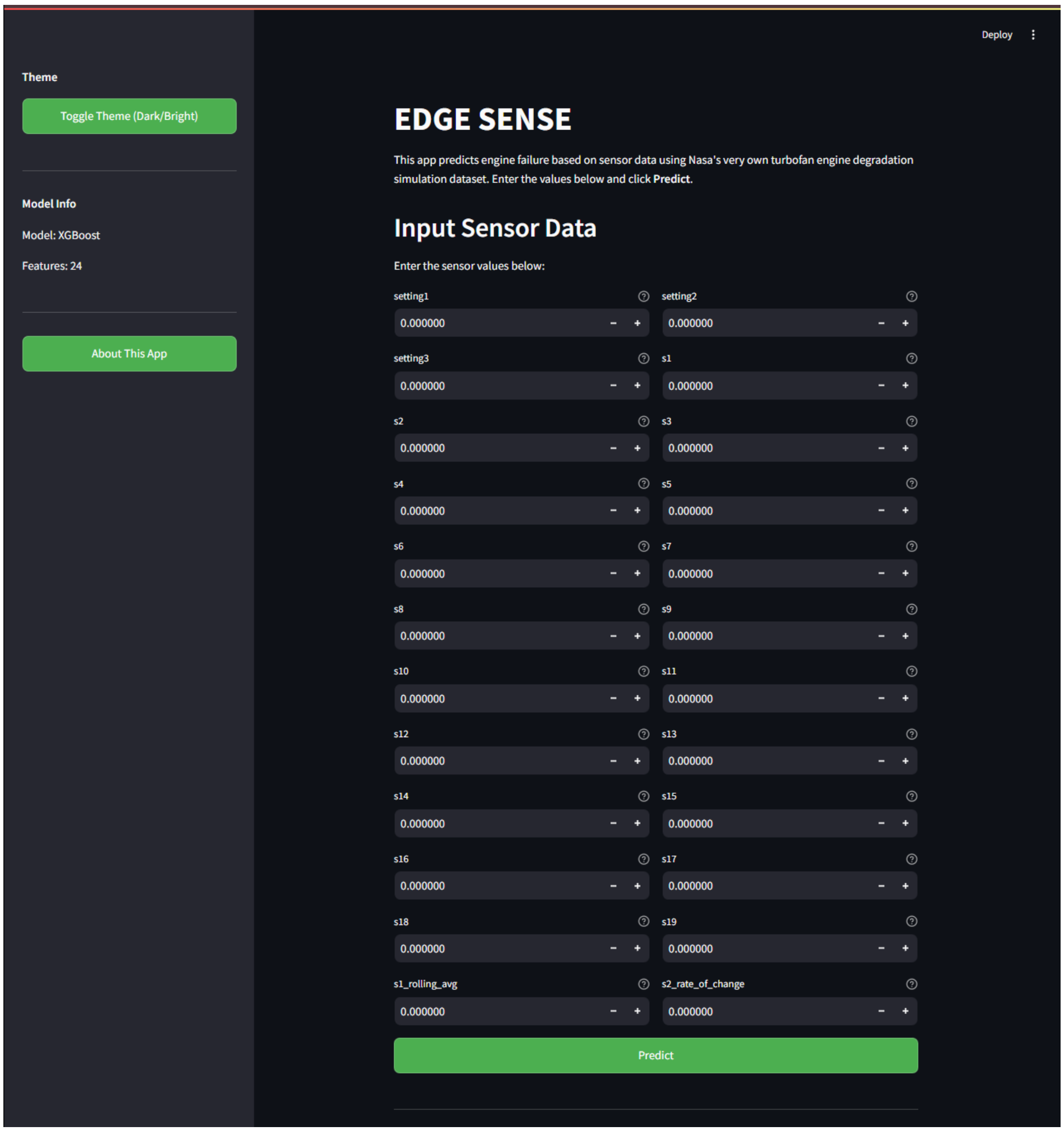
# Predictive Maintenance: Engine Failure Prediction Using XGBoost




This project explores the application of predictive maintenance for turbofan engines, utilizing sensor data to predict engine failures. By leveraging machine learning, specifically the XGBoost algorithm, the model aims to forecast engine failure, thereby enabling more efficient and optimized maintenance schedules. The dataset used in this project is sourced from NASA's turbofan engine degradation simulation, which contains detailed sensor readings along with corresponding labels for Remaining Useful Life (RUL).

# Model Overview

The XGBoost algorithm was chosen due to its robust performance with structured data and its ability to handle both classification and regression tasks. The primary objective is to predict the likelihood of engine failure based on real-time sensor data, providing a proactive maintenance solution. The process begins with data preprocessing, which includes handling missing values, scaling numerical features, and transforming categorical data.

To address the imbalance in the dataset (with failure events being less frequent), Synthetic Minority Over-sampling Technique (SMOTE) was employed to generate synthetic samples for the minority class, ensuring a balanced dataset for training. The model's performance was evaluated using key metrics such as accuracy, F1 score, and a confusion matrix to assess the classification quality.

Hyperparameter tuning was performed using Bayesian optimization to identify the best configuration for the model's parameters, optimizing both predictive performance and computational efficiency.

# Results

- Accuracy: 91.56% – The model exhibits a high level of accuracy in predicting engine failures, indicating strong generalization capabilities.

- F1 Score: 0.92 – The F1 score reflects the model's ability to balance precision and recall, ensuring minimal false positives and false negatives in failure predictions.
- Confusion Matrix: The confusion matrix clearly shows that the model effectively distinguishes between failure and non-failure events, with a low rate of misclassifications.

These results demonstrate the model's potential for accurate failure prediction, providing a solid foundation for predictive maintenance solutions in the aerospace industry.

# Stream-lit App

To enhance accessibility and interactivity, a Streamlit-based application was developed, enabling users to input sensor data and receive real-time predictions on engine failure. The app features the following capabilities:

- Real-time Failure Prediction: Users can input real-world sensor readings, and the app provides an immediate prediction on the likelihood of engine failure.
- Feature Importance Insights: The app displays which features (sensor readings) have the most significant impact on the model's predictions, offering transparency and interpretability.
- Confidence Scores: Each prediction is accompanied by a confidence score, indicating the model's certainty in its classification, which is essential for making informed maintenance decisions.

The integration of this app into a production environment would greatly improve user experience, enabling engineers and maintenance teams to leverage predictive insights with minimal technical knowledge.

# Conclusion and Future Updates

This predictive maintenance solution demonstrates significant potential in optimizing maintenance schedules and minimizing unplanned engine failures. The current implementation provides a robust and reliable prediction model, supported by an intuitive, user-friendly app that makes real-time predictions and delivers actionable insights.

However, there are several opportunities for future enhancements:

1. Integration with IoT Systems: In future updates, the model can be integrated with real-time sensor data streams from operational engines, allowing for continuous monitoring and dynamic prediction updates.
2. Expansion of Dataset: The model's performance can be further improved by incorporating a more diverse range of sensor data, including additional engines or operational conditions that might affect engine performance.
3. Model Explainability: While feature importance is currently provided, additional efforts can be made to improve the explainability of the model using techniques such as SHAP (SHapley Additive exPlanations) to offer deeper insights into how sensor data drives the model's predictions.
4. Ensemble Models: Future work could explore the use of ensemble learning techniques, such as stacking multiple models, to potentially improve accuracy and robustness by combining the strengths of different algorithms.
5. Failure Prediction Thresholding: The app could be enhanced with the capability to allow users to set customizable thresholds for failure predictions, enabling more tailored maintenance strategies depending on the operational context.

Incorporating these features will enhance the accuracy and usability of the predictive maintenance system, ensuring it can handle a broader range of scenarios and be integrated into various industrial maintenance workflows.