State of Palestine
AN-Najah National University
Engineering and Information Technology
Computer science

دولة فلسطين
جامعة النجاح الوطنية
كلية الهندسة وتكنولوجيا المعلومات
قسم علم الحاسوب

Machine learning

Final document

Ali hayek 12220103

[redacted]

Supervisors: Dr. adnan salman

summary:

The project focuses on predicting energy production based on **weather conditions** using Python. Libraries such as **pandas, numpy, matplotlib, scikit-learn, and meteostat** were utilized for data **processing, visualization, and machine learning modeling.** Various models, including **Linear Regression, Decision Trees, Random Forest, and Stacking Regressor,** were employed. The project demonstrated how weather factors like **temperature and wind speed influence energy production.**

Among all models, the Voting Regressor achieved the highest accuracy, indicating the effectiveness of combining multiple models for robust predictions. This project emphasizes the importance of leveraging machine learning in renewable energy forecasting to enhance resource management and grid planning.

introduction:

With the growing reliance on renewable energy, accurate forecasting of energy production has become essential for efficient grid management and resource optimization. This project aims to develop a machine learning-based system to analyze weather data and predict energy production. By understanding the relationship between weather conditions and energy output, stakeholders can make informed decisions, improve energy efficiency, and ensure sustainable development. The project underscores the potential of data-driven approaches in addressing global energy challenges.

Dataset:

Source:

The dataset was sourced from the **Meteostat** library, which provides comprehensive weather data, including temperature, wind speed, and solar radiation.

Features:

The dataset includes:

- Daily temperature averages.
- Solar radiation levels.
- Humidity percentages.

Preprocessing Steps:

- Handling missing values using imputation techniques.
- Converting date features into month-based categories for seasonal analysis.
- Normalizing numerical features for better performance in machine learning models.

Methodology:

1. **Workflow:**
   - Import and clean the data.
   - Extract significant features like average temperature
   - Split the data into training and testing sets.
2. **Models Used:**
   - **Linear Regression:** For understanding linear relationships between features and energy production.
   - **Decision Trees:** To analyze non-linear patterns.
   - **Random Forest:** To improve accuracy and reduce overfitting.
   - **Voting Regressor:** For combining multiple models to achieve robust predictions.
   - **Stacking Regressor:** To enhance performance using ensemble methods.
3. **Evaluation Metrics:**
   - Root Mean Squared Error (RMSE): Measures prediction errors.
   - R2: Evaluates the proportion of variance explained by the model.

Results and Discussion:

Results:

- The Voting Regressor achieved the best performance with the lowest RMSE and highest R2.
- Graphical representations, such as scatter plots and bar graphs, effectively illustrated model predictions.

Challenges:

- Outliers in weather data affected some models, such as Linear Regression.
- Ensemble models like Stacking Regressor were complex to interpret.

Limitations:

- The dataset relied on weather data, which may not capture all factors influencing energy production.
- Seasonal variations could be better accounted for by incorporating additional data sources.

Conclusion and References:

- **Conclusion:**
  - The project successfully demonstrated the use of machine learning for predicting energy production based on weather conditions.
  - Voting Regressor outperformed other models, showcasing the strength of ensemble approaches.
- **Future Work:**
  - Incorporate more granular data, such as hourly weather reports, for improved predictions.
  - Explore advanced techniques like deep learning for capturing complex patterns.
- **References:**
  - Python libraries: pandas, numpy, scikit-learn, meteostat.
  - SMA device for Northern Electricity Distribution Company