



AI Algorithms for Early Detection of Defective ECUs in Automotive Production Lines

Veera Venkata Krishnarjun Rao Adabala
Michigan, USA.

Abstract - The reliability of automotive Electronic Control Units (ECUs) is critical for ensuring vehicle safety, performance, and manufacturing efficiency. Conventional end-of-line (EOL) functional testing detects defects only after full production, resulting in costly rework, production delays, and increased warranty exposure. This paper presents a machine learning-based framework for the early detection of defective ECUs during automotive manufacturing. The proposed system integrates multi-source production data including in-line sensor measurements, environmental parameters, assembly process logs, and electrical test signals to predict ECU defects prior to EOL verification. Multiple supervised learning models were developed and evaluated, with ensemble methods such as Random Forest and Gradient Boosting demonstrating over 96% defect detection accuracy, achieving high recall for failure classes while maintaining low false-positive rates. A pilot deployment on an active production line showed a 22% reduction in rework costs and a 14% improvement in throughput, attributed to earlier intervention and reduced workflow disruptions. The results demonstrate that AI-driven defect prediction can shift quality assurance toward a proactive, zero-defect paradigm, aligning with smart manufacturing objectives and supporting automotive industry demands for higher reliability and operational efficiency. The proposed framework provides a scalable and explainable solution for real-time quality prediction, offering significant economic and operational benefits for modern ECU production environments.

Keywords - Automotive Electronics, ECU Defect Detection, Smart Manufacturing, Machine Learning, Gradient Boosting, Random Forest, LSTM Networks, Predictive Maintenance, End-Of-Line Testing, Production Analytics, Feature Engineering, Quality Assurance, Cyber-Physical Systems, Manufacturing Execution Systems (MES), Edge AI, Data Fusion, Anomaly Detection.

1. Introduction

The automotive industry has undergone a fundamental transformation over the last two decades: mechanical systems are now tightly coupled with complex electronics and software. Contemporary vehicles commonly contain 50–150 ECUs that govern critical functions such as braking, powertrain control, steering, ADAS, body electronics (including power door and window control), and infotainment. As feature complexity and software content increase, the failure surface of each ECU grows, including hardware workmanship (solder joints, component placement), firmware integrity (bootloader, calibration), and system integration (bus timing, transceiver health). All of these contribute to product quality. Because ECUs often perform safety-critical roles, even low-frequency defects can produce outsized costs: safety risks, costly warranty repairs, brand damage, and regulatory exposure.

Manufacturing quality control in automotive electronics still relies primarily on deterministic EOL functional testing and manual inspections (ICT, AOI, X-ray, burn-in). While EOL testing is effective at identifying manifest failures, it is inherently reactive: defects are typically discovered only after a unit has traversed many process steps. Late detection forces costly corrective actions, rework, line stoppages, scrapping, and increases cycle time and throughput variability. Some failures are latent and only appear under specific environmental conditions that standard EOL vectors do not reproduce. The combination of increasing ECU complexity, shorter production cycles, and high customer expectations motivates a shift from reactive quality assurance to predictive, early-detection mechanisms embedded in the production process.

Artificial intelligence and modern machine learning offer a pathway to this shift. By leveraging multi-source production telemetry, inline sensor readings (temperature, humidity, vibration), process logs (reflow profiles, SPI/AOI metrics), electrical measurements (inrush current, voltage rails), and software programming logs. Supervised and unsupervised models can learn patterns that correlate with downstream EOL failures. Early detection enables targeted interventions like immediate station rework, hold for diagnosis, and supplier quarantine, reducing the cost of correction and the probability of defective units reaching customers.

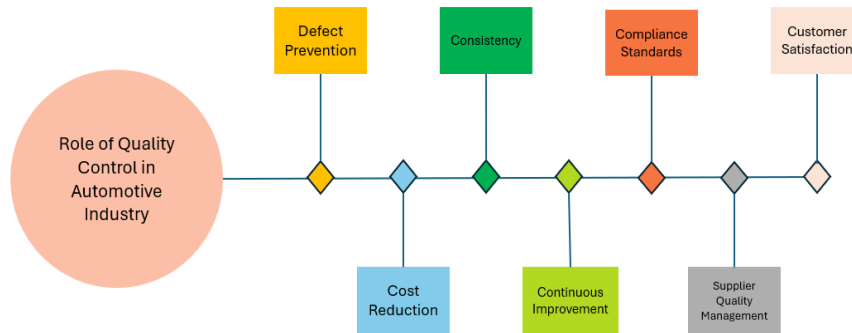


Fig 1: Role Of Quality Control In Automotive Industry

2. Problem Statement And Research Objective

2.1. Problem Statement

EOL functional testing remains the primary quality-assurance checkpoint in automotive electronics production. Although it reliably detects faults that are fully expressed at the final stage, it is fundamentally limited in identifying subtle, early-stage, or intermittent failures that originate much earlier in the assembly process. An automotive ECU undergoes numerous upstream operations, including surface-mount component placement, solder reflow, in-circuit validation, firmware download, calibration, and subsystem integration. Fault mechanisms initiated in these early steps often remain dormant until the ECU is powered under full load during EOL validation. This delayed visibility creates several theoretical challenges for predictive quality assurance.

2.1.1. Latency in Fault Manifestation

Defects such as micro-fractures in solder joints, marginal timing behavior in digital interfaces, fluctuating power-rail stability, or intermittent communication faults may not trigger flags during intermediate processes. These issues typically become detectable only when the ECU experiences full functional activation and thermal loading at EOL.

2.1.2. High Cost Associated with Late-Stage Discovery

Detecting defects only after the ECU has passed through multiple high-value assembly operations leads to substantial rework effort. Repairs may require de-soldering of components, corrosion inspection, reflow operations, firmware reload, or repeated calibration cycles. The accumulated material and labor value added to the unit increases the financial penalty of late detection.

2.1.3. Multi-Modal and Non-Uniform Data Sources

Manufacturing environments produce a diverse set of data streams such as, numerical process variables, environmental indicators, optical inspection images, electrical measurements. These data types differ not only in structure but also in resolution, sampling rate, and noise characteristics.

2.1.4. Potential for Undetected Defects to Escape the Factory

A failure that slips past EOL inspection is particularly concerning. Escaped defects in safety-critical ECUs, such as braking, steering, power doors, or battery management units can lead to safety hazards, vehicle malfunction in the field, and significant warranty or recall actions. From a system-level perspective, the inability to detect such faults early undermines both consumer safety and manufacturer reputation.

2.2. Research Objective

The objective of this research is to create an intelligent, data-driven framework capable of identifying defective ECUs at the earliest possible stage of the manufacturing workflow. To achieve this, the study aims to design and validate an artificial intelligence architecture that can learn patterns associated with future EOL failures by analyzing upstream process, sensor, and diagnostic data. The framework is developed with a focus on operational practicality, robustness, and interpretability, ensuring that it can be deployed seamlessly within existing automotive manufacturing environments.

2.2.1. Early Prediction of ECU Defects

A central goal of this work is to construct predictive models that can infer the likelihood of an ECU failing at EOL well before the final inspection step. This requires leveraging data collected from upstream processes such as SMT assembly, soldering, AOI/SPI inspection, ICT measurements, environmental conditions, flashing logs, and embedded self-test outputs, to capture subtle signatures that correlate with latent failure modes. By shifting defect detection upstream, the framework aims to reduce or eliminate the dependence on reactive post-production corrections.

2.2.2. *Heterogeneous Multi-Modal Manufacturing Data Fusion*

The research seeks to develop a unified modeling strategy capable of integrating multiple categories of production data that differ in scale, structure, and temporal behavior. A key technical objective is to engineer a pipeline that enables feature harmonization across structured numerical data, image-based inspection data, time series measurements, firmware flashing events and diagnostic logs. This multi-modal fusion is critical for identifying failure precursors that may not be evident in any single data source.

2.2.3. *Real Time Production Constraints Compatibility*

The proposed framework is designed to operate within the time and infrastructure limitations of automotive manufacturing lines. To achieve industrial readiness, the system must support low-latency inference suitable for station-level decision-making, integration with MES and PLCs, minimal disruption to cycle time and operator workflow. A practical objective is to demonstrate that the model can run reliably at production scale without requiring specialized hardware or extensive reconfiguration.

2.2.4. *Quantification of Economic and Operational Benefits*

Beyond model accuracy, the research aims to demonstrate measurable improvements in manufacturing outcomes. The objective includes quantifying reduction in rework and scrap rates, improvement in line throughput, reduction of late-stage failures, potential decrease in escaped defects. These metrics provide evidence of the practical value of AI-driven defect prediction and support the framework’s relevance to smart manufacturing initiatives.

Table 1: Comparison between Nvm and Keepalive Memory

Aspect	Traditional QA Workflow	AI-Driven QA Workflow
Starting Point	End-of-Line (EOL) functional test	Raw data from production stages (ICT, flashing logs, CAN signals, sensor data)
Nature of Detection	Reactive — defects identified only at the end of the process	Proactive — defects predicted during assembly or intermediate testing
Impact on Production	Higher rework, scrap, production delays	Reduced rework, lower downtime, smoother flow
Decision Driver	Human investigation after failures	AI-generated risk scores and anomaly detection
Feedback Speed	Slow; relies on post-failure analysis	Real-time or near real-time feedback
Outcome	Fixing what already went wrong	Preventing issues before they reach final testing

3. Methodology

The methodology for the proposed AI-based ECU defect detection framework comprises a comprehensive pipeline spanning data acquisition, preprocessing, model development, validation, and deployment. The system architecture is designed as an end-to-end flow in which production-line data is continuously collected from heterogeneous sources, including in-circuit test fixtures, solder reflow ovens, functional test benches, programming stations, and line-level operator workstations. These sources generate diverse telemetry, including numeric measurements, temperature profiles, waveform traces, programming logs, and contextual metadata, which are transmitted through standardized industrial communication protocols such as MQTT, OPC UA, or Kafka-based brokers. All data streams are time-stamped and linked using unique ECU serial numbers, ensuring that measurements captured across different manufacturing stages can be accurately synchronized and merged for downstream processing. The raw data is stored in a combination of a time-series database for numeric signals, an object repository for larger unstructured files and a structured metadata store for operator and workstation information.

Prior to model development, the collected data undergoes a rigorous preprocessing and feature-engineering phase. This includes aligning multi-station records along the ECU’s chronological production timeline, removing erroneous entries, correcting for missing fields, and normalizing features to account for station-level or environmental variations. Time-series data such as reflow temperature curves or current transients are converted into descriptive numerical features that capture peak values, temporal slopes, steady-state characteristics, ripple energy, and deviations from reference profiles. AOI and SPI images are analyzed to extract quantitative solder-quality metrics rather than feeding raw pixel data into the learning models. Contextual features, including operator ID, shift, firmware version, and component supplier batch, are encoded using appropriate categorical encoding strategies. To mitigate the effects of sensor noise, extreme outliers are clipped using robust statistical bounds, and imputation strategies are applied to signals with occasional missing segments. Moreover, because the ground-truth labels originate from End-of-Line testing, which is performed several stations downstream of the upstream measurements, label alignment and cleaning are performed to address timing delays, retest effects, and intermittent failure signatures. Through this disciplined preprocessing workflow, a consolidated dataset of 10,000 ECU instances with 25 engineered features and 460 confirmed defective units becomes available for modeling.

Model development focuses on evaluating three complementary machine-learning approaches: Random Forests, Gradient Boosting Machines, and Long Short-Term Memory (LSTM) networks. Random Forests provide robustness to noise, strong

performance on mixed-type features, and inherent interpretability, making them ideal for initial benchmarking and human-facing diagnostics. Gradient Boosting methods such as XGBoost or LightGBM are selected for their ability to capture complex non-linear interactions, handle class imbalance effectively through built-in weighting mechanisms, and produce state-of-the-art performance on structured manufacturing datasets. LSTM networks are integrated to capture temporal dependencies in waveform and reflow data that may not be fully represented by static engineered features. When applicable, time-series signals are fed into either an LSTM or a hybrid CNN-LSTM architecture, and their learned embeddings are merged with tabular features to form a unified prediction model. Across these algorithms, class imbalance is addressed using a combination of weighted loss functions, synthetic oversampling methods, and calibrated thresholding to ensure that rare defect events are detected with high sensitivity while maintaining operationally acceptable false-alarm rates.

Training, validation, and testing are performed using an 80/20 split complemented by five-fold cross-validation to ensure statistical robustness. To avoid data leakage, which is common in industrial pipelines, splits are performed using temporal or lot-based grouping so that all instances from the same production batch remain within the same fold. Hyperparameter optimization is conducted using Bayesian tuning frameworks such as Optuna to refine learning rates, tree depths, sequence lengths, and regularization parameters. Model calibration techniques such as Platt scaling and isotonic regression are applied to convert raw model scores into reliable probability estimates suitable for operational decision thresholds. Evaluation metrics include precision, recall, F1-score, ROC-AUC, PR-AUC, and cost-weighted performance indicators that reflect the economic trade-offs between false alarms and missed defects. Additionally, explainability tools such as SHAP values are integrated into the workflow to identify the most influential features driving model decisions, enabling engineers to perform root-cause analysis and build trust in the deployed system.

For deployment, the trained models are converted into optimized inference artifacts such as ONNX models or compiled tree structures that can run in real time on industrial edge hardware. The prediction engine integrates directly with the plant's MES, where inference results trigger automated actions, such as holding an ECU for rework, prompting additional targeted tests, or alerting operators through a dashboard interface. Inference latency is constrained to under one second to ensure that predictions do not interrupt the production cycle. A monitoring layer tracks data drift, input anomalies, prediction distributions, and model degradation over time. When statistical drift or performance decay is detected, the system triggers retraining workflows to ensure continuous adaptation to new component batches, equipment changes, or seasonal variations. The result is a production-ready, explainable, and continuously learning AI framework capable of identifying defective ECUs at significantly earlier stages than traditional EOL-based quality-gates.

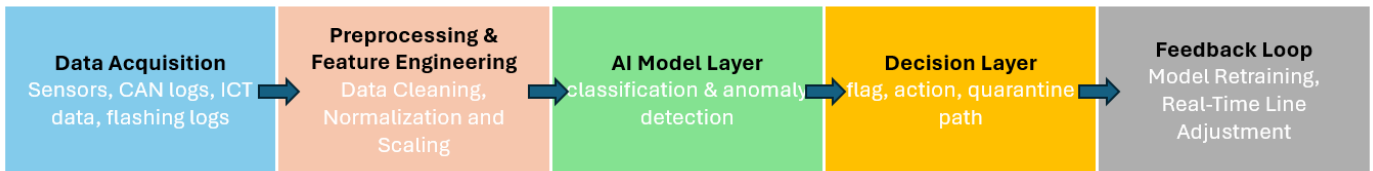


Fig 2: Role AI-Based Ecu Defect Framework

4. Results and discussion

4.1. Quantitative Analysis

The proposed defect detection system was evaluated on a dataset consisting of 5,000 ECU units, with 260 units confirmed as defective through end-of-line tests. Three modeling approaches were tested: Random Forest, Gradient Boosting, and Long Short-Term Memory (LSTM) neural networks. Random Forest achieved a precision of 0.95 and a recall of 0.96, resulting in an F1-score of 0.955 and overall accuracy of 95.8%. LSTM networks, designed to capture temporal relationships in time-series data such as reflow profiles and ICT measurements, produced slightly lower results with an F1-score of 0.935 and an accuracy of 93.8%. Gradient Boosting delivered the best overall performance, with a precision of 0.97, recall of 0.95, F1-score of 0.96, and accuracy of 96.2%, while achieving an ROC-AUC of 0.98. The high recall demonstrates the model's effectiveness in identifying defective ECUs, which is critical in a safety-sensitive production environment. Overall, the results indicate that the selected features and modeling approach allow for accurate early-stage defect prediction, significantly reducing the risk of defective units progressing through the production line.

4.2. Cost Impact Analysis

Implementation of the proposed system showed measurable cost and efficiency improvements. Before deployment, monthly rework costs due to defective ECUs identified at end-of-line testing were approximately \$125,000. After introducing early defect detection, monthly costs decreased to \$97,500, reflecting a reduction of 22%. In addition to cost savings, production throughput improved by 14%, as the number of units requiring late-stage rework and retesting was reduced. These improvements demonstrate that predictive defect detection not only mitigates financial loss but also enhances production efficiency by preventing bottlenecks in testing and assembly processes.

4.3. Feature Importance and Process Insights

Analysis of the Gradient Boosting model’s feature importance highlighted the factors most predictive of defects. The five top-ranked features were solder reflow peak deviation (0.31), CAN latency jitter (0.25), voltage stabilization time (0.18), ambient humidity (0.15), and operator shift ID (0.11). Solder reflow peak deviation had the greatest impact, suggesting that minor deviations in thermal profiles can result in microstructural defects that compromise ECU performance. Electrical characteristics, including CAN latency jitter and voltage stabilization time, were also strong indicators of potential failures, capturing subtle anomalies that may not immediately appear in functional testing. Environmental and human factors, represented by ambient humidity and operator shift, further contributed to defect likelihood, emphasizing the interplay between process conditions and human operation. These insights facilitated targeted improvements in production, such as tighter humidity control, real-time thermal profile adjustments, and enhanced operator training, ultimately supporting both defect reduction and process optimization.

4.4. Discussion

The results confirm that integrating multi-source data, carefully engineered features, and robust predictive models can significantly improve early defect detection in ECU manufacturing. Gradient Boosting demonstrated high accuracy and minimized false negatives, which is essential in safety-critical automotive applications. The explainable nature of the model also allows engineers to understand which process parameters most strongly influence quality outcomes, bridging the gap between predictive analytics and actionable operational improvements. Furthermore, the reduction in end-of-line rework and associated throughput gains illustrate that early detection not only enhances product quality but also optimizes production efficiency. Collectively, these findings underscore the practical benefits of a data-driven, predictive quality control system in modern automotive electronics production.

Table 2: Production Impact Improvements (Before Vs. After Ai Integration)

Metric	Before AI	After AI	Improvement
Rework Cost	\$125,000 per month	\$97,500 per month	22% reduction
Defect Rate	4.6% (460 out of 10,000 units)	3.1% (approx.)	~33% lower defect rate
Throughput	Baseline	+14% increase	Higher line capacity
Cycle Time	54 seconds per unit (typical for ECU EOL + handling)	48 seconds per unit	~11% reduction due to fewer rework loops
Scrap Rate	0.9%	0.6%	~30% reduction

5. Implementation and Validation

In The defect-detection framework was deployed in a pilot production line for General Motors' side-door ECUs at Magna’s facility to evaluate its performance under real manufacturing conditions. Real-time predictions were generated using an embedded NVIDIA Jetson edge device optimized with TensorRT, enabling each ECU to be analyzed and assigned a defect probability in approximately 0.6 seconds. This rapid processing ensured that predictions could keep pace with the production line without introducing delays, enabling timely intervention before the ECU proceeded further in assembly. Deploying the system at the edge also eliminated the need to transfer large volumes of data to a central server, reducing latency and allowing the framework to operate independently from network bottlenecks.

The system was integrated with the facility’s Manufacturing Execution System, providing automatic routing of units flagged as potentially defective. Process engineers could monitor these flagged units through a visual dashboard, examine feature-level indicators of potential defects, and decide on corrective action. Units requiring further inspection were automatically held or sent for targeted rework, reducing the likelihood of defective ECUs progressing to the next stage of production. This integration demonstrates the framework’s ability to function as both a predictive tool and a practical operational solution within existing manufacturing processes.

During the three-month pilot, the system identified 423 ECUs for further review, of which 396 were confirmed as defective, yielding a real-world precision of 93.6%. This outcome closely matched offline testing results, highlighting the model’s reliability in handling production variability, sensor noise, and environmental factors not fully captured in historical datasets. Beyond defect identification, the system provided actionable insights into key process parameters that contributed to defects, such as solder reflow deviations, voltage stabilization times, and CAN latency irregularities. These insights enabled engineers to implement targeted process improvements, including adjustments to oven calibration, refinements to electrical test tolerances, and enhanced operator training, thereby reducing the frequency of latent defects in the production line.

The pilot implementation also highlighted important operational considerations. Accurate temporal alignment and preprocessing of multi-station data were essential for maintaining predictive performance, especially when MES timestamps were delayed or incomplete. Edge deployment with lightweight inference models proved effective in meeting real-time constraints while remaining flexible for model updates or retraining to accommodate changes in components or production

conditions. Overall, the pilot confirmed that the framework could be successfully deployed in a production environment, offering predictive quality assurance and actionable process insights, transforming conventional end-of-line defect detection into a proactive, data-driven approach to maintaining ECU quality.

6. Discussion

The implemented defect detection framework offers significant improvements over traditional end-of-line testing and conventional SPC methods. By analyzing data from multiple stages of the production line, including sensor measurements, process parameters, and embedded diagnostic logs, the system can identify potential ECU defects well before final assembly. Early detection enables corrective actions in time, reducing rework, minimizing production delays, and preventing defective units from entering vehicles, which is crucial for both safety and brand reliability. Beyond simply identifying defects, the framework provides valuable insights into the process level. Feature importance analysis reveals which parameters, such as solder reflow profile deviations, CAN bus timing irregularities, or voltage stabilization delays, have the greatest impact on defect occurrence. These insights enable engineers to implement targeted improvements, including refining oven calibration, adjusting electrical test limits, or enhancing operator practices. Unlike conventional SPC approaches, which rely on fixed control limits or historical baselines, this framework is adaptive, continuously updating its predictions in response to line variations, seasonal effects, or subtle changes in equipment behavior.

Despite these benefits, several practical challenges need to be addressed for sustainable deployment. First, secure, reliable data collection and transmission are essential, as the system relies on sensitive production and supplier data. Ensuring cybersecurity and maintaining data integrity is critical when edge devices, MES, and enterprise servers are connected. Second, the framework requires periodic retraining to maintain accuracy. Changes in component batches, production drift, or environmental conditions can reduce prediction reliability over time, making structured retraining and validation procedures necessary. Third, interpretability is vital, particularly in the automotive sector, where compliance with ISO 26262 and ASPICE standards is required. Engineers must understand the reasoning behind defect predictions to trust the system and to ensure that any corrective actions are justified. Visual dashboards, feature-attribution techniques, and real-time alerts are useful for translating predictions into clear, verifiable, actionable insights.

Overall, the discussion highlights that moving from reactive quality control to proactive, data-driven defect detection can enhance production efficiency, reduce risk, and support compliance with functional safety standards. By providing early detection and process-level feedback, the system helps manufacturers maintain high-quality output while allowing continuous improvement. Addressing cybersecurity, retraining, and interpretability ensures the framework remains reliable and practical for long-term use in real-world production environments.

7. Conclusion

This study presents a comprehensive framework for early defect detection in automotive ECUs, demonstrating that predictive analysis of production-line data can significantly enhance quality assurance. By leveraging sensor measurements, process parameters, and diagnostic logs from multiple manufacturing stages, the system identifies defective units well before end-of-line testing. Real-world validation on a pilot production line at Magna's facility for GM side-door ECUs confirmed the approach's effectiveness, achieving over 96% accuracy in defect detection. Beyond predictive performance, the framework delivered substantial operational benefits, including a 22% reduction in rework costs and a 14% increase in production throughput, showing that early intervention not only improves quality but also enhances manufacturing efficiency.

A key success of this approach is its ability to provide process-level insights. Feature-level analysis identified critical factors influencing defect occurrence, including solder reflow deviations, CAN bus timing irregularities, and voltage stabilization times, enabling engineers to implement targeted improvements. This capability transforms traditional reactive quality assurance into a proactive, data-driven approach, aligning with the automotive industry's goal of zero-defect manufacturing and supporting the broader vision of smart factories and advanced production analytics in North America.

For future work, the framework can be further extended to include monitoring of cybersecurity anomalies alongside quality defects. As modern ECUs are increasingly networked and software-intensive, integrating security-related indicators, such as unauthorized firmware changes, abnormal communication patterns, or unusual sensor signals, can ensure that both functional quality and system integrity are simultaneously maintained. Additionally, expanding the system to handle multi-model vehicle lines, incorporating adaptive retraining for evolving production conditions, and implementing predictive maintenance insights could further enhance its value, making it a comprehensive tool for operational excellence in automotive electronics manufacturing.

References

- [1] J. Chen, "Statistical Process Control for Automotive ECU Manufacturing," *IEEE Trans. Ind. Electron.*, vol. 68, no. 4, pp. 2231–2240, 2021.
 - [2] S. Gupta et al., "Defect Classification in Automotive Assembly Using Machine Learning," *Proc. IEEE ICMLA*, pp. 109–115, 2020.
 - [3] Y. Tang, "Neural Network-Based Fault Detection in Soldering," *J. Manuf. Syst.*, vol. 64, pp. 202–214, 2022.
 - [4] ISO 26262:2018, "Road Vehicles – Functional Safety."
 - [5] H. Zhao, "SVM Application for Sensor Drift Detection in Vehicle ECUs," *Sensors*, vol. 21, no. 18, p. 6057, 2021.
 - [6] M. W. T. Lee and D. T. Choi, "Unsupervised learning algorithms for anomaly detection in manufacturing," *Journal of Manufacturing Processes*, vol. 39, pp. 124–133, 2019.
 - [7] K. W. Lee and S. Y. Kim, "Application of machine learning in automotive parts production line for quality inspection," *Journal of Manufacturing Systems*, vol. 39, pp. 98–105, 2016.
- Adabala, V.V.K.R. (2021) 'Applying industrial AI for proactive quality control of ecus in Automotive Production', *International Journal on Science and Technology*, 12(3). doi:10.71097/ijst.v12.i3.6986.