

## Predicting the Lead Time of Auto Parts Orders in the Supply Chain Using Machine Learning

**Faezeh Zamani** 

Ph.D. Student of Industrial Management Group,  
Department of Economics and Management, Science  
and Research Branch, Islamic Azad University,  
Tehran, Iran

**Ahmad Ebrahimi** \*

Assistant Professor of Industrial Management &  
Technology Group, Department of Economics and  
Management, Science and Research Branch, Islamic  
Azad University, Tehran, Iran

**Roya Soltani** 

Assistant Professor of Industrial Engineering Group,  
Department of Engineering, Khatam University,  
Tehran, Iran

**Babak Farhang  
Moghaddam** 

Associate Professor, Institute Management and  
Planning Studies, Tehran, Iran

### Abstract

This research aims to investigate the effective factors in predicting lead time (LT) and create a predictive model of LT to improve sustainability and resilience for Kanban orders in the lean supply chain (LSC). The study follows the data mining (DM) method, and the dataset includes 103023 observations from the Kanban system, which were extracted in compliance with the requirements of the dataset quality indicators in the period 1402/6 to 1402/11. First, indicators affecting the LT of orders were extracted. Process mining was used to identify influential variables in high-variance processes to improve performance and accuracy. A stepwise analysis approach was used to select features for the model fitting stage. Also, tuning the parameters of non-parametric approaches

\* Corresponding Author: Ahmad.Ebrahimi@srbiau.ac.ir

**How to Cite:** Zamani, F., Ebrahimi, A., Soltani, R., Farhang Moghaddam, B. (2025). Predicting the Lead Time of Auto Parts Orders in the Supply Chain Using Machine Learning, *Journal of Business Intelligence Management Studies*, 13(52), 87-123. DOI: 10.22054/ims.2025.80935.2491

was used. The predictive model uses Multiple Linear Regression, Multiple with curvature, Lasso, Elastic Net, Boosted Decision Tree, Bootstrap Random Forest, K-Nearest Neighbor, and Boosted Multi-Layer Perceptron. The performance of the fitted regression models has been confirmed using  $R^2$ , RASE, and validation of the results and model. The results showed that the logistical features are effective in LT, and the Boosted Multi-Layer Perceptron is the best for predicting orders' LT with an accuracy of 96% and an error of 5.84. Using the model's predictive capability for new data in the Kanban system, the results obtained within four months have been used. The improvements from using DM capabilities in the Kanban system all express the significant impact of combining lean and machine learning (ML) tools to empower and resilient Lean Supply Chain Management (LSCM).

## **1. Introduction**

The main problem in this research is identifying the factors that effectively predict the LT of orders in the LSC, choosing the best ML algorithm for predicting the exact LT, and how process mining can effectively identify the most repeatable variables in the main variants and investigate how DM can reduce waste in LSC.

Despite classification studies on risk, disruption, and delay prediction in the literature, to our knowledge, fewer articles were found regarding the use of DM to predict the accurate LT of orders in the LSC with logistical features. Also, according to researchers, DM is considered a tool to overcome the limitations of lean tools and strengthen their performance. However, the studies corresponding to the executive case did not observe the results and improvements from the ML application in predicting the LT of orders.

Therefore, in this research, in terms of innovation, 1) machine learning has been used to accurately predict the LT of Kanban orders, considering logistical factors, 2) Process mining has been used in the identification stage of influential variables, 3) The results and improvements obtained from predicting the LT of orders regarding risk reduction and sustainability improvement have been examined and compared.

## **Research Question(s)**

The main questions in this research are specified as follows:

1. What factors affect LT's prediction in the lean supply chain?

3. How do we predict the LT in the lean supply chain?
4. How can DM effectively reduce waste in the lean supply chain?

## **2. Literature Review**

Regarding the issue's importance and urgency, transparency and accurate prediction of the LT have reduced risk and improved sustainability and resilience in the LSC. These effects are significant in both theoretical and operational dimensions, such as reducing logistic costs, safety stock, working capital, stoppage, level of inventories, storage cost, energy consumption, and risk. After reviewing the literature, the most relevant articles in the field of ML are listed in Table2.

## **3. Methodology**

This research is practical from the objective point of view, and from the data point of view, it is quantitative. This study includes four main processes: 1) reviewing the literature and data collection, 2) research method and pre-processing, 3) model construction, and 4) model evaluation and results (Jayanti, 2022 & Wasesa). First, influential variables were extracted by reviewing the literature. Then, the dataset was extracted from the Kanban system in compliance with the requirements of the data set quality indicators from 6/1402 to 11/1402. Then, process mining was used to identify the features with the most repeatability in the main variants, and finally, influential variables were extracted through brainstorming. An integrated stepwise analysis approach has been used to select features. The predictive model uses MLR, curvature, Lasso, Elastic Net, Boosted DT, Bootstrap RF, KNN, and Boosted Multi-Layer Perceptron. The parameters of non-parametric approaches are tuned to improve forecasting performance and accuracy. In this research, evaluation and validation are the main criteria for evaluating the model's predictive power, and error and accuracy indices have been used together. Therefore, the performance of the fitted regression models using  $R^2$  and RASE evaluation indices and validation of the results and the model are confirmed.

## **4. Results**

After fitting the regression models, for each row of test data, predict the LT and compare it with the actual values of the LT; then, to identify the best model,  $R^2$ , RASE, and model comparison approaches are used.

The results show that the Boosted Multi-Layer Perceptron, with one hidden layer, five activation functions, and a learning rate of 0.1, has the highest accuracy at 96% and the lowest root average square error at 5.84, compared to other fitted models.

### **5. Discussion and Conclusion**





The obtained results show that the identified independent variables are related to customer factors (safety stock), manufacturer factors (inspection status, quality point), logistic factors (vehicle, distance), part factors (name, part-expert), and order factors (number of holidays, Kanban issue date) are effective on the LT. As the selected model in this research, the regression model of the Boosted Multi-Layer Perceptron has the highest  $R^2$  and the lowest RASE criteria. Process mining is practical and helpful in identifying the main variants. By using the model's predictive capability for new data in the Kanban order issuing system within four months, the improvements all express the significant impact of combining lean tools and ML to empower LSCM. The practical implications of this research can guide managers in implementing practices with lean tools, improving sustainability, eliminating waste, and being more competitive in the current challenging business environment. Academics can benefit from the present study because it provides ML practices that can be further tested and validated.

This research generalizes and develops the use of DM as a decision-making support tool in predicting the LT to overcome the limitations of lean tools, and it can improve the efficiency and stability of the LSC and reduce the risk. While this research provides valuable insights, it also has limitations, including the lack of data on influential variables identified in the literature. In implementing this research, there are suggestions for future research that examine factors such as production capacity, weather, and location conditions and deep learning to fit more reliable and accurate results and investigate prescriptive analyses to optimize the LT of orders based on the fitted regression models, the design of the experiment and using the profiler's capabilities.

**Keywords:** Machine Learning, Regression, Lean Supply Chain Management, Kanban, Lead Time.



## پیش‌بینی زمان انتظار سفارشات قطعات خودرو در زنجیره تأمین با استفاده از یادگیری ماشین

- فائزه زمانی 
- دانشجوی دکتری رشته مدیریت صنعتی، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران
- احمد ابراهیمی \*
- استادیار گروه مدیریت صنعتی و تکنولوژی، دانشگاه آزاد اسلامی، واحد علوم و تحقیقات، تهران، ایران، تهران، ایران
- رؤیا سلطانی 
- استادیار گروه مهندسی صنایع، دانشگاه خاتم، تهران، ایران
- بابک فرهنگ مقدم 
- دانشیار موسسه مطالعات برنامه‌ریزی و مدیریت، تهران، ایران

### چکیده

هدف این پژوهش بررسی عوامل مؤثر در پیش‌بینی زمان انتظار و ایجاد مدل پیش‌بینانه زمان انتظار سفارشات کانبان به جهت بهبود پایداری و تاب‌آوری در زنجیره تأمین ناب می‌باشد. برای دستیابی به این هدف، مطالعه از روش داده‌کاوی پیروی می‌کند، مجموعه داده‌ها شامل ۱۰۳۰۲۳ مشاهده، از سیستم کانبان واکسترانت زنجیره تأمین با رعایت الزامات شاخص‌های کیفیت دیتاست در بازه ۱۴۰۲/۶ تا ۱۴۰۲/۱۱ استخراج شده است. ابتدا شاخص‌های مؤثر بر زمان انتظار سفارشات استخراج شده است و به جهت بهبود عملکرد و دقت پیش‌بینی، از فرآیند کاوی جهت شناسایی متغیرهای پرتکرار و تأثیرگذار در واریانت‌های اصلی و سپس در مرحله برآزش مدل، از رویکرد تحلیل گام‌به‌گام تلفیقی جهت انتخاب ویژگی‌ها و از تنظیم پارامتر رویکردهای رگرسیونی ناپارامتریک استفاده شده است. مدل پیش‌بینانه با استفاده از مدل‌های رگرسیونی خطی چندمتغیره، چندمتغیره دارای انحنای لاسو، الاستیک‌نت، درخت تصمیم تقویتی، جنگل تصادفی بوت‌استرپ، k-نزدیک‌ترین همسایه، شبکه عصبی تقویتی برآزش داده شده است. عملکرد مدل‌های رگرسیونی برآزش شده با استفاده از شاخص‌های ارزیابی  $R^2$ ، RASE و اعتبارسنجی نتایج و مدل تأیید شده است. نتایج نشان

داد که عوامل لجستیکی در زمان انتظار سفارشات مؤثر بوده و الگوریتم شبکه عصبی تقویت شده بهترین مدل در پیش‌بینی زمان انتظار سفارشات با دقت ۹۶ درصد و با خطا ۵/۸۴ است. سپس قابلیت پیش‌بینی مدل برای دیتاهای جدید در سیستم صدور سفارشات کانبان به کار گرفته شده است، نتایج و بهبودهای حاصل از بهره‌گیری قابلیت‌های داده‌کاوی در سیستم کانبان همگی بیان‌گر تأثیر معنی‌دار ترکیب ابزار ناب و یادگیری ماشین به جهت توانمندسازی و تاب‌آوری زنجیره تأمین ناب می‌باشد.

**کلیدواژه‌ها:** یادگیری ماشین، رگرسیون، مدیریت زنجیره تأمین ناب، کانبان، زمان انتظار.

## مقدمه

شفافیت و پیش‌بینی دقیق زمان انتظار سفارشات کانبان در زنجیره تأمین ناب منجر به تاب‌آوری، کاهش هزینه‌های عملیاتی و ارتقا سطح پایداری می‌شود (Alsadi *et al.*, 2021; Schneckenreither *et al.*, 2022) و عدم قطعیت در برآورد زمان تحویل تأمین‌کنندگان منجر به ابطال سفارشات صادره کانبان، تأخیر در تأمین سفارشات و افزایش ریسک توقف خطوط تولید می‌گردد و جریمه‌های گزافی را به شرکت‌های تأمین‌کننده تحمیل می‌نماید (Gabellini *et al.*, 2024; Singh and Soni, 2019). لذا برای پرکردن این خلأ، سازمان‌ها با استفاده از یادگیری ماشین به شناسایی فاکتورهای پرمخاطره و موقعیت‌های پرریسک می‌پردازند، به این معنا که در زمان مواجهه زنجیره تأمین با چالش‌های زمان، هزینه و محدودیت‌های منابع، یادگیری ماشین به‌عنوان یک رویکرد برتر به مدیریت ریسک‌ها می‌پردازد (Baryannis *et al.*, 2019; Pozzi *et al.*, 2022; Yang *et al.*, 2023). با استناد به پیوند قوی و هم‌افزایی بین تولید ناب و صنعتی<sup>۱</sup>، می‌توان با بهره‌گیری از داده‌کاوی جهت پیش‌بینی هر چه دقیق‌تر زمان انتظار و تلفیق آن با رویکرد تولید ناب، به جهت کاهش یا حذف اتلاف‌ها در سطح زنجیره تأمین ناب اقدام نمود (Dillinger *et al.*, 2022; Pozzi *et al.*, 2022) و هر چه بیشتر به اهداف اقتصادی، اجتماعی، محیطی (پایداری) در مدیریت زنجیره تأمین ناب دست یافت (Maware and Parsley, 2023). از آنجایی سازمان‌ها به دنبال کاهش زمان انتظار به‌منظور کاهش هزینه‌های تولید خود هستند، زمان انتظار کوتاه‌تر به‌عنوان یک منبع اصلی دستیابی به مزیت رقابتی بالقوه، می‌تواند به بهینه‌سازی در زنجیره تأمین داخلی و پایداری بیشتر نیز کمک کند (Ivanov and Jaff, 2017). امروزه اکوسیستم زنجیره تأمین دیجیتال از پیامدهای تحول و پویایی ناشی از دیجیتالی شدن بهره‌مند گردیده است و برای کلیه ذینفعان در این حوزه، این یک پیشرفت چشمگیر است و یادگیری ماشین در هسته این تحول بزرگ قرار

---

۱. انقلاب صنعتی چهارم، یا Industry 4.0، نسل چهارم انقلاب صنعتی است که بر پایه استفاده گسترده از فناوری‌های نوین و پیشرفته در تولید و زنجیره‌های تأمین بنا شده است.

دارد (Mohamed-Iliasse et al., 2022). لذا هدف این پژوهش شناسایی و غربال کردن عوامل مؤثر، تعیین رابطه بین متغیرهای تأثیرگذار و ایجاد مدل پیش‌بینی زمان انتظار سفارشات به منظور بهبود پایداری و تاب‌آوری در زنجیره تأمین ناب می‌باشد. ضرورت پیش‌بینی زمان انتظار در زنجیره تأمین ناب از دو بعد نظری (ادبیات موضوع) و بعد عملیاتی (مصاحبه با متخصصین صنعت) مطرح است که در جدول شماره ۱ خلاصه شده است.

جدول ۱. جمع‌بندی اثرات پیش‌بینی زمان انتظار از بعد نظری و عملیاتی

منابع	کاهش توقف خط	افزایش بهره‌وری	پایداری	مدیریت ریسک	کاهش هزینه	کاهش آلاینده‌گی	کاهش مصرف انرژی	کاهش سرمایه در گردش	کاهش ذخیره احتیاطی	کاهش هزینه انبارش	کاهش سطح موجودی‌ها	کاهش هزینه لجستیک	دینفعان	بعد
(Alnahhal et al., 2021; Bassiouni et al., 2024; Gabellini et al., 2024;	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓	تأمین‌کننده	نظری
(Alnahhal et al., 2021; Bassiouni et al., 2024; Gabellini et	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	زنجیره تأمین	
(Burggräf et al., 2021; Ivanov and Jaff, 2017; Rokoss et al., 2024)	✓		✓	✓	✓		✓	✓	✓	✓	✓		مشتری	



منابع	کاهش توقف خط	افزایش بهره‌وری	پایداری	مدیریت ریسک	کاهش هزینه	کاهش آلودگی	کاهش مصرف انرژی	کاهش سرمایه در گردش	کاهش ذخیره احتیاطی	کاهش هزینه انبارش	کاهش سطح موجودی‌ها	کاهش هزینه لجستیک	دینفعان	بعد
(Alnahhal et al., 2021; ER and MOSAWI, 2022)							✓						جامعه	
							✓						محیط	
مصاحبه با متخصصین صنعت خودرو	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓	تأمین‌کننده	✓
	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	زنجیره تأمین	✓
	✓												مشتری	✓
	✓						✓	✓	✓				جامعه	
			✓				✓	✓					محیط	

با مطالعه پیشینه پژوهش (جدول شماره ۲)، مشاهده گردید که علیرغم پژوهش‌های کلاسه‌بندی<sup>۱</sup> در خصوص تعیین ریسک، پیش‌بینی اختلال در زنجیره تأمین و ریسک تأخیر، مطالعات کمتری به بهره‌گیری از یادگیری ماشین جهت پیش‌بینی دقیق زمان انتظار سفارشات کاتبان در زنجیره تأمین ناب با لحاظ عوامل جغرافیایی پرداخته‌اند. همچنین در پیشینه پژوهش کمتر از فرآیند کاوی به جهت شناسایی متغیرهای تأثیرگذار در فرآیندهای با واریانس بالا استفاده شده است. لذا در این پژوهش به لحاظ نوآوری (۱) به بهره‌گیری از یادگیری ماشین جهت پیش‌بینی دقیق زمان انتظار سفارشات کاتبان با لحاظ عوامل

## 1. Classification

لجستیکی پرداخته شده است. ۲) از فرآیند کاوی در مرحله شناسایی متغیرهای تأثیرگذار با بیشترین تکرار در واریانت‌های اصلی علاوه بر سایر رویکردهای کیفی استفاده شده است. ۳) نتایج و بهبودهای حاصل از پیش‌بینی زمان انتظار سفارشات در خصوص کاهش ریسک و بهبود پایداری در یک سیستم ناب مورد بررسی و مقایسه قرار گرفته است. لذا این پژوهش با ارائه یک چارچوب عملی جهت پر کردن این شکاف پژوهشی در زنجیره تأمین ناب قابل تعمیم و ضروری است.

با توجه به ضرورت بیان شده، سؤالات پژوهش به شرح ذیل تعریف می‌شود:

- ۱) عوامل مؤثر بر محاسبه زمان انتظار سفارشات در زنجیره تأمین ناب چیست؟
- ۲) زمان انتظار سفارشات در زنجیره تأمین ناب را چگونه می‌توان پیش‌بینی کرد؟
- ۳) داده کاوی چگونه در جهت کاهش اتلاف‌ها در زنجیره تأمین ناب مؤثر واقع می‌شود؟

### پیشینه پژوهش

#### مبانی نظری پژوهش

تولید ناب اصالتاً از شرکت تویوتا و در پایان جنگ جهانی دوم- که با کمبود منابع روبرو شدند- باهدف افزایش بهره‌وری (یا کاهش مصرف منابعی که منجر به ایجاد ارزش افزوده نمی‌شوند) و کاهش هزینه تولید به واسطه حذف هرگونه اتلاف و با استقرار سیستم تولید تویوتا<sup>۱</sup> پدید آمده است (Maware and Parsley, 2023; Qureshi *et al.*, 2023).

زنجیره تأمین یک تغییر پارادایم را در نتیجه صنعتی ۴,۰ و ادغام تکنولوژی‌های جدید از قبیل اینترنت اشیاء، رایانش ابری، ناب ۴,۰، هوش مصنوعی، تجزیه و تحلیل پیش‌بینانه در سیستم‌های تولیدی تجربه نموده است. صنعتی ۴,۰ اجرای ناب ۴,۰ را ممکن ساخته تا منجر به کاهش زمان انتظار، مواد و منابع به جهت کاهش اتلاف‌ها گردد. ناب ۴,۰ و صنعتی ۴,۰ هر دو به دنبال تقویت بهره‌وری و انعطاف‌پذیری فرآیندهای تولیدی و خدماتی هستند. ابزار ناب در زنجیره تأمین به منظور بهینه‌سازی فعالیت‌ها به کار گرفته شده و منجر به کاهش

---

1. Toyota Production System

اتلاف‌ها، بهبود کیفیت، کاهش هزینه‌ها، افزایش انعطاف‌پذیری و پایداری در سطح زنجیره می‌گردد (Garcia-Buendia *et al.*, 2021; Qureshi *et al.*, 2023).

زمان انتظار<sup>۱</sup> از دیدگاه تولیدکننده، زمان سپری‌شده بین ثبت سفارش و زمان تحویل واقعی آن است. زمان انتظار در میان مهم‌ترین شاخص‌های عملکرد شرکت‌ها قرار دارد که سازمان‌های در پی به حداقل رساندن آن و کاهش سطح موجودی خود هستند تا بتوانند محصول درست را در کم‌ترین زمان ممکن به دست مشتریان خود رسانده و انتظارات مشتریان را با لحاظ نوسانات تقاضا برآورده سازند. غالب رویکردهای برنامه‌ریزی و زمان‌بندی تولید بر پایه زمان انتظار است، بدین ترتیب کارایی این رویکردها به شدت تحت تأثیر صحت در پیش‌بینی زمان انتظار می‌باشد و انعطاف‌پذیری زمان انتظار امری مهم تلقی می‌شود (Burggräf *et al.*, 2021; de Oliveira *et al.*, 2021).

یادگیری ماشین زیرشاخه‌ای از هوش مصنوعی و علوم کامپیوتر است که بر مبنای یادگیری آماری به جهت استخراج دانش، روندها، الگوهای پنهان در مجموعه داده‌هاست، در حالت کلی به دو دسته تکنیک یادگیری با نظارت و بدون نظارت تقسیم می‌شود. یادگیری با نظارت به دنبال پیش‌بینی پارامترهای خروجی بر پایه متغیرهای ورودی و دیاست آموزشی است که یکی از بنیادی‌ترین رویکردهای آن مدل‌های رگرسیون است (Bassiouni *et al.*, 2024; Gong *et al.*, 2023).

### مرور مطالعات گذشته

با مرور پیشینه پژوهش، روش‌های بکارگرفته شده در خصوص پیش‌بینی زمان انتظار در حوزه‌های تحلیل آماری و شبیه‌سازی و یادگیری ماشین دسته‌بندی گردیده است؛ که مرتبط‌ترین مقالات حوزه یادگیری ماشین در جدول شماره ۲ آمده است:

جدول ۲. خلاصه مطالعات گذشته

منابع	هدف	ویژگی	انتخاب ویژگی <sup>۱</sup>	الگوریتم	دستاورد/ نتایج
(Öztürk <i>et al.</i> , 2006)	تخمین زمان انتظار تولیدات سفارشی	۲۶	مبتنی بر درخت	Tree	مقایسه عملکرد با ۴ رویکرد در ادبیات
(Alenezi <i>et al.</i> , 2008)	پیش‌بینی زمان از طریق زمان چرخه سفارش	۲	-	SVR ANN	مقایسه عملکرد ماشین بردار پشتیبان نسبت به شبکه عصبی
(Pfeiffer <i>et al.</i> , 2016)	تخمین زمان انتظار تولید	۱۲	-	Tree	مقایسه عملکرد درخت نسبت به الگوریتم‌های خطی
(Gyulai <i>et al.</i> , 2018)	پیش‌بینی زمان انتظار در محیط فروشگاه‌های بر اساس ویژگی‌هایی از محصول	۷	مبتنی بر درخت	Regression Tree SVR	مقایسه عملکرد جنگل تصادفی نسبت به سایر الگوریتم‌ها
(Lingitz <i>et al.</i> , 2018)	پیش‌بینی زمان انتظار با استفاده از الگوریتم‌های یادگیری ماشین	۴۱	-	LM-Ridge RT-RF- SVM- KNN- ANN	مقایسه عملکرد الگوریتم‌های مختلف یادگیری ماشین
(Baryannis <i>et al.</i> , 2019)	پیش‌بینی ریسک‌های زنجیره تأمین	۱۵	آزمون واریانس، درخت	SVM DT-RDT	پیش‌بینی ریسک
(Brintrup <i>et al.</i> , 2020)	پیش‌بینی سفارشات دیر هنگام در زنجیره تأمین	۱۹	طوفان فکری	RF-SVM- LOGISTIC	پیش‌بینی اختلال در زنجیره
(Schneckenreither <i>et al.</i> , 2021)	پیش‌بینی زمان انتظار سفارش	۴۰	آزمون همبستگی	ANN	پیش‌بینی زمان انتظار داینامیک سفارش
(de Oliveira <i>et al.</i> , 2021)	پیش‌بینی زمان انتظار خرید در زنجیره تأمین	۸	-	LR SVM-RF KNN-MLP	مقایسه عملکرد ۵ الگوریتم،

1 Feature Selection

منابع	هدف	ویژگی	انتخاب ویژگی <sup>۱</sup>	الگوریتم	دستاورد/ نتایج
	داروسازی				لحاظ شاخص جغرافیایی
(Welsing et al., 2021b)	ترکیب فرآیند کاوی و یادگیری ماشین در فرآیندهای با واریانس بالا		فرآیند کاوی	DT SVM Naïve Bayes	بالا بردن صحت پیش‌بینی در فرآیندهای با واریانس بالا
(Steinberg et al., 2023)	پیش‌بینی تحویل دیر هنگام تأمین‌کننده	۱۲	آزمون همبستگی	LR-SVR-DT RF-AB-GB MLP	پیش‌بینی تأخیر تحویل
(Bassiouni et al., 2024)	تعیین وضعیت سفارشات در زنجیره تأمین	۵۲	شبکه عصبی	RT-RF KNN-ANN SVM	تعیین اولویت ریسک، لحاظ شاخص‌های جغرافیایی
(Gabellini et al., 2024)	پیش‌بینی ریسک تأخیر تحویل در زنجیره تأمین	۲۰	شبکه عصبی	Deep Learning	برآورد دقیق میزان تأخیر، لحاظ شاخص‌های اقتصادی
(Rokoss et al., 2024)	تعیین زمان تحویل سفارش	۴۶	SHAP <sup>۱</sup>	XGB-ANN SVM-RF DT-LR	پیش‌بینی تاریخ تحویل

بررسی ادبیات نشان داد که در برخی از مطالعات از جمله النزی و همکاران (۲۰۰۸)، پفیفیر و همکاران (۲۰۱۶)، لینگیتز و همکاران (۲۰۱۸) و دالیوریا و دیگران (۲۰۲۱) صرفاً به مقایسه عملکرد الگوریتم‌ها بدون انتخاب ویژگی پرداخته شده است. در مطالعات اوزترک و دیگران (۲۰۰۶) و گیولی و دیگران (۲۰۱۸) با لحاظ انتخاب ویژگی‌ها، عملکرد الگوریتم‌های مختلف مقایسه شده است. در پژوهش‌های باریانس و دیگران (۲۰۱۹)، برینترپ و دیگران (۲۰۲۰) و باسیونی و دیگران (۲۰۲۴) از طریق الگوریتم‌های کلاسه‌بندی به پیش‌بینی اختلال و ریسک در زنجیره تأمین پرداخته شده است؛ اما فقط در مطالعات

1. Shapley Additive Explanation (SHAP) value

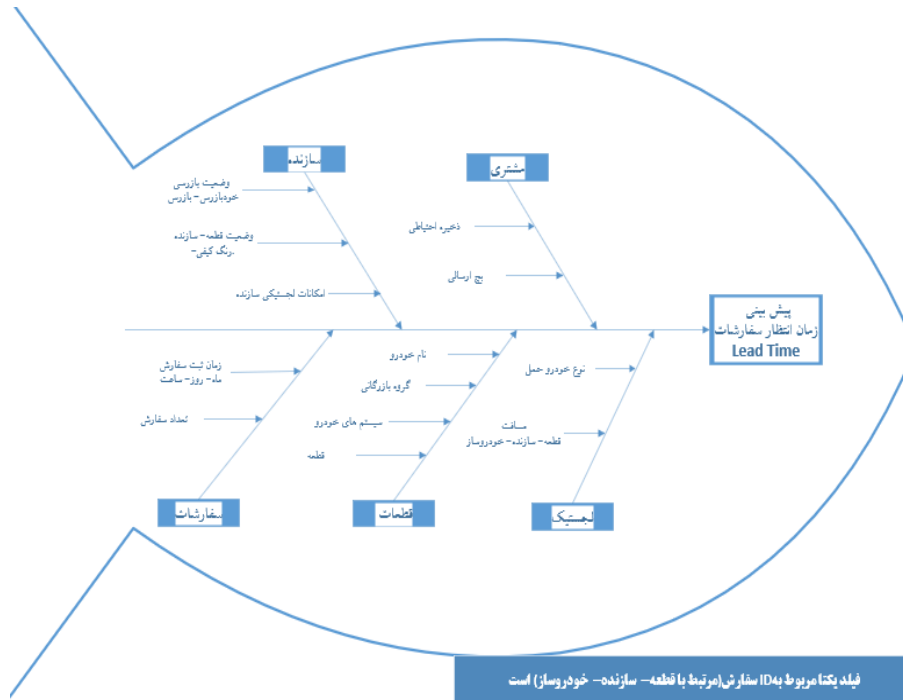


متغیر	باریانس و دیگران (۲۰۱۹)	اسچنکنزینر و دیگران (۲۰۲۱)	دالیوریا و دیگران (۲۰۲۱)	استینرگ و دیگران (۲۰۲۳)	اوزترک و دیگران (۲۰۲۱)	ایوانو و جف (۲۰۱۷)	لینگیز و دیگران (۲۰۱۸)	باسونی و دیگران (۲۰۲۴)	برینترپ و دیگران (۲۰۲۰)	ولسینگ و دیگران (۲۰۲۱)	النزی و دیگران (۲۰۰۸)	پفیسر و دیگران (۲۰۱۶)	روکوس و دیگران (۲۰۲۴)	برگراف و دیگران (۲۰۲۱)	گابلینی و دیگران (۲۰۲۴)
موقعیت		✓						✓							
ظرفیت	✓	✓			✓		✓				✓	✓			
محصول		✓	✓	✓					✓	✓		✓			
قطعه	✓	✓	✓	✓	✓				✓	✓		✓	✓		
گروه			✓						✓	✓					
گروه													✓		
ایراتور		✓			✓							✓			
وضعیت	✓			✓						✓					
نوع حمل				✓				✓							
زمان	✓		✓	✓	✓		✓	✓	✓			✓			
بج سائز				✓	✓				✓			✓			
ذخیره												✓			
متغیرهای															✓
زمان	✓								✓				✓		

در مرحله بعدی، با استخراج جدول رویداد<sup>۱</sup> و از طریق فرآیند کاوی، واریانت‌های پرتکرار شناسایی گردیده و سپس از طریق برگزاری ۸ جلسه مصاحبه با متخصصین صنعت، برگزاری جلسات طوفان فکری و تحلیل علل اصلی، ۱۸ متغیر تأثیرگذار شناسایی و نمودار استخوان ماهی جهت نمایش نقشه ذهنی (شکل شماره ۱) ترسیم گردیده است. چارچوب این پژوهش بر اساس ساخت مدل‌های پیش‌بینانه مبتنی بر دیتاست جمع‌آوری شده در سطح زنجیره تأمین گروه خودروسازی سایپا پیاده‌سازی شده و نتایج آن جهت بررسی عوامل مؤثر و پیش‌بینی زمان انتظار سفارشات در زنجیره تأمین ناب قابل استفاده و تعمیم می‌باشد.

## 1. Event Log

شکل ۱. نمودار استخوان ماهی عوامل



گروه خودروسازی سایپا با ۵ سایت تولیدی و ظرفیت تولید سالانه ۹۵۰ هزار خودرو در حال فعالیت می باشد و گستره کاملی از دسته بندی محصولات از سواری تا تجاری را پوشش می دهد. زنجیره تأمین گروه سایپا در چندین لایه تعریف شده است و مدیریت زنجیره تأمین توسط شرکت سازه گستر سایپا صورت می پذیرد که با بیش از ۱۰۰۰ تأمین کننده در کشور در سطوح مختلف همکاری می نماید. شرکت سازه گستر سایپا از ابزار کانبان جهت کنترل موجودی و مدیریت زنجیره تأمین ناب در سطح گروه بهره گرفته است.

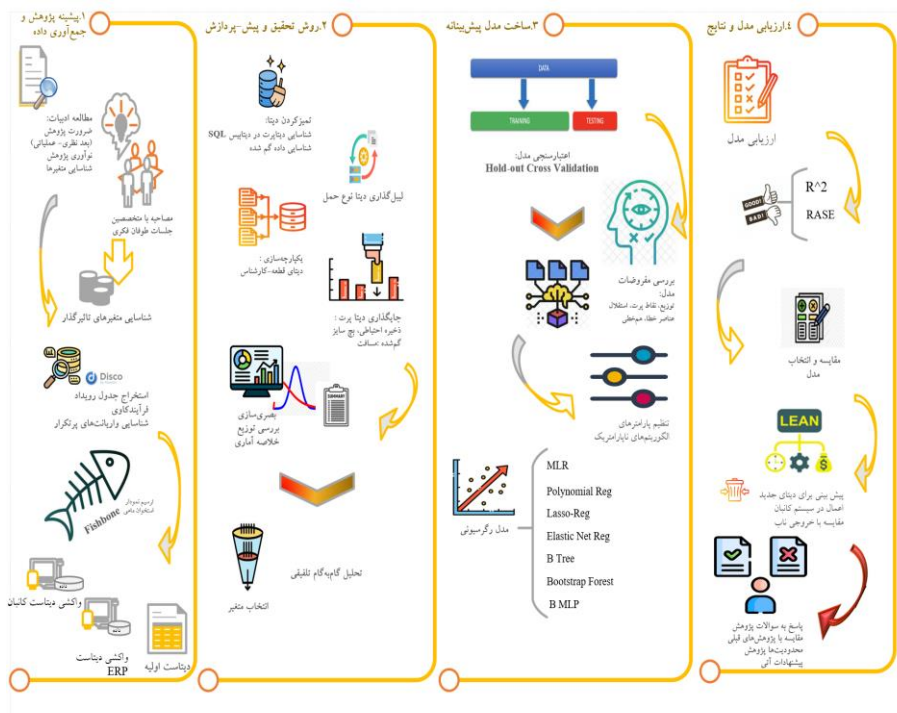
مراحل اجرای پژوهش (شکل شماره ۲) شامل ۴ فرآیند اصلی (۱) بررسی پیشینه پژوهش و جمع آوری دیتا (۲) روش تحقیق و پیش پردازش (۳) ساخت مدل پیش بینانه (۴) ارزیابی مدل و نتایج می باشد (Jayanti and Wasesa, 2022). جهت اجرای این پژوهش و جمع آوری داده از اطلاعات موجود در پایگاه داده سیستم های ERP، کانبان و اکسترانت



پیش‌بینی زمان انتظار سفارشات قطعات خودرو در زنجیره تأمین...؛ زمانی و همکاران | ۱۰۳

زنجیره تأمین گروه خودروسازی سایپا با رعایت الزامات شاخص‌های کیفیت دیتاست (تمامیت، خودسازگاری، به‌موقع بودن، محرمانگی، صحت، استانداردسازی، بی‌طرفی، سهولت) استفاده گردیده است (Gong et al., 2023).

شکل ۲. مراحل اجرای پژوهش



جامعه آماری این تحقیق شامل سفارشات ثبت‌شده در سیستم کانبان (به تفکیک شماره فنی - سازنده - خودروساز) برای سازندگان دارای صفات مشترک (نظیر وجود قراردادفعال، درصد تخصیص) است. بر همین اساس نمونه آماری شامل دیتای مرتبط به سفارشات ثبت‌شده کانبان در بازه ۰۶/۱۴۰۲ لغایت ۱۱/۱۴۰۲ در زنجیره تأمین گروه خودروسازی سایپا می‌باشد. پیش از اقدام به برازش دیتا، پیش‌پردازش داده‌ها به جهت اطمینان از این‌که دیتای خام سازگار، تمیز و مفید باشد ضروری است، لذا از تکنیک‌های مختلفی از جمله تمیز کردن داده‌ها (شناسایی داده پرت، داده گم‌شده)، بصری‌سازی،

انتخاب متغیرها استفاده شده است. انتخاب ویژگی را می‌توان به‌عنوان فرآیند شناسایی ویژگی‌های مرتبط و حذف ویژگی‌های غیرمرتبط و تکراری باهدف مشاهده زیرمجموعه‌ای از ویژگی‌ها که مسئله را به‌خوبی و با حداقل کاهش درجه کارایی تشریح می‌کنند تعریف کرد (GONG ET AL., 2023). لذا به‌جهت شناسایی و صحت‌گذاری متغیرهای مؤثر علاوه بر بررسی ادبیات موضوع (جدول شماره ۳)، از فرآیند کاوی در نرم‌افزار DISCO و برگزاری جلسات مصاحبه و طوفان فکری، استفاده‌شده و سپس در مرحله برازش مدل نیز از تحلیل گام‌به‌گام تلفیقی<sup>۱</sup> برای انتخاب متغیرهای تأثیرگذار در پیش‌بینی زمان انتظار بهره‌گرفته شده است. در این پژوهش، به‌جهت ساخت مدل پیش‌بینانه از رگرسیون استفاده شده است؛ رگرسیون شامل الگوریتم‌هایی از یادگیری با نظارت ماشین می‌شود که مقادیر متغیر وابسته را بر اساس مقادیر متغیرهای مستقل پیش‌بینی می‌کند. اساساً یک معادله ریاضی که متغیر وابسته را به‌عنوان تابعی از متغیرهای مستقل تعریف می‌نماید (IFRAZ ET AL., 2023; KUMAR AND SINGH, 2022). به‌منظور پیش‌بینی متغیر پیوسته زمان انتظار سفارشات از الگوریتم‌های رگرسیونی مندرج در جدول ۴ با استفاده از ابزار JMP استفاده شده است.

جدول ۴. مدل رگرسیونی به‌کارگرفته شده و خصوصیات آن‌ها

نام روش	parametric	Non-Parametric	خطی	غیر خطی
Multiple Linear Regression	✓		✓	
Polynomial Regression	✓			✓
Lasso-Regression	✓		✓	
Elastic Net Regression		✓	✓	
Boosted Tree		✓		✓
Bootstrap Forest		✓		✓
K-Nearest Neighborhood		✓		✓
Boosted Multi-Layer Perceptron		✓		✓

#### 1. Mixed Selection

معمول‌ترین شیوه رگرسیون، متد رگرسیون حداقل مربعات چندمتغیره کلاسیک<sup>۱</sup> است. مدل رگرسیون خطی چندگانه به شکل زیر می‌باشد:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_k x_{ki} + \varepsilon_i \quad \text{معادله ۱}$$

به طوری که  $y_i \in R$  متغیر وابسته،  $x_{i1}, \dots, x_{ki} \in R$  مجموعه متغیرهای مستقل،  $\beta_0, \beta_1, \dots, \beta_k \in R$  مجموعه پارامترها و  $\varepsilon_i \sim N(0, \sigma^2)$  باقیمانده<sup>۱</sup> یا خطا باشد.  $\beta_0$  عرض از مبدأ و  $\beta_1$  برابر میانگین  $y$  وقتی که متغیر  $X$  به اندازه یک واحد افزایش داشته باشد، است. با فرض مدل خطی، پارامترهای رگرسیون با استفاده از متد حداقل مربعات برای مینیمم نمودن خطا به دست می‌آیند (Rahman et al., 2023).

در صورتی که در مدل خطی چندگانه، الگویی غیرمعمول و غیرتصادفی در باقیمانده‌ها دیده شود، نسبت به برازش مدل رگرسیونی چندگانه دارای انحنا اقدام خواهیم نمود. در مدل رگرسیون چندگانه دارای انحنا رابطه بین متغیر پیش‌بینی کننده و متغیر هدف به صورت چند جمله‌ای از درجه  $n$  (برای متغیرهای  $x$ ) مدل شده است، اما پارامترهای معادله رگرسیونی همچنان درجه یک می‌باشند. رگرسیون لاسو یک مدل از نوع رگرسیون خطی بر پایه عملگر انتخاب و انقباض کم‌ترین قدر مطلق است که در آن ضرایب به صورت تنک تخمین زده می‌شود. این روش بدین دلیل مفید است که تمایل به تعداد محدودتری از متغیرهای مستقل و تخمین مقادیر کوچک‌تر برای ضرایب دارد (Rahman et al., 2023; Tao, 2023). رگرسیون الاستیک نت دو مدل رگرسیونی (لاسو و ریج) را با هم ترکیب کرده و معایب آن‌ها را حذف نموده و جایگزین مطمئنی برای آن‌هاست.

درخت تصمیم تقویت گرادیان یک مدل یادگیری اثرکلی برپایه درخت تصمیم است. گرادیان تقویتی با یک حدس از پاسخ شروع می‌شود سپس گرادیان محاسبه شده و یک مدل برازش می‌شود که در آن تابع هزینه مینیمم گردد. مدل جاری برپایه یک پارامتر نرخ یادگیری<sup>۱</sup> به مدل قبلی اضافه خواهد شد. جنگل تصادفی یک تکنیک اثرکلی است که چندین درخت تصمیم را شامل می‌شود، انتخاب تصادفی هر معیار به عنوان گره منجر به

دستیابی به تنوع در ساخت مدل می شود و متغیرها به صورت تصادفی انتخاب خواهند شد نه بر اساس سودمندترین متغیر. این روش از یک استراتژی به عنوان کیسه بندی<sup>۱</sup> استفاده کرده که درخت های زیاد و عمیقی به یکدیگر متصل خواهند شد. جنگل تصادفی متد درخت تصادفی را  $m$  بار از طریق بردار مقادیر ورودی تکرار و سپس میانگین مقادیر نتایج به عنوان مقدار پیش بینی شده در نظر گرفته می شود. مشخصه و مزیت اصلی این رویکرد استفاده از شاخص جینی<sup>۲</sup> به عنوان معیار انتخاب ویژگی و اعتبارسنجی داخلی است که شامل رتبه بندی متغیرهای مستقل می گردد (Kumar and Singh, 2022; Tao, 2023).

الگوریتم  $k$ -نزدیک ترین همسایه به دلیل معیارهای ساده و قابلیت پیش بینی آن برای الگوهای پیچیده غیرخطی بکار گرفته می شود. این پیش بینی به واسطه تعیین نقاط مشابه بین نقاط داده در فضای ویژه ارائه می گردد. در این روش از تابع فاصله اقلیدسی<sup>۳</sup> و تنظیم کردن<sup>۴</sup> پارامتر ( $K$ ) استفاده می شود (Shapi *et al.*, 2021; Singh and Soni, 2019).

شبکه های عصبی، ساختارهای شبکه ای بسیار سازمان یافته هستند که لایه های آن بر طبق سیستم عصبی در مغز انسان قرار گرفته اند. شبکه عصبی شامل ۳ لایه ورودی، پنهان و خروجی می شود. نرون های ورودی ابتدا اطلاعات ویژگی های ورودی را دریافت و سپس به نرون های لایه پنهان ارسال می نمایند، لایه پنهان این داده ها را با توابع فعال سازی پردازش کرده و نتایج را به لایه پنهان بعدی ارسال می کند و این مراحل ادامه می یابد تا داده ها به نرون های لایه خروجی برسد. شبکه به وسیله تنظیم وزن و بایاس بین لایه ها آموزش داده شده و خطا تنظیم گردیده و خروجی های مرتبط با ورودی ها تولید می شود، مجموعه این فرآیند پیش انتشار<sup>۵</sup> نامیده می شود. مهم ترین ویژگی شبکه عصبی قابلیت آن در ساخت مدل غیرخطی، توانایی در یادگیری و تعمیم دهی و تاب آوری خطا است (Ifraz *et al.*, 2023).

- 
1. Bagging
  2. Gini Index
  3. Euclidean distance function
  4. tune
  5. forward propagation

در مرحله ارزیابی مدل، یافتن مدلی با بهترین ویژگی‌ها که قابلیت تعمیم اطلاعات را داشته باشند، مسئله‌ای مهم است و کار چالش‌برانگیز که بتواند دقت پیش‌بینی رضایت‌بخشی را ایجاد کند. در این پژوهش ارزیابی و اعتبارسنجی<sup>۱</sup>، معیارهای اصلی برای ارزیابی توان پیش‌بینی مدل محسوب می‌شوند. هدف از انتخاب مدل، تعیین سطح مناسبی از بایاس و واریانس<sup>۲</sup> برای دستیابی به دقت بالاتر و جلوگیری از بیش‌برازش یا کم‌برازش<sup>۳</sup> است. دقت یک معیار بسیار مهم است، اما برای ارزیابی مدل کافی نیست؛ بنابراین هنگام ارزیابی مدل از شاخص‌های خطا و دقت به‌صورت توأمان استفاده شده است. پس از برازش هر مدل به جهت ارزیابی، مقدار پیش‌بینی شده توسط الگوریتم با مقدار واقعی آن در مجموعه دیتای تست مورد مقایسه قرار رفته (اعتبارسنجی نتایج) و خطای آن با معیارهای مرتبط مقایسه شده است.

لذا در این پژوهش از معیارهای  $R^2$ ، RASE برای ارزیابی مدل استفاده شده است.

$R^2$ : ضریب تعیین یا همان مربع ضریب همبستگی است.

$$R^2 = \frac{SS_{\text{model}}}{SS_{\text{Total}}} = 1 - \frac{SS_{\text{Error}}}{SS_{\text{Total}}}$$

معادله ۲

در تشریح رابطه مذکور، تغییرپذیری کل در متغیر پاسخ ( $SS_{\text{Total}}$ ) به دو بخش تقسیم می‌شود:  $SS_{\text{model}}$  تغییرپذیری که توسط مدل رگرسیونی توضیح داده شده و  $SS_{\text{Error}}$  تغییرپذیری که توسط مدل توضیح داده نشده است. در حالت کلی هر چه مقدار ضریب  $R^2$  به یک نزدیک‌تر باشد، تغییرپذیری که مدل تشریح نموده است بیشتر خواهد بود. RASE: ریشه میانگین مربعات خطاست.

$$RASE = \sqrt{\frac{SSE}{n}}$$

معادله ۳

- 
1. Evaluation & Validation
  2. Bias & Variance
  3. Overfitting & Underfitting
  4. Coefficient of Determination
  5. Root Average Squared Error

در خصوص چند مدل با ضرایب تعیین و ضرایب تعیین اصلاح شده تقریباً یکسان، مدلی انتخاب خواهد شد که RASE کوچک تری داشته باشد. برای اعتبارسنجی مدل نیز از رویکرد hold-out Cross-validation و تقسیم دیتا به دو بخش آموزش (۷۵٪) و تست (۲۵٪) بهره گرفته شده است.

### یافته‌ها

دیتاست اولیه به شرح جدول شماره ۵ می‌باشد. پس از مرحله حذف وضعیت‌های ابطال و منقضی، تمیز کردن داده‌ها، بررسی نقاط پرت، اصلاح دیتاهای گم شده و نادرست، اصلاح فرمت داده، لیبلینگ، تعداد ۱۰۳۰۲۳ رکورد سفارشات احصاء گردیده است که نمایش بصری دیتا در شکل شماره ۳ نمایش داده شده است.

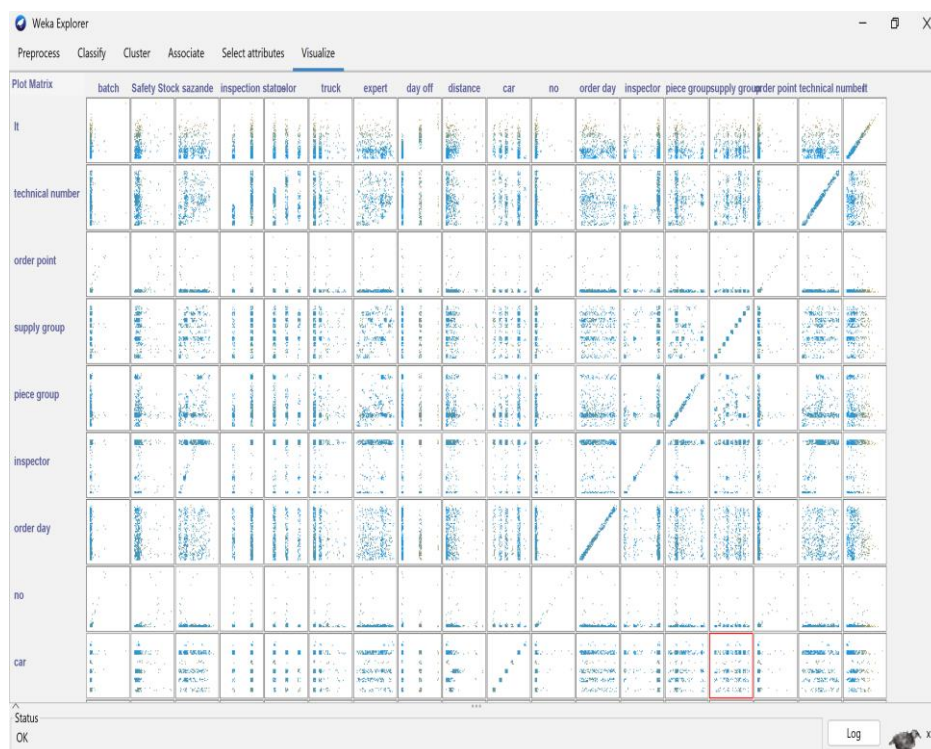
جدول ۵. دیتاست اولیه

زنجیره تأمین گروه خودروسازی سایپا			
تعداد رکورد: ۱۰۸۴۹۵	بازه جمع‌آوری دیتا: ۱۴۰۲/۶ لغایت ۱۴۰۲/۱۱		
مشتری	۵ کلاس	متغیرهای گسسته	
تأمین‌کننده	۳۳۱ کلاس		
خودرو	۶ کلاس		
قطعه	۲۷۰۲ کلاس		
گروه تأمین	۹ کلاس		
گروه تکنولوژی	۵۶ کلاس		
قطعه- کارشناس	۱۲۹ کلاس		
وضعیت بازرسی	۲ کلاس		
بازرس	۱۶۸ بازرس		
رنگ کیفی	۳ کلاس		
نوع حمل	۱۲ کلاس		
تاریخ صدور کانبان	۱۳۹ کلاس		
مسافت	به تفکیک قطعه-سازنده- محل تحویل		متغیرهای پیوسته
بیج سایز	تعداد چیدمان در هر پالت		
ذخیره احتیاطی	ذخیره احتیاطی به ساعت		
تعداد روز تعطیل	روز تعطیل در بازه سفارش تا رسیدن محموله		

پیش‌بینی زمان انتظار سفارشات قطعات خودرو در زنجیره تأمین...؛ زمانی و همکاران | ۱۰۹

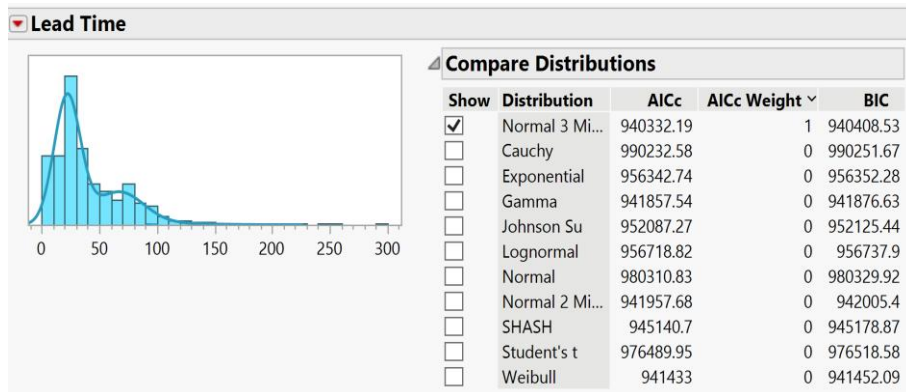
زنجیره تأمین گروه خودروسازی سایپا		
تعداد رکورد: ۱۰۸۴۹۵	بازه جمع‌آوری دیتا: ۱۴۰۲/۶ لغایت ۱۴۰۲/۱۱	
تعداد رسید شده	تعداد رسید شده	
	زمان انتظار	متغیر هدف

شکل ۳. بصری‌سازی



سپس در پژوهش به بررسی توزیع و خلاصه آماری متغیر هدف، پرداخته شده است (شکل شماره ۳ و ۴). همان‌طور که نمایش داده‌شده، متغیر هدف دارای توزیع نرمال ۳ متغیره بوده و چولگی به سمت راست دارد.

شکل ۴. نمایش توزیع متغیر هدف



شکل ۵. خلاصه آماری

Quantiles			Summary Statistics	
100.0%	maximum	297	Mean	38.144987
99.5%		135	Std Dev	28.184578
97.5%		103	Std Err Mean	0.0878101
90.0%		80	Upper 95% Mean	38.317094
75.0%	quartile	53	Lower 95% Mean	37.97288
50.0%	median	29	N	103023
25.0%	quartile	20	N Missing	0
10.0%		8		
2.5%		3		
0.5%		2		
0.0%	minimum	1		

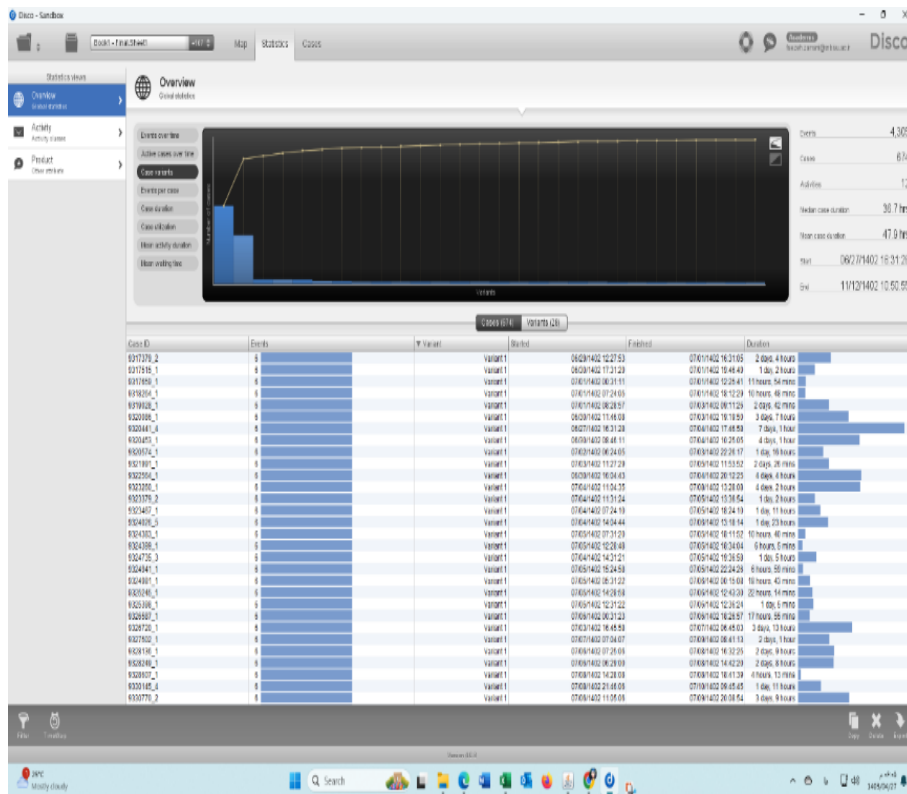
در ادامه جهت شناسایی و صحت‌گذاری متغیرهای مؤثر علاوه بر بررسی سوابق پژوهش (جدول شماره ۳) و برگزاری جلسات طوفان فکری با متخصصین صنعت، از فرآیند کاوی با استفاده از فیلترهای مبتنی بر فعالیت و مسیرها در نرم‌افزار DISCO استفاده شده است و واریانتهای دارای بیشترین تکرار استخراج که در شکل شماره ۶ نمایش داده شده است. در مرحله بعد، در فاز برآزش مدل رگرسیونی از طریق تحلیل گام‌به‌گام تلفیقی که به‌صورت هم‌زمان از رویکردهای روبه‌جلو و روبه‌عقب بهره‌می‌گیرد برای شناسایی متغیرهایی که دارای بالاترین سطح تأثیرگذاری در متغیر هدف می‌باشند بهره‌گرفته شده



پیش‌بینی زمان انتظار سفارشات قطعات خودرو در زنجیره تأمین...؛ زمانی و همکاران | ۱۱۱

است. سپس متغیرها بر اساس معیارهای ارزش<sup>۱</sup> و ( $P\text{-Value} < 0.05$ ) رتبه‌بندی شده‌اند (شکل ۷ شماره).

شکل ۶. فرآیندکاوی جهت شناسایی فرآیندهای دارای واریانس بالا



## 1. Logworth

شکل ۷. تحلیل گام به گام تلفیقی جهت انتخاب متغیر

Effect Summary		
Source	Logworth	PValue
تعداد روز تعطیل	38661.08	0.00000
تاریخ صدور کانبان	16455.54	0.00000
نام سازنده	7486.315	0.00000
کارشناس قطعه	3862.779	0.00000
نام خودرو	1500.609	0.00000
مسافت	416.142	0.00000
Safety Stock	310.485	0.00000
خودروی حمل	181.922	0.00000
وضعیت بازرسی	30.251	0.00000
رنگ کیفی	20.060	0.00000

سپس، برازش مدل‌های رگرسیونی مطابق با متدهای مذکور (جدول شماره ۴) صورت پذیرفته است. لازم به ذکر است جهت برازش مدل رگرسیونی خطی چند متغیره و مدل رگرسیونی دارای انحنا ابتدا مفروضات مرتبط با مدل رگرسیونی شامل: (۱) فرض مناسب و واقعی بودن رابطه خطی، (۲) فرض برابری واریانس خطا، (۳) فرض نرمال بودن عناصر خطا از طریق آزمون اندرسون-دارلینگ<sup>۱</sup>، (۴) فرض وجود استقلال یا عدم همبستگی عناصر خطا از طریق آزمون دورین-واتسون<sup>۲</sup>، (۵) فرض وجود نقاط پرت و نقاط پرت بالقوه از طریق آماره Cook's D و (۶) فرض هم خطی<sup>۳</sup> یا همبستگی متغیرهای مستقل از طریق فاکتور تورم واریانس<sup>۴</sup> به تفکیک مورد بررسی و تحلیل قرار گرفته است. در ادامه از رویکردهای رگرسیونی لاسو، الاستیک نت (بر پایه جریمه) در محاسبه ضرایب مدل‌های رگرسیونی چند گانه برای حل مسائل هم خطی استفاده شده است. همچنین به جهت بهینه‌سازی مقادیر عددی پارامترها در رویکردهای رگرسیونی ناپارامتریک از مقادیر و روش‌های بهینه‌سازی مندرج در جدول شماره ۶ استفاده شده است.

- 
- 1 Anderson-Darling
  - 2 Durbin-Watson
  - 3 Multicollinearity
  - 4 Variance Inflation Factor (VIF)

جدول ۶. بهینه‌سازی پارامترهای الگوریتم رگرسیون

الگوریتم	پارامتر	بهینه‌سازی
Lasso-Regression	Response Distribution= Normal Number of Grid points=150 Min Penalty Fraction=0.0001 Grid Scale=Square Root Initial Displayed Solution=Best Fit	Solution Path Chart Lambda Penalty=0.66
Elastic Net Regression	Response Distribution= Normal Elastic Net Alpha=0.99 Number of Grid points=150 Min Penalty Fraction=0.0001 Grid Scale=Square Root Initial Displayed Solution=Best Fit	Solution Path Chart Lambda Penalty=0.67
Boosted Tree	Number of Layers= 1000 Splits per Tree=19 Learning Rate=0.255 Min Size Split=10 Random Seed=1234	Cumulative Validation Chart (RSquare Validation-Number of Layers) Number of Layers= 1000
Bootstrap Forest	Number of Trees in the Forest=1000 Number of Terms Sampled per Split=7 Bootstrap Sample Rate=1 Min Splits per Tree=10 Max Splits per Tree=2000 Min size Split=103 Random Seed=1234	Cumulative Validation Chart (RSquare Validation-Number of Layers)
K-Nearest Neighborhood	$1 \leq K \leq K_{\max} = 320$ $K_{\max} = \sqrt{N} = 320$ Random Seed=1234	Model Selection Chart (RASE-K) K=5
Boosted Multi-Layer Perceptron	Hidden Layers=1 First (TanH=3, Linear=1, Guassian=1) Number of Models=60 Learning Rate=0.1	Try and Error

پس از برازش مدل‌های رگرسیونی با استفاده از داده‌های آموزشی نسبت به پیش‌بینی زمان انتظار سفارشات برای هر ردیف سفارشات در بخش داده‌های تست و مقایسه آن با مقادیر

واقعی زمان انتظار بهره گرفته شده، سپس برای شناسایی بهترین مدل از شاخص های  $R^2$ ، RASE و رویکرد مقایسه مدل<sup>۱</sup> (جدول شماره ۷ و شکل ۸) استفاده شده است.

جدول ۷. مقایسه شاخص های مدل رگرسیونی دیتای تست

نام روش	$R^2$	RASE
Multiple Linear Regression	۰/۹۵	۶/۲۱
Multiple Linear Regression (with Interaction)	۰/۹۵	۶/۲۱
Lasso-Regression	۰/۹۵	۶/۳۰
Elastic Net Regression	۰/۹۵	۷/۰۰
Boosted Tree	۰/۹۵	۶/۱۲
Bootstrap Forest	۰/۹۵	۶/۳۴
K-Nearest Neighborhood	۰/۸۸	۹/۶۶
Boosted Multi-Layer Perceptron	۰/۹۶	۵/۸۴

شکل ۸. رویکرد مقایسه مدل در بخش دیتای تست از نرم افزار JMP

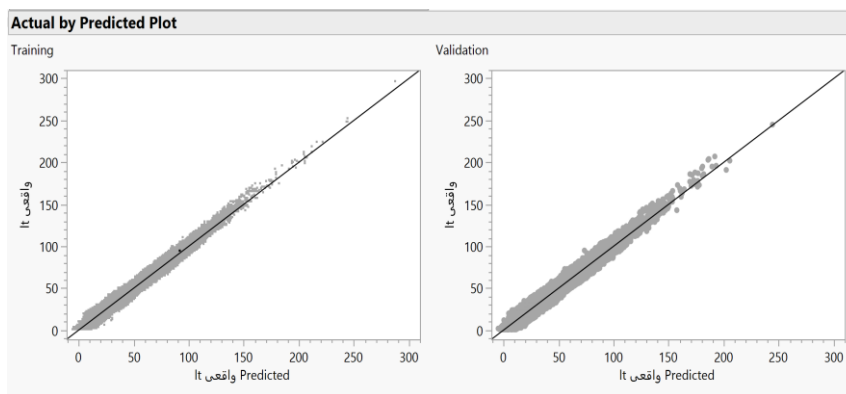
Model Comparison Validation=1									
Predictors									
Target	Predictors								
It واقعی	Pred Formula It واقعی (LS-total)	Fit Least Squares							
	Pred Formula It واقعی (LS.Cross-total)	Fit Least Squares							
	It واقعی Prediction Formula(Lasso-total)	Fit Generalized Lasso							
	It واقعی Prediction Formula(Elastic-total)	Fit Generalized Elastic Net							
	It واقعی Predictor(BTree-total)	Boosted Tree							
	It واقعی Predicted(BF-total)	Bootstrap Forest							
	Predicted It 5 واقعی (KNN=5-total)	K Nearest Neighbors							
	Predicted It واقعی (Boosting-MLP-total)	Neural							
Measures of Fit for It واقعی									
Predictor	Creator	.2	.4	.6	.8	RSquare	RASE	AAE	Freq
Predicted It واقعی (Boosting-MLP-total)	Neural					0.9559	5.8448	4.8934	25518
It واقعی Predictor(BTree-total)	Boosted Tree					0.9517	6.1163	4.7521	25518
Pred Formula It واقعی (LS.Cross-total)	Fit Least Squares					0.9503	6.2068	5.2858	25518
Pred Formula It واقعی (LS-total)	Fit Least Squares					0.9502	6.2097	5.3015	25518
It واقعی Prediction Formula(Lasso-total)	Fit Generalized Lasso					0.9487	6.3018	5.3540	25518
It واقعی Predicted(BF-total)	Bootstrap Forest					0.9480	6.3432	5.0195	25518
It واقعی Prediction Formula(Elastic-total)	Fit Generalized Elastic Net					0.9366	7.0046	5.8107	25518
Predicted It 5 واقعی (KNN=5-total)	K Nearest Neighbors					0.8794	9.6643	6.9730	25518

نتایج مقایسه ای جدول ۷ و شکل شماره ۸ نشان می دهد که الگوریتم رگرسیونی شبکه عصبی تقویتی با یک لایه پنهان و پنج تابع فعال سازی (مندرج در جدول شماره ۶) و نرخ یادگیری (۰/۱) دارای ۶۰ تکرار، بالاترین دقت پیش بینی را با ۰/۹۶ و کم ترین ریشه میانگین

## 1. Model Comparison

مربعات خطا را با ۵/۸۴ نسبت به سایر مدل‌های برازش شده داراست؛ بنابراین مدل رگرسیونی برازش شده در شکل شماره ۹، به‌عنوان مدل منتخب در این پژوهش می‌باشد که علاوه بر دقت بالاتر نسبت به مدل‌های دیگر، از نظر معیار خطا جایگاه بالاتری نسبت به سایر مدل‌ها دارد. لازم به ذکر است که پارامترهای این مدل پیش‌بینانه از طریق سعی و خطا بهینه‌سازی گردیده است.

شکل ۹. مدل رگرسیونی برازش شده از طریق شبکه عصبی تقویتی



سپس از طریق به‌کارگیری قابلیت پیش‌بینی مدل برای صدور سفارشات جدید، نتایج به‌دست آمده توسعه یافته و بدین ترتیب عملکرد روش پیشنهادی در بازه چهارماهه مورد ارزیابی قرار گرفته است. بهبودهای حاصل (جدول شماره ۸)، همگی بیان‌گر تأثیر معنی‌دار ترکیب رویکرد ناب و ابزار یادگیری ماشین به‌جهت توانمندسازی زنجیره تأمین ناب در سطح گروه خودروسازی می‌باشد.

جدول ۸. کاهش اتلاف‌ها ناشی از ترکیب به‌کارگیری ناب و داده‌کاوی

اتلاف	شرح دستاورد	درصد بهبود	ارزش ریالی (میلیارد ریال)
زمان	کاهش میانگین LT	۳۵	-
	کاهش توقفات خط	۱۹	-
سرمایه در گردش	کاهش ارزش موجودی قطعات	۴۹	۱۸۱
هزینه انبارش	فضای اشغال‌شده در انبارها	۳۹	۸

اتلاف	شرح دستاورد	درصد بهبود	ارزش ریالی (میلیارد ریال)
هزینه لجستیک	تعداد پالت در گردش	۶۱	۳
	هزینه ساخت پالت جدید	۱۰۰	۲۵
هزینه موجودی	کاهش سطح ذخیره احتیاطی	۲۸	۴۱۰
ریسک تأمین	کاهش ابطال	۱۴/۱	-
	کاهش انقضاء	۱۳/۲۶	-

### بحث و نتیجه گیری

هدف این پژوهش شناسایی و غربال کردن عوامل مؤثر، تعیین رابطه بین متغیرهای تأثیرگذار و ایجاد مدل پیش‌بینانه زمان انتظار سفارشات کانبان به منظور بهبود پایداری و تاب‌آوری در زنجیره تأمین ناب می‌باشد. در خصوص پاسخگویی به سؤال اول پژوهش، نتایج به دست آمده نشان می‌دهد که متغیرهای مستقل شناسایی شده مرتبط با عوامل مشتری (ذخیره احتیاطی)، عوامل سازنده (وضعیت بازرسی، رنگ کیفی)، عوامل لجستیکی (خودروی حمل، مسافت)، عوامل قطعه (نام خودرو، کارشناس قطعه) و عوامل سفارش (تعداد روز تعطیل، تاریخ صدور کانبان) بر زمان انتظار سفارشات مؤثر می‌باشند. در پاسخ به سؤال دوم، پس از بررسی مفروضات، انتخاب متغیرهای تأثیرگذار، بهینه‌سازی پارامترهای مدل، برازش مدل‌های رگرسیونی، نتایج مدل برازش شده به تفصیل مورد بررسی قرار گرفته و مدل رگرسیونی شبکه عصبی تقویتی به عنوان مدل منتخب در این پژوهش بالاترین میزان دقت پیش‌بینی ۹۶٪ و کم‌ترین ریشه میانگین مربعات خطا ۸۴/۵ را دارا می‌باشد. در پاسخ به سؤال سوم، بهبودهای حاصل از به کارگیری قابلیت پیش‌بینی مدل برای صدور سفارشات جدید کانبان (جدول شماره ۸)، همگی بیان‌گر تأثیر معنی‌دار ترکیب رویکرد ناب و ابزار یادگیری ماشین به جهت توانمندسازی زنجیره تأمین ناب در سطح گروه خودروسازی می‌باشد. این مسئله همچنین تأییدکننده یافته‌های سایر مقالات پژوهشی در خصوص رابطه معنی‌دار بین تکنولوژی صنعتی ۴.۰ (ساخت مدل‌های پیش‌بینانه) و تکنیک‌های تولید ناب برای انجام محاسبات کانبان به منظور کاهش اتلاف‌ها می‌باشد (ALSADI ET AL., 2022; DILLINGER ET AL., 2022; ELAFRI ET

AL., 2022; GARCIA-BUENDIA ET AL., 2021; MAWARE AND (PARSLEY, 2023; POZZI ET AL., 2022

بنابر یافته‌های این پژوهش، ادعای ولسینگ و همکاران (۲۰۲۱) در خصوص قابلیت فرآیند کاوی در شناسایی متغیرهای پرتکرار واریانت‌های اصلی به‌منظور بالابردن صحت پیش‌بینی داده کاوی، مورد تأیید قرار گرفته است (Welsing et al., 2021). همچنین استفاده از مدل‌های یادگیری ماشینی‌عنوانان یک ابزار پشتیبان تصمیم‌سازی، برای غلبه بر محدودیت‌های تولید ناب و تقویت عملکرد زنجیره تأمین ناب قابل تعمیم و توسعه می‌باشد و مدیران ارشد این امکان را دارند که تاب‌آوری و پایداری زنجیره تأمین ناب را بهبود بخشیده و ریسک تأمین را به حداقل رسانند. همچنین دانشگاهیان می‌توانند با بهره‌گیری از یکپارچگی فرآیند کاوی و داده کاوی و ترکیب ابزارهای ناب با یادگیری ماشین، تعمیم‌پذیری این مدل پیش‌بینانه را مورد آزمایش و اعتبارسنجی قرار دهند. درحالی‌که این پژوهش بینش‌های ارزشمندی در زنجیره تأمین ناب ارائه می‌کند، با محدودیت‌هایی نیز مواجه بوده است از جمله: نبود دیتا در خصوص برخی متغیرهای تأثیرگذار شناسایی شده در بخش ادبیات، موانع در جمع‌آوری اطلاعات در سطح زنجیره تأمین، وجود داده‌های پرت، دیتاهای گم‌شده و احتیاط در تعمیم یافته‌ها از جمله محدودیت‌های این پژوهش می‌باشد. در اجرای این روش تحقیق، پیشنهادهایی برای تحقیقات آتی وجود دارد که عبارت‌اند از: اول، با جمع‌آوری و بررسی سایر عوامل از جمله ظرفیت تولید و تعداد شیفت سازنده، شرایط جوی و مکانی منطبق بر طول و عرض جغرافیایی و با گسترش مدل‌های یادگیری ماشین از جمله یادگیری عمیق می‌توان به نتایج مطمئن‌تر و دقیق‌تری دست یافت. دوم، بر اساس سطوح مشخص شده عوامل، روابط بین متغیرهای تأثیرگذار مستقل و متغیر هدف و مدل‌های رگرسیونی برآزش شده، از طریق طراحی آزمایش<sup>۱</sup> و با استفاده از قابلیت‌های پروفایلر<sup>۲</sup> نسبت به تحلیل‌های تجویزی<sup>۳</sup> در حوزه بهینه‌سازی زمان انتظار سفارشات اقدام

---

1. Design of Experiment

2. Profiler

3. Prescriptive Analysis

گردد. سوم، به منظور شناسایی بهترین مدل پیش‌بینانه از مقایسه مدل برازش شده با رویکرد فرآیند کاوی با سایر رویکردهای متداول انتخاب متغیر استفاده گردد.

### سپاسگزاری

مقاله حاضر مستخرج از پایان‌نامه دکتری مدیریت صنعتی بوده و در اینجا فرصت را مغتنم شمرده تا از داوران محترم دانشکده و همچنین داوران گرامی مجله مطالعات مدیریت کسب و کار هوشمند که با بیان نظرات ارزشمند خود به هر چه بهتر شدن این پژوهش کمک شایانی نمودند، تشکر و قدردانی گردد.

### تعارض منافع

این پژوهش فاقد تعارض منافع است.

### ORCID

Faezeh Zamani



<https://orcid.org/0009-0003-7422-0823>

Ahmad Ebrahimi



<https://orcid.org/0000-0002-5373-7466>

Roya Soltani



<https://orcid.org/0000-0002-1473-5337>

Babak Farhang Moghadam



<https://orcid.org/0000-0002-5120-9037>



## References

1. Alenezi, A., Moses, S.A. and Trafalis, T.B. (2008), "Real-time prediction of order flowtimes using support vector regression", *Computers & Operations Research*, Vol. 35 No. 11, pp. 3489–3503, <https://doi.org/10.1016/j.cor.2007.01.026>
2. Alnahhal, M., Ahrens, D. and Salah, B. (2021), "Dynamic Lead-Time Forecasting Using Machine Learning in a Make-to-Order Supply Chain", *Applied Sciences*, Vol. 11 No. 21, <https://doi.org/10.3390/app112110105>
3. Alsadi, J., Antony, J., Mezher, T., Jayaraman, R. and Maalouf, M. (2022), "Lean and Industry 4.0: A Bibliometric Analysis, Opportunities for Future Research Directions", *Quality Management Journal*, Taylor and Francis Ltd., <https://doi.org/10.1080/10686967.2022.2144785>
4. Baryannis, G., Dani, S. and Antoniou, G. (2019), "Predicting supply chain risks using machine learning: The trade-off between performance and interpretability", *Future Generation Computer Systems*, Vol. 101, pp. 993–1004, <https://doi.org/10.1016/j.future.2019.07.059>
5. Bassiouni, M.M., Chakraborty, R.K., Sallam, K.M. and Hussain, O.K. (2024), "Deep learning approaches to identify order status in a complex supply chain", *Expert Systems with Applications*, Vol. 250, p. 123947 <https://doi.org/10.1016/j.eswa.2024.123947>
6. Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P. and McFarlane, D. (2020), "Supply chain data analytics for predicting supplier disruptions: a case study in complex asset manufacturing", *International Journal of Production Research*, Taylor & Francis, Vol. 58 No. 11, pp. 3330–3341, <https://doi.org/10.1080/00207543.2019.1685705>
7. Burggräf, P., Wagner, J., Heinbach, B. and Steinberg, F. (2021), "Machine Learning-Based Prediction of Missing Components for Assembly – a Case Study at an Engineer-to-Order Manufacturer", *IEEE Access*, Vol. 9, pp. 105926–105938, DOI:10.1109/ACCESS.2021.3075620
8. Dillinger, F., Bergermeier, J. and Reinhart, G. (2022), "Implications of Lean 4.0 Methods on Relevant Target Dimensions: Time, Cost, Quality, Employee Involvement, and Flexibility", *Procedia CIRP*, Vol. 107, Elsevier B.V., pp. 202–208, <https://doi.org/10.1016/j.procir.2022.04.034>
9. Elafri, N., Tappert, J., Rose, B. and Yassine, M. (2022), "Lean 4.0:

Synergies between Lean Management tools and Industry 4.0 technologies", *IFAC-PapersOnLine*, Vol. 55, Elsevier B.V., pp. 2060–2066, <https://doi.org/10.1016/j.ifacol.2022.10.011>

10. ER, C.H. and MOSAWI, T. Al. (2022), "Effects of Big Data Analytics on Sustainable Manufacturing: A Comparative Study Analysis", *Chinese Journal of Urban and Environmental Studies*, World Scientific Publishing Co., Vol. 10 No. 04, p. 2250022, <https://doi.org/10.1142/S2345748122500221>
11. Gabellini, M., Civolani, L., Calabrese, F. and Bortolini, M. (2024), "A Deep Learning Approach to Predict Supply Chain Delivery Delay Risk Based on Macroeconomic Indicators: A Case Study in the Automotive Sector", *Applied Sciences (Switzerland)*, Multidisciplinary Digital Publishing Institute (MDPI), Vol. 14 No. 11, <https://doi.org/10.3390/app14114688>
12. Garcia-Buendia, N., Moyano-Fuentes, J., Maqueira-Marín, J.M. and Cobo, M.J. (2021), "22 Years of Lean Supply Chain Management: a science mapping-based bibliometric analysis", *International Journal of Production Research*, Taylor and Francis Ltd., <https://doi.org/10.1080/00207543.2020.1794076>
13. Gong, Y., Liu, G., Xue, Y., Li, R. and Meng, L. (2023), "A survey on dataset quality in machine learning", *Information and Software Technology*, Vol. 162, p. 107268, doi: <https://doi.org/10.1016/j.infsof.2023.107268>
14. Gyulai, D., Pfeiffer, A., Nick, G., Gallina, V., Sihm, W. and Monostori, L. (2018), "Lead time prediction in a flow-shop environment with analytical and machine learning approaches", *IFAC-PapersOnLine*, Vol. 51 No. 11, pp. 1029–1034, doi: <https://doi.org/10.1016/j.ifacol.2018.08.472>
15. İfraz, M., Aktepe, A., Ersöz, S. and Çetinyokuş, T. (2023), "Demand forecasting of spare parts with regression and machine learning methods: Application in a bus fleet", *Journal of Engineering Research*, Vol. 11 No. 2, p. 100057, doi: <https://doi.org/10.1016/j.jer.2023.100057>
16. Ivanov, A. and Jaff, T. (2017), "Manufacturing Lead Time Reduction and Its Effect on Internal Supply Chain", in Campana, G., Howlett, R.J., Setchi, R. and Cimatti, B. (Eds.), *Sustainable Design and Manufacturing 2017*, Springer International Publishing, Cham, pp. 398–407, DOI:10.1007/978-3-319-57078-5\_38
17. Jayanti, L.P.S.D. and Wasesa, M. (2022), "Application of Predictive Analytics To Improve The Hiring Process In A Telecommunications Company", *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer*

*Dan Teknologi Informasi*, Universitas Islam Negeri Sultan Syarif Kasim Riau, Vol. 8 No. 1, DOI: <http://dx.doi.org/10.24014/-coreit.v8i1.16915>

18. Kumar, C. and Singh, B. (2022), "A Comparative Study of Machine Learning Regression Approach on Dental Caries Detection", *Procedia Computer Science*, Vol. 215, pp. 519–528, <https://doi.org/10.1016/j.procs.2022.12.054>
19. Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihm, W. and Monostori, L. (2018), "Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer", *Procedia CIRP*, Vol. 72, pp. 1051–1056, <https://doi.org/10.1016/j.procir.2018.03.148>
20. Maware, C. and Parsley, D.M. (2023), "Can Industry 4.0 Assist Lean Manufacturing in Attaining Sustainability over Time? Evidence from the US Organizations", *Sustainability (Switzerland)*, MDPI, Vol. 15 No. 3, <https://doi.org/10.3390/su15031962>
21. Mohamed-Iliasse, M., Loubna, B. and Abdelaziz, B. (2022), "Machine Learning in Supply Chain Management: A Systematic Literature Review", *International Journal of Supply and Operations Management*, Kharazmi University, Vol. 9 No. 4, pp. 398–416, DOI:10.22034/ijssom.2021.109189.2279
22. de Oliveira, M.B., Zucchi, G., Lippi, M., Cordeiro, D.F., da Silva, N.R. and Iori, M. (2021), "Lead Time Forecasting with Machine Learning Techniques for a Pharmaceutical Supply Chain", *International Conference on Enterprise Information Systems, ICEIS - Proceedings*, Vol. 1, Science and Technology Publications, Lda, pp. 634–641, DOI:10.5220/0010434406340641
23. Öztürk, A., Kayalığıl, S. and Özdemirel, N.E. (2006), "Manufacturing lead time estimation using data mining", *European Journal of Operational Research*, Vol. 173 No. 2, pp. 683–700, <https://doi.org/10.1016/j.ejor.2005.03.015>
24. Pfeiffer, A., Gyulai, D., Kádár, B. and Monostori, L. (2016), "Manufacturing Lead Time Estimation with the Combination of Simulation and Statistical Learning Methods", *Procedia CIRP*, Vol. 41, pp. 75–80, doi: <https://doi.org/10.1016/j.procir.2015.12.018>
25. Pozzi, R., Cannas, V.G. and Ciano, M.P. (2022), "Linking data science to lean production: a model to support lean practices", *International Journal of Production Research*, Taylor and Francis Ltd., Vol. 60 No. 22, pp. 6866–6887, <https://doi.org/10.1080/00207543.2021.1946192>
26. Qureshi, K.M., Mewada, B.G., Kaur, S. and Qureshi, M.R.N.M. (2023), "Assessing Lean 4.0 for Industry 4.0 Readiness Using PLS-SEM

- towards Sustainable Manufacturing Supply Chain", *Sustainability (Switzerland)*, MDPI, Vol. 15 No. 5, <https://doi.org/10.3390/su15053950>
27. Rahman, Md.S., Ghosh, T., Aurna, N.F., Kaiser, M.S., Anannya, M. and Hosen, A.S.M.S. (2023), "Machine learning and internet of things in industry 4.0: A review", *Measurement: Sensors*, Vol. 28, p. 100822, doi: <https://doi.org/10.1016/j.measen.2023.100822>
28. Rokoss, A., Syberg, M., Tomidei, L., Hülsing, C., Deuse, J. and Schmidt, M. (2024), "Case study on delivery time determination using a machine learning approach in small batch production companies", *Journal of Intelligent Manufacturing*, <https://doi.org/10.1007/s10845-023-02290-2>
29. Schneckenreither, M., Haeussler, S. and Gerhold, C. (2021), "Order release planning with predictive lead times: a machine learning approach", *International Journal of Production Research*, Taylor & Francis, Vol. 59 No. 11, pp. 3285–3303, <https://doi.org/10.1080/00207543.2020.1859634>
30. Shapi, M.K.M., Ramli, N.A. and Awal, L.J. (2021), "Energy consumption prediction by using machine learning for smart building: Case study in Malaysia", *Developments in the Built Environment*, Vol. 5, p. 100037, <https://doi.org/10.1016/j.dibe.2020.100037>
31. Singh, S. and Soni, U. (2019), "Predicting Order Lead Time for Just in Time production system using various Machine Learning Algorithms: A Case Study", *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 422–425, DOI: 10.1109/CONFLUENCE.2019.8776892
32. Steinberg, F., Burggräf, P., Wagner, J., Heinbach, B., Saßmannshausen, T. and Brintrup, A. (2023), "A novel machine learning model for predicting late supplier deliveries of low-volume-high-variety products with application in a German machinery industry", *Supply Chain Analytics*, Vol. 1, p. 100003, <https://doi.org/10.1016/j.sca.2023.100003>
33. Tao, S. (2023), Predicting BMW Stock Price Based on Linear Regression, LSTM, and Random Forest Regression, *BCP Business & Management EMFRM*, Vol. 2022, DOI: <https://doi.org/10.54691/bcpbm.v38i.3712>
34. Welsing, M., Maetschke, J., Thomas, K., Gützlaff, A., Schuh, G. and Meusert, S. (2021), "Combining Process Mining and Machine Learning for Lead Time Prediction in High Variance Processes", in Behrens, B.-A., Brosius, A., Hintze, W., Ihlenfeldt, S. and Wulfsberg, J.P. (Eds.), *Production at the Leading Edge of Technology*, Springer

- Berlin Heidelberg, Berlin, Heidelberg, pp. 528–537
35. Yang, M., Lim, M.K., Qu, Y., Ni, D. and Xiao, Z. (2023), "Supply chain risk management with machine learning technology: A literature review and future research directions", *Computers & Industrial Engineering*, Vol. 175, p. 108859, <https://doi.org/10.1016/j.cie.2022.108859>

**استناد به این مقاله:** زمانی، فائزه، ابراهیمی، احمد، سلطانی رؤیا، فرهنگ مقدم، بابک. (۱۴۰۴). پیش‌بینی زمان انتظار سفارشات قطعات خودرو در زنجیره تأمین با استفاده از یادگیری ماشین، مطالعات مدیریت کسب و کار هوشمند، ۱۳(۵۲)، ۸۷-۱۲۳. DOI: 10.22054/ims.2025.80935.2491



Journal of Business Intelligence Management Studies is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License..