WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

An Explainable Deep Learning Framework for Predicting Production Line Downtime Risks in New Vehicle Launch Projects: A Case Study of the Iranian Automotive Industry

Fereshteh larti

M.Sc. in Industrial Management, Specialization in Project Management, University of Tehran, Faculty of Industrial Management and Technology,Tehran.Iran

Fereshteh.larti1@gmail.com

ABSTRACT :

New vehicle launch projects are continuously exposed to various risks, including production line downtime, due to their technical complexity, extensive dependence on supply chains, and uncertainties during the transition phase to mass production. Although traditional risk management methods are effective in identifying critical factors, they have limitations in analyzing the complex, nonlinear, and multidimensional relationships embedded in manufacturing data. Therefore, this study aims to propose a conceptual framework based on explainable deep learning for predicting production line downtime risks in new vehicle launch projects, with a specific focus on the conditions of the Iranian automotive industry.This research is a systematic literature review and conceptual framework development study. Initially, relevant studies related to automotive project risk management, smart manufacturing, machine learning, deep learning, and explainable artificial intelligence were reviewed and analyzed. Subsequently, based on the research gaps identified in the literature, a proposed architecture consisting of four main layers was developed: the data acquisition and integration layer, the data processing and feature extraction layer, the predictive deep learning model layer, and the explainable artificial intelligence layer.The review of previous studies indicates that deep learning models, particularly LSTM and Transformer architectures, have significant capabilities in analyzing time-series and multi-source manufacturing data and predicting critical events. However, the main limitation of these models is the lack of transparency in their decision-making processes. Incorporating explainable artificial intelligence methods, such as SHAP and LIME, can maintain predictive accuracy while revealing the factors contributing to increased production line downtime risks for industrial managers. By integrating equipment, quality, production, and supply chain data, the proposed framework enables the development of an early warning system for production risk management.The proposed framework can serve as a novel approach for transitioning from reactive risk management toward proactive risk management in the Iranian automotive industry. The application of explainable deep learning not only improves production downtime prediction capabilities but also enhances the reliability of intelligent systems and supports managerial decision-making in new product launch projects.

Keywords: Deep Learning; Explainable Artificial Intelligence; Production Line Downtime Prediction; Production Risk Management; Iranian Automotive Industry.

1. Introduction

The automotive industry is recognized as one of the most complex manufacturing sectors worldwide due to its extensive interdependencies among product design, supply chains, manufacturing processes, quality control, and project management. In new vehicle launch projects, the transition phase from product development to mass production, known as **New Product Introduction (NPI)**, is considered one of the most critical stages of the product lifecycle. Any disruption in component supply, equipment performance, assembly processes, or quality control activities can result in production line downtime, increased operational costs, and delays in product market introduction (Cooper, 2019). Therefore, the ability to predict production downtime risks at an early stage has become a fundamental requirement for smart manufacturing management.Traditional production risk management approaches, such as Failure Mode and Effects Analysis (FMEA), Fault Tree Analysis (FTA), and risk assessment matrices, primarily rely on expert knowledge and limited historical data. Although these methods remain widely used in manufacturing industries, they have limited capability in identifying nonlinear relationships, temporal dependencies, and hidden patterns within large-scale industrial datasets, particularly in complex and dynamic environments such as automotive production lines (Carvalho et al., 2019). The development of Industry 4.0 technologies and increased access to manufacturing data have created new opportunities for applying artificial intelligence and machine learning approaches to predict and mitigate operational risks.Deep learning, as an advanced branch of artificial intelligence, has the capability to automatically extract features from complex and multidimensional data. It has been widely applied in areas such as predictive maintenance, equipment failure diagnosis, and manufacturing process optimization (Ucar et al., 2024). Models such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Transformer-based architectures are capable of analyzing temporal data and identifying patterns preceding failures or production disruptions. However, one of the major challenges in adopting deep learning

models in industrial environments is the black-box nature of many of these models and the difficulty in interpreting their decisions by production managers and industrial experts (Arrieta et al., 2020).

In this context, Explainable Artificial Intelligence (XAI) has emerged as an approach to improve the transparency of intelligent models. Methods such as **SHAP** and **LIME** enable the identification of factors influencing model outputs and provide understandable explanations of algorithmic decisions, thereby increasing users' trust in artificial intelligence-based prediction systems (Molnar, 2022).In the automotive industry, production line downtime is considered one of the most costly operational events and may result from various factors, including equipment failures, component shortages, supplier quality fluctuations, process errors, and supply chain coordination problems. The complexity of these factors makes it impossible to predict production downtime based solely on individual indicators; instead, it requires simultaneous analysis of diverse production, quality, maintenance, and supply chain data.In Iran, the automotive industry faces specific challenges in production risk management due to factors such as partial dependence of the supply chain on limited sources, fluctuations in supply conditions, restricted access to certain technologies, and pressure to reduce the time required for launching new products. Many operational decisions are still based on managerial experience and post-event analyses, whereas artificial intelligence-based approaches can enable proactive risk prediction.Therefore, the main research problem addressed in this study is the lack of an integrated framework that can utilize the capabilities of deep learning to predict the probability of production line downtime while also providing interpretable explanations regarding the factors influencing these predictions for industrial decision-makers.The importance of developing such a framework in the Iranian automotive industry can be examined from several perspectives. First, reducing unplanned downtime can improve production line efficiency, decrease downtime-related costs, and enhance the performance of new product launch projects. Second, explainable artificial intelligence models can bridge the gap between advanced data-driven technologies and the operational knowledge of factory managers. Previous studies have demonstrated that AI-based predictive maintenance systems can improve equipment reliability and reduce operational costs; however, model transparency and explainability remain essential conditions for the widespread adoption of these technologies in industrial environments (Carvalho et al., 2019; Ucar et al., 2024). Therefore, developing a customized framework aligned with the characteristics of the Iranian automotive industry can represent an important step toward smart manufacturing transformation.The main objective of this article is to present a conceptual framework based on explainable deep learning for predicting production line downtime risks in new vehicle launch projects. This study is a review-based research article conducted to analyze the current state of knowledge and identify existing research gaps. The review process follows the **PRISMA framework**.

2. Theoretical Foundations and Literature Review

2.1. Risk Management in New Vehicle Launch Projects

New vehicle launch projects are considered among the most complex types of industrial projects because they require simultaneous coordination among product design activities, process engineering, supply chain development, production equipment preparation, quality control, and the transition to mass production. Unlike established production processes, new product introduction stages are associated with a higher level of uncertainty, exposing organizations to various risks, including component supply delays, production process instability, quality-related issues, and insufficient equipment readiness (Cooper, 2019). Therefore, effective risk management during this stage plays a critical role in achieving project objectives related to time, cost, and quality.

Project risk management generally involves the processes of identifying, analyzing, responding to, and monitoring risks that may affect the achievement of project objectives (Project Management Institute [PMI], 2021). In the automotive industry, this concept has become increasingly important due to the complex structure of supply chains and the strong interdependencies between suppliers and assembly lines. Modern vehicles typically consist of thousands of components manufactured by an extensive network of suppliers; consequently, disruption in one part of this network can create cascading effects, leading to production interruptions or reduced manufacturing capacity (Ivanov, 2021).

Traditionally, risk management in automotive projects has relied on tools such as **Failure Mode and Effects Analysis (FMEA)**, **Advanced Product Quality Planning (APQP)**, and **Production Part Approval Process (PPAP)**. These approaches were developed to identify potential problems before the beginning of mass production and remain important components of automotive quality management systems (AIAG, 2024). However, the main limitation of these methods is their dependence on expert-based qualitative assessments, limited historical data, and static analysis. As a result, they have limited capability in predicting nonlinear behaviors and dynamic interactions within complex manufacturing environments.

In recent years, risk management in the automotive industry has shifted from reactive approaches toward proactive and data-driven strategies. The emergence of Industry 4.0, Industrial Internet of Things (IIoT), smart manufacturing systems, and big data analytics has enabled organizations to collect and analyze real-time information related to equipment performance, product quality, and supply chain conditions (Frank et al., 2019). This transformation allows organizations to predict potential disruptions and implement corrective actions before significant losses occur, rather than responding after production failures.

One of the critical areas in managing risks associated with new vehicle launch projects is production line downtime risk. Production interruptions may result from interconnected factors, including machine failures, raw material shortages, component nonconformities, design changes, workforce-related issues, and logistics disruptions. Previous studies have demonstrated that analyzing these factors simultaneously using traditional approaches is challenging, whereas artificial intelligence-based methods can improve prediction capabilities and decision-making processes (Carvalho et al., 2019).

Accordingly, recent research has increasingly focused on applying machine learning and deep learning models for evaluating production and supply chain risks. Machine learning algorithms are capable of identifying hidden relationships among multiple variables and detecting patterns that precede disruptions (Jahin et al., 2024). For instance, analyzing historical downtime records, maintenance data, component quality information, and supplier performance indicators can enable the development of predictive models capable of estimating the probability of production line failures before they occur.

Despite the advantages of artificial intelligence applications, one of the major challenges associated with implementing these technologies in industrial environments is model interpretability. Many advanced deep learning models achieve high predictive accuracy; however, their decision-making processes remain unclear to human users. In automotive projects, decision-makers need to understand why a model predicts a high probability of production

downtime and which factors contribute most significantly to this risk. Therefore, integrating deep learning approaches with **Explainable Artificial Intelligence (XAI)** methods has emerged as an important research direction (Arrieta et al., 2020).

2.2. Production Line Downtime Prediction

Production line downtime is one of the most significant operational challenges in complex manufacturing industries, particularly in the automotive sector, because any interruption in production flow can increase operational costs, reduce equipment efficiency, delay product delivery, and disrupt new vehicle launch schedules. In modern manufacturing environments, downtime is not only caused by equipment failures but also results from complex interactions among technical, human, process-related, and supply chain factors. Therefore, contemporary production management approaches have moved toward developing predictive systems that enable organizations to identify the likelihood of disruptions before they occur and implement appropriate corrective actions (Zonta et al., 2020).Historically, production downtime management mainly relied on reactive maintenance strategies and scheduled preventive maintenance programs. In reactive maintenance, repair activities were performed after failures occurred, which often resulted in extended downtime periods and increased emergency maintenance costs. Preventive maintenance attempted to reduce failure probability through scheduled interventions; however, this approach was not always capable of accurately identifying actual equipment conditions or sudden changes in production processes (Lei et al., 2018). The development of Industry 4.0 technologies and real-time data acquisition capabilities has provided the foundation for predictive maintenance and intelligent production downtime prediction.

Production downtime prediction refers to the use of historical and real-time data to estimate the probability of future disruptions and identify the factors influencing their occurrence. This process is typically based on the analysis of various data sources, including equipment failure records, downtime duration, sensor measurements, product quality indicators, maintenance information, process conditions, and supply chain data (Carvalho et al., 2019). The primary objective of this approach is to transform production management from a reactive system into a proactive system capable of identifying operational risks before they develop into critical events.

In the automotive industry, predicting production line downtime is particularly important due to the continuous and highly coordinated nature of assembly processes. Automotive assembly lines consist of interconnected workstations, meaning that the failure of a critical workstation can affect the entire production flow. Furthermore, new vehicle launch projects involve higher levels of uncertainty because production processes have not yet achieved full stability, and design modifications, component supply problems, and quality fluctuations can increase the likelihood of downtime (Cooper, 2019). Therefore, effective downtime prediction requires models capable of analyzing complex relationships among multiple variables.

In recent years, machine learning and deep learning approaches have received considerable attention as effective tools for production downtime prediction. Machine learning algorithms, including decision trees, support vector machines, random forests, and XGBoost, are capable of extracting patterns from manufacturing data and have been applied in equipment failure diagnosis and system performance prediction (Zhao et al., 2019). However, when dealing with highly complex, multidimensional, and time-dependent industrial data, these methods may require extensive feature engineering.

Deep learning has played an important role in developing advanced production downtime prediction systems due to its ability to automatically learn features and process large-scale datasets. Recurrent neural network architectures, particularly **Long Short-Term Memory (LSTM)** and **Gated Recurrent Unit (GRU)** models, are highly suitable for analyzing time-series production data because they can capture long-term dependencies within operational patterns (Hochreiter & Schmidhuber, 1997). Furthermore, Transformer-based models using the **Attention Mechanism** have demonstrated strong capabilities in analyzing complex and multi-source datasets (Vaswani et al., 2017).

Despite the high accuracy of deep learning models, one of the main challenges in applying these technologies within industrial environments is the lack of transparency in their decision-making processes. In applications such as production downtime prediction, managers require not only warnings regarding potential failures but also explanations of the underlying causes of increased risk. For example, it is necessary to determine whether the predicted downtime is associated with declining equipment performance, reduced component quality, or changes in production processes. In this context, explainable artificial intelligence methods, including **SHAP**, **LIME**, and attention-based mechanisms, can provide insights into the contribution of each variable to model decisions (Molnar, 2022).

2. Theoretical Foundations and Literature Review

2.1. Risk Management in New Vehicle Launch Projects

New vehicle launch projects are considered among the most complex types of industrial projects because they require simultaneous coordination among product design activities, process engineering, supply chain development, production equipment preparation, quality control, and the transition to mass production. Unlike established production processes, new product introduction stages are associated with a higher level of uncertainty, exposing organizations to various risks, including component supply delays, production process instability, quality-related issues, and insufficient equipment readiness (Cooper, 2019). Therefore, effective risk management during this stage plays a critical role in achieving project objectives related to time, cost, and quality.Project risk management generally involves the processes of identifying, analyzing, responding to, and monitoring risks that may affect the achievement of project objectives (Project Management Institute [PMI], 2021). In the automotive industry, this concept has become increasingly important due to the complex structure of supply chains and the strong interdependencies between suppliers and assembly lines. Modern vehicles typically consist of thousands of components manufactured by an extensive network of suppliers; consequently, disruption in one part of this network can create cascading effects, leading to production interruptions or reduced manufacturing capacity (Ivanov, 2021).Traditionally, risk management in automotive projects has relied on tools such as Failure Mode and Effects Analysis (FMEA), Advanced Product Quality Planning (APQP), and Production Part Approval Process (PPAP). These approaches were developed to identify potential problems before the beginning of mass production and remain important components of automotive quality management systems (AIAG, 2024). However, the main limitation of these methods is their dependence on expert-based qualitative

assessments, limited historical data, and static analysis. As a result, they have limited capability in predicting nonlinear behaviors and dynamic interactions within complex manufacturing environments.

In recent years, risk management in the automotive industry has shifted from reactive approaches toward proactive and data-driven strategies. The emergence of Industry 4.0, Industrial Internet of Things (IIoT), smart manufacturing systems, and big data analytics has enabled organizations to collect and analyze real-time information related to equipment performance, product quality, and supply chain conditions (Frank et al., 2019). This transformation allows organizations to predict potential disruptions and implement corrective actions before significant losses occur, rather than responding after production failures.One of the critical areas in managing risks associated with new vehicle launch projects is production line downtime risk. Production interruptions may result from interconnected factors, including machine failures, raw material shortages, component nonconformities, design changes, workforce-related issues, and logistics disruptions. Previous studies have demonstrated that analyzing these factors simultaneously using traditional approaches is challenging, whereas artificial intelligence-based methods can improve prediction capabilities and decision-making processes (Carvalho et al., 2019).Accordingly, recent research has increasingly focused on applying machine learning and deep learning models for evaluating production and supply chain risks. Machine learning algorithms are capable of identifying hidden relationships among multiple variables and detecting patterns that precede disruptions (Jahin et al., 2024). For instance, analyzing historical downtime records, maintenance data, component quality information, and supplier performance indicators can enable the development of predictive models capable of estimating the probability of production line failures before they occur.Despite the advantages of artificial intelligence applications, one of the major challenges associated with implementing these technologies in industrial environments is model interpretability. Many advanced deep learning models achieve high predictive accuracy; however, their decision-making processes remain unclear to human users. In automotive projects, decision-makers need to understand why a model predicts a high probability of production downtime and which factors contribute most significantly to this risk. Therefore, integrating deep learning approaches with **Explainable Artificial Intelligence (XAI)** methods has emerged as an important research direction (Arrieta et al., 2020).

2.2. Production Line Downtime Prediction

Production line downtime is one of the most significant operational challenges in complex manufacturing industries, particularly in the automotive sector, because any interruption in production flow can increase operational costs, reduce equipment efficiency, delay product delivery, and disrupt new vehicle launch schedules. In modern manufacturing environments, downtime is not only caused by equipment failures but also results from complex interactions among technical, human, process-related, and supply chain factors. Therefore, contemporary production management approaches have moved toward developing predictive systems that enable organizations to identify the likelihood of disruptions before they occur and implement appropriate corrective actions (Zonta et al., 2020).Historically, production downtime management mainly relied on reactive maintenance strategies and scheduled preventive maintenance programs. In reactive maintenance, repair activities were performed after failures occurred, which often resulted in extended downtime periods and increased emergency maintenance costs. Preventive maintenance attempted to reduce failure probability through scheduled interventions; however, this approach was not always capable of accurately identifying actual equipment conditions or sudden changes in production processes (Lei et al., 2018). The development of Industry 4.0 technologies and real-time data acquisition capabilities has provided the foundation for predictive maintenance and intelligent production downtime prediction.Production downtime prediction refers to the use of historical and real-time data to estimate the probability of future disruptions and identify the factors influencing their occurrence. This process is typically based on the analysis of various data sources, including equipment failure records, downtime duration, sensor measurements, product quality indicators, maintenance information, process conditions, and supply chain data (Carvalho et al., 2019). The primary objective of this approach is to transform production management from a reactive system into a proactive system capable of identifying operational risks before they develop into critical events.In the automotive industry, predicting production line downtime is particularly important due to the continuous and highly coordinated nature of assembly processes. Automotive assembly lines consist of interconnected workstations, meaning that the failure of a critical workstation can affect the entire production flow. Furthermore, new vehicle launch projects involve higher levels of uncertainty because production processes have not yet achieved full stability, and design modifications, component supply problems, and quality fluctuations can increase the likelihood of downtime (Cooper, 2019). Therefore, effective downtime prediction requires models capable of analyzing complex relationships among multiple variables.

In recent years, machine learning and deep learning approaches have received considerable attention as effective tools for production downtime prediction. Machine learning algorithms, including decision trees, support vector machines, random forests, and XGBoost, are capable of extracting patterns from manufacturing data and have been applied in equipment failure diagnosis and system performance prediction (Zhao et al., 2019). However, when dealing with highly complex, multidimensional, and time-dependent industrial data, these methods may require extensive feature engineering.Deep learning has played an important role in developing advanced production downtime prediction systems due to its ability to automatically learn features and process large-scale datasets. Recurrent neural network architectures, particularly Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models, are highly suitable for analyzing time-series production data because they can capture long-term dependencies within operational patterns (Hochreiter & Schmidhuber, 1997). Furthermore, Transformer-based models using the Attention Mechanism have demonstrated strong capabilities in analyzing complex and multi-source datasets (Vaswani et al., 2017).Despite the high accuracy of deep learning models, one of the main challenges in applying these technologies within industrial environments is the lack of transparency in their decision-making processes. In applications such as production downtime prediction, managers require not only warnings regarding potential failures but also explanations of the underlying causes of increased risk. For example, it is necessary to determine whether the predicted downtime is associated with declining equipment performance, reduced component quality, or changes in production processes. In this context, explainable artificial intelligence methods, including SHAP, LIME, and attention-based mechanisms, can provide insights into the contribution of each variable to model decisions (Molnar, 2022).

Explainable Artificial Intelligence

Explainable Artificial Intelligence (XAI) refers to a set of methods and techniques developed to enhance the transparency, trustworthiness, and

interpretability of artificial intelligence models. The primary objective of XAI is to present the decision-making processes of complex AI models in a manner that is understandable to human users (Arrieta et al., 2020).

This issue is particularly important in industrial applications, where AI-based decisions may directly influence production continuity or shutdown, maintenance scheduling, and resource allocation. Explainability methods are generally categorized into two main groups: **intrinsic methods** and **post-hoc methods**. In intrinsic methods, the model is designed from the outset to be interpretable, as is the case with certain decision-tree models. In contrast, post-hoc methods seek to provide explanations for the outputs of complex models, such as deep neural networks, after they have been trained (Molnar, 2022).

One of the most widely used XAI techniques is the **SHAP (SHapley Additive exPlanations)** algorithm, which is based on game theory and determines the contribution of each input variable to the model's output. In manufacturing environments, SHAP can identify which factors—such as increased equipment temperature, reduced component quality, or supplier delays—have the greatest influence on the probability of production line stoppages (Lundberg & Lee, 2017).

Another popular approach is **LIME (Local Interpretable Model-agnostic Explanations)**, which explains the decision of a complex model for a specific instance by constructing simple local surrogate models (Ribeiro et al., 2016).

The integration of deep learning and explainable artificial intelligence can provide a powerful framework for production risk management. Within such a framework, the deep learning model is responsible for predicting the probability of production line stoppages, while the XAI module identifies the factors contributing to these predictions. In addition to improving predictive accuracy, this combination enhances industrial experts' trust in intelligent systems and facilitates managerial decision-making.

Previous Studies

One of the primary research areas has been the application of machine learning algorithms for predictive maintenance and equipment failure detection. Carvalho et al. (2019), in a systematic review, examined the use of machine learning methods in predictive maintenance and demonstrated that algorithms such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), Random Forests, and other data-driven approaches possess considerable capability in predicting industrial equipment failures. However, the study emphasized that most previous research focused on predicting failures of individual pieces of equipment rather than analyzing multifactorial risks such as complete production line stoppages.

In the field of deep learning, Zhao et al. (2019) investigated the application of deep neural networks for machinery health monitoring. Their findings showed that deep learning models outperform many traditional methods in fault diagnosis and equipment condition prediction due to their ability to extract complex features from sensor data. Models such as **Convolutional Neural Networks (CNNs)**, **Long Short-Term Memory (LSTM)** networks, and **Autoencoders** have been widely applied in vibration signal analysis, time-series data processing, and anomaly detection in manufacturing systems.

Building on this research trajectory, more recent studies have explored various deep learning architectures in manufacturing environments. The review conducted by Khelifati et al. (2024) demonstrated that deep learning-based models, particularly recurrent neural networks and hybrid architectures, play a significant role in the development of next-generation predictive maintenance systems. The study emphasized that the selection of an appropriate deep learning architecture should depend on the type of data, the structure of the production process, and the prediction objective.

Other studies have examined the broader application of artificial intelligence in smart manufacturing. Wuest et al. (2016), in their review of machine learning applications in manufacturing, showed that intelligent algorithms can be utilized in areas such as quality control, machinery performance prediction, process optimization, and production decision management. The study argued that the primary value of machine learning in manufacturing lies in its ability to transform operational data into actionable knowledge that improves decision-making.

With the advancement of Industrial Internet of Things (**IIoT**)–based manufacturing systems, more recent research has focused on real-time data analytics and model interpretability. Presciuttini et al. (2024), in a systematic review of machine learning applications using IoT-generated manufacturing data, found that models such as Random Forest and LSTM are among the most frequently employed techniques in industrial data analysis. Nevertheless, the study identified model interpretability as one of the major research gaps in smart manufacturing, as many highly accurate models provide limited explainability for industrial users.

In the field of Explainable Artificial Intelligence, recent research has increasingly focused on integrating powerful predictive models with explanation techniques. Arrieta et al. (2020) argued that the widespread adoption of AI systems in sensitive industrial domains requires a balance between predictive accuracy and model explainability. Techniques such as SHAP and LIME enable the quantification of the contribution of each input variable to the model output and can enhance experts' trust in intelligent systems.

Regarding the direct application of XAI in maintenance and manufacturing, Dereci and Tuzkaya (2024) proposed an explainable AI model for predictive maintenance and spare-parts management. Their findings indicated that although machine learning models can provide highly accurate predictions, explaining the factors influencing model decisions is essential for improving maintenance-related decision-making. Furthermore, recent studies on explainable predictive maintenance suggest the need for approaches that not only predict failures but also consider the type and level of explanation required by different industrial stakeholders (Pashami et al., 2023).

Despite these advancements, a review of previous studies reveals several important gaps. First, a substantial portion of existing research focuses on failures of individual equipment rather than providing a comprehensive framework for predicting the risk of complete production line stoppages. Second, many deep learning models have been developed primarily to maximize predictive accuracy, while limited attention has been given to explaining their decisions. Third, relatively few studies have incorporated the specific conditions associated with new vehicle launch projects, including high uncertainty, process changes, and supply chain risks, into predictive models.

Proposed Framework Architecture

Given the increasing complexity of new vehicle launch projects and the multitude of factors influencing production line stoppages, a predictive framework must be capable of collecting multi-source manufacturing data, extracting hidden patterns, predicting the probability of disruptions, and ultimately explaining the causes of risk to industrial decision-makers. Accordingly, the proposed framework is designed based on four main layers: **the Data and Information Integration Layer, the Data Processing and Feature Engineering Layer, the Predictive Deep Learning Model Layer, and the Explainable AI and Managerial Decision Support Layer.**

The first component of the proposed architecture is the **Data Layer**, which is responsible for collecting and integrating information related to manufacturing processes. In new vehicle launch projects, the risk of production line stoppages is rarely caused by a single factor; rather, it results from the interaction of multiple internal and external variables. Therefore, the proposed framework should enable the simultaneous utilization of diverse data sources.

The data sources incorporated within this layer include:

- Production equipment data
- Production quality data
- Supply chain data
- Production planning data
- Maintenance and repair data

At this stage, raw data from various sources are extracted and subsequently cleaned, standardized, and synchronized in time before being prepared for further processing.

Industrial datasets often contain noise, missing values, class imbalance, and differences in measurement scales. Consequently, data preprocessing is essential before feeding the information into the deep learning model.

The primary activities within this layer include:

- Removal of outliers and data-entry errors
- Missing data imputation
- Normalization of numerical variables
- Transformation of qualitative variables into machine-processable formats
- Generation of new features from time-series data

At this stage, indicators such as trends in production stoppages, quality degradation rates, equipment performance decline, and supplier risk levels can be extracted. These features enable the model to better identify hidden relationships among the various factors influencing production line disruptions.

The core component of the proposed framework is a **deep learning model** responsible for predicting the probability of production line stoppages. Given the temporal nature of manufacturing data, a hybrid architecture based on **LSTM and Transformer networks** is proposed.

Within this architecture, the **LSTM component** is employed to analyze temporal behavior and identify gradual trends. For example, a progressive increase in machine vibration, a decline in output quality, or longer maintenance durations may signal a higher likelihood of production interruptions.

The **Transformer component** is utilized to identify complex relationships among multiple variables. This module can analyze dependencies between factors such as equipment conditions, component quality, and supply conditions.

The model outputs may include:

- Probability of production line stoppage within a future time horizon
- Risk level (low, medium, high)
- Estimated time of disruption occurrence
- Prioritization of high-risk production lines or equipment

One of the key distinctions between the proposed framework and conventional predictive models is the incorporation of an **Explainability Layer**. The purpose of this layer is to transform the outputs of complex deep learning models into information that is understandable for production managers and industrial experts.

Within this layer, methods such as **SHAP** are used to quantify the contribution of each factor to the model's prediction. For example, the system may determine that:

- Reduced quality of supplied components is the primary risk factor;
- Increased equipment repair time is the second most influential factor;
- Changes in the production schedule constitute the third factor.

LIME is employed to provide local explanations for specific predictions. For instance, if the model predicts a high probability of disruption for a particular production line, LIME can identify the specific factors responsible for that prediction.

This capability enables decision-makers not only to receive risk warnings but also to understand the underlying causes of those risks and select appropriate corrective actions.

The final component of the proposed architecture is the **Decision Support Module**, which transforms the model's predictions and explanatory outputs into actionable managerial recommendations.

The outputs of this module may include:

- Early warning notifications for production stoppages
- Recommendations for prioritizing corrective actions
- Identification of critical risk factors
- Suggested maintenance plans or alternative sourcing strategies
- Project risk reports for senior management

This layer enables the system to evolve from a purely predictive tool into a comprehensive **decision support system**.

Application of the Framework in the Iranian Automotive Industry

The Iranian automotive industry, as one of the largest manufacturing sectors in the country, possesses a complex structure comprising an extensive supplier network, multi-stage assembly lines, quality control processes, and numerous product development projects. In such an environment, new vehicle launch projects face a high level of uncertainty because the transition from the design and prototyping phase to mass production requires precise coordination among engineering, production, quality, procurement, suppliers, and logistics departments. Any disruption within this chain can lead to increased ramp-up time, reduced production capacity, and assembly-line stoppages. Therefore, the use of an intelligent framework for predicting and managing production stoppage risks can play a significant role in enhancing the stability of manufacturing processes.

The framework proposed in this study has been designed to align with the characteristics of the Iranian automotive industry and can be utilized as a risk prediction system for new product launch projects. Within this framework, operational data from various sections of the factory and supply chain are collected and analyzed by a deep learning model after preprocessing to predict the probability of production line stoppages. The primary distinction between this approach and traditional risk management methods lies in its focus on the simultaneous analysis of multiple factors and its ability to provide warnings before disruptions occur.

Implementing the proposed framework requires the establishment of an integrated data infrastructure capable of consolidating information from different organizational units. In automotive manufacturing plants, the required data may be obtained from the following sources:

a) Production and Assembly Data

This includes daily production rates, operation cycle times, workstation downtime, actual line capacity, deviations from production schedules, and operator performance information. These data enable the identification of abnormal trends in production line performance.

b) Equipment and Maintenance Data

Information related to machinery failures, Mean Time Between Failures (MTBF), Mean Time To Repair (MTTR), maintenance records, and equipment sensor data can be utilized to predict stoppages caused by equipment failures.

c) Product Quality Data

In new vehicle launch projects, quality issues are highly likely during the early stages of production. Information regarding defect rates, rework, returned components, process nonconformities, and inspection results can serve as predictive risk indicators.

d) Supply Chain Data

The automotive industry's dependence on a broad supplier network makes component supply issues a significant cause of production interruptions. Information regarding delivery delays, variations in component quality, supplier capacity, and inventory levels can be incorporated into risk prediction models.

For example, during the introduction of a new vehicle model, the assembly line may encounter multiple challenges during the initial pilot-production phase. Suppose that system data indicate the following within a specific period:

- Lead time for a critical component has increased.
- The rate of component nonconformities has risen.
- The number of stoppages involving a particular assembly machine is increasing.
- Emergency maintenance durations have become longer.

In traditional approaches, each of these issues might be analyzed separately, and corrective actions would often be taken only after a production stoppage occurs. In contrast, the proposed framework enables the deep learning model to analyze these variables simultaneously and predict an increased probability of production line stoppage. Subsequently, the Explainable Artificial Intelligence module identifies the most influential risk factors and determines which variables contribute most significantly to the model's prediction.

For instance, the system may generate the following report:

- Probability of production line stoppage during the upcoming week: **High**
- Primary risk factor: Reduced performance of supplier **X**
- Secondary risk factor: Increased failure rate of assembly equipment **Y**
- Recommended action: Increase safety stock of the component and perform preventive maintenance on the equipment

Such outputs transform managerial decision-making from a reactive approach to a predictive and proactive one.

The implementation of the proposed framework can provide several benefits for Iranian automotive manufacturers:

- Reduction of unplanned production stoppages through early identification of risk factors and implementation of corrective actions before disruptions occur.
- Improved management of new vehicle launch projects through continuous monitoring of risk indicators, enabling project managers to identify production ramp-up issues more rapidly.
- Enhanced transparency in managerial decisions through the application of XAI methods, allowing AI-generated decisions to be interpreted and understood by experts.
- Strengthened managerial decision-making through data-driven analysis while still leveraging the experience and expertise of production specialists.

Despite its potential advantages, implementing the framework also presents challenges. The most significant challenge is the quality and integration of industrial data. Many manufacturing organizations operate with fragmented information systems, and data are often stored in different formats. Furthermore, establishing a culture of data-driven decision-making, training personnel, and integrating existing production systems with intelligent infrastructures are critical requirements for the successful adoption of this approach.

Case Study Scenario

To evaluate the practical applicability of the proposed framework within the Iranian automotive industry, this section presents a case study scenario focused on predicting production line stoppage risk during the launch phase of a new vehicle. The scenario is based on the general characteristics of new product development and launch projects in the automotive sector and aims to demonstrate how an explainable deep learning framework can identify early warning signs of production disruptions.

Assume that an Iranian automotive manufacturer is transitioning a new vehicle from pilot production to mass production. During this stage, assembly-line capacity is gradually increased while process-related issues, component quality problems, equipment performance challenges, and supplier coordination difficulties may lead to unexpected production interruptions.

Under conventional circumstances, production and project teams rely on daily production reports, problem-solving meetings, and expert analyses to identify issues. However, many risks exhibit warning signs within operational data before evolving into complete production stoppages, and these signals may remain undetected through traditional analyses. The objective of the proposed framework is to identify these indicators and provide early warnings.

In this scenario, the framework collects data from four primary domains:

- Assembly-line production data
- Equipment and maintenance data
- Product quality data
- Supply chain data

During the first stage, data collected from production, maintenance, quality, and supplier systems are stored in an integrated database. Data cleaning procedures, including the removal of unrealistic values, imputation of missing data, and variable standardization, are then performed.

During the second stage, risk-related features are extracted, including:

- Gradual increases in assembly-station stoppages
- Rising rework rates
- Declining supplier performance
- Increasing equipment repair times

These features are then fed into the deep learning model.

In the third stage, a hybrid deep learning architecture such as an **LSTM–Transformer model** is trained using historical data to learn patterns associated with production stoppages. The model output represents the probability of production line stoppage within a specified time horizon.

For example, the model may generate the following prediction:

Indicator	**Predicted Value**
Probability of line stoppage within the next 7 days	82%
Risk level	High
Critical station	Powertrain Assembly

After the risk prediction is generated, the Explainable AI module determines the underlying causes of the elevated risk. Suppose the model predicts a high probability of assembly-line stoppage. A SHAP analysis may produce the following results:

Influencing Factor	**Contribution to Risk Increase**
Delay in critical component supply	35%
Increased assembly equipment failures	28%
Higher component nonconformity rate	22%
Reduced safety stock inventory	15%

Based on these results, production managers can implement targeted corrective actions such as:

- Increasing inventory levels of critical components
- Performing preventive maintenance on the affected equipment
- Conducting supplier quality audits
- Adjusting production schedules

The implementation of the proposed framework in this case study can yield several benefits for the new vehicle launch project:

1. **Reduced response time to risks:** The organization can utilize predictive alerts instead of waiting for disruptions to occur.
2. **Improved managerial decision-making:** Managers gain visibility into both the probability of production stoppages and the underlying causes of risk.
3. **Enhanced production stability during product launch:** Initial production issues can be detected and resolved more rapidly.
4. **Strengthened organizational knowledge management:** Historical data on production interruptions and corrective actions can be transformed into a valuable knowledge base for future projects.

The case study demonstrates that the explainable deep learning framework can serve as an effective decision-support tool for risk management in new vehicle launch projects. By integrating production, quality, equipment, and supply-chain data, the framework can predict production line stoppages and reveal the key risk factors through explainability techniques. Such an approach can play a significant role in advancing the Iranian automotive industry toward intelligent, predictive, and data-driven manufacturing.

Discussion and Conclusion

Digital transformation in manufacturing industries, particularly in the automotive sector, has increased the necessity of adopting intelligent approaches for managing operational risks. New vehicle launch projects are inherently exposed to production line stoppage risks due to the complexity of product development processes, extensive supply chain dependencies, frequent process changes, and uncertainties associated with production ramp-up stages. The literature reviewed in this study indicates that although traditional risk management methods remain useful for the initial identification and assessment of risks, they have limited capability in predicting nonlinear relationships among multiple factors within dynamic and complex manufacturing environments. Consequently, the transition toward predictive and data-driven models has emerged as a necessity for enhancing production resilience.

The findings of this study suggest that deep learning can play an effective role in predicting production line stoppage risks. Unlike classical methods, which typically rely on manual feature selection and predefined relationships among variables, deep learning models can extract complex patterns from production, quality, maintenance, and supply-chain data. This capability is particularly valuable in new vehicle launch projects, where early warning

signs of risk may be hidden within large volumes of distributed operational data and may not be easily identifiable through conventional analytical techniques.

The proposed framework is built upon the integration of three key concepts: **Smart Manufacturing, Deep Learning, and Explainable Artificial Intelligence (XAI).** The deep learning component is responsible for analyzing historical and real-time data and predicting the likelihood of production line stoppages, while the explainability layer reveals the rationale behind the model's decisions to industrial users. This feature is especially important in manufacturing environments because merely providing a numerical prediction, such as a "high probability of production stoppage," is insufficient for managers. Decision-makers must also understand which factors contributed to the increased risk and what corrective actions should be taken.

In the Iranian automotive industry, the importance of such a framework is amplified by the unique characteristics of its production and supply networks. Dependence on multiple suppliers, fluctuations in component availability, quality-related challenges during the early stages of new product production, and the need to reduce time-to-market have made risk management a critical success factor in automotive projects. The proposed framework can establish connections among production, equipment, quality, and supply-chain data, providing a comprehensive perspective on project risks and enabling managers to implement preventive actions before disruptions occur.

One of the most significant conceptual contributions of this study is its emphasis on shifting from a **"failure prediction"** perspective toward a **"production stoppage risk prediction"** perspective. Many previous studies have focused primarily on predicting failures of individual pieces of equipment, whereas production line stoppages in automotive manufacturing are typically the result of interactions among multiple simultaneous factors. For example, a delay in component delivery combined with declining component quality and increasing equipment repair times can substantially increase the likelihood of production interruptions. Future predictive models should therefore be capable of analyzing such multidimensional relationships at the level of the entire production system.

Despite the potential advantages of the proposed framework, its implementation within industrial environments faces several challenges. The most important challenge is the availability of high-quality, integrated, and reliable data. Artificial intelligence models can only achieve optimal performance when input data are accurate, continuous, and standardized. In addition, the development of digital infrastructures, integration of existing manufacturing systems, workforce training, and the establishment of a data-driven decision-making culture are essential prerequisites for the successful deployment of such systems.

From a research perspective, this study provides a foundation for the development of more advanced predictive frameworks. Future research may integrate the proposed architecture with technologies such as **Digital Twins**, **Federated Learning**, and **Agent-Based Intelligent Systems** to enable large-scale risk prediction and management. Furthermore, empirical validation of the framework using real-world automotive production data and comparative analyses of different deep learning architectures could constitute valuable directions for future investigations.

The results of this study indicate that the integration of deep learning and Explainable Artificial Intelligence offers an effective approach for the proactive management of production line stoppage risks in new vehicle launch projects. By providing predictive capabilities, identifying key influencing factors, and supporting managerial decision-making, the proposed framework can serve as a strategic tool for guiding the Iranian automotive industry toward intelligent, flexible, and data-driven manufacturing.

Recommendations for Future Research

Based on the findings of this study, the following directions are recommended for future research:

1. Development and empirical validation of the proposed framework using real-world data from Iranian automotive manufacturing plants.
2. Comparative evaluation of different deep learning architectures, including LSTM, Transformer, and hybrid models.
3. Development of multi-level predictive models capable of simultaneously analyzing equipment, quality, and supply-chain risks.
4. Application of Digital Twin technology for simulating production line stoppage scenarios.
5. Development of intelligent decision-support systems that not only predict risks but also recommend optimal corrective actions.

REFERENCES

- AIAG. (2024). *Advanced Product Quality Planning (APQP) and Control Plan Reference Manual*. Automotive Industry Action Group.
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., et al. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012
- Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., et al. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering, 137*, 106024. https://doi.org/10.1016/j.cie.2019.106024
- Cooper, R. G. (2019). The drivers of success in new-product development. *Industrial Marketing Management, 76*, 36–47. https://doi.org/10.1016/j.indmarman.2018.07.005
- Dereci, U., & Tuzkaya, G. (2024). An explainable artificial intelligence model for predictive maintenance and spare parts optimization. *Supply Chain Analytics, 8*, 100078. https://doi.org/10.1016/j.sca.2024.100078
- Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. *International Journal of Production Economics, 210*, 15–26. https://doi.org/10.1016/j.ijpe.2019.01.004
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
- Ivanov, D. (2021). Supply chain viability and the COVID-19 pandemic: A conceptual and formal generalisation of four major adaptation

strategies. *International Journal of Production Research, 59*(12), 3535–3552. https://doi.org/10.1080/00207543.2021.1890852

- Jahin, M. A., Naife, S. A., Saha, A. K., & Mridha, M. F. (2024). AI in supply chain risk assessment: A systematic literature review and bibliometric analysis. *Artificial Intelligence Review*. https://doi.org/10.1007/s10462-024-10700-9
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415
- Khelifati, A., et al. (2024). A survey of deep learning-driven architecture for predictive maintenance. *Engineering Applications of Artificial Intelligence, 133*, 108285. https://doi.org/10.1016/j.engappai.2024.108285
- Kusiak, A. (2018). Smart manufacturing. *International Journal of Production Research, 56*(1–2), 508–517. https://doi.org/10.1080/00207543.2017.1351644
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436–444. https://doi.org/10.1038/nature14539
- Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., & Nandi, A. K. (2018). Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing, 138*, 106587. https://doi.org/10.1016/j.ymssp.2018.12.026
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems, 30*.
- Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.).
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71
- Pashami, S., Nowaczyk, S., Fan, Y., et al. (2023). Explainable predictive maintenance. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*.
- Presciuttini, A., Cantini, A., Costa, F., & Portioli-Staudacher, A. (2024). Machine learning applications on IoT data in manufacturing operations and their interpretability implications: A systematic literature review. *Journal of Manufacturing Systems, 74*, 477–486. https://doi.org/10.1016/j.jmsy.2024.04.012
- Project Management Institute. (2021). *A Guide to the Project Management Body of Knowledge (PMBOK Guide) – Seventh Edition*. Project Management Institute.
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778
- Tao, F., Zhang, M., Liu, Y., & Nee, A. Y. C. (2018). Digital twin driven prognostics and health management for complex equipment. *CIRP Annals, 67*(1), 169–172. https://doi.org/10.1016/j.cirp.2018.04.055
- Ucar, A., Karakose, M., & Kırımça, N. (2024). Artificial Intelligence for Predictive Maintenance Applications: A review. *Applied Sciences, 14*(2), 898. https://doi.org/10.3390/app14020898
- Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*.
- Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. *Journal of Manufacturing Systems, 48*, 144–156. https://doi.org/10.1016/j.jmsy.2018.01.003
- Wuest, T., Weimer, D., Irgens, C., & Thoben, K. D. (2016). Machine learning in manufacturing: Advantages, challenges, and applications. *Production & Manufacturing Research, 4*(1), 23–45. https://doi.org/10.1080/21693277.2016.1192517
- 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
- https://doi.org/10.1016/j.ymssp.2018.05.050
- Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. (2020). Predictive maintenance in the Industry 4.0: A systematic literature review. *Computers & Industrial Engineering, 150*, 106889. https://doi.org/10.1016/j.cie.2020.106889