Machine Learning for Wind Turbine Blade Failure Prediction: Analysing Vibration Data Effects on Maintenance Costs

Renishkumar Mavani
mavani20963@hs-Ansbach.de
Campus Feuchtwangen, Hochschule Ansbach
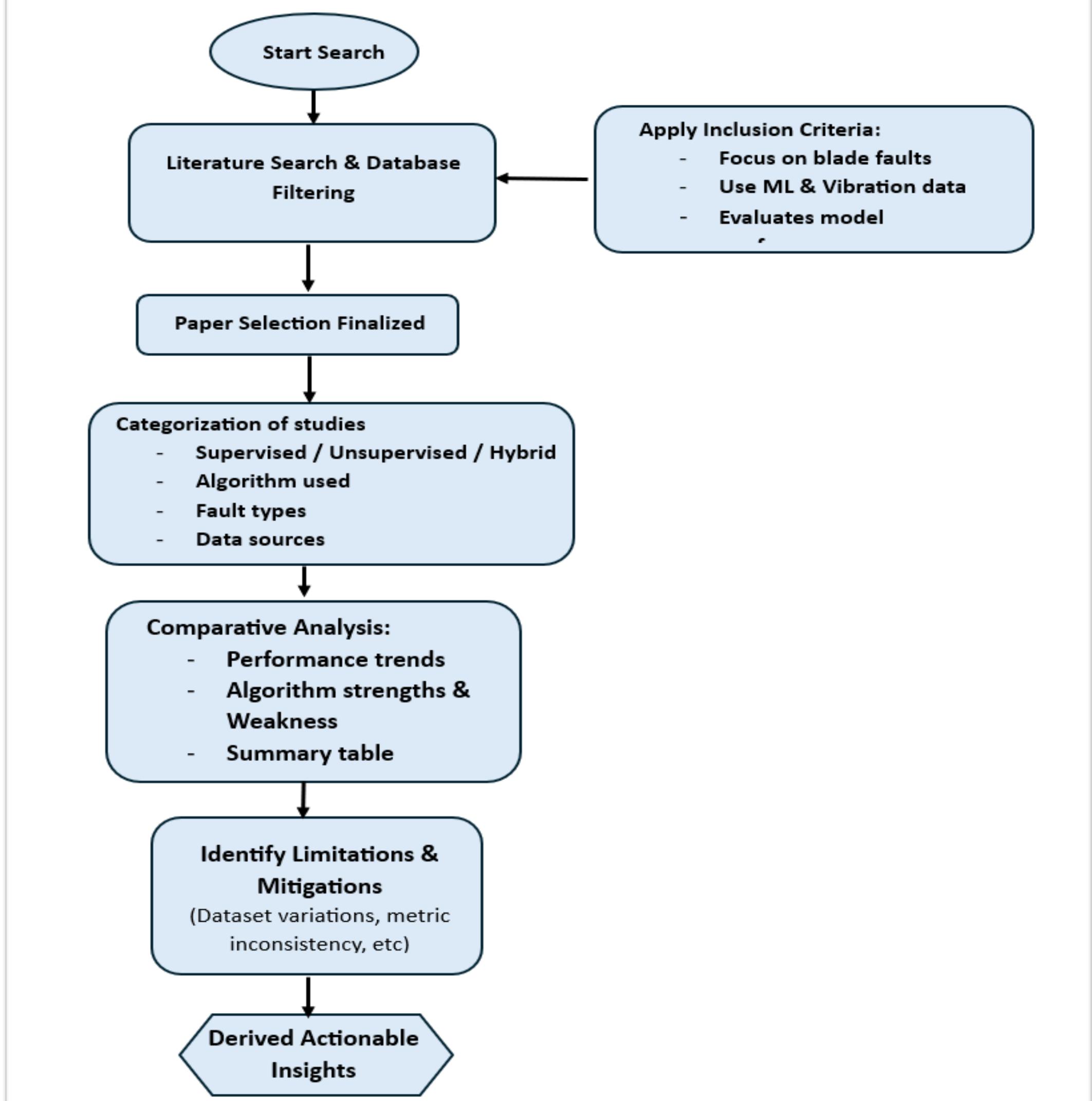
Abstract

Wind turbine blade failures, though rare, can cause costly repairs and significant downtime. This study explores how machine learning techniques applied to blade vibration data can predict failures early and thereby reduce maintenance costs. A structured literature review compares state-of-the-art ML models for blade health monitoring based on criteria such as fault prediction accuracy, false alarm rate, and maintenance cost impact. The goal is to identify which ML approaches most effectively detect blade faults in advance and how their deployment can optimize maintenance strategies. Preliminary findings suggest that supervised models achieve high fault detection accuracy, but their practical value depends on controlling false positives to avoid unnecessary inspections. Emerging approaches like deep learning and ensemble methods show promise in capturing complex vibration patterns, while unsupervised techniques help detect anomalies with minimal labeled data. In a predictive maintenance context, effective blade-failure prediction enables operators to schedule repairs during planned maintenance windows, minimizing unplanned outages. This review ultimately provides a comparative analysis of ML models and guidelines for leveraging them to enhance wind farm reliability and maintenance cost-efficiency.

Introduction

Wind turbine blades are vital for energy generation but are prone to fatigue, material wear, and extreme weather. Though failures are rare, they can cause severe damage and revenue loss. Traditional maintenance methods like scheduled inspections are inefficient, prompting a shift toward predictive maintenance. Machine Learning enables early fault detection by analyzing vibration data to identify cracks or imbalances. This study aims to identify the most effective ML techniques for early fault detection to optimize maintenance and reduce costs.

Methods
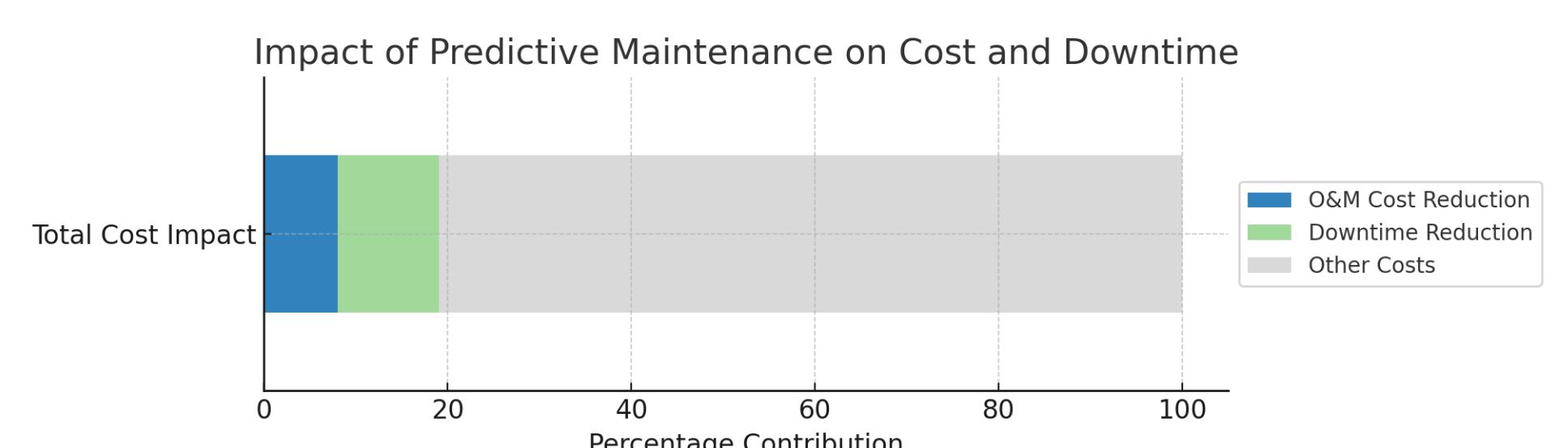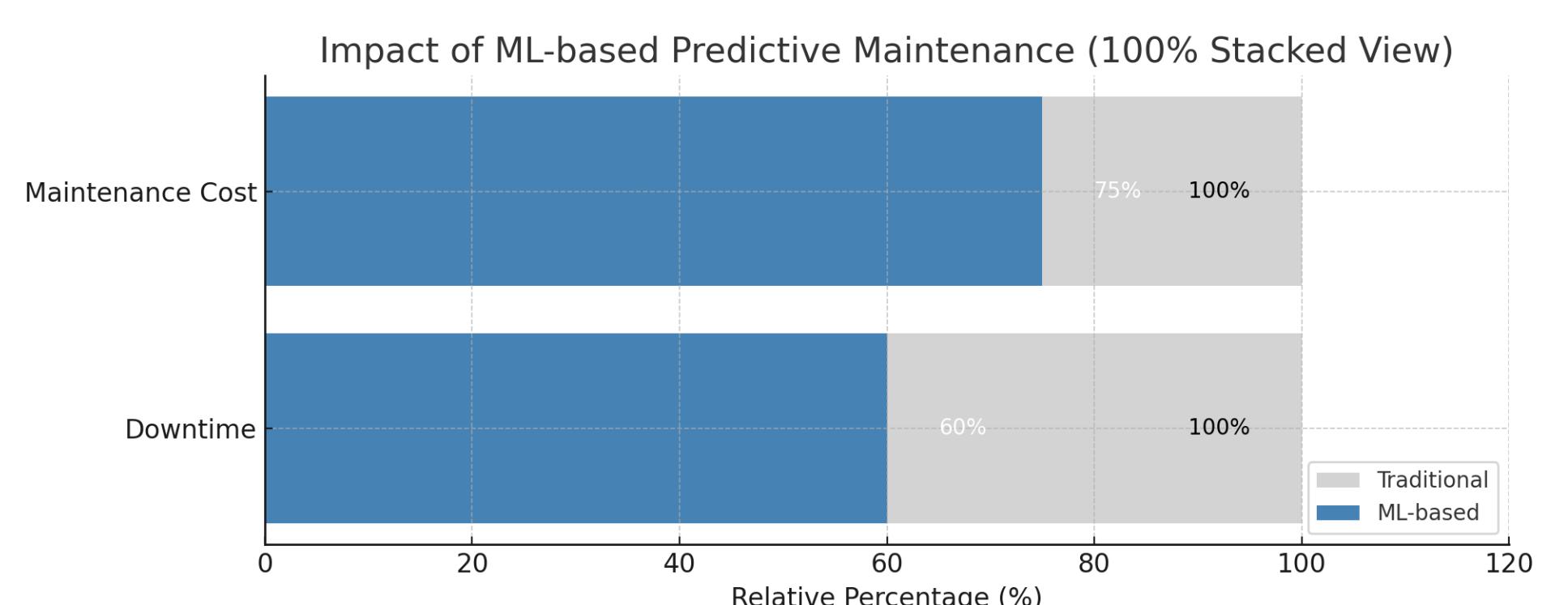
Results

Model Comparison:

ML Model	Typical Accuracy	False Alarm Rate	Maintenance Cost
Support Vector Machine (SVM)	80–95% accuracy in fault classification.	Requires careful tuning to avoid moderate false alarms, especially with noisy data.	High accuracy reduces missed failures, but false alarms can cause unnecessary inspections.
Random Forest	Handles complex features with 80-97% accuracy.	Low false positive bias (if well-trained).	Fast, robust prediction; fewer false alarms (if trained well) lead to warranted maintenance.
K-Nearest Neighbours (KNN)	Accuracy varies (60-97%) depending on data and feature selection.	Feature selection reduces false alarms from overlapping clusters.	Low computational cost, cost-effective maintenance if tuned for high accuracy.
Deep Learning	Typically, >90% accuracy (large data); up to 99% reported.	Very low false alarms with extensive training; risk if overfitting or unseen data (needs filters).	High initial cost (data/compute); potential for long-term ROI via early detection, but false alarms can be costly (needs tuning).
Unsupervised Machine Learning Model	no fixed accuracy; detects novel faults. High precision in separating normal/abnormal data.	Potential for false alarms or missed faults due to threshold; calibrate to turbine baseline.	Catches unforeseen issues (requires tuning); prevents major failures if calibrated, but false alarms possible if not. Use with other methods for cost-efficiency.

Discussion & Conclusion

Conditions	Recommended Model (Best to Least)
High-noise environment	Random Forest → Transformer → CNN → SVM → Logistic Regression → KNN → Decision Tree
Limited data availability	Logistic Regression → Support Vector Machines → Decision Tree → KNN → Random Forest → CNN → Transformer
Predictive analytics (Early)	Transformer → CNN → Random Forest → Support Vector Machines → KNN → Logistic Regression → Decision Tree
Cost-sensitive (low false alarm)	Logistic Regression → Support Vector Machines → Random Forest → Transformer → CNN → KNN → Decision Tree
High interpretability	Decision Tree → Logistic Regression → KNN → Random Forest RF → SVM → CNN → Transformer

Impact of ML-based Predictive Maintenance (100% Stacked View)

Impact of Predictive Maintenance on Cost and Downtime

References

- Turnbull, A., & Carroll, J. (2021). Cost Benefit of Implementing Advanced Monitoring and Predictive Maintenance Strategies for Offshore Wind Farms. *Energies*, *14*(16), 4922. https://doi.org/10.3390/en14164922
- Bocewicz, G., Frederiksen, R. D., Nielsen, P., & Banaszak, Z. (2024). Integrated preventive–proactive–reactive offshore wind farms maintenance planning. *Annals of Operations Research*. https://doi.org/10.1007/s10479-024-05951-4