Anomaly Detection in the FSAE Athena EV using Dynamic Bayesian Networks

Tommaso Petrelli
Master's Degree in Artificial Intelligence, University of Bologna
tommaso.petrelli2@studio.unibo.it

April 5, 2025

Abstract

Formula SAE[1] (FSAE) is an international university competition focused on designing and producing high-performance race cars, evaluated on multiple factors, including cost, engineering efficiency, and safety. This study investigates inverter faults in the UniBo Motorsport[2] team's Athena electric vehicle (EV) using Dynamic Bayesian Network (DBN) trained on real-world telemetry data from test sessions. To assess its effectiveness in fault detection, we implemented a log-likelihood probability scoring function. While the DBN successfully captured normal behavioural patterns, it struggled to distinguish between normal and anomalous time series, highlighting challenges in applying DBNs for inverter fault detection.

Introduction

Domain

The powertrain and Battery Management System (BMS) of electric FSAE race cars are complex, requiring insights on error occurrence for better resource allocation to prevent or mitigate future failures. In collaboration with the powertrain sub-division, this study focuses on inverter faults using key signals including wheel speed, motor and inverter temperatures, and battery pack current flow. We employ a Dynamic Bayesian Network, which has been successfully used in neuroscience (Rajapakse and Zhou 2007) and engineering (Codetta-Raiteri and Portinale 2015), offering an interpretable representation of uncertain knowledge without hidden variables (Concha and Pedro 2014).

Aim

This project develops a two time-slice Dynamic Bayesian Network (2-TBN) modeling normal behavior of Athena, exploring data discretization techniques and implementing a log-likelihood scoring function to assess the model's ability to distinguish between normal and anomalous observations in telemetry data.

[1] https://www.fsaeonline.com/
[2] https://motorsport.unibo.it/

Method

Data preprocessing utilized Python libraries including Pandas[3] and NumPy[4]. Two discretization approaches were explored: threshold-based binning using expert-defined cutoffs and K-Means clustering with optimal bin numbers determined by Silhouette Score. The pgmpy library (Ankur and Johannes 2024) was employed to construct and analyze the DBN, focusing on Markov blankets, local independencies, and conditional probability distributions (CPDs). Additional analyses included correlation analysis, distribution analysis using KDE plots and boxplots, and Granger causality tests[5]. Our custom log-likelihood scoring function was used to evaluate the model's ability to detect behavioural deviations.

Results

The DBN structure significantly impacted the model's learning capability. Although the model yielded higher log-likelihood scores for normal datasets, test results showed anomalous data split between true negatives and false negatives, suggesting normal and anomalous sequences were too similar for clear separation.

Model

The DBN structure (Figure 1) was designed based on both domain expertise and insights derived from the analysis of discrete factors. The final model comprises 20 nodes and 40 directed edges, where each node represents a discrete random variable corresponding to a specific component of Athena. Threshold-based binning was employed to discretize continuous variables, with special attention to preserving the vehicle's left-right symmetry. The network topology reflects plausible causal relationships following heat and current flow. Core variables like `(BatteryCurrent_A, 0)` have numerous outgoing edges, while peripheral variables like `(InverterSpeed_RearLeft_RPM, 0)` serve primarily as effect nodes. The model was trained using Maximum Likelihood Estimation to learn CPDs from normal data.

[3] https://pandas.pydata.org/
[4] https://numpy.org/
[5] https://w.wiki/DiSk

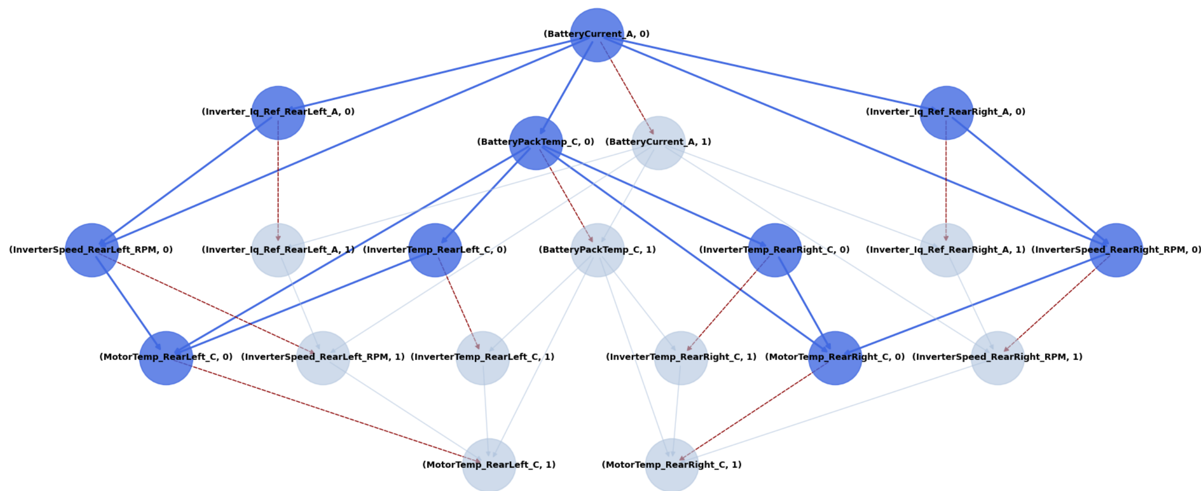

Figure 1: Structure of the proposed Dynamic Bayesian Network

Analysis

Experimental setup

Analyses focused on CPDs, discrete factors, Markov blankets and local independencies using pgmpy. Log-likelihood contributions of individual CPDs provided diagnostics for model behaviour. The model was evaluated on unseen data with a log-likelihood threshold defining normal versus anomalous classifications.

Results

The experimental results showed that the DBN successfully captured normal behaviours and learned CPDs without random guessing. However, despite positive training results, anomaly detection performance on new samples was limited, as shown in Figure 2. The system's complexity, the DBN's simplicity, and substantial overlap between normal and anomalous data hindered reliable distinction between classes.

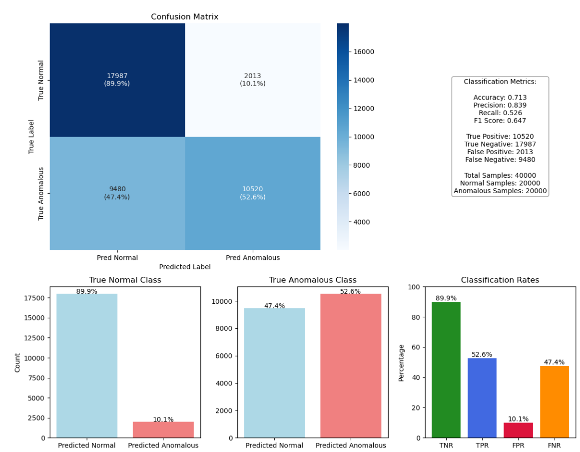

Figure 2: Classification results on new telemetry data

Conclusion

The model achieved log-likelihood scores of $-40,235.4719$ for normal data and $-196,581.9639$ for anomalous data (Figure 3), demonstrating effective learning of normal patterns despite lacking perfect separation. These limitations stem from dataset challenges rather than the 2-TBN model itself. Overall, while the DBN modelled normal behaviour successfully, more sophisticated modelling and richer data would be necessary for reliable inverter fault detection in the Athena EV.

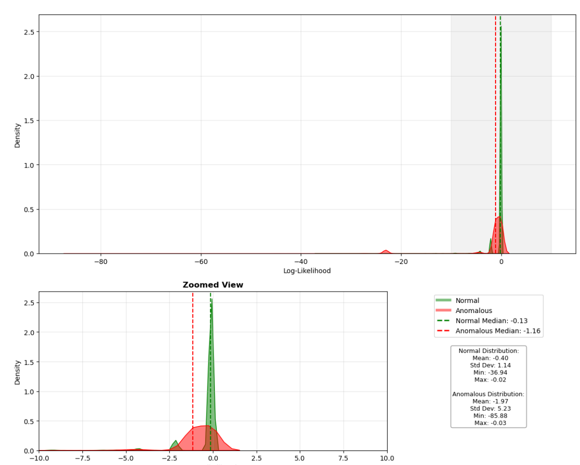

Figure 3: Log-likelihood distributions

Links to external resources

The developed code for this project is available on a GitHub repository[6]. To request data access contact the author of this report by e-mail.

[6]https://github.com/petrello/AnomalyDetectionFSAE

References

Ankur, A., and Johannes, T. 2024. pgmpy: A python toolkit for bayesian networks. *Journal of Machine Learning Research* 25(265):1–8.

Codetta-Raiteri, D., and Portinale, L. 2015. Dynamic bayesian networks for fault detection, identification, and recovery in autonomous spacecraft. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 45(1):13–24.

Concha, B., and Pedro, L. 2014. Discrete bayesian network classifiers: A survey. *ACM Computing Surveys* 47(1):1–43.

Rajapakse, J., and Zhou, J. 2007. Learning effective brain connectivity with dynamic bayesian networks. *NeuroImage* 37:749–760.