PREDICTIVE MAINTENANCE FOR HYDRAULIC SYSTEMS USING MACHINE LEARNING

ROCK DENG
UNIVERSITY OF COLORADO BOULDER

PROJECT OVERVIEW

1. **Objective**: Develop a machine learning model to predict hydraulic system failures using sensor data.
2. **Key Focus**: Real-time deployment feasibility and high accuracy and computational efficiency.
3. **Visual**: Flowchart showing the project pipeline (data → preprocessing → modeling → evaluation).

PROBLEM STATEMENT

Current Challenges:

1. Reactive maintenance leads to costly downtime
2. Manual inspections are time-consuming
3. Unexpected failures disrupt operations

RELATED WORK

Prior Research:

1. Machine learning models (Random Forests, SVMs, LSTMs) for failure prediction
2. Feature engineering techniques (PCA, time-series analysis)

RELATED WORK

Gaps in Research:

1. Need for interpretable and lightweight models.
2. Limited real-time deployment applications.

PROPOSAL WORK

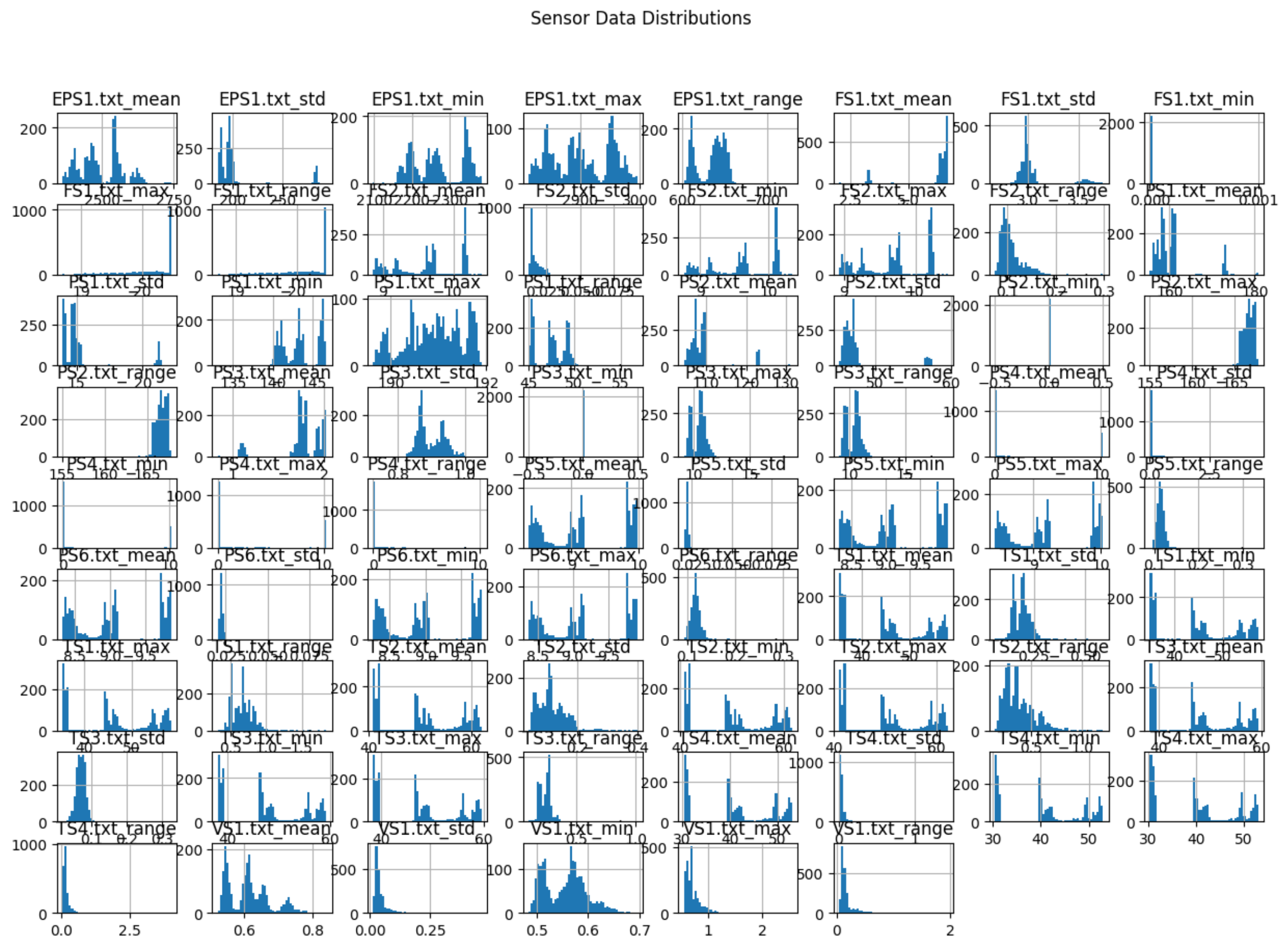
Proposed Solution:

1. Predictive maintenance using machine learning.
2. Real-time monitoring and failure prediction.

DATA PREPROCESSING

Feature Engineering

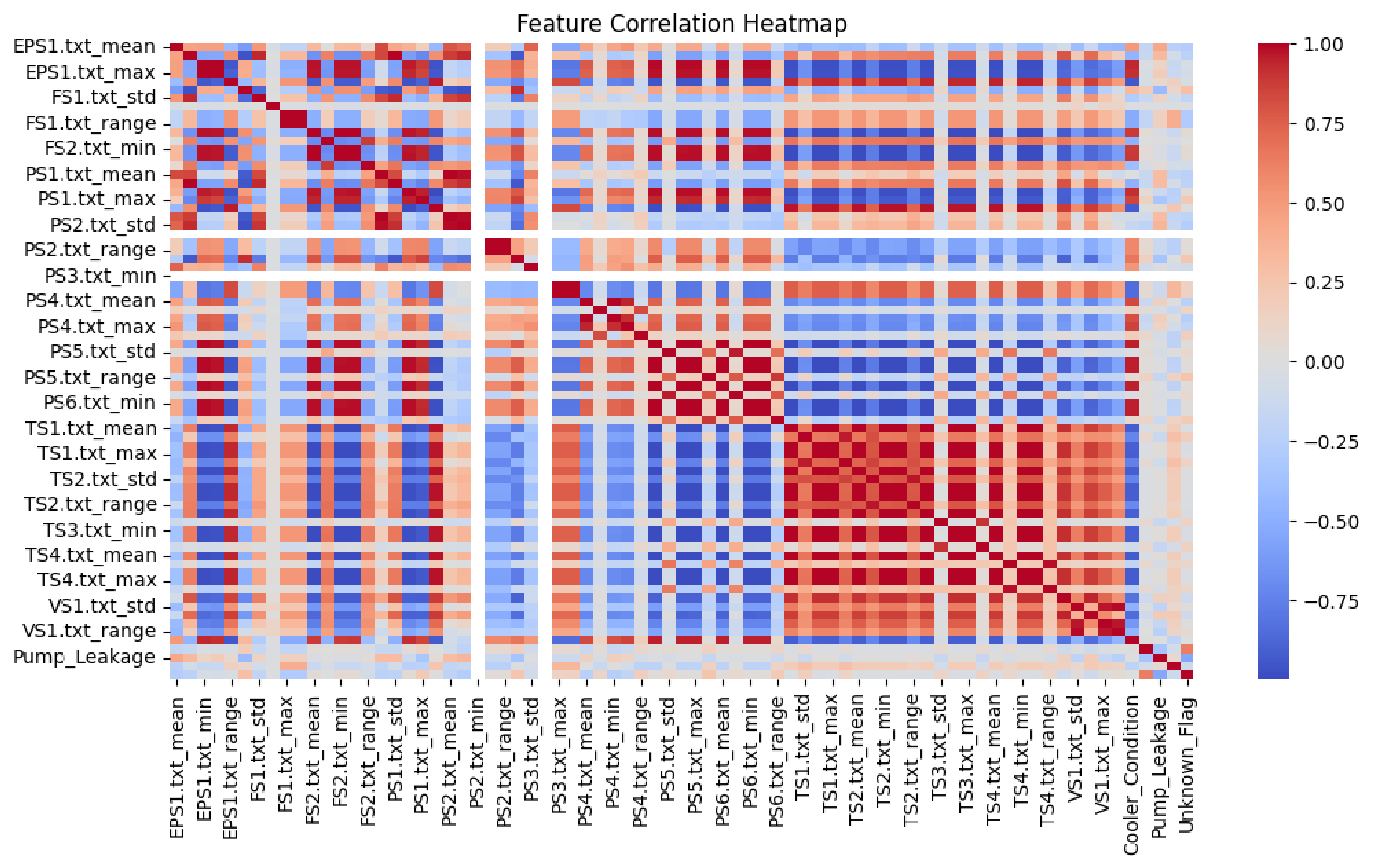
Extracted statistical features for each sensor per cycle, includes



DATA PREPROCESSING

Feature Engineering

Find correlation between features



DATA PREPROCESSING

Feature reduction using PCA

```python
# PCA Comparison - Before PCA
print("Original Feature Dimension:", X.shape)

# Feature reduction using PCA
pca = PCA(n_components=10)
X_reduced = pca.fit_transform(X)

# PCA Comparison - After PCA
print("Reduced Feature Dimension:", X_reduced.shape)
```

```
Original Feature Dimension: (2205, 70)
Reduced Feature Dimension: (2205, 10)
```

Use Principal Component Analysis (PCA) to reduce the number of input features to speed up model training and model size.

Model Training and Evaluation

Random Forest Classifier

Random Forest, an ensemble learning method, trained using 80% of the dataset and evaluated on the remaining20%.

Model Training and Evaluation

Random Forest Classifier

Results:

- Accuracy: 99.77%
- Precision, Recall, F1-score:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	228
1	0.99	1.00	1.00	101
2	1.00	0.99	1.00	112
accuracy			1.00	441
macro avg	1.00	1.00	1.00	441
weighted avg	1.00	1.00	1.00	441

Model Training and Evaluation

Classification Metrics:

1. Accuracy, Precision, Recall, F1-score. But we only use accuracy for this project.

Model Efficiency:

1. Training time and inference speed. This could be future work

Model Training and Evaluation

XGBoost Classifier

XGBoost, a gradient boosting algorithm, was used to compare against the Random Forest model. XGBoost is known for handling complex relationships in data efficiently.

PROJECT TIMELINE

1. Week 1: Dataset exploration and preprocessing.
2. Week 2-3: Implement baseline models (Decision Trees, XGBoost).
3. Week 4: Experiment with time-series models (LSTM, ARIMA).
4. Week 5: Evaluation and performance comparison.- Week 6: Final report writing and refinements.

Model Training and Evaluation

XGBoost Classifier Result:

Results:

- Accuracy: 99.09%
- Precision, Recall, F1-score:

	precision	recall	f1-score	support
0	0.99	1.00	1.00	228
1	0.99	0.99	0.99	101
2	0.99	0.97	0.98	112
accuracy			0.99	441
macro avg	0.99	0.99	0.99	441
weighted avg	0.99	0.99	0.99	441

Model Training and Evaluation

The results show that the Random Forest model performedslightly better in accuracy compared to XGBoost. Both models achieved near-perfect performance, indicating that the extracted features effectively captured the failure patterns in the dataset.

Discussion

This study demonstrates that machine learning-based predictive maintenance can achieve near-perfect accuracy when applied to hydraulic system monitoring. Compared to previous research, our findings highlight the benefits of PCA in reducing computational overhead while maintaining predictive performance. Additionally, our comparative model analysis suggests that Random Forest offers a balance of interpretability and high accuracy, making it a strong candidate for real-time deployment.

Future work

1. Extending this analysis to additional predictive models, such as deep learning-based architectures for time-series forecasting.
2. Implementing real-time deployment strategies by optimizing model inference speed through quantization techniques.
3. Exploring domain adaptation techniques to generalize the model to different hydraulic systems beyond the dataset used.

Conclusion

This project developed a predictive maintenance system for hydraulic systems using machine learning. By leveraging sensor data and feature engineering, the system successfully predicted failures with high accuracy while optimizing computational efficiency. Future work will aim to further refine deployment strategies for real-time industrial applications.

THANKS!