# Automated Anomaly Detection for Predictive Maintenance

## Description of Design Choices

The project focuses on building an anomaly detection model for predictive maintenance. Key design choices include:

1. Data Preprocessing: Handling missing values and outliers to ensure data quality.
2. Exploratory Data Analysis: Understanding patterns in the data and identifying correlations between features.
3. Model Selection: Using a Random Forest Classifier for its robustness and ability to handle imbalanced datasets.
4. Hyperparameter Tuning: Optimizing the model for better accuracy and generalizability.
5. Deployment Plan: Outlining steps to integrate the trained model into a production environment.

## Performance Evaluation

The model achieved over 93% accuracy on the test dataset. Key performance metrics include:

- Precision, Recall, and F1-Score: Highlighting the model's effectiveness in detecting anomalies.
- Confusion Matrix: Showcasing true positives, false positives, and other classifications.

The optimized model (via GridSearchCV) improved performance by selecting the best hyperparameters for Random Forest.

## Discussion of Future Work

Future improvements could include:

1. Using additional algorithms like XGBoost or Neural Networks to compare performance.
2. Implementing a real-time anomaly detection pipeline.
3. Extending the dataset to cover more diverse scenarios.
4. Automating the retraining process with new data.

These enhancements aim to improve accuracy, scalability, and deployment efficiency.