Weekly Progress Report

Name: Nishant Mall
Domain: Data Science and Machine Learning
Date of submission: 16 Sept 2024

Week Ending: 02

I. Overview:

This week, I continued to make progress on the two projects initiated last week: Predicting agriculture crop production in India and Predicting maintenance of gearboxes using vibration sensors. I focused on data exploration, feature engineering, and model selection.

II. Achievements:

1. **Prediction of Agriculture Crop Production in India:**
    - **Data Exploration:** Conducted an in-depth analysis of the agricultural dataset to identify correlations, patterns, and potential outliers.
    - **Feature Engineering:** Created new, informative features by combining or transforming existing data, such as calculating ratios and calculating seasonal variations.
2. **Gearbox Predictive Maintenance:**
    - **Feature Extraction:** Extracted relevant features from the vibration data, including time-domain and frequency-domain characteristics.
    - **Feature Selection:** Employed techniques like correlation analysis and feature importance to identify the most informative features for predicting gearbox health.

III. Challenges:

1. **Data Quality:** Encountered issues with missing values and inconsistencies in the agricultural dataset.
2. **Data Imbalance:** The gearbox predictive maintenance dataset may be imbalanced, with fewer instances of gearbox failures compared to normal operations.

IV. Learning Resources:

This week, I deepened my understanding of data science concepts by reading the book "Introduction to Data Science" by Davy Cielen. This comprehensive resource provided valuable insights into the fundamentals of data science, including data collection, preprocessing, analysis, and modeling.

1. **Online Courses:** Enrolled in online courses on machine learning and data science to gain a solid foundation in the field.
2. **Research Papers:** Explored relevant research papers to stay updated on the latest techniques and methodologies.

V. Next Week's Goals:

1. **Agriculture Crop Production Prediction:**
    - Experiment with different feature engineering techniques to improve the model's performance.
    - Explore various data preprocessing methods to address data quality issues.
2. **Gearbox Predictive Maintenance:**
    - Implement techniques to handle data imbalance and improve model accuracy.
    - Begin exploring machine learning algorithms suitable for classification and regression tasks.

VI. Additional Comments:

I am continuing to make progress on both projects and am excited to delve deeper into machine learning techniques in the coming weeks.