**GROUP 14**

**PHASE 5 CAPSTONE PROJECT PROPOSAL**

**Predictive Maintenance in Manufacturing**

**Problem Statement:** Imagine a large manufacturing plant that produces automotive parts. This plant relies heavily on several complex machines, including CNC (Computer Numerical Control) machines, robotic arms, and conveyor belts. Unexpected machine breakdowns can lead to significant production delays, increased costs, and reduced product quality. To mitigate these risks, the plant decides to implement a predictive maintenance system.

**Data Collection:** The plant collects various types of data from its machines, including:

1. **Operational Data:**
   * Machine operating hours
   * Temperature, vibration, and pressure readings
   * Usage patterns and load conditions
2. **Maintenance Logs:**
   * Historical maintenance records
   * Types of maintenance performed (e.g., routine checks, repairs)
   * Time taken for repairs
3. **Failure Events:**
   * Historical records of machine failures
   * Types of failures (mechanical, electrical, etc.)
   * Downtime associated with each failure
4. **Environmental Conditions:**
   * Humidity and ambient temperature in the manufacturing area
   * Maintenance schedules and employee shifts

**Data Analysis:** Using this data, the data science team at the plant performs several analyses:

1. **Data Preprocessing:**
   * Clean and preprocess the data, handling missing values and normalizing different types of measurements.
2. **Feature Engineering:**
   * Create features such as the number of operating hours since the last maintenance, averages of temperature/vibration over time, and rolling averages for operational conditions.
3. **Model Development:**
   * Use machine learning algorithms (e.g., Random Forest, Gradient Boosting, or Neural Networks) to predict the likelihood of a machine failure based on the collected features.
4. **Validation:**
   * Split the dataset into training and testing sets to validate the model's accuracy. Use metrics such as precision, recall, and F1-score to evaluate performance.

**Implementation:** Once the predictive model is validated, the plant implements a dashboard for the maintenance team. This dashboard includes:

* **Real-Time Monitoring:** Continuous tracking of machine health indicators.
* **Alerts:** Notifications when a machine shows signs of potential failure, allowing for proactive maintenance.
* **Maintenance Scheduling:** Recommendations for maintenance based on predictive insights, optimizing schedules to minimize disruption.

**Results:**

1. **Reduced Downtime:** By predicting failures before they occur, the plant experiences a significant reduction in unplanned downtime.
2. **Cost Savings:** Predictive maintenance helps avoid costly emergency repairs and extends the lifespan of machinery.
3. **Improved Efficiency:** The maintenance team can prioritize tasks based on urgency and impact, leading to smoother operations.
4. **Enhanced Safety:** With fewer machine failures, the risk of workplace accidents decreases, creating a safer environment for employees.

**Conclusion:** This predictive maintenance project not only helps the manufacturing plant optimize its operations but also illustrates the value of data-driven decision-making in real-world scenarios. By leveraging historical data and advanced analytics, the plant can enhance productivity and sustainability in its manufacturing processes.

**Data Source:**

For a predictive maintenance project in manufacturing, one of the best datasets we will use is the **NASA Turbofan Engine Degradation Simulation Data Set**.

**Dataset Details:**

**Dataset Name:** NASA Turbofan Engine Degradation Simulation Data Set  
**Source:** NASA Prognostics Center of Excellence (PCoE)  
**Link:** [NASA Turbofan Engine Dataset](https://data.nasa.gov/dataset/NASA-Turbofan-Engine-Degradation-Simulation-Dataset-1/ptmt-fh57)

**Key Features:**

* **Operational Data:** Includes various operating conditions and settings, such as temperature, pressure, and rotational speeds of the engines.
* **Degradation States:** Contains simulated failure modes and degradation levels for each engine, allowing you to model the relationship between operational parameters and failure occurrences.
* **Time Series Data:** The dataset is structured in a time-series format, providing continuous data over time, which is essential for predictive maintenance analysis.

**Advantages:**

* **Realistic Simulation:** The data simulates real-world engine conditions and failures, making it applicable for practical predictive maintenance scenarios.
* **Comprehensive:** It includes a rich set of features, allowing for deep exploration of how different parameters affect engine performance and failure.
* **Well-Documented:** The dataset comes with detailed documentation, making it easier to understand the context and use it effectively.

**Alternative Datasets:**

If you're looking for other options, consider the following:

1. **SECOM Data Set** (UCI Machine Learning Repository)
   * Focuses on manufacturing quality control and includes sensor readings related to a manufacturing process.
   * [SECOM Dataset](https://archive.ics.uci.edu/ml/datasets/SECOM)
2. **Turbofan Engine Degradation Simulation Data Set** (UCI Machine Learning Repository)
   * A simpler version of the NASA dataset, good for initial experimentation.
   * [UCI Turbofan Engine Dataset](https://archive.ics.uci.edu/ml/datasets/Turbofan+Engine+Degradation+Simulation+Data+Set)
3. **Predictive Maintenance Dataset** (Kaggle)
   * Contains various datasets related to predictive maintenance across different industries.
   * Kaggle Predictive Maintenance Dataset

These datasets provide a solid foundation for developing and testing predictive maintenance models in a manufacturing context.