"Project Proposal"

Machine Learning

"Machine Failure Detection"

Problem Statement:

As we all know that in industrial settings, unexpected machinery failures lead to costly downtime, high maintenance expenses and moreover safety risks. This project aims to develop a machine learning-based system for predicting machine failures. By analyzing historical sensor data and operational logs, the system will detect early signs of potential failures, allowing for proactive maintenance. This approach seeks to minimize downtime, reduce maintenance costs, and enhance operational efficiency and safety.

Significance of the problem in the Pakistani context:

In Pakistan, the industrial sector is a critical component of the economy, contributing significantly to GDP and employment. However, many industries face challenges related to unexpected machinery failures, which lead to substantial financial losses and production delays. A machine learning-based failure detection system can significantly improve productivity, reduce maintenance costs, enhance worker safety, and optimize energy usage. This project addresses a critical need, promoting economic stability, safety, and sustainability in Pakistan's industrial sector.

Scope of the problem:

The problem of unexpected machinery failures spans various industries in Pakistan, including manufacturing, textiles, agriculture, and energy. The breadth and depth of this issue are characterized by:

- **Industrial Impact**: Widespread productivity losses and increased costs.
- **Economic Consequences**: SMEs face crippling financial burdens due to unplanned maintenance.
- **Safety Risks**: Frequent failures pose safety hazards to workers, leading to injuries and lower productivity.
- **Energy Efficiency**: Inefficient machinery increases energy consumption, worsening the energy crisis and operational costs.

Limitations:

- **Data Availability**: Many industries may lack comprehensive historical data for training machine learning models.
- **Technological Barriers**: Limited access to advanced technologies and skilled personnel.
- **Initial Costs**: High upfront investment for developing and deploying the system.
- **Integration Challenges**: Difficulty integrating new systems with existing infrastructure.

Addressing this problem through machine learning can significantly enhance productivity, safety, and economic stability despite these limitations.

Dataset Information:

Data has been collected from Kaggle, a reputable platform known for high-quality datasets. This dataset includes comprehensive sensor readings and operational logs from industrial machinery, crucial for predicting failures. Kaggle datasets are reliable, well-documented, and versatile, covering various machinery types, making them ideal for developing accurate predictive models for machine failure detection.

ML Approach & expected Algorithms:

The Machine Learning approach that will be used in this project will be supervised learning for classification. Specifically, the algorithm being implemented is k-Nearest Neighbors (kNN) for classification. This algorithm is utilized to predict machine failures based on features extracted from sensor data and operational logs. We will also be using preprocessing steps such as data cleaning, visualization using seaborn and matplotlib, label encoding for categorical variables, and model evaluation using accuracy score.

Relevant Industry & Partner Name:

Asfar Jan –Data/Software Engineer at Afiniti

This Industry partner is providing us with the relevant and necessary guidance and mentorship that will be efficient to make this semester project.

Methodology to solve the problem:

To address this issue we will first collect and preprocess the data from Kaggle to ensure that it's clean and standardized. Next, we will conduct exploratory data analysis (EDA) will identify key features and patterns, followed by feature engineering. Machine learning models will be trained and evaluated using metrics like accuracy and ROC-AUC. The best model will be optimized and

deployed for real-time monitoring and maintenance recommendations. The system will be tested in real-world settings and continuously updated to maintain accuracy. This methodology aims to create an effective machine learning-based system for predicting machinery failures, enhancing efficiency, reducing costs, and improving safety.

Group Members Name:

Mohammad Ashar – cs211083

Ameen Uddin Zain U Din – cs211066

Shaikh Ahmed – cs211095

Mohammad Fariz Farooqui – cs211080

Hasnain Ashraf – cs211115