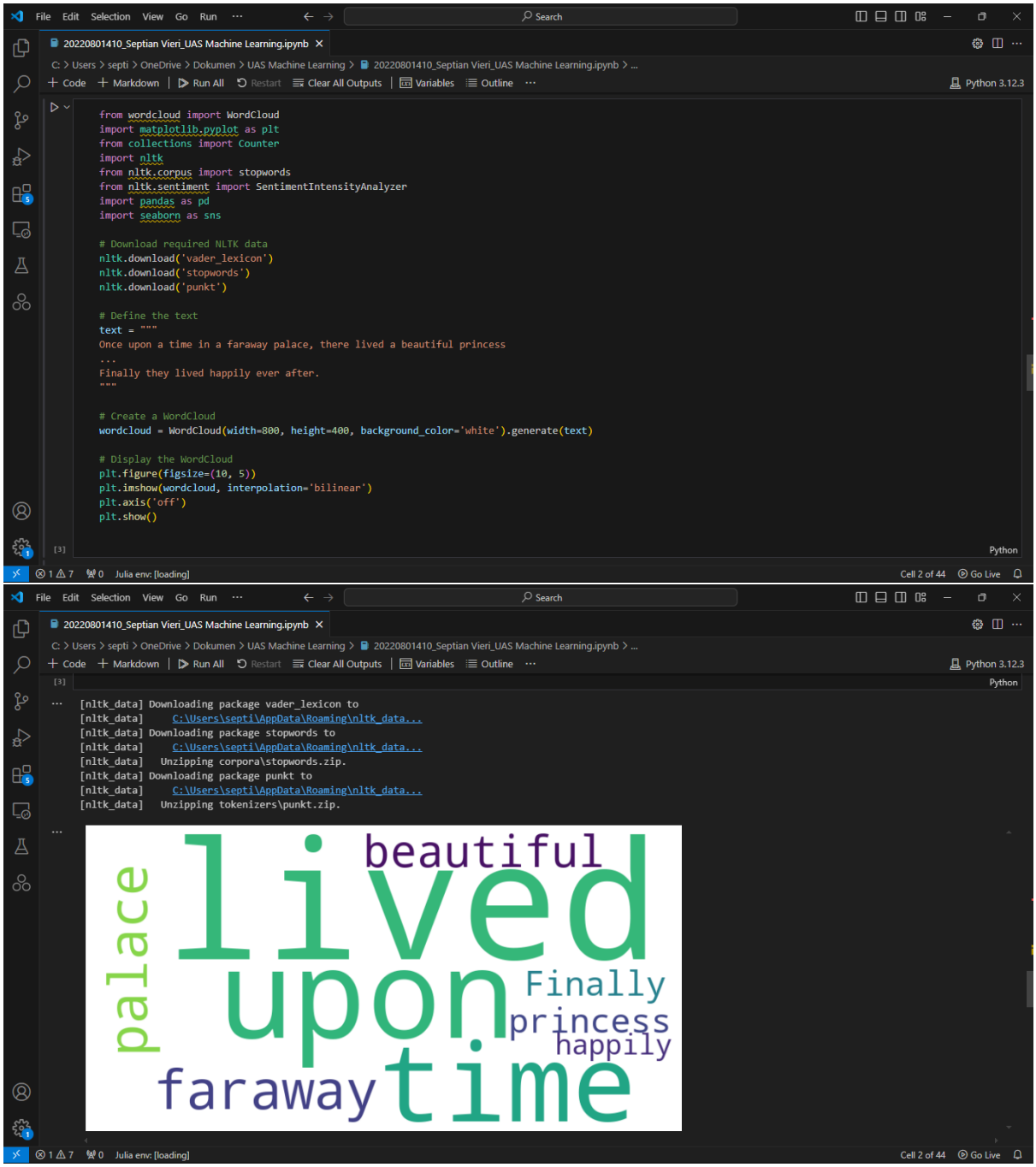

File Edit Selection View Go Run
Search
20220801410_Septian Vieri_UAS Machine Learning.ipynb
C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...
Code Markdown Run All Restart Clear All Outputs Variables Outline
Python 3.12.3
Jenis-jenis Unsupervised Learning
A.Clustering (Pengelompokan):
K-Means Clustering: Mengelompokkan data ke dalam k kelompok berdasarkan jarak terdekat ke pusat cluster.
Hierarchical Clustering: Membangun hierarki kelompok dengan menggabungkan atau membagi cluster.
DBSCAN (Density-Based Spatial Clustering of Applications with Noise): Mengelompokkan titik data berdasarkan kepadatan titik di sekitarnya.
B.Association (Asosiasi):
Apriori Algorithm: Mengidentifikasi aturan asosiasi dalam data, seperti menemukan produk yang sering dibeli bersama dalam transaksi ritel.
Eclat Algorithm: Alternatif dari Apriori, menggunakan metode intersection untuk menemukan itemset yang sering muncul.
C.Dimensionality Reduction (Reduksi Dimensionalitas):
Principal Component Analysis (PCA): Mengurangi jumlah dimensi dalam data dengan mengubahnya menjadi komponen utama yang memaksimalkan varians.
t-Distributed Stochastic Neighbor Embedding (t-SNE): Digunakan untuk visualisasi data dengan mereduksi dimensi tinggi menjadi dua atau tiga dimensi.
3.Sebutkan beberapa tools yang di gunakan untuk mengolah machine learning
A.Tools untuk Pemrosesan dan Analisis Data
Pandas: Library Python yang kuat untuk manipulasi dan analisis data.
NumPy: Library fundamental untuk komputasi ilmiah dengan Python, terutama untuk operasi pada array.
SciPy: Library untuk komputasi ilmiah dan teknik, yang mencakup modul untuk optimasi, aljabar linier, integrasi, dan statistik.
Dask: Library untuk parallel computing di Python yang memungkinkan manipulasi array dan dataframe besar.
Julia env: [loading]
Cell 2 of 44
Go Live


File Edit Selection View Go Run
Search
20220801410_Septian Vieri_UAS Machine Learning.ipynb
C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...
Code Markdown Run All Restart Clear All Outputs Variables Outline
Python 3.12.3
Dask: Library untuk parallel computing di Python yang memungkinkan manipulasi array dan dataframe besar.
B.Tools untuk Visualisasi Data
Matplotlib: Library plotting yang sangat dapat disesuaikan untuk membuat grafik statis, animasi, dan interaktif di Python.
Seaborn: Library visualisasi statistik yang dibangun di atas Matplotlib, yang menyediakan antarmuka tingkat tinggi untuk menggambar grafik yang menarik.
Plotly: Library untuk membuat grafik interaktif, termasuk plot 3D, grafis web, dan dashboard.
Altair: Library untuk membuat visualisasi statistik yang dideklarasikan, mendukung deskripsi visualisasi dengan JSON.
C.Tools untuk Pembelajaran Mesin
Scikit-Learn: Library pembelajaran mesin yang mudah digunakan untuk Python yang mencakup banyak algoritma untuk klasifikasi, regresi, clustering, dan reduksi dimensionalitas.
TensorFlow: Platform open-source untuk pembelajaran mesin yang dikembangkan oleh Google, digunakan untuk membangun dan melatih model deep learning.
Keras: API tingkat tinggi untuk pembangunan dan pelatihan model neural network, sering digunakan bersama dengan TensorFlow.
PyTorch: Library pembelajaran mesin open-source yang dikembangkan oleh Facebook, banyak digunakan untuk penelitian dan aplikasi deep learning.
XGBoost: Implementasi efisien dari algoritma gradient boosting untuk tugas klasifikasi dan regresi.
LightGBM: Framework gradient boosting yang dioptimalkan untuk kecepatan dan performa, cocok untuk dataset besar.
D.Tools untuk Natural Language Processing (NLP)
NLTK (Natural Language Toolkit): Platform untuk bekerja dengan data bahasa manusia, menyediakan tool untuk tokenisasi, stemming, tagging, parsing, dan lainnya.
SpaCy: Library NLP open-source yang cepat dan sangat dapat diandalkan, cocok untuk aplikasi produksi.
Gensim: Library untuk pemodelan topik dan pengolahan teks, termasuk word2vec.
E.Tools untuk Deployment dan Produksi
1 7 0 Julia env: [loading]
Cell 2 of 44 Go Live


File Edit Selection View Go Run
Search
20220801410_Septian Vieri_UAS Machine Learning.ipynb
C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...
Code Markdown Run All Restart Clear All Outputs Variables Outline
Python 3.12.3
E.Tools untuk Deployment dan Produksi
Flask: Framework mikro untuk membangun aplikasi web dengan Python, sering digunakan untuk membuat API untuk model machine learning.
FastAPI: Framework modern dan cepat untuk membangun API dengan Python, mendukung asynchronous request handling.
Docker: Platform untuk mengembangkan, mengirim, dan menjalankan aplikasi di dalam container, mempermudah deployment model machine learning.
Kubernetes: Sistem orkestrasi container untuk automasi deployment, scaling, dan pengelolaan aplikasi containerized.
TensorFlow Serving: Sistem untuk serving model machine learning yang mendukung deployment model TensorFlow.
F.Tools untuk Eksperimen dan Manajemen Model
MLflow: Platform open-source untuk manajemen siklus hidup machine learning, termasuk pelacakan eksperimen, penyimpanan model, dan deployment model.
Weights & Biases (W&B): Platform untuk pelacakan eksperimen, visualisasi performa model, dan kolaborasi tim dalam proyek machine learning.
TensorBoard: Tool untuk memvisualisasikan dan memantau eksperimen pembelajaran mesin yang datang dengan TensorFlow.
G.Tools untuk Pemrosesan Data Besar (Big Data)
Apache Hadoop: Framework untuk pemrosesan data besar dalam lingkungan terdistribusi.
Apache Spark: Engine pemrosesan data besar yang cepat dan umum untuk pemrosesan data besar, termasuk modul MLlib untuk machine learning.
Apache Kafka: Platform streaming yang digunakan untuk membangun pipeline data real-time dan aplikasi streaming.
H.Tools untuk A/B Testing dan Eksperimen
Optimizely: Platform eksperimentasi untuk A/B testing dan pengoptimalan pengalaman pengguna.
Google Optimize: Tool A/B testing dan personalisasi yang terintegrasi dengan Google Analytics.
SOAL STUDI KASUS
1 7 0 Julia env: [loading]
Cell 2 of 44 Go Live


File Edit Selection View Go Run
Search
20220801410_Septian Vieri_UAS Machine Learning.ipynb
C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...
Code Markdown Run All Restart Clear All Outputs Variables Outline
Python 3.12.3
SOAL STUDI KASUS
Periksalah nim anda misalkan nim anda adalah 20190801067. Silahkan cek 2 digit terakhir apakah berakhiran genap atau ganjil, karena disini berakhiran angka 67 dan masuk kategori ganjil maka anda mengerjakan soal ganjil. Kumpulkan juga hasil running dan penjelasan terkait step yang dilakukan dalam pengolahan text tersebut.
Soal untuk NIM Genap
Code Markdown
Add Markdown Cell
Dengan menggunakan Jupyter Notebook lakukan pengolahan text dari url berikut ini https://shared.djambred.my.id/uas-nim-genap.txt olah text tersebut sehingga dapat mengeluarkan hasil wordcloud, barplot frekuensi kata yang sering digunakan, plot terkait sentiment feel text tersebut.
!pip install wordcloud matplotlib nltk seaborn
[2]
Python
Collecting wordcloud
Downloading wordcloud-1.9.3-cp312-cp312-win_amd64.whl.metadata (3.5 kB)
Requirement already satisfied: matplotlib in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (3.9.0)
Collecting nltk
Downloading nltk-3.8.1-py3-none-any.whl.metadata (2.8 kB)
Collecting seaborn
Downloading seaborn-0.13.2-py3-none-any.whl.metadata (5.4 kB)
Requirement already satisfied: numpy>=1.6.1 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from wordcloud) (1.26.4)
Requirement already satisfied: pillow in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from wordcloud) (10.3.0)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (1.2.1)
Requirement already satisfied: cycler>=0.10 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (4.53.0)
Requirement already satisfied: kiwisolver>=1.3.1 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (1.4.5)
Requirement already satisfied: packaging>=20.0 in c:\users\septi\appdata\roaming\python\python312\site-packages (from matplotlib) (24.1)
1 7 0 Julia env: [loading]
Cell 2 of 44 Go Live

File Edit Selection View Go Run
Search
20220801410_Septian Vieri_UAS Machine Learning.ipynb
C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...
Code Markdown Run All Restart Clear All Outputs Variables Outline
Python 3.12.3
!pip install wordcloud matplotlib nltk seaborn
[2]
Python
Collecting wordcloud
Downloading wordcloud-1.9.3-cp312-cp312-win_amd64.whl.metadata (3.5 kB)
Requirement already satisfied: matplotlib in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (3.9.0)
Collecting nltk
Downloading nltk-3.8.1-py3-none-any.whl.metadata (2.8 kB)
Collecting seaborn
Downloading seaborn-0.13.2-py3-none-any.whl.metadata (5.4 kB)
Requirement already satisfied: numpy>=1.6.1 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from wordcloud) (1.26.4)
Requirement already satisfied: pillow in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from wordcloud) (10.3.0)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (1.2.1)
Requirement already satisfied: cycler>=0.10 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (4.53.0)
Requirement already satisfied: kiwisolver>=1.3.1 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (1.4.5)
Requirement already satisfied: packaging>=20.0 in c:\users\septi\appdata\roaming\python\python312\site-packages (from matplotlib) (24.1)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (3.1.2)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\septi\appdata\local\programs\python\python312\lib\site-packages (from matplotlib) (2.9.0.post0)
Collecting click (from nltk)
Downloading click-8.1.7-py3-none-any.whl.metadata (3.0 kB)
Collecting joblib (from nltk)
Downloading joblib-1.4.2-py3-none-any.whl.metadata (5.4 kB)
Collecting regex>=2021.8.3 (from nltk)
Downloading regex-2024.7.24-cp312-cp312-win_amd64.whl.metadata (41 kB)
0.0/41.5 kB ? eta -:--:--
10.2/41.5 kB ? eta -:--:--
41.5/41.5 kB 666.1 kB/s eta 0:00:00
0.0/78.3 kB ? eta -:--:--
78.3/78.3 kB 4.3 MB/s eta 0:00:00
1 7 0 Julia env: [loading]
Cell 2 of 44 Go Live


File Edit Selection View Go Run
Search
20220801410_Septian Vieri_UAS Machine Learning.ipynb
C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...
Code Markdown Run All Restart Clear All Outputs Variables Outline
Python 3.12.3
from wordcloud import WordCloud
import matplotlib.pyplot as plt
from collections import Counter
import nltk
from nltk.corpus import stopwords
from nltk.sentiment import SentimentIntensityAnalyzer
import pandas as pd
import seaborn as sns
# Download required NLTK data
nltk.download('vader_lexicon')
nltk.download('stopwords')
nltk.download('punkt')
# Define the text
text = """
Once upon a time in a faraway palace, there lived a beautiful princess
...
Finally they lived happily ever after.
"""
# Create a WordCloud
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(text)
# Display the WordCloud
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.show()
[3]
Python
Julia env: [loading]
Cell 2 of 44
Go Live

File Edit Selection View Go Run
Search
20220801410_Septian Vieri_UAS Machine Learning.ipynb
C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...
Code Markdown Run All Restart Clear All Outputs Variables Outline
Python 3.12.3
[3]
Python
[nltk_data] Downloading package vader_lexicon to
[nltk_data] C:\Users\septi\AppData\Roaming\nltk_data...
[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\septi\AppData\Roaming\nltk_data...
[nltk_data] Unzipping corpora\stopwords.zip.
[nltk_data] Downloading package punkt to
[nltk_data] C:\Users\septi\AppData\Roaming\nltk_data...
[nltk_data] Unzipping tokenizers\punkt.zip.
beautiful
lived
palace
upon
Finally
princess
happily
faraway
time
Julia env: [loading]
Cell 2 of 44
Go Live

File Edit Selection View Go Run ··· Search

20220801410_Septian Vieri_UAS Machine Learning.ipynb

C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...

+ Code + Markdown | Run All Restart Clear All Outputs | Variables Outline ··· Python 3.12.3

```
# Tokenize the text
words = nltk.word_tokenize(text.lower())

# Remove stop words
stop_words = set(stopwords.words('english'))
filtered_words = [word for word in words if word.isalnum() and word not in stop_words]

# Get word frequencies
word_freq = Counter(filtered_words)
common_words = word_freq.most_common(10)

# Create a DataFrame for visualization
df = pd.DataFrame(common_words, columns=['Word', 'Frequency'])

# Create a bar plot for word frequencies
plt.figure(figsize=(10, 5))
sns.barplot(x='Word', y='Frequency', data=df)
plt.title('Top 10 Most Common Words')
plt.show()
```

[4] Python

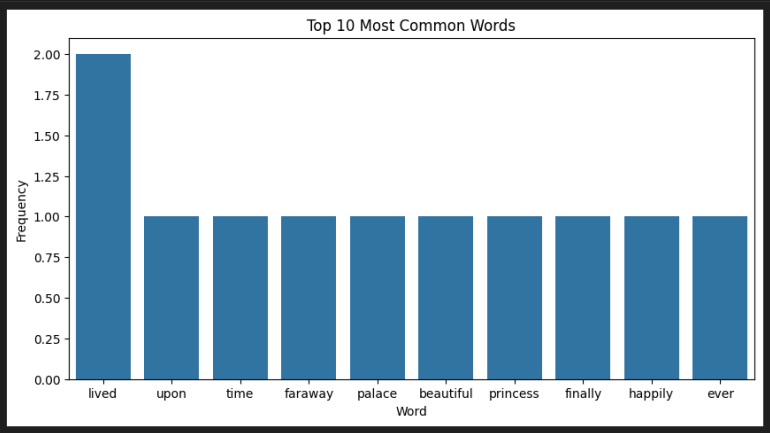
⊗ 1 ⚠ 7 0 Julia env: [loading] Cell 2 of 44 Go Live

File Edit Selection View Go Run ··· Search

20220801410_Septian Vieri_UAS Machine Learning.ipynb

C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...

+ Code + Markdown | Run All Restart Clear All Outputs | Variables Outline ··· Python 3.12.3




```
# Sentiment analysis
sia = SentimentIntensityAnalyzer()
sentiment_scores = sia.polarity_scores(text)
```

⊗ 1 ⚠ 7 0 Julia env: [loading] Cell 2 of 44 Go Live

File Edit Selection View Go Run ··· Search

20220801410_Septian Vieri_UAS Machine Learning.ipynb

C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...

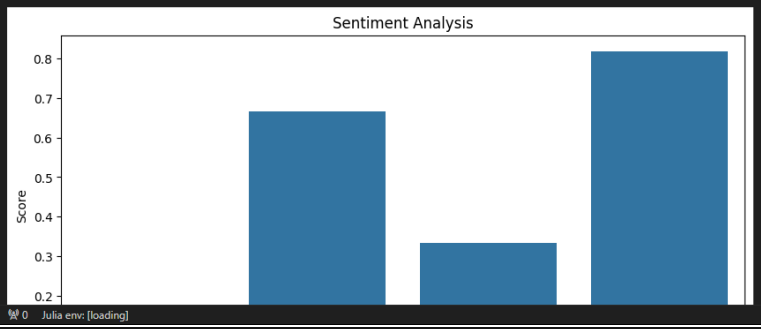
+ Code + Markdown | Run All | Restart | Clear All Outputs | Variables | Outline ··· Python 3.12.3

+ Code + Markdown

```
# Sentiment analysis
sia = SentimentIntensityAnalyzer()
sentiment_scores = sia.polarity_scores(text)

# Display sentiment scores
sentiment_df = pd.DataFrame(list(sentiment_scores.items()), columns=['Sentiment', 'Score'])
plt.figure(figsize=(10, 5))
sns.barplot(x='Sentiment', y='Score', data=sentiment_df)
plt.title('Sentiment Analysis')
plt.show()
```

[5] Python



⊗ 1 ⚠ 7 0 Julia env: [loading] Cell 2 of 44 Go Live

File Edit Selection View Go Run ··· Search

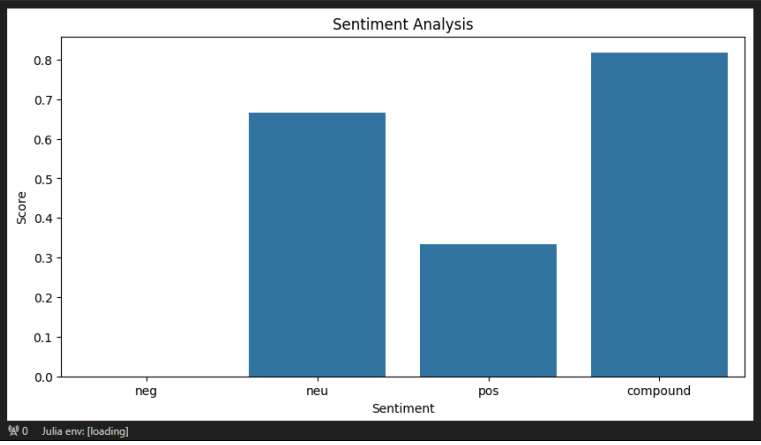
20220801410_Septian Vieri_UAS Machine Learning.ipynb

C: > Users > septi > OneDrive > Dokumen > UAS Machine Learning > 20220801410_Septian Vieri_UAS Machine Learning.ipynb > ...

+ Code + Markdown | Run All | Restart | Clear All Outputs | Variables | Outline ··· Python 3.12.3

```
sentiment_df = pd.DataFrame(list(sentiment_scores.items()), columns=['Sentiment', 'Score'])
plt.figure(figsize=(10, 5))
sns.barplot(x='Sentiment', y='Score', data=sentiment_df)
plt.title('Sentiment Analysis')
plt.show()
```

[5] Python



⊗ 1 ⚠ 7 0 Julia env: [loading] Cell 2 of 44 Go Live