Data Analytics Capstone Synopsis

Transfer Learning for Fault Detection in Industrial Predictive Maintenance Systems

Student Name: Sundaresh Prasanna Chandran

Institution: Walsh College

Course: QM640 – Data Analytics Capstone (Winter 2025)

Mentor: Anoop Ragolu

Introduction

Industrial systems increasingly rely on sensor-driven monitoring to ensure operational reliability, safety, and cost efficiency. Predictive maintenance (PdM) has emerged as a critical paradigm enabling early fault detection and prevention of catastrophic failures. Traditional approaches often focus on Remaining Useful Life (RUL) estimation; however, these approaches suffer from scarcity of labeled failure data and heterogeneity across machines.

Fault detection offers a scalable alternative and can be formulated as a binary classification problem. Models trained on historical machines frequently fail to generalize to new machines due to domain shift caused by sensor calibration differences, operating regimes, manufacturing tolerances, and environmental conditions.

Transfer learning mitigates this issue by reusing knowledge from source domains and adapting it to target domains using limited labeled data. This study investigates transfer learning strategies for fault detection using multivariate sensor time-series data from turbofan engines.

Scope and Objectives

This study focuses on binary fault detection using multivariate sensor time-series data from turbofan engines.

Objectives:

1. Quantify performance degradation due to domain shift.

2. Evaluate fine-tuning strategies.

3. Compare Random Forest, CNN, and LSTM models.

4. Analyze label efficiency.

Research Questions

RQ1: How does the performance of a fault detection model degrade when transferred directly from a large source dataset to smaller target datasets without adaptation?

RQ2: To what extent does fine-tuning improve fault detection performance compared to direct transfer?

RQ3: Which model architecture (Random Forest, CNN, or LSTM) demonstrates the best robustness under domain shift?

RQ4: What is the minimum amount of labeled target data required to achieve acceptable fault detection performance?

Research Hypotheses

H01: There is no statistically significant difference in fault detection accuracy between the source-trained model and its performance on the target datasets without adaptation.

H11: There is a statistically significant decrease in fault detection accuracy under direct transfer.

H02: Fine-tuning does not significantly improve performance.

H12: Fine-tuning significantly improves performance.

H03: Model robustness is equal across architectures.

H13: At least one model architecture is significantly more robust.

H04: Increasing labeled data does not improve performance.

H14: Increasing labeled data improves performance until saturation.

Sample Size Calculation

A power analysis for two-proportion comparison is applied using $\alpha = 0.05$ and power = 0.80. The required sample size is approximately 550 samples per group, yielding a total of 1,100 labeled samples, which is satisfied by the NASA C-MAPSS dataset.

Data Description

Dataset: NASA C-MAPSS Turbofan Engine Dataset

Sensors: 21 continuous sensors

Operating conditions: 3

Fault label definition:

Fault = RUL $\leq$ 30 cycles

Healthy = RUL > 30 cycles

Training subset: FD002

Testing subsets: FD001, FD003, FD004

Data Dictionary

unit_id: Engine identifier (integer)

cycle: Time step index (integer)

op_setting_1–3: Operating conditions (float)

sensor_1–sensor_21: Sensor measurements (float)

fault_label: Binary fault indicator (0/1)

GitHub Repository Structure

```
project-root/
├── data/
│   ├── raw/
│   ├── processed/
│   └── labels/
├── notebooks/
├── models/
├── src/
│   ├── preprocessing.py
│   ├── training.py
│   ├── transfer_learning.py
│   └── evaluation.py
├── results/
└── README.md
```

Analytic Approach

RQ1: Direct model transfer evaluation using accuracy, F1-score, ROC-AUC.

RQ2: Fine-tuning classification layers using limited labeled target data.

RQ3: Model comparison using paired statistical tests.

RQ4: Learning curve analysis using 1%, 5%, 10%, and 20% labeled target data.

Evaluation Metrics

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)

F1-score = 2 × (Precision × Recall) / (Precision + Recall)

ROC-AUC = Area under ROC curve

Recommendation and Application

Target users:

Reliability Engineers

Maintenance Engineers

Applications:

Cross-machine fault monitoring

Reduced labeling costs

Faster PdM deployment

Improved operational reliability

Threats to Validity

Internal validity: Hyperparameter selection and preprocessing choices.

External validity: Dataset is simulated and may not generalize to all industries.

Construct validity: Fixed RUL threshold may not reflect true fault onset.

Conclusion validity: Temporal dependence in time-series data may violate independence assumptions.

References

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

Ganin, Y., & Lempitsky, V. (2015). Unsupervised domain adaptation by backpropagation. Proceedings of the 32nd International Conference on Machine Learning (ICML), 1180–1189.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. Mechanical Systems and Signal Processing, 104, 799–834. https://doi.org/10.1016/j.ymssp.2017.11.016

He, Z., Shao, H., Zhong, X., & Zhao, X. (2020). Transfer learning for intelligent fault diagnosis: A survey. Measurement, 152, 107326. https://doi.org/10.1016/j.measurement.2019.107326

Widodo, A., & Yang, B. S. (2007). Support vector machine in machine condition monitoring and fault diagnosis. Mechanical Systems and Signal Processing, 21(6), 2560–2574. https://doi.org/10.1016/j.ymssp.2006.12.007

Zhang, W., Peng, G., Li, C., Chen, Y., & Zhang, Z. (2017). A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. Sensors, 17(2), 425. https://doi.org/10.3390/s17020425

Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. Mechanical Systems and Signal Processing, 115, 213–237. https://doi.org/10.1016/j.ymssp.2018.05.050

Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. Pattern Recognition Letters, 119, 3–11. https://doi.org/10.1016/j.patrec.2018.02.010

Saxena, A., Goebel, K., Simon, D., & Eklund, N. (2008). Damage propagation modeling for aircraft engine run-to-failure simulation. Proceedings of the IEEE International Conference on Prognostics and Health Management, 1–9.