ML-based Forecasting and Anomaly Detection

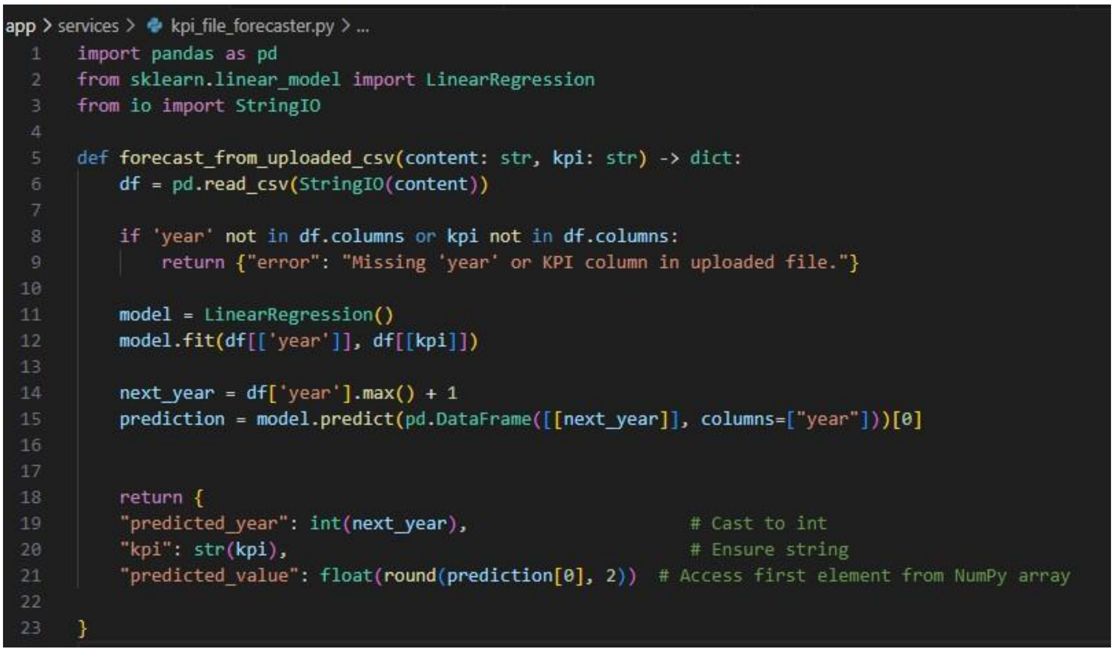
Objective: Analyze uploaded CSV files and predict future trends or irregularities.

Forecasting

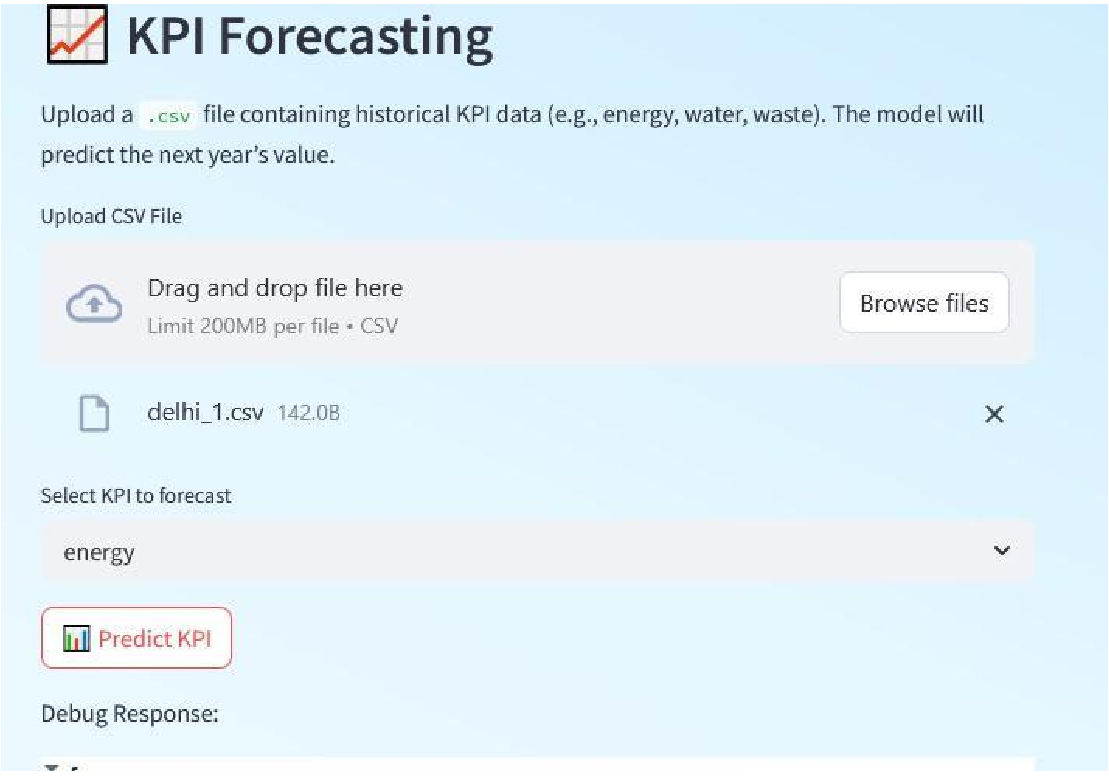
1.Use Linear Regression in kpi_file_forecaster.py

2.Predict water/energy use based on past data

3.Display forecast on dashboard



```python
import pandas as pd
from sklearn.linear_model import LinearRegression
from io import StringIO

def forecast_from_uploaded_csv(content: str, kpi: str) -> dict:
    df = pd.read_csv(StringIO(content))

    if 'year' not in df.columns or kpi not in df.columns:
        return {"error": "Missing 'year' or KPI column in uploaded file."}

    model = LinearRegression()
    model.fit(df[['year']], df[[kpi]])

    next_year = df['year'].max() + 1
    prediction = model.predict(pd.DataFrame([[next_year]], columns=["year"]))[0]


    return {
    "predicted_year": int(next_year),                     # Cast to int
    "kpi": str(kpi),                                      # Ensure string
    "predicted_value": float(round(prediction[0], 2))  # Access first element from NumPy array

}
```



KPI Forecasting

Upload a `.csv` file containing historical KPI data (e.g., energy, water, waste). The model will predict the next year's value.

Upload CSV File

Drag and drop file here

Limit 200MB per file • CSV

delhi_1.csv 142.0B

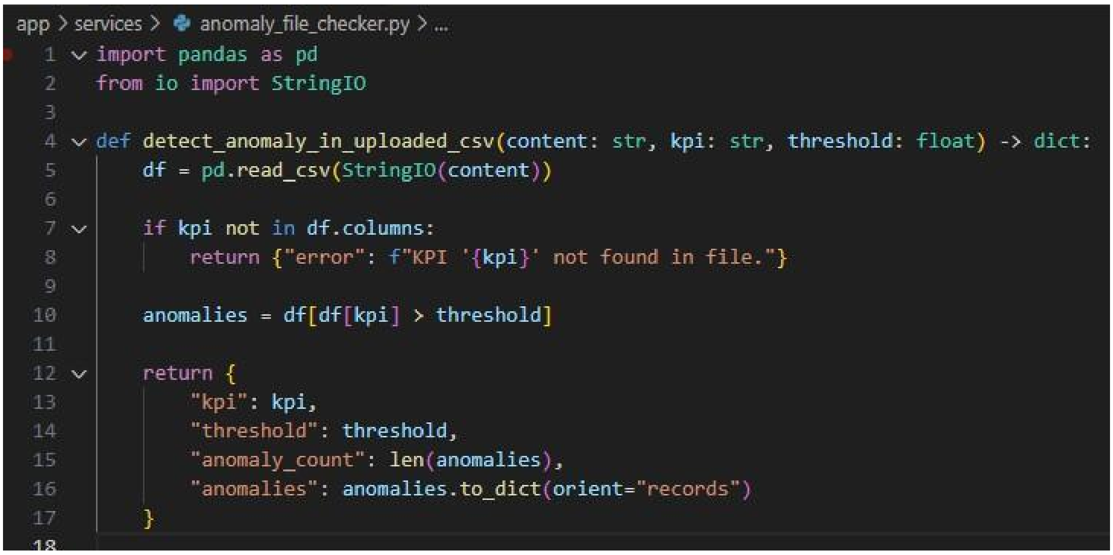
Select KPI to forecast: energy

Debug Response:

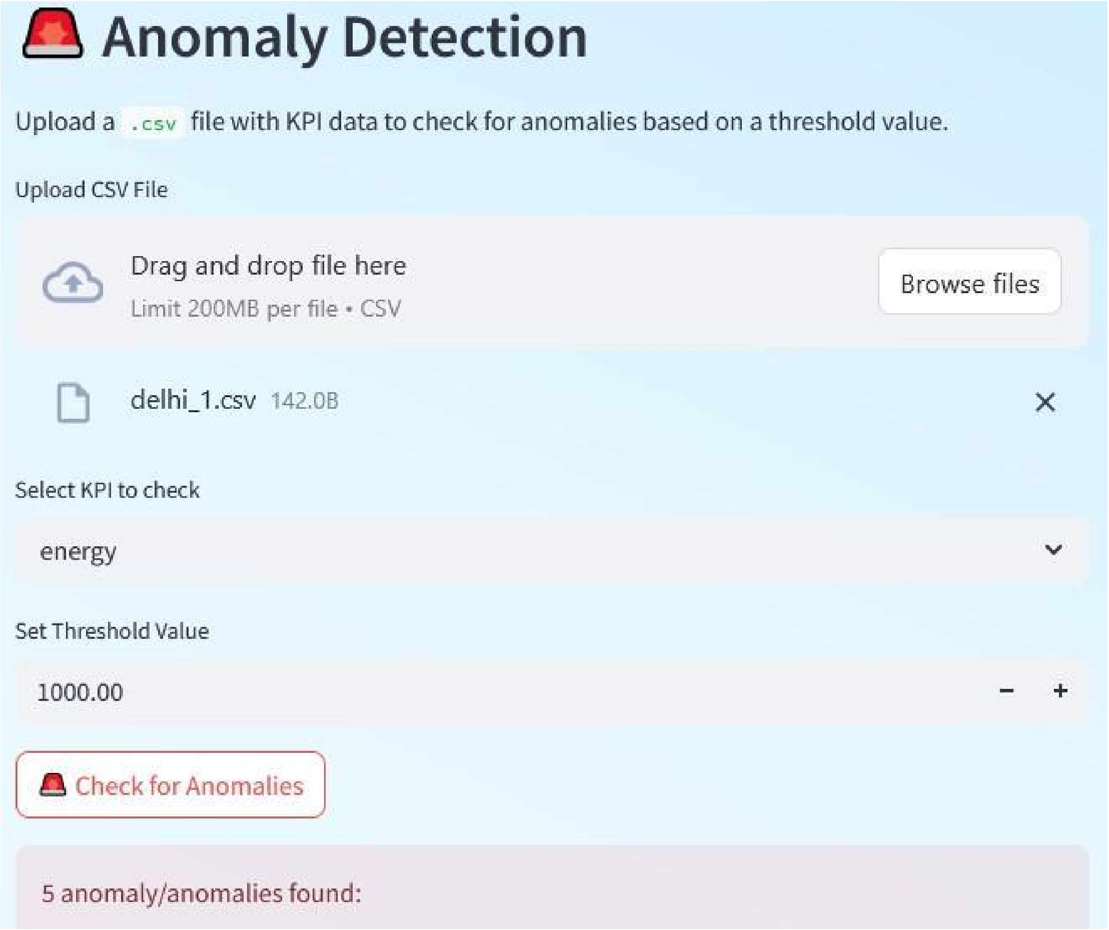
Anomaly Detection

1.anomaly_file_checker.py flags abnormal spikes

2.Display results in tabular or colored badge format



```python
import pandas as pd
from io import StringIO

def detect_anomaly_in_uploaded_csv(content: str, kpi: str, threshold: float) -> dict:
    df = pd.read_csv(StringIO(content))

    if kpi not in df.columns:
        return {"error": f"KPI '{kpi}' not found in file."}

    anomalies = df[df[kpi] > threshold]

    return {
        "kpi": kpi,
        "threshold": threshold,
        "anomaly_count": len(anomalies),
        "anomalies": anomalies.to_dict(orient="records")
    }
```



🚨 Anomaly Detection

Upload a `.csv` file with KPI data to check for anomalies based on a threshold value.

Upload CSV File

Drag and drop file here

Limit 200MB per file • CSV

delhi_1.csv 142.0B

Select KPI to check

energy

Set Threshold Value

1000.00

5 anomaly/anomalies found: