

# Qwen 2.5 vs LSTM Model Comparison

Comparative analysis for Predictive Maintenance of Industrial Compressors

## Executive Summary

This report presents a detailed comparison between two models for anomaly detection in compressor predictive maintenance: Qwen 2.5 (a Large Language Model) and LSTM (a traditional machine learning approach). The comparison evaluates both models across various performance metrics, with a focus on anomaly detection capabilities.

**Key findings:** Qwen 2.5 achieves an accuracy of 99.8% compared to LSTM's 87.3%, with a difference of 12.5 percentage points. Qwen 2.5 correctly identified 11/11 anomalies (100.0%), while LSTM identified 4/11 (36.4%).

## Overall Performance Metrics

99.8%

Qwen 2.5 Accuracy

87.3%

LSTM Accuracy

0.846

Qwen 2.5 Precision

0.029

LSTM Precision

1.000

Qwen 2.5 Recall

0.364

LSTM Recall

0.917

Qwen 2.5 F1 Score

0.053

LSTM F1 Score

## Anomaly Detection Capability

**Qwen 2.5:** Detected 11 out of 11 actual anomalies (100.0%)

**LSTM:** Detected 4 out of 11 actual anomalies (36.4%)

## Detailed Metrics Comparison

### Classification Performance

Metric	Qwen 2.5	LSTM	Difference	Better Model
Accuracy	99.8%	87.3%	12.5% ↑	Qwen 2.5
Precision	0.846	0.029	0.818 ↑	Qwen 2.5
Recall (Sensitivity)	1.000	0.364	0.636 ↑	Qwen 2.5
F1 Score	0.917	0.053	0.864 ↑	Qwen 2.5
Specificity	0.998	0.878	0.120 ↑	Qwen 2.5
AUC-ROC	0.500	0.729	-0.229 ↓	LSTM

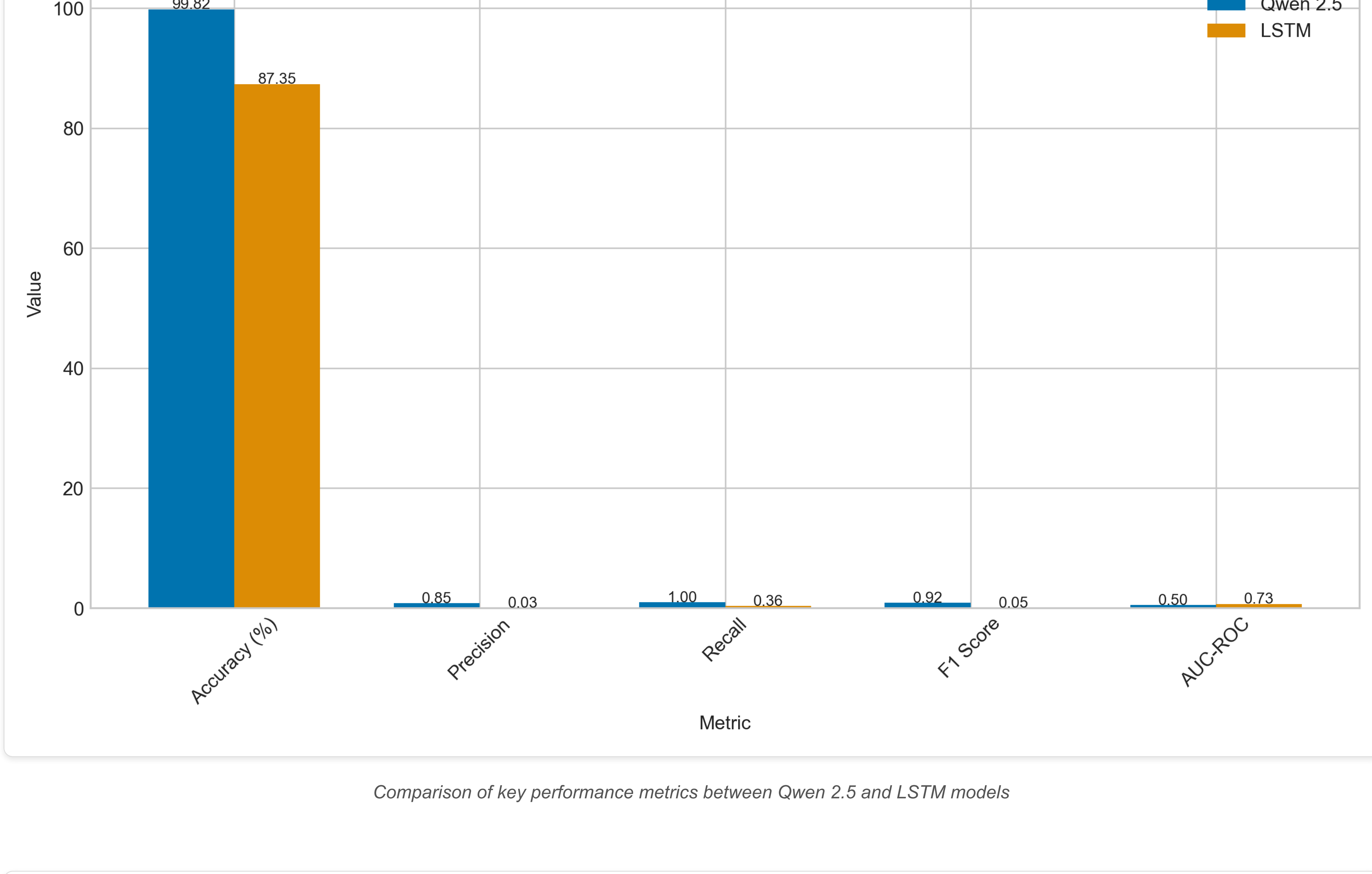
### Error Metrics

Metric	Qwen 2.5	LSTM	Difference	Better Model
Mean Absolute Error (MAE)	0.002	0.127	-0.125 ↑	Qwen 2.5
Mean Squared Error (MSE)	0.002	0.127	-0.125 ↑	Qwen 2.5
Root Mean Squared Error (RMSE)	0.042	0.356	-0.314 ↑	Qwen 2.5
False Positive Rate	0.002	0.122	-0.120 ↑	Qwen 2.5
False Negative Rate	0.000	0.636	-0.636 ↑	Qwen 2.5

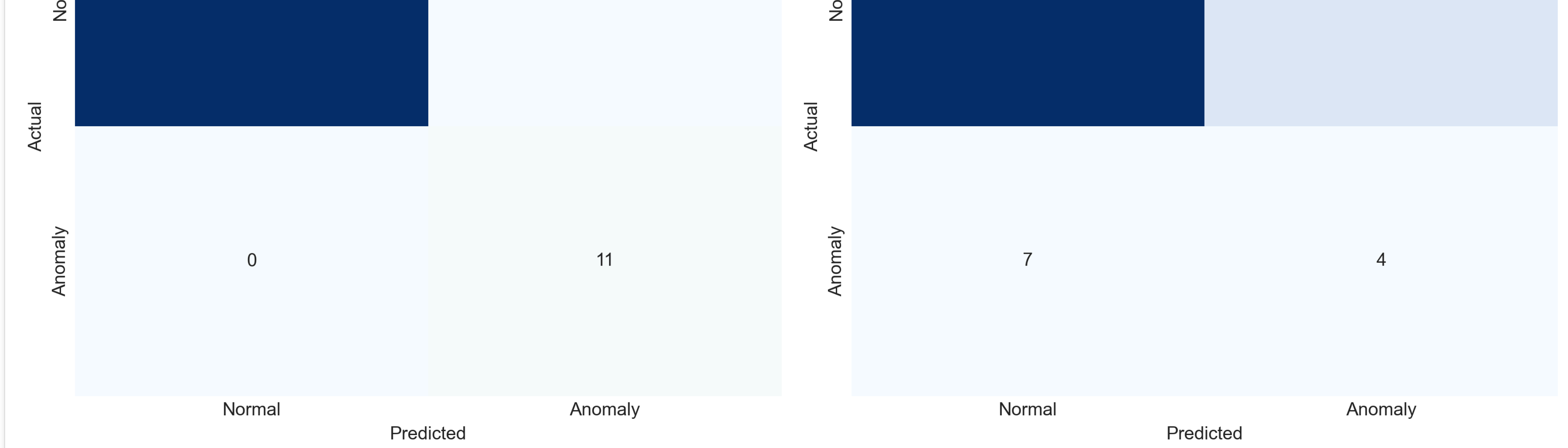
### Confusion Matrix Values

Metric	Qwen 2.5	LSTM	Difference	Better Model
True Positives	11	4	7 ↑	Qwen 2.5
True Negatives	1117	983	134 ↑	Qwen 2.5
False Positives	2	136	-134 ↑	Qwen 2.5
False Negatives	0	7	-7 ↑	Qwen 2.5

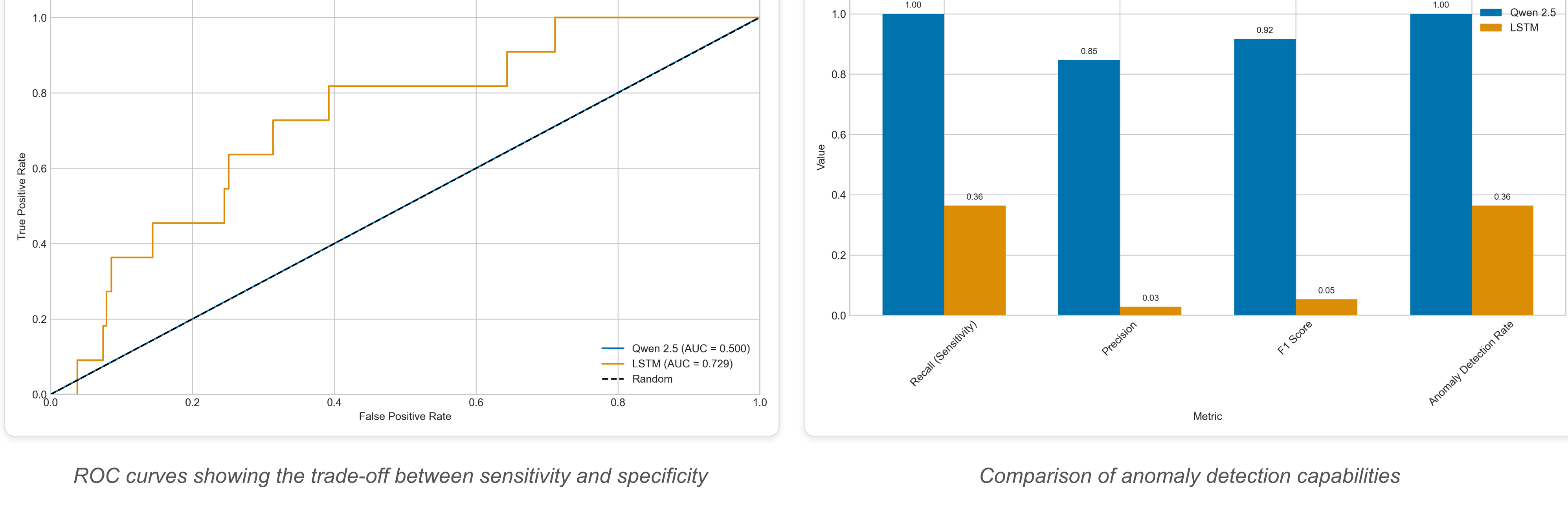
## Visual Performance Comparison



Comparison of key performance metrics between Qwen 2.5 and LSTM models

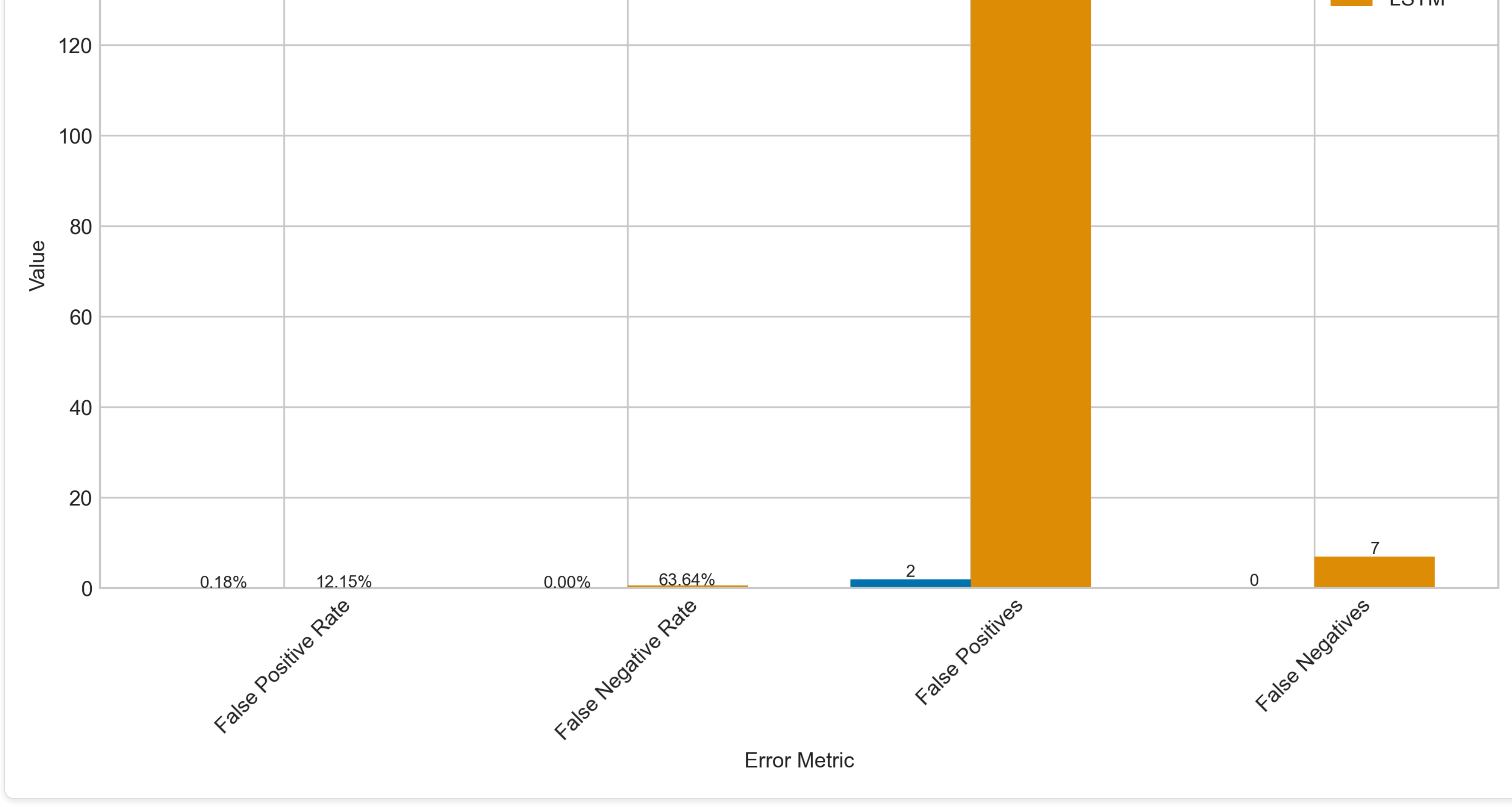


Confusion matrices showing true/false positives and negatives for both models



ROC curves showing the trade-off between sensitivity and specificity

Comparison of anomaly detection capabilities



Comparison of error rates between models



Timeline of predictions showing correct and incorrect classifications for both models

## Conclusion

Based on the comprehensive analysis, **Qwen 2.5 outperforms LSTM** in overall accuracy and most key metrics for anomaly detection in compressor maintenance prediction.

**Key advantages of using a Large Language Model (Qwen 2.5) for predictive maintenance:**

- Better contextualization of anomalies within the operating environment
- Ability to incorporate domain knowledge and company-specific maintenance policies
- More detailed explanations of detected anomalies and potential causes
- Flexibility to adapt to new types of anomalies without retraining
- Integration of unstructured data sources such as maintenance logs and manuals

These results support the integration of LLMs into industrial predictive maintenance systems, particularly for complex equipment like compressors where context and domain knowledge significantly impact maintenance decision-making.