



# Development of a Deep Reinforcement Learning Algorithm in a Dynamic Cellular Manufacturing System Considering Order Rejection, Case Study: Stone Paper Factory

Mostafa Jafari<sup>1</sup>, Amir Hossein Akbari<sup>\*2</sup>

1. Associate Professor, Department of Industrial Engineering, Iran University of Science & Technology, Tehran, Iran, Email: jafari@iust.ac.ir
2. Corresponding Author, PhD Student, Department of Industrial Engineering, Iran University of Science & Technology, Tehran, Iran, Email: akbari\_amir@ind.iust.ac.ir

Article Info	ABSTRACT
<p><b>Article type:</b> Research Article</p> <p><b>Article history:</b> Received 7 - 12 - 2024 Received in revised form 7 - 01 - 2025 Accepted 18 - 2 - 2025 Published online 18 - 3 - 2025</p> <p><b>Keywords:</b> Deep reinforcement learning, cellular manufacturing system, order acceptance and scheduling, genetic algorithm</p>	<p>In this research, a deep reinforcement learning algorithm is proposed for the cellular manufacturing system problem considering the costs of delay and rejection of orders. Orders with different characteristics including revenue, lead time, delivery date, and delay cost are dynamically entered into the system at different times. Due to the limited capacity of the system, it is not possible to accept all orders and some of them must be rejected at the time of entry to enable timely execution of other orders. A mathematical model with two objectives of maximizing profit and minimizing the number of rejected orders is presented and a deep reinforcement learning algorithm is used to solve this problem. The proposed algorithm is compared with the algorithms available in the literature in different categories of example problems and real problems and its efficiency is proven. The results show a 36.3% advantage in profit and 13.87% in the number of accepted orders. Also, by accepting 1% more orders, the profit decreases by 2.7% on average. By accepting more orders, the decision maker can attract more customers and increase the system's competitive advantage in the long term.</p>

Cite this article: .....



© The Author(s)  
DOI: <https://doi.org/>

Publisher: University of Qom

## توسعه الگوریتم یادگیری تقویتی عمیق در سیستم تولید سلولی پویا با در نظر گرفتن رد سفارشات، مطالعه موردی: کارخانه کاغذ سنگی

مصطفی جعفری<sup>1</sup>، امیر حسین اکبری<sup>2\*</sup>

1. دانشیار، دانشکده مهندسی صنایع، دانشگاه علم و صنعت، تهران، ایران؛ رایانامه: jafari@iust.ac.ir  
2. نویسنده مسئول، دانشجوی دکتری، دانشکده مهندسی صنایع، دانشگاه علم و صنعت، تهران، ایران؛ رایانامه: akbari\_amir@ind.iust.ac.ir

اطلاعات مقاله	چکیده
<b>نوع مقاله:</b> مقاله پژوهشی <b>تاریخ دریافت:</b> 1403/09/17 <b>تاریخ بازنگری:</b> 1403/10/17 <b>تاریخ پذیرش:</b> 1403/11/30 <b>تاریخ انتشار:</b> 1403/12/30	در این تحقیق، یک الگوریتم یادگیری تقویتی عمیق برای مسئله سیستم تولید سلولی با در نظر گرفتن هزینه‌های تأخیر و رد سفارشات پیشنهاد شده است. سفارشات با ویژگی‌های مختلف شامل در آمد، زمان انجام، موعد تحویل و هزینه تأخیر به صورت پویا و در زمان‌های مختلف وارد سیستم می‌شوند. با توجه به ظرفیت محدود سیستم، امکان پذیرش تمامی سفارشات وجود ندارد و برخی از آنها باید در زمان ورود رد شوند تا امکان انجام به موقع سایر سفارشات فراهم شود. یک مدل ریاضی با دو هدف بیشینه‌سازی سود و کمینه‌سازی تعداد سفارشات رد شده ارائه شده است و برای حل این مسئله، از یک الگوریتم یادگیری تقویتی عمیق استفاده شده است. الگوریتم پیشنهادی در دسته‌های مختلفی از مسائل نمونه‌ای و مسائل واقعی با الگوریتم‌های موجود در ادبیات مقایسه شده و کارایی آن به اثبات رسیده است. نتایج نشان‌دهنده برتری 36.3 درصدی در سود و 13.87 درصدی در تعداد سفارشات پذیرفته شده است. همچنین، با پذیرش 1 درصد سفارش بیشتر، میزان سود به‌طور متوسط 2.7 درصد کاهش می‌یابد. تصمیم‌گیرنده می‌تواند با پذیرش سفارشات بیشتر، مشتریان بیشتری جذب کرده و مزیت رقابتی سیستم را در بلندمدت افزایش دهد.
<b>کلیدواژه‌ها:</b> یادگیری تقویتی عمیق، سیستم تولیدی سلولی، پذیرش و زمان بندی سفارشات، الگوریتم ژنتیک	
<b>استناد:</b> .....	

## 1) مقدمه

سازمان‌ها امروزه با محیط‌های پویایی مواجه هستند که در آن تغییرات سریع در خواسته‌های مشتریان امری معمول است. برای اینکه بنگاه‌ها بتوانند در بازار رقابتی با سایر شرکت‌ها دوام بیاورند، ناگزیر به بازنگری در سیستم تولید خود هستند. در سیستم‌های پویا، هر دوره می‌تواند با تقاضاهای متفاوتی همراه باشد. سیستم تولید سلولی (Cellular Manufacturing System) توانایی مقابله با بسیاری از ضعف‌های موجود در سیستم‌های سنتی را دارد و می‌تواند پاسخگوی نیازهای یک محیط تولید پویا باشد (هوشیار و همکاران 2014). استفاده از سیستم تولید سلولی مزایای متعددی به همراه دارد که از جمله مهم‌ترین آن‌ها می‌توان به کاهش زمان و هزینه تولید، افزایش انعطاف‌پذیری، کنترل بهتر فرآیند تولید، کاهش هزینه‌های گردش مالی، و کاهش میزان کار در جریان اشاره کرد (رفیعی، ربانی و همکاران، 1395). در سیستم‌های تولید سلولی پویا (Dynamic CMS)، فاکتورهایی مانند تقاضا، زمان پردازش، و دسترسی ماشین‌ها می‌تواند دستخوش تغییر شوند (پاپایون و ویلسون 2020). این تحقیق به بررسی یک سیستم تولید سلولی پویا چندهدفه پرداخته که در آن هزینه تأخیر و رد سفارش لحاظ شده است. سفارشات در زمان‌های مختلف وارد سیستم می‌شوند و ممکن است برخی از آن‌ها پذیرفته نشوند. سفارش‌های دریافتی شامل ویژگی‌هایی همچون درآمد، هزینه تأخیر، زمان پردازش، و موعد تحویل هستند. همچنین، ماشین‌ها با اثر زوال مواجه‌اند و راندمان آن‌ها کاهش یافته، که این امر باعث افزایش زمان پردازش می‌شود. پذیرش تمامی سفارشات می‌تواند منجر به تأخیر در تحویل آن‌ها، نارضایتی مشتریان و در نهایت خروج آن‌ها از سیستم تولید شود (ژاو و یو، 2017). در عوض، تنها بخشی از سفارشات برای پردازش انتخاب می‌شوند. توجه به تصمیم‌گیری یکپارچه برای پذیرش سفارش و برنامه‌ریزی تولید، اخیراً اهمیت زیادی یافته است (اسلوتیک، 2011). در شرایطی که تعداد سفارشات پیشنهادی بیشتر از ظرفیت سیستم باشد، لازم است بخشی از آن‌ها پذیرفته و بخشی دیگر رد شوند تا سود کلی افزایش یابد (هربوت و همکاران، 2007). با این حال، رد سفارشات ممکن است به کاهش وفاداری مشتریان منجر شود. تصمیم به رد سفارشات به منظور کسب سود بیشتر معمولاً یک استراتژی کوتاه‌مدت محسوب می‌شود، اما برای رقابت با سایر شرکت‌ها، سازمان‌ها باید مشتریان بیشتری را جذب کنند. بنابراین، پذیرش تعداد بیشتری از سفارشات می‌تواند به‌عنوان یک هدف مهم در نظر گرفته شود. در این پژوهش، اهداف اصلی شامل بهینه‌سازی سود و افزایش تعداد سفارشات پذیرفته‌شده است. مسئله سیستم تولید سلولی پویا با رد سفارش به‌عنوان یک مسئله چندهدفه، در دسته مشکلات NP-hard قرار می‌گیرد. در اغلب موارد، برای حل چنین مسائل پیچیده‌ای، از روش‌های نوآورانه نظیر الگوریتم‌های فراابتکاری و قوانین اعزام استفاده می‌شود (لو، لیو و همکاران، 2012). قوانین اعزام می‌توانند برای واکنش به رویدادهای پویا به کار گرفته شوند، اما یکی از محدودیت‌های این روش‌ها عدم تضمین بهینه‌سازی محلی است. فرآیند تصمیم‌مارکوف (MDP) به‌عنوان ابزاری توانمند می‌تواند بر این محدودیت غلبه کند و در هر لحظه تصمیم‌گیری، بهترین توزیع قوانین را ارائه دهد. در این فرآیند، یک عامل هوشمند وجود دارد که می‌تواند پس از وقوع تغییرات یا اختلالات، با تحلیل وضعیت موجود سیستم تولید، مناسب‌ترین اقدام را انتخاب کند. از روش‌های قدرتمند برای پیاده‌سازی فرآیند تصمیم‌مارکوف، یادگیری تقویتی (Reinforcement Learning) است که قابلیت حل مسائل مختلفی از جمله پذیرش سفارش، زمان‌بندی پویا و مدیریت سیستم تولید سلولی پویا را داراست (دینگ و همکاران، 2020).

در این پژوهش، یک مدل ریاضی معرفی شده که اهداف آن شامل بیشینه‌سازی سود (تعریف شده به‌عنوان درآمد حاصل از سفارشات پذیرفته‌شده منهای هزینه‌های تأخیر و جابجایی بین‌سلولی و درون‌سلولی) و بیشینه‌سازی تعداد سفارشات پذیرفته‌شده است. برای حل این مسئله، از روشی مبتنی بر یادگیری تقویتی عمیق استفاده شده است. این روش، تصمیم‌گیری در مورد پذیرش سفارشات و زمان‌بندی عملیات آن‌ها را با لحاظ اثر زوال ماشین‌ها امکان‌پذیر می‌سازد. همچنین، عملکرد این روش با استفاده از داده‌های واقعی یک کارخانه کاغذ سنگی در شهر گیلان (ایران) مورد ارزیابی قرار گرفته است. نوآوری‌های این پژوهش شامل موارد زیر است:

1. ارائه مسئله سیستم تولید سلولی پویا با اهداف کوتاه‌مدت (بیشینه‌سازی سود) و بلندمدت (بیشینه‌سازی تعداد سفارشات پذیرفته‌شده)، با در نظر گرفتن رد سفارشات، هزینه‌های تأخیر، و اثر زوال سفارشات که در زمان‌های مختلف وارد سیستم می‌شوند و ممکن است برخی از آن‌ها رد شوند.
  2. توسعه یک روش یادگیری تقویتی عمیق برای حل مسئله پیشنهاد شده، شامل تصمیم‌گیری در مورد پذیرش سفارشات و زمان‌بندی عملیات آن‌ها با در نظر گرفتن اثر زوال.
  3. تحلیل عملکرد روش پیشنهادی با استفاده از یک مطالعه موردی واقعی مرتبط با یک کارخانه کاغذ سنگی در گیلان (ایران)، جهت مقایسه نتایج داده‌های واقعی و روش ارائه‌شده.
- ادامه تحقیق به شرح زیر است. بخش 2، مروری جامع از ادبیات مرتبط با مشکل ارائه می‌کند. بخش 3، یک مدل ریاضی برای مسئله ایجاد می‌کند. بخش 4، یک الگوریتم فراابتکاری برای حل آن طراحی می‌کند. بخش 5، مسئله را از طریق مثال عددی تحلیل می‌کند. بحث، نتیجه‌گیری و مراجع به ترتیب در بخش‌های ششم و هفتم بررسی می‌شوند.

## 2) پیشینه تحقیق

این بخش مطالعات انجام شده در ادبیات تحقیق را برای مسائل سیستم تولید سلولی، پذیرش و زمان‌بندی سفارشات و الگوریتم یادگیری تقویتی عمیق را ارائه می‌کند.

سیستم تولید سلولی به‌عنوان یکی از سیستم‌های کارآمد در صنعت امروزی، در سال‌های اخیر نقش برجسته‌ای ایفا کرده است. جبل عاملی و همکاران (2008) مسئله تشکیل سلول را با در نظر گرفتن قابلیت اطمینان دستگاه‌ها مطرح کرده و یک مدل برنامه‌ریزی خطی برای حل آن ارائه کردند. مهدوی و همکاران (1389) مسائل مرتبط با تشکیل سلول و برنامه‌ریزی تولید را در یک محیط پویا و یکپارچه بررسی کردند. چانگ و همکاران (2011) یک سیستم تولید سلولی چندهدفه را با لحاظ خرابی ماشین‌ها معرفی کردند. اعلائی و داوودپور (1391) یک سیستم تولید سلولی پویا را در محیط زنجیره تأمین ارائه داده و یک مدل ریاضی برای آن طراحی کردند. همچنین، اعلائی و داوودپور (1395) با توجه به امکانات تولید و محدودیت‌های زمانی، یک سیستم تولید سلولی پویا را مطالعه و یک مدل ریاضی برای آن توسعه دادند. لیو و همکاران (2018) یک سیستم تولید سلولی پویا را در محیط زنجیره تأمین با هدف حداقل‌سازی هزینه‌های کل عملیات بررسی کرده و از الگوریتم گرگ خاکستری برای حل آن بهره گرفتند. ربانی و همکاران (2019) نیز یک سیستم تولید سلولی پویا را با در نظر گرفتن قابلیت اطمینان ماشین‌ها ارائه دادند و برای آن یک مدل چندهدفه جدید طراحی کردند.

چو و همکاران (2019) مسئله آموزش متقابل را با اثرات یادگیری و فراموشی مورد مطالعه قرار دادند. مطهری و همکاران (2023) با توجه به زمان بندی سفارشات خانوادگی، یک مسئله سیستم تولید سلولی را مطرح کرده و مدلی چندهدفه برای آن ارائه دادند. گلی و همکاران (2021) مسئله تشکیل سلول را با لحاظ زمان بندی سفارشات و عوامل انسانی بررسی کردند. شفیع گل و همکاران (2021) یک سیستم تولید سلولی پویا را با در نظر گرفتن زمان بندی سفارشات، برنامه ریزی تولید، و MRP پیشنهاد دادند.

مسئله پذیرش و زمان بندی سفارشات در پژوهش های مختلفی مورد بررسی قرار گرفته است. به عنوان نمونه، سلمان و همکاران (2010)، لین و یانگ (2013)، چن و همکاران (2014)، و ژانگ و همکاران (2014) به این موضوع پرداخته اند. زندیه و رومانی (1396) مسئله پذیرش سفارش را مورد بررسی قرار دادند، در حالی که گرامی پور و همکاران (1396) با در نظر گرفتن مسئله پذیرش و زمان بندی سفارشات، الگوریتم های اکتشافی و شاخه و کران را برای حل آن طراحی کردند. یو و ژانگ (2017) یک مسئله پذیرش و زمان بندی سفارشات دو معیاره را در محیط ماشین های موازی مطرح کردند و آن را با استفاده از الگوریتم های تقریبی تحلیل نمودند. وانگ و همکاران (2019) این مسئله را در یک برنامه ریزی تولید ترکیبی شامل تولید بر اساس سفارش و تولید برای انبار در محیط کارگاهی بررسی کردند. سیلوا و همکاران (2018) مدل ریاضی مبتنی بر شاخص زمان قوس را برای این مسئله ارائه داده و از الگوریتم های دقیق برای حل آن استفاده کردند. سروستانی و همکاران (2019) برای حداکثر سازی سود، به پذیرش و زمان بندی سفارشات همراه با انتخاب تأمین کننده پرداختند. یآوری و همکاران (2020) یک مسئله پذیرش و زمان بندی سفارشات دو مرحله ای مونتاژی را بررسی کرده و الگوریتم اکتشافی برای حل آن پیشنهاد دادند. رحمان و همکاران (2019) نیز یک مسئله پذیرش و زمان بندی سفارشات بلادرنگ را در محیط کارگاه جریان یافته مطالعه کرده و راه حل اکتشافی برای آن ارائه کردند. وانگ و همکاران (2020) چندین مسئله زمان بندی چندوظیفه ای را با در نظر گرفتن اثر زوال و ارتقای کارایی مطرح کردند. یآوری و اکبری (2023) یک مسئله پذیرش و زمان بندی سفارشات دوهدفه را تحلیل کردند که اهداف آن شامل حداکثر سازی سود و افزایش تعداد سفارشات پذیرفته شده (سطح خدمات) بود و برای حل آن از یک الگوریتم اکتشافی بهره گرفتند. آن و همکاران (2023) با لحاظ تعمیر و نگهداری پیشگیرانه، به بررسی مسئله پذیرش و زمان بندی سفارشات بلادرنگ در محیط کارگاه انعطاف پذیر پرداختند و الگوریتم تکاملی برای حل آن طراحی کردند. لی و همکاران (2023) نیز این مسئله را با توجه به برنامه ریزی تولید و حمل و نقل گسترش داده و الگوریتم های دقیق و تقریبی برای حل آن ارائه دادند.

محققان مختلف از یادگیری تقویتی و یادگیری تقویتی عمیق برای حل مسائل زمان بندی بهره برده اند. چن و همکاران (2010) مسئله ای مرتبط با زمان بندی پویا را با هدف حداکثر سازی سود مطرح کرده و روشی مبتنی بر قوانین برای حل آن طراحی کردند. سالز و همکاران (2017) یک مسئله زمان بندی را بررسی کرده و یک روش اکتشافی ارائه دادند که در آن محصولات هوشمند ماشین مناسب را انتخاب کرده و یادگیری تقویتی عمیق وظیفه برنامه ریزی سفارشات را انجام می داد. لنگ و همکاران (2023) به مطالعه مسئله پذیرش و زمان بندی سفارشات با هدف افزایش سود پرداخته و از یک الگوریتم یادگیری تقویتی دو گانه برای حل آن استفاده کردند. ژانگ و همکاران (2021) مسئله پذیرش و زمان بندی سفارشات را با لحاظ محدودیت ظرفیت و مهلت تحویل برای حداکثر سازی سود بررسی کرده و روشی بر پایه یادگیری

تقویتی عمیق توسعه دادند. مارکا و همکاران (2019) مسئله پذیرش و زمان‌بندی سفارشات را در یک سیستم تولید سفارشی، همراه با تعیین قیمت و زمان تحویل، معرفی کرده و از یادگیری تقویتی برای حل آن بهره گرفتند. مطالعات دیگری نیز در این زمینه صورت گرفته است که شامل مقالات میرحسینی و همکاران (2017)، شهرابی و همکاران (2017)، شیو و همکاران (2018)، و وانگ و هو (2020) می‌شود. با توجه به این پژوهش‌ها، یادگیری تقویتی عمیق به‌عنوان یک الگوریتم کارآمد برای تحلیل مسائل زمان‌بندی با ویژگی‌های مختلف (مانند رد سفارش و سایر خصوصیات) در محیط‌های گوناگون (مانند سیستم تولید سلولی، سیستم تولید کارگاهی، و سیستم‌های تولید خطی) شناخته شده است. در مطالعات مذکور، به ظرفیت محدود سیستم تولید و اهداف بلندمدت توجه نشده است و فرض بر این بوده که تمامی سفارشات در سیستم تولید سلولی پردازش می‌شوند. این رویکرد، که تنها بر اهداف کوتاه‌مدت متمرکز است، می‌تواند به افزایش نارضایتی مشتریان، هزینه‌های تأخیر، و کاهش وفاداری مشتریان منجر شود. در نظر گرفتن چنین مسئله‌ای با لحاظ هزینه تأخیر در سیستم تولید سلولی می‌تواند منجر به افزایش رضایت مشتری، ارتقای سود و تضمین بقای سازمان شود. یادگیری تقویتی عمیق قادر است تصمیم‌گیری در سیستم‌های تولید را بهبود بخشد و به‌عنوان یک روش مؤثر برای دستیابی به اهداف بلندمدت در این زمینه مورد استفاده قرار گیرد.

### (3) روش تحقیق

یک سیستم تولید سلولی پویا با سلول‌ها و ماشین‌های متعدد در نظر گرفته شده است که در آن تصمیماتی درباره تشکیل سلول‌ها و تخصیص ماشین‌ها به سلول‌ها باید اتخاذ شود. عملکرد ماشین‌ها به مرور زمان به دلیل اثر زوال کاهش می‌یابد، که این موضوع منجر به افزایش زمان پردازش عملیات می‌شود. سفارشات جدید به‌صورت متوالی و پویا وارد سیستم می‌شوند و نیاز به تصمیم‌گیری درباره پذیرش یا رد آن‌ها وجود دارد. هر سفارش پذیرفته شده شامل چندین عملیات است که باید به‌طور کامل پردازش شود. انتخاب مناسب‌ترین ماشین برای انجام هر عملیات یکی از تصمیمات کلیدی در این سیستم است. هر سفارش دارای ویژگی‌هایی نظیر درآمد، هزینه تأخیر، و تاریخ سررسید است، و هر عملیات زمان پردازش متفاوتی بر روی ماشین‌های مختلف دارد. اهداف این مسئله عبارتند از: حداکثرسازی سود، که شامل درآمد سفارشات پذیرفته شده منهای هزینه تأخیر و هزینه‌های مرتبط با انتقال بین سلولی و درون سلولی است، و همچنین حداکثرسازی تعداد سفارشات پذیرفته شده، به‌طوری که تعداد سفارشات رد شده به حداقل برسد.

### (3.1) مدل ریاضی

در این قسمت نمادها، پارامترهای مسئله و متغیرهای تصمیم ارائه شده و مدل ریاضی ارائه شده است.

جدول (1) اندیس‌ها، پارامترها و متغیرها

اندیس‌ها		پارامترها	
$c$	سلول $c=1 \dots C$	$j_p$	عملیات های سفارشات $p, j=1 \dots n_i$
$p$	سفارش $p=1 \dots P$	$t$	زمان $t=1 \dots T$
$m$	ماشین $m=1 \dots M$		
$R_p$	درآمد حاصل از انجام سفارش $p$	$C_p^T$	هزینه تاخیر سفارش $p$

$C^{in}$	هزینه جابجایی درون سلولی	$C^{out}$	هزینه جابجایی برون سلولی
$A_{jpm}$	توانایی انجام ماشین $m$ برای عملیات $j$ قطعه $p$	$t_{jpm}$	زمان انجام عملیات $j$ سفارش $p$ در ماشین
$d_p$	موعد تحویل سفارش $p$		
متغیرها			
$N_{mc}$	منغیر باینری اگر ماشین به سلول تخصیص داده شود	$X_{jpmct}$	اگر عملیات $j$ سفارش $p$ در ماشین $m$ در سلول $c$ در دوره $t$ انجام شود
$XC_{out}$	هزینه جابجایی برون سلولی	$TF_p$	هزینه تاخیر سفارش $p$
$XC_{in}$	هزینه جابجایی درون سلولی	$COT_{jp}$	زمان اتمام عملیات $j$ سفارش $p$
$ST_{jp}$	زمان شروع عملیات $j$ سفارش $p$	$Y_p$	متغیر باینری اگر سفارش $p$ پذیرفته شود

بعد از ارائه اندیس‌ها، متغیرها و پارامترها، مدل ریاضی مسئله ارائه می‌شود.

$$\text{Max } Z1 = \sum_p R_p Y_p - \sum_p C_p^t TF_p - C_{in} XC_{in} - C_{out} XC_{out}$$

$$\text{Max } Z2 = \sum_p Y_p$$

$$\sum_{c,t} X_{jpmct} \leq A_{jpm} \quad \forall j, p, m \quad (1)$$

$$\sum_{j,p,t} X_{jpmct} \leq N_{mc}, \quad N_{mc} \leq 1 \quad \forall m, c, t \quad (2)$$

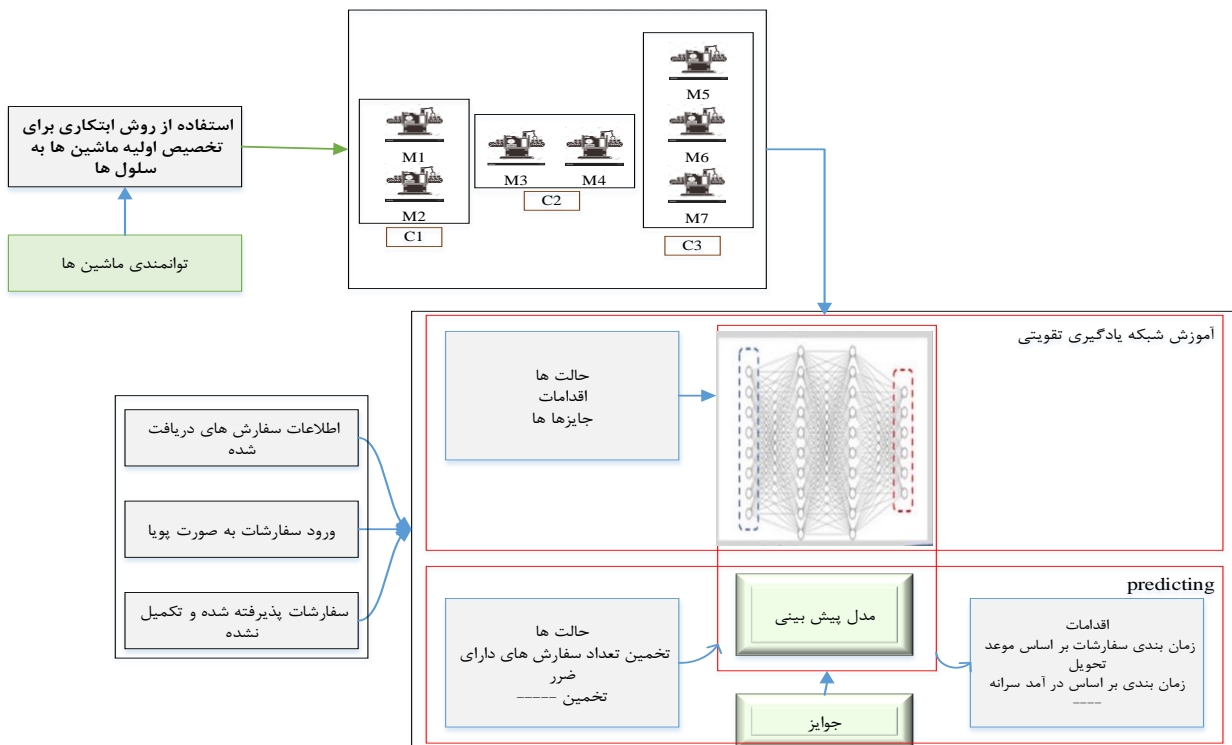
$$COT_{jp} = ST_{jp} + \sum_{m,c,t} X_{jpmct} \times t_{jpm} \times (1 + \alpha_{mt}) \quad \forall j, p \quad (3)$$

$$TF_p = \max(COT_p - d_p, 0) \quad \forall j, p \quad (4)$$

تابع هدف اول تابع هدف اول، سود را به حداکثر می‌رساند، جایی که سود، درآمد سفارش‌های پذیرفته شده منهای هزینه تأخیر و هزینه‌های بین سلولی و درون سلولی است. تابع هدف دوم تعداد سفارشات پذیرفته شده را به حداکثر می‌رساند. رد سفارشات برخی از مشتریان به دلیل محدودیت‌های داخلی سیستم تولید باعث دلسردی آنها از سیستم تولید و عدم مراجعه آنها به سازمان در آینده می‌شود. محدودیت اول برای تخصیص عملیات‌ها به ماشین‌هایی است که توانایی انجام آن را دارند. محدودیت دوم برای تخصیص مناسب ماشین‌ها به سلول‌ها است. محدودیت سوم زمان شروع و پایان عملیات‌ها را محاسبه می‌کند. محدودیت چهارم نیز برای محاسبه زمان تأخیر محدودیت‌ها است.

### 3.2 الگوریتم حل

این بخش یک رویکرد ابتکاری برای حل مسئله ارائه می‌دهد. در مرحله نخست، ماشین‌ها با استفاده از یک روش اکتشافی به سلول‌های مشخصی تخصیص داده می‌شوند. در مرحله دوم، یک الگوریتم یادگیری تقویتی عمیق طراحی شده است که وظایف زیر را انجام می‌دهد: زمان‌بندی سفارش‌های ورودی، اختصاص عملیات سفارش‌های پذیرفته شده به مناسب‌ترین ماشین‌ها، انتخاب بهینه‌ترین سفارش‌ها برای پذیرش، و زمان‌بندی مجدد عملیات سفارش‌های باقی‌مانده.



شکل 1 رویه روش حل

### 3.2.1 روش ابتکاری برای تخصیص ماشین به سلول‌ها

ابتدا میزان توانایی هر ماشین برای انجام کارها، با استفاده از فرمول زیر محاسبه می‌شود.

$$AT_m = \sum_{p,m} A_{jpm} \quad (5)$$

سپس ماشین با توانایی زیاد به صورت تصادفی به سلول‌ها تخصیص داده می‌شود. بعد از اینکه هر سلول یک ماشین دریافت کرد، مجدد از کمترین توانایی به همان ترتیب قبلی، ماشین‌ها به سلول‌ها تخصیص داده می‌شوند. این رویه تا تخصیص تمامی ماشین‌ها به سلول‌ها ادامه دارد.

### 3.2.2 الگوریتم یادگیری تقویتی

این الگوریتم با هدف زمان‌بندی سفارش‌های ورودی، تخصیص عملیات سفارش‌های پذیرفته‌شده به مناسب‌ترین ماشین‌ها، پذیرش سفارش‌های بهینه و باززمان‌بندی عملیات سفارش‌های باقی‌مانده پیشنهاد شده است. الگوریتم یادگیری تقویتی توانایی حل مسائل متنوع با تصمیم‌گیری‌های متوالی را به‌طور کارآمد دارد (اوترلو و ویرینگ 2012). در این الگوریتم، وضعیت‌های مختلف تعریف می‌شوند، برای هر وضعیت اقدامات ممکن مشخص می‌گردند و به هر اقدام یک مقدار پاداش اختصاص داده می‌شود.

#### حالت‌ها (St)

در سیستم‌های تولید، ویژگی‌هایی مانند تعداد ماشین‌ها، تعداد شغل‌ها یا عملیات، و موارد مشابه را می‌توان به‌عنوان مشخصه‌های حالت در الگوریتم‌های یادگیری تقویتی تعریف کرد (شهرابی و همکاران 2017، شیو و همکاران 2018). این مقاله شامل تصمیم‌گیری‌های مختلفی است، از جمله پذیرش یا رد سفارش‌ها، انتخاب ماشین‌ها، و زمان‌بندی سفارش‌ها.



بنابراین، پنج حالت برای توصیف جزئیات سیستم تولید در نظر گرفته شده است. مقدار هر حالت در بازه  $[0,1]$  قرار دارد (لو 2020). جدول زیر، حالت‌های مرتبط با این مسئله را ارائه می‌دهد.

### جدول 2) حالت‌ها و جزئیات آنها

ردیف	حالت	توضیحات
1	تخمین تعداد سفارش‌هایی که هزینه تأخیر آنها بیشتر از درآمد آنها است.	تخمین تعداد سفارش‌هایی که هزینه تأخیر آنها بیشتر از درآمد آنها است به شناسایی سفارش‌هایی که باید رد شوند کمک می‌کند.
2	تخمین تعداد سفارش‌هایی که هزینه تأخیر بیش از 80 درصد درآمد است.	هزینه‌های دیگری مانند جابجایی درون سلولی و برون سلولی و ... وجود دارد که این حالت همان حالت 1 است اما نرخ سفارش‌هایی که هزینه تأخیر آن‌ها بیش از 80 درصد سود آن سفارش‌ها باشد محاسبه می‌شود.
3	تعداد واقعی سفارش‌هایی که هزینه تأخیر بیشتر از درآمد است.	تعداد واقعی سفارش‌هایی که هزینه تأخیر آنها بیشتر از سود آنها است به تعیین وضعیت سیستم و تصمیم‌گیری بهتر در مورد پذیرش یا رد سفارشات جدید کمک می‌کند.
4	تعداد واقعی سفارش‌هایی که هزینه تأخیر بیش از 0.8 درآمد است.	هزینه‌های دیگری مانند جابجایی درون سلولی و برون سلولی و ... نیز وجود دارد که این حالت مشابه حالت قبل است، میزان سفارشات که هزینه تأخیر آنها بیش از 80 درصد درآمد محاسبه می‌شود.
5	نرخ تکمیل عملیات‌ها	درصد سفارشات پذیرفته شده و عملیات‌های انجام شده آنها

### اقدام‌ها (At)

برای بهینه‌سازی وضعیت‌های مختلف تولید، لازم است قوانین متعددی اعمال شوند تا بهترین پاسخ ممکن برای تمامی زیرمسئله‌های مسئله اصلی به دست آید. در این مقاله، سیستم تولید سلولی پویا شامل سه زیرمسئله است:

- پذیرش یا رد سفارش‌های دریافتی

- زمان‌بندی سفارشات پذیرفته شده (که هر سفارش شامل عملیات مختلف است)

- تخصیص عملیات سفارشات پذیرفته شده به بهترین ماشین

تصمیم‌گیرنده باید بهترین اقدام را برای هر یک از این زیرمسئله‌ها تعیین کند. در رویکرد یادگیری تقویتی عمیق پیشنهادی، چهار اقدام برای اتخاذ بهترین تصمیمات ارائه شده است. تمامی این قوانین با هدف کاهش هزینه کل و افزایش تعداد سفارشات پذیرفته شده طراحی شده‌اند. در مدل پیشنهادی، ماشین‌های مختلفی وجود دارند و زمان پردازش عملیات در هر ماشین متفاوت است. در نتیجه  $\bar{t}_{jp}$  را به عنوان زمان پردازش یک عملیات در همه ماشین‌ها در نظر می‌گیریم که  $\bar{t}_{jp} = \frac{\sum m t_{jpm}}{M}$ . همچنین به دلیل اثر زوال ماشین‌ها،  $\bar{t}_{jp}$  مطابق با آن اصلاح می‌شود. اقدام‌های مسئله به شرح زیر است.

### جدول 3) اقدام‌ها و جزئیات آنها

ردیف	اقدام	توضیحات
1	زمان بندی سفارشات دریافت شده بر حسب درآمد واحد	درآمد سفارش دریافتی بر مجموع زمان مورد نیاز برای پردازش آنها تقسیم، و درآمد واحد هر سفارش محاسبه می‌شود. سپس سفارشات از بیشتری درآمد به کمترین درآمد زمان بندی می‌شوند. سپس سود هر سفارش در زمان بندی محاسبه می‌شود. سپس با بیشترین ضرر حذف می‌شود و سود هر سفارش مجدد محاسبه می‌شود. این رویه تا حذف تمامی سفارشات دارای ضرر تکرار می‌یابد. سپس عملیات‌های هر سفارش ب حسب ترتیب سفارشات، به اولین ماشین خالی تخصیص داده می‌شود.

زمان بندی	سفارشات از زودترین موعد تحویل به دیرترین موعد تحویل چیدمان می شوند. سپس سود هر سفارش در زمان
2	بندهای محاسبه می شود. سپس با بیشترین ضرر حذف می شود. و سود هر سفارش مجدد محاسبه می شود این رویه تا حذف تمامی سفارشات دارای ضرر تکرار می یابد. سپس عملیات های هر سفارش ب حسب ترتیب سفارشات، به اولین ماشین خالی تخصیص داده می شود.
زمان بندی	درآمد سفارش دریافتی بر مجموع زمان مورد نیاز برای پردازش آنها تقسیم، و درآمد واحد هر سفارش محاسبه می
3	شود. سپس سفارشات از بیشتری درآمد به کمترین درآمد زمان بندی می شوند. سپس سود هر سفارش در زمان بندی محاسبه می شود. سپس یک سفارش ضرر ده به صورت تصادفی حذف می شود و سود هر سفارش مجدد محاسبه می شود. این رویه تا حذف تمامی سفارشات دارای ضرر تکرار می یابد. سپس عملیات های هر سفارش بر حسب ترتیب سفارشات، به اولین ماشین خالی تخصیص داده می شود.
زمان بندی	سفارشات از زودترین موعد تحویل به دیرترین موعد تحویل چیدمان می شوند. سپس سود هر سفارش در زمان
4	بندهای محاسبه می شود. سپس یک سفارش ضرر ده به صورت تصادفی حذف می شود و سود هر سفارش مجدد محاسبه می شود. این رویه تا حذف تمامی سفارشات دارای ضرر تکرار می یابد. سپس عملیات های هر سفارش بر حسب ترتیب سفارشات، به اولین ماشین خالی تخصیص داده می شود.

نحوه انتخاب اقدام ها به صورت تصادفی است.

### جایزه ها (Rt)

میزان جایزه برابر با اختلاف حالت جدید با حالت قبلی تابع هدف وزنی مسئله پیشنهادی است.

$$r_t = (1 - \beta) \frac{obj1(t+1) - obj1(t)}{obj1(t+1)} + \beta \frac{obj2(t+1) - obj2(t)}{P(t+1)}$$

$$\beta = \frac{\text{شده رد سفارشات تعداد}}{\text{شده دریافت سفارشات کل تعداد}}$$

### ساختار شبکه

شبکه عصبی عمیق مورد استفاده در این تحقیق دارای 3 لایه پنهان علاوه بر لایه های ورودی و خروجی است. لایه ورودی شامل حالت ها است، لایه خروجی شامل اقدامات بهینه است. همچنین در هر لایه مخفی 5 گره وجود دارد. تابع فعال سازی در لایه های ورودی، مخفی و خروجی به ترتیب "sotmax"، "tanh"، "tansig" هستند. تعداد دوره های تکرار برای آموزش برابر با 5000 عدد، نرخ یادگیری از 10-2 تا 10-3، نرخ  $\gamma$  برابر با 0.9 و اندازه بافر N نیز برابر با 1000 است.

### چارچوب کلی روش آموزشی

نقطه تصمیم t به عنوان رسیدن سفارش جدید یا تکمیل یک عملیات تعریف می شود. تکنیک آموزشی استفاده شده بر اساس (لو 2020) است.

### 5 یافته ها

این بخش عملکرد الگوریتم پیشنهادی را نشان می دهد و پارامترهای مختلف در الگوریتم را بررسی می کند.

#### 5.1 تولید داده

در این بخش، چندین نمونه مسئله مختلف برای بررسی عملکرد مدل و روش حل پیشنهادی بررسی شده است. جدول 4 اندازه های مختلف نمونه های مشکل را نشان می دهد.

جدول 4) جزئیات ابعاد مورد برای بررسی کارایی الگوریتم پیشنهادی

عملیات ها	ماشین	سفارشات	سلول ها	ابعاد کوچک
2	5	500	4	ابعاد کوچک
6	10	1000	8	ابعاد متوسط
10	15	2000	12	ابعاد بزرگ

برخی از داده ها (درآمد، موعد تحویل، زمان پردازش عملیات ها و نرخ رسیدن سفارش) به شرح زیر معرفی می شوند.

- درآمد به طور تصادفی از توزیع لگ نرمال با میانگین 0 و انحراف استاندارد 1 ایجاد شد است (بیکوب 2012)
- زمان پردازش و هزینه تاخیر به صورت تصادفی و در بازه  $[2, 12]$  و  $[5, 15]$ ، ایجاد شده اند (دلگشایی و علی 2020، رحیمی و همکاران 2020).
- موعد تحویل از توزیع یکنواخت و از تابع  $[(P_T)(1 - \tau + \frac{R}{2}), (P_T)(1 - \tau - \frac{R}{2})]$  تولید شدند. در این تابع مقدار  $P_T = \sum_{j,p,m} time_{jpm}$  است و مقدار  $\tau$  یک عدد رندوم در بازه  $[0.3, 0.9]$  است (تورس و همکاران 2013).
- نرخ دریافت سفارشات از کارخانه تولید کاغذ سنگی الهام گرفته شده است.

## 5.2 نتایج محاسباتی

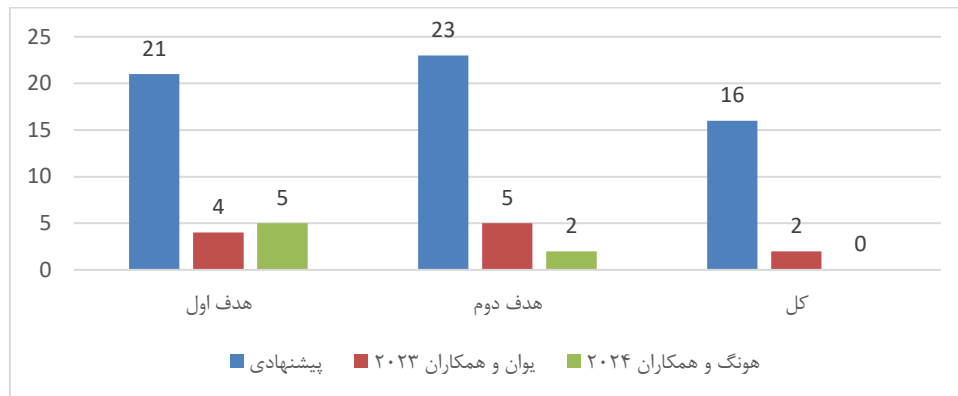
مسئله سیستم تولید سلولی پویا با استفاده از الگوریتم یادگیری تقویتی عمیق تاکنون در تحقیقات پیشین بررسی نشده است. برای ارزیابی الگوریتم پیشنهادی، مسائل مشابه در این حوزه با روش های پیشنهادی ترکیب شده و مورد استفاده قرار گرفته است. با حذف تخصیص ماشین ها به سلول ها و جابجایی های بین سلولی و درون سلولی، مسئله پیشنهادی به یک سیستم تولید کارگاهی (JSSP) با رد سفارشات شباهت پیدا می کند. بنابراین، الگوریتم پیشنهادی با روش های (یوان و همکاران 2023) و (هونگ و همکاران 2024) مقایسه شده است. در هر دو الگوریتم پیشنهادی، نحوه تخصیص ماشین ها به سلول ها مشابه با الگوریتم پیشنهادی تنظیم شده است. 30 نمونه مسئله در 10 نوبت با تمامی الگوریتم ها اجرا و نتایج آن در جدول 5 ارائه شده است.

جدول شماره 5) نتایج الگوریتم پیشنهادی در مقایسه با دیگر الگوریتم ها

اندازه	مثال	الگوریتم پیشنهادی		یوان و همکاران (2023)		هونگ و همکاران (2024)	
		هدف 1	هدف 2	هدف 1	هدف 2	هدف 1	هدف 2
کوچک	1	18263	92.3	17405	88.66	18237	88.91
	2	15883	91.9	15362	92.96	15264	92.54
	3	16231	91.2	16030	86.52	15712	86.29
	4	16619	90.5	16250	85.84	16350	89.19
	5	19305	91	19249	85.99	18315	91.49
	6	16252	90.9	15352	88.13	16424	87.73

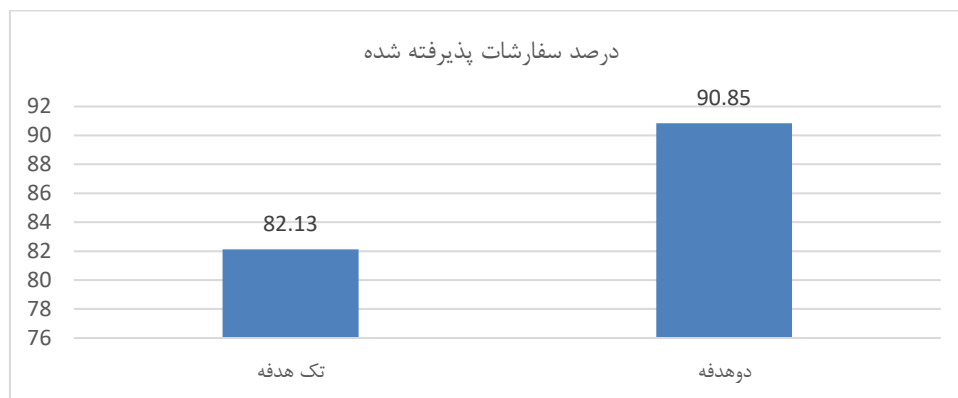
اندازه	مثال	الگوریتم پیشنهادی				هونگ و همکاران (2024)	
		هدف 1	هدف 2	هدف 1	هدف 2	هدف 1	هدف 2
متوسط	7	14758	89.9	14079	87.16	14422	89.82
	8	19628	91.5	19761	88.39	18543	89.06
	9	19613	91.5	18638	86.47	19827	88.32
	10	17186	92.1	16885	88.59	17403	91.07
	1	25232	89.8	23885	85.83	24306	89.79
	2	24000	90.2	23263	90	24178	86.69
	3	23855	91	24012	91.16	23609	90.03
	4	25729	92	25425	90.99	24769	90.81
	5	24565	91.8	24695	86.93	23725	90.88
	6	27301	92.2	26512	88.36	26378	88.21
بزرگ	7	27257	89.7	27096	90.24	26685	89.88
	8	25985	91.6	25390	87.84	26133	88.42
	9	24912	89.4	23524	84.78	23701	88.7
	10	25106	89.9	24652	85.76	24539	90.7
	1	37844	90.6	36353	90.51	37625	88.05
	2	35267	89.9	34742	88.35	34047	88.61
	3	39098	89.4	39501	90.16	36928	89.73
	4	37284	90	37176	88.91	36128	87.86
	5	36111	89.6	34768	86.54	35649	86.13
	6	36737	92.2	34885	93.4	35470	91.96
تعداد بهترین جواب به دست آمده	هدف 1	21	4	5	2	0	
	هدف 2	23	5	2	2	0	
	کل	16	2	2	2	0	

با توجه به جدول 5، روش حل پیشنهادی در مجموع 30 مثال حل شده، در تابع هدف اول و دوم به ترتیب در 21 و 23 مسئله، جواب بهتری را به دست آورده است. همچنین در 16 مثال، هر دو تابع هدف اول و دوم در مسئله پیشنهادی جواب بهتری نسبت به دیگر روش ها به دست آمده است. در واقع، جواب های کارا در مسئله پیشنهادی 16 عدد هستند و جواب های به دست آمده توسط دو روش دیگر نسبت به این جواب ها مغلوب هستند. این شرایط نشان از کارآمدی روش حل پیشنهادی دارد. اختلاف روش حل پیشنهادی در تابع هدف اول و دوم نسبت به روش یوان و همکاران به ترتیب 2.55 و 2.64 درصد و نسبت به روش هونگ و همکاران 2.34 و 1.9 درصد بهتر است. شکل 2 این برتری را نشان می دهد.



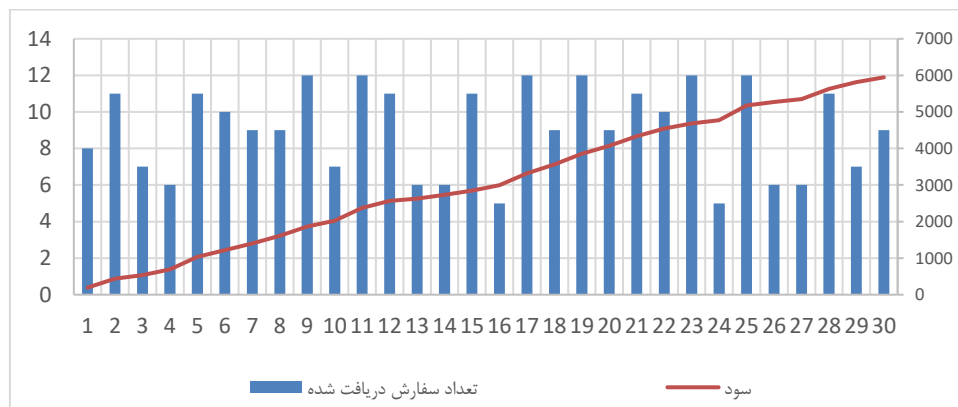
شکل 2) تعداد جواب کارای به دست آمد توسط الگوریتم پیشنهادی در مقایسه با دیگر الگوریتم‌ها

شکل 3 میزان تأثیر تابع هدف دوم بر مسئله را نشان می‌دهد. با توجه به اینکه سود یک هدف کوتاه‌مدت است و رد سفارشات باعث نارضایتی مشتریان می‌شود، همچنین سازمان‌ها باید برای افزایش مزیت رقابتی و بقای بلندمدت خود برنامه‌ریزی کنند، لذا هدف دوم می‌تواند با پذیرش سفارشات بیشتر و افزایش رضایت مشتریان این امر را محقق سازد. پذیرش سفارشات بیشتر با کاهش سود همراه است. این کاهش سود می‌تواند به عنوان یک سرمایه‌گذاری برای آینده تلقی شود که در نهایت منجر به افزایش توان رقابتی سازمان در مقابل دیگر سازمان‌ها شود.



شکل 3) درصد سفارشات پذیرفته در مسئله تک هدفه و دو هدفه

مسئله دو هدفه نسبت به مسئله تک هدفه 8.72 درصد سفارشات بیشتری را پذیرفته است، اما میزان سود 6.85 درصد کاهش یافته است. به طور کلی، به ازای افزایش 1 درصدی در تعداد سفارشات پذیرفته‌شده، میزان سود 0.78 درصد کاهش می‌یابد. با توجه به اینکه مسئله پیشنهادی به صورت پویا حل می‌شود، با ورود سفارشات جدید، روش حل پیشنهادی اجرا شده و جواب بهینه به دست می‌آید. شکل زیر میزان بهبود تابع هدف را با ورود سفارشات جدید نشان می‌دهد. این حالت برای یک دوره 30 روزه با 387 سفارش بررسی شده است. این شرایط از کارخانه کاغذ سنگی الهام گرفته شده است. شکل 4 ورود سفارشات در هر دوره و میزان سود حاصل از آنها را نشان می‌دهد.



شکل 4) تعداد سفارش دریافت شده و میزان سود حاصل شده

### 5.3 مفاهیم نظری

این مطالعه به طور قابل توجهی به درک مسائل زمان بندی پویا در سیستم های تولید سلولی چندهدفه کمک کرده است. اولین و مهم ترین ارمان این تحقیق، گسترش مدل های موجود به ویژه در زمینه پذیرش سفارشات، زمان بندی و تخصیص ماشین ها در محیط های پویا است. بسیاری از مدل های پیشین تمرکز بیشتری بر مسائل زمان بندی ثابت یا ساده تر داشته اند، در حالی که این مطالعه با اضافه کردن عواملی مانند رد سفارشات، کاهش عملکرد ماشین ها و تغییرات پویا در سیستم، به واقعیت های پیچیده تر و نزدیک تر به شرایط دنیای واقعی پرداخته است.

این تحقیق همچنین بر اهمیت تصمیمات پویا در سیستم های تولید سلولی تأکید می کند، جایی که توازن بین حداکثر کردن سود و پذیرش سفارشات می تواند عملکرد کلی سیستم را تحت تأثیر قرار دهد. یافته ها نشان می دهد که در نظر گرفتن اهداف کوتاه مدت (حداکثر کردن سود) و بلندمدت (پذیرش سفارشات بیشتر) هم زمان می تواند در بهینه سازی عملکرد سیستم مؤثر باشد. استفاده از یادگیری تقویتی عمیق (DRL) به عنوان یک روش حل، افق های جدیدی را برای حل مسائل پیچیده و پویا در زمان بندی فراهم می آورد. این روش نشان داده که توانایی مدیریت عدم قطعیت ها و تغییرات در شرایط واقعی تولید را دارد و بر لزوم استفاده از تکنیک های زمان بندی تطبیقی در محیط هایی که شرایط تولید آن ها به طور مداوم تغییر می کند، تأکید می کند.

این مطالعه همچنین اهمیت تصمیمات عملیاتی مانند تخصیص ماشین ها و زمان بندی سفارشات را در دستیابی به عملکرد مطلوب در سیستم های پویا برجسته می کند. استفاده از پیشرفت های تکنولوژیکی مانند یادگیری ماشین و یادگیری تقویتی در این تحقیق، چارچوب نظری بهینه سازی سیستم های تولید را گسترش می دهد و زمینه های جدیدی را برای تحقیقات آینده در این حوزه فراهم می آورد. علاوه بر این، این تحقیق نشان می دهد که سیاست های رد سفارشات می تواند تأثیر زیادی بر رضایت مشتری و رقابت پذیری سازمان ها داشته باشد و تصمیمات پذیرش سفارشات، علی رغم کاهش کوتاه مدت سود، می تواند به عنوان سرمایه گذاری برای آینده و تقویت موقعیت رقابتی در بلندمدت عمل کند.

در نتیجه، این تحقیق پیامدهای نظری قابل توجهی در زمینه زمان‌بندی پویا با اهداف چندگانه در سیستم‌های تولید سلولی دارد و زمینه را برای تحقیقات بیشتر در استفاده از الگوریتم‌های پیشرفته مانند یادگیری تقویتی عمیق در بهینه‌سازی سیستم‌های تولید فراهم می‌کند.

#### 5.4 مفاهیم کاربردی

این تحقیق پیامدهای عملی قابل توجهی برای مدیران و تصمیم‌گیرندگان در صنایع تولیدی دارد. نتایج این مطالعه می‌تواند به‌ویژه برای شرکت‌هایی که با چالش‌های پویا و متغیر در زمان‌بندی سفارشات و تخصیص ماشین‌ها روبه‌رو هستند، مفید باشد. استفاده از الگوریتم یادگیری تقویتی عمیق (DRL) در این تحقیق نشان می‌دهد که این تکنیک می‌تواند به‌طور مؤثر در بهبود تصمیمات زمان‌بندی و پذیرش سفارشات در شرایط متغیر و پویا به کار رود. بنابراین، مدیران می‌توانند از این روش برای افزایش بهره‌وری سیستم‌های تولیدی و بهینه‌سازی استفاده از منابع خود بهره‌برداری کنند. یکی از کاربردهای عملی مهم این تحقیق، امکان استفاده از مدل پیشنهادی برای برنامه‌ریزی دقیق‌تر در محیط‌های تولیدی است که با مشکلاتی مانند کاهش کارایی ماشین‌ها، تغییرات ناگهانی در سفارشات و محدودیت‌های ظرفیت مواجه هستند. این مدل می‌تواند به شرکت‌ها کمک کند تا تصمیمات بهتری در زمینه پذیرش یا رد سفارشات اتخاذ کنند، که می‌تواند به رضایت بیشتر مشتریان، کاهش هزینه‌های تأخیر و افزایش سود منجر شود. در واقع، پذیرش سفارشات بیشتر و زمان‌بندی دقیق‌تر می‌تواند به افزایش وفاداری مشتریان و تقویت مزیت رقابتی سازمان‌ها کمک کند. علاوه بر این، این مطالعه بر لزوم استفاده از رویکردهای پویا در مدیریت تولید تأکید می‌کند. بسیاری از صنایع با شرایطی روبه‌رو هستند که در آن‌ها پیش‌بینی دقیق آینده دشوار است و باید توانایی تطبیق با تغییرات سریع را داشته باشند. به همین دلیل، شرکت‌ها می‌توانند از روش‌های یادگیری ماشین و یادگیری تقویتی برای ایجاد سیستم‌های خود تطبیق‌پذیر استفاده کنند که قادر به واکنش سریع به شرایط متغیر بازار باشند.

در نهایت، این تحقیق برای صنایع مختلف، از جمله تولیدات سفارش‌محور، صنایع خودروسازی، و صنایع الکترونیک که با چالش‌های مشابه در زمان‌بندی و تخصیص منابع مواجه هستند، کاربردهای عملی زیادی دارد. مدیران این صنایع می‌توانند از مدل‌های مشابه برای بهبود بهره‌وری و کاهش هزینه‌های عملیاتی استفاده کنند و همچنین در برنامه‌ریزی برای سرمایه‌گذاری‌های بلندمدت و بهبود رابطه با مشتریان، به کار گیرند.

#### 6 نتیجه‌گیری

این تحقیق یک روش نوآورانه مبتنی بر یادگیری تقویتی عمیق (DRL) را برای حل مسائل زمان‌بندی پویا در سیستم‌های تولید سلولی پیشنهاد کرد. نتایج تجربی نشان داد که الگوریتم پیشنهادی در مقایسه با روش‌های سنتی در بهینه‌سازی دو تابع هدف، یعنی حداکثرسازی سود و حداکثرسازی تعداد سفارشات پذیرفته‌شده، عملکرد بهتری دارد. در حالی که پذیرش

بیشتر سفارشات ممکن است با کاهش سود همراه باشد، این کاهش سود به عنوان یک سرمایه گذاری بلندمدت در نظر گرفته می شود که در نهایت موجب افزایش رضایت مشتری و مزیت رقابتی سازمان خواهد شد. با این حال، این تحقیق محدودیت هایی نیز دارد که می تواند زمینه ساز تحقیقات آینده باشد. یکی از محدودیت ها، استفاده از مدل های شبیه سازی برای ارزیابی عملکرد الگوریتم است که ممکن است نتایج در محیط های واقعی کمی متفاوت باشد. همچنین، تحقیق حاضر تنها به یک نوع سیستم تولید سلولی و شرایط خاص آن پرداخته است. بنابراین، توسعه روش های پیشنهادی برای سایر انواع سیستم های تولید، از جمله سیستم های تولید کارگاهی یا خطی، می تواند از جهت های مختلفی مفید باشد.

پیشنهادات برای تحقیقات آینده عبارتند از:

1. گسترش مدل های پیشنهادی به سایر سیستم های تولید: تحقیق آینده می تواند به بررسی عملکرد الگوریتم در انواع مختلف سیستم های تولید مانند تولید کارگاهی و خطی پرداخته و نتایج آن را با سیستم های سلولی مقایسه کند.
  2. استفاده از داده های واقعی: برای ارزیابی بهتر عملکرد الگوریتم، می توان از داده های واقعی از کارخانه ها و خطوط تولید استفاده کرد تا نتایج مدل در شرایط واقعی بررسی شود.
  3. بررسی عوامل محیطی و شرایط متغیر: در آینده، می توان مدل هایی توسعه داد که شرایط محیطی، مانند نوسانات بازار یا تغییرات غیرمنتظره در سفارشات، را در نظر بگیرند.
  4. بهینه سازی با الگوریتم های هیبریدی: ترکیب الگوریتم های یادگیری تقویتی عمیق با سایر الگوریتم های بهینه سازی مانند الگوریتم های فراابتکاری می تواند به افزایش دقت و کارایی مدل ها کمک کند.
- در نهایت، این تحقیق نشان می دهد که یادگیری تقویتی عمیق می تواند ابزاری موثر برای حل مسائل پیچیده و پویا در سیستم های تولید باشد و به بهبود کارایی و رضایت مشتری کمک کند.

## منابع و مراجع

- Aalaei, A., & Davoudpour, H. (2012, November). Designing a mathematical model for integrating dynamic cellular manufacturing into supply chain system. In *AIP Conference Proceedings* (Vol. 1499, No. 1, pp. 239-246). American Institute of Physics. <https://doi.org/10.1063/1.4768994>
- Aalaei, A., & Davoudpour, H. (2016). TWO BOUNDS FOR INTEGRATING THE VIRTUAL DYNAMIC CELLULAR MANUFACTURING PROBLEM INTO SUPPLY CHAIN MANAGEMENT. *Journal of Industrial & Management Optimization*, 12(3). [10.3934/jimo.2016.12.907](https://doi.org/10.3934/jimo.2016.12.907)
- An, Y., Chen, X., Gao, K., Zhang, L., Li, Y., & Zhao, Z. (2023). A hybrid multi-objective evolutionary algorithm for solving an adaptive flexible job-shop rescheduling problem with real-time order acceptance and condition-based preventive maintenance. *Expert systems with applications*, 212, 118711. <https://doi.org/10.1016/j.eswa.2022.118711>
- Arkat, J., Hosseiniabadi Farahani, M., & Hosseini, L. (2012). Integrating cell formation with cellular layout and operations scheduling. *The International Journal of Advanced Manufacturing Technology*, 61, 637-647. <https://doi.org/10.1007/s00170-011-3733-4>
- Azadeh, A., Ravanbakhsh, M., Rezaei-Malek, M., Sheikhalishahi, M., & Taheri-Moghaddam, A. (2017). Unique NSGA-II and MOPSO algorithms for improved dynamic cellular manufacturing systems considering human factors. *Applied Mathematical Modelling*, 48, 655-672. <https://doi.org/10.1016/j.apm.2017.02.026>



- Bílková, D. (2012). Lognormal distribution and using L-moment method for estimating its parameters. *International journal of mathematical models and methods in applied sciences*, 6(1), 30-44.
- Bouazza, W., Sallez, Y., & Beldjilali, B. (2017). A distributed approach solving partially flexible job-shop scheduling problem with a Q-learning effect. *IFAC-PapersOnLine*, 50(1), 15890-15895. <https://doi.org/10.1016/j.ifacol.2017.08.2354>
- Chen, C., Yang, Z., Tan, Y., & He, R. (2014). Diversity controlling genetic algorithm for order acceptance and scheduling problem. *Mathematical Problems in Engineering*, 2014(1), 367152.
- Chen, X., Hao, X., Lin, H. W., & Murata, T. (2010, August). Rule driven multi objective dynamic scheduling by data envelopment analysis and reinforcement learning. In *2010 IEEE International Conference on Automation and Logistics* (pp. 396-401). IEEE. DOI: [10.1109/ICAL.2010.5585316](https://doi.org/10.1109/ICAL.2010.5585316)
- Chu, X., Gao, D., Cheng, S., Wu, L., Chen, J., Shi, Y., & Qin, Q. (2019). Worker assignment with learning-forgetting effect in cellular manufacturing system using adaptive memetic differential search algorithm. *Computers & Industrial Engineering*, 136, 381-396. <https://doi.org/10.1016/j.cie.2019.07.028>
- Chung, S. H., Wu, T. H., & Chang, C. C. (2011). An efficient tabu search algorithm to the cell formation problem with alternative routings and machine reliability considerations. *Computers & Industrial Engineering*, 60(1), 7-15. <https://doi.org/10.1016/j.cie.2010.08.016>
- Delgoshaei, A., & Ali, A. (2020). A hybrid ant colony optimization and simulated annealing algorithm for multi-objective scheduling of cellular manufacturing systems. *International Journal of Applied Metaheuristic Computing (IJAMC)*, 11(3), 1-40. DOI: [10.4018/IJAMC.2020070101](https://doi.org/10.4018/IJAMC.2020070101)
- Ding, Z., Huang, Y., Yuan, H., & Dong, H. (2020). Introduction to reinforcement learning. *Deep reinforcement learning: fundamentals, research and applications*, 47-123. [https://doi.org/10.1007/978-981-15-4095-0\\_2](https://doi.org/10.1007/978-981-15-4095-0_2)
- Geramipour, S., Moslehi, G., & Reisi-Nafchi, M. (2017). Maximizing the profit in customer's order acceptance and scheduling problem with weighted tardiness penalty. *Journal of the Operational Research Society*, 68(1), 89-101.
- Goli, A., Tirkolaee, E. B., & Aydın, N. S. (2021). Fuzzy integrated cell formation and production scheduling considering automated guided vehicles and human factors. *IEEE transactions on fuzzy systems*, 29(12), 3686-3695. DOI: [10.1109/TFUZZ.2021.3053838](https://doi.org/10.1109/TFUZZ.2021.3053838)
- Hammami, N. E. H., Lardeux, B., B. Hadj-Alouane, A., & Jridi, M. (2024). Design and calibration of a DRL algorithm for solving the job shop scheduling problem under unexpected job arrivals. *Flexible Services and Manufacturing Journal*, 1-32. <https://doi.org/10.1007/s10696-024-09540-2>
- Herasymovych, M., Märka, K., & Lukason, O. (2019). Using reinforcement learning to optimize the acceptance threshold of a credit scoring model. *Applied Soft Computing*, 84, 105697. <https://doi.org/10.1016/j.asoc.2019.105697>
- Herbots, J., Herroelen, W., & Leus, R. (2007). Dynamic order acceptance and capacity planning on a single bottleneck resource. *Naval Research Logistics (NRL)*, 54(8), 874-889.
- Houshyar, A. N., Leman, Z., Moghadam, H. P., Ariffin, M. K. A. M., Ismail, N., & Iranmanesh, H. (2014, June). Literature review on dynamic cellular manufacturing system. In *IOP conference series: materials science and engineering* (Vol. 58, No. 1, p. 012016). IOP Publishing. DOI [10.1088/1757-899X/58/1/012016](https://doi.org/10.1088/1757-899X/58/1/012016)
- Huang, J. P., Gao, L., & Li, X. Y. (2024). An end-to-end deep reinforcement learning method based on graph neural network for distributed job-shop scheduling problem. *Expert Systems with Applications*, 238, 121756. <https://doi.org/10.1016/j.eswa.2023.121756>
- Jabal Ameli, M. S., Arkat, J., & Barzinpour, F. (2008). Modelling the effects of machine breakdowns in the generalized cell formation problem. *The International Journal of Advanced Manufacturing Technology*, 39, 838-850. <https://doi.org/10.1007/s00170-007-1269-4>
- Leng, J., J. Guo, H. Zhang, K. Xu, Y. Qiao, P. Zheng and W. Shen (2023). "Dual deep reinforcement learning agents-based integrated order acceptance and scheduling of mass individualized prototyping." *Journal of Cleaner Production* 427: 139249. <https://doi.org/10.1016/j.jmsy.2011.03.004>
- Li, F., S. Xu and Z. Xu (2023). "New exact and approximation algorithms for integrated production and transportation scheduling with committed delivery due dates and order acceptance." *European Journal of Operational Research* 306(1): 127-140. <https://doi.org/10.1016/j.ejor.2013.07.032>
- Lin, S.-W. and K.-C. Ying (2013). "Increasing the total net revenue for single machine order acceptance and scheduling problems using an artificial bee colony algorithm." *Journal of the Operational Research Society* 64(2): 293-311. <https://doi.org/10.1007/s00170-007-1269-4>

- Liu, C., J. Wang and J. Y.-T. Leung (2018). "Integrated bacteria foraging algorithm for cellular manufacturing in supply chain considering facility transfer and production planning." *Applied Soft Computing* 62: 602-618. <https://doi.org/10.1007/s00170-007-1269-4>
- Lou, P., Q. Liu, Z. Zhou, H. Wang and S. X. Sun (2012). "Multi-agent-based proactive–reactive scheduling for a job shop." *The International Journal of Advanced Manufacturing Technology* 59: 311-324.
- Luo, S. (2020). "Dynamic scheduling for flexible job shop with new job insertions by deep reinforcement learning." *Applied Soft Computing* 91: 106208. <https://doi.org/10.1016/j.ejor.2013.07.032>
- Mahdavi, I., A. Aalaei, M. M. Paydar and M. Solimanpur (2010). "Designing a mathematical model for dynamic cellular manufacturing systems considering production planning and worker assignment." *Computers & Mathematics with Applications* 60(4): 1014-1025. <https://doi.org/10.1016/j.jmsy.2011.03.004>
- Mirhoseini, A., H. Pham, Q. V. Le, B. Steiner, R. Larsen, Y. Zhou, N. Kumar, M. Norouzi, S. Bengio and J. Dean (2017). Device placement optimization with reinforcement learning. *International Conference on Machine Learning*, PMLR. <https://doi.org/10.1007/s00170-007-1269-4>
- Motahari, R., Z. Alavifar, A. Z. Andaryan, M. Chipulu and M. Saberi (2023). "A multi-objective linear programming model for scheduling part families and designing a group layout in cellular manufacturing systems." *Computers & Operations Research* 151: 106090. <https://doi.org/10.1016/j.jmsy.2011.03.004>
- Nie, L., L. Gao, P. Li and X. Li (2013). "A GEP-based reactive scheduling policies constructing approach for dynamic flexible job shop scheduling problem with job release dates." *Journal of Intelligent Manufacturing* 24: 763-774.
- Og, C., F. S. Salman and Z. B. Yalçın (2010). "Order acceptance and scheduling decisions in make-to-order systems." *International Journal of Production Economics* 125(1): 200-211. <https://doi.org/10.1007/s00170-007-1269-4>
- Ou, J. and X. Zhong (2017). "Bicriteria order acceptance and scheduling with consideration of fill rate." *European Journal of Operational Research* 262(3): 904-907. <https://doi.org/10.1016/j.jmsy.2011.03.004>
- Papaiouannou, G. and J. M. Wilson (2010). "The evolution of cell formation problem methodologies based on recent studies (1997–2008): Review and directions for future research." *European journal of operational research* 206(3): 509-521. <https://doi.org/10.1016/j.jmsy.2011.03.004>
- Potts, C. N. and L. N. Van Wassenhove (1985). "A branch and bound algorithm for the total weighted tardiness problem." *Operations research* 33(2): 363-377. <https://doi.org/10.1007/s00170-007-1269-4>
- Rabbani, M., H. Farrokhi-Asl and M. Ravanbakhsh (2019). "Dynamic cellular manufacturing system considering machine failure and workload balance." *Journal of Industrial Engineering International* 15(1): 25-40. <https://doi.org/10.1016/j.cie.2018.03.039>
- Rafiei, H., M. Rabbani, H. Gholizadeh and H. Dashti (2016). "A novel hybrid SA/GA algorithm for solving an integrated cell formation–job scheduling problem with sequence-dependent set-up times." *International Journal of Management Science and Engineering Management* 11(3): 134-142. <https://doi.org/10.1016/j.cie.2018.03.039>
- Rahimi, V., J. Arkat and H. Farughi (2020). "A vibration damping optimization algorithm for the integrated problem of cell formation, cellular scheduling, and intercellular layout." *Computers & Industrial Engineering* 143: 106439. <https://doi.org/10.1016/j.cie.2018.03.039>
- Rahman, H. F., M. N. Janardhanan and I. E. Nielsen (2019). "Real-time order acceptance and scheduling problems in a flow shop environment using hybrid GA-PSO algorithm." *IEEE Access* 7: 112742-112755. <https://doi.org/10.1016/j.cie.2018.03.039>
- Ruiz-Torres, A. J., Paletta, G., & Pérez, E. (2013). Parallel machine scheduling to minimize the makespan with sequence dependent deteriorating effects. *Computers & Operations Research*, 40(8), 2051-2061. <https://doi.org/10.1016/j.cor.2013.02.018>
- Sarvestani, H. K., Zadeh, A., Seyfi, M., & Rasti-Barzoki, M. (2019). Integrated order acceptance and supply chain scheduling problem with supplier selection and due date assignment. *Applied Soft Computing*, 75, 72-83. <https://doi.org/10.1016/j.asoc.2018.10.045>
- Shafiee-Gol, S., Kia, R., Tavakkoli-Moghaddam, R., Kazemi, M., & Kamran, M. A. (2021). Integration of parts scheduling, MRP, production planning and generalized fixed-charge transportation planning in the design of a dynamic cellular manufacturing system. *RAIRO-Operations Research*, 55, S1875-S1912. <https://doi.org/10.1051/ro/2020062>
- Shahrabi, J., Adibi, M. A., & Mahootchi, M. (2017). A reinforcement learning approach to parameter estimation in dynamic job shop scheduling. *Computers & Industrial Engineering*, 110, 75-82. <https://doi.org/10.1016/j.cie.2017.05.026>
- Shiue, Y. R., Lee, K. C., & Su, C. T. (2018). Real-time scheduling for a smart factory using a reinforcement learning approach. *Computers & Industrial Engineering*, 125, 604-614. <https://doi.org/10.1016/j.cie.2018.03.039>

- Silva, Y. L. T., Subramanian, A., & Pessoa, A. A. (2018). Exact and heuristic algorithms for order acceptance and scheduling with sequence-dependent setup times. *Computers & operations research*, 90, 142-160. <https://doi.org/10.1016/j.cor.2017.09.006>
- Slotnick, S. A. (2011). Order acceptance and scheduling: A taxonomy and review. *European Journal of Operational Research*, 212(1), 1-11. <https://doi.org/10.1016/j.ejor.2010.09.042>
- Tarhan, İ., & Oğuz, C. (2021). Generalized order acceptance and scheduling problem with batch delivery: Models and metaheuristics. *Computers & Operations Research*, 134, 105414. <https://doi.org/10.1016/j.cor.2021.105414>
- Wiering, M. A., & Van Otterlo, M. (2012). Reinforcement learning. *Adaptation, learning, and optimization*, 12(3), 729. <https://doi.org/10.1016/j.cor.2021.105414>
- Wang, C., & Hu, Q. (2020). Knowledge sharing in supply chain networks: Effects of collaborative innovation activities and capability on innovation performance. *Technovation*, 94, 102010. <https://doi.org/10.1016/j.technovation.2017.12.002>
- Wang, T., Baldacci, R., Lim, A., & Hu, Q. (2018). A branch-and-price algorithm for scheduling of deteriorating jobs and flexible periodic maintenance on a single machine. *European Journal of Operational Research*, 271(3), 826-838. <https://doi.org/10.1016/j.ejor.2018.05.050>
- Wang, Y., Wang, J. Q., & Yin, Y. (2020). Multitasking scheduling and due date assignment with deterioration effect and efficiency promotion. *Computers & Industrial Engineering*, 146, 106569. <https://doi.org/10.1016/j.cie.2020.106569>
- Wang, Z., Qi, Y., Cui, H., & Zhang, J. (2019). A hybrid algorithm for order acceptance and scheduling problem in make-to-stock/make-to-order industries. *Computers & Industrial Engineering*, 127, 841-852. <https://doi.org/10.1016/j.cie.2018.11.021>
- Wang, Z., Qi, Y., Cui, H., & Zhang, J. (2019). A hybrid algorithm for order acceptance and scheduling problem in make-to-stock/make-to-order industries. *Computers & Industrial Engineering*, 127, 841-852. <https://doi.org/10.1016/j.cie.2018.11.021>
- Wu, C. C., Hsu, P. H., & Lai, K. (2011). Simulated-annealing heuristics for the single-machine scheduling problem with learning and unequal job release times. *Journal of Manufacturing Systems*, 30(1), 54-62. <https://doi.org/10.1016/j.jmsy.2011.03.004>
- Yavari, M., & Akbari, A. H. (2023). Service level and profit maximisation in order acceptance and scheduling problem with weighted tardiness. *International Journal of Industrial and Systems Engineering*, 43(3), 331-362. <https://doi.org/10.1504/IJISE.2023.129138>
- Yavari, M., Marvi, M., & Akbari, A. H. (2020). Semi-permutation-based genetic algorithm for order acceptance and scheduling in two-stage assembly problem. *Neural Computing and Applications*, 32, 2989-3003. <https://doi.org/10.1007/s00521-019-04027-w>
- Yuan, E., Cheng, S., Wang, L., Song, S., & Wu, F. (2023). Solving job shop scheduling problems via deep reinforcement learning. *Applied Soft Computing*, 143, 110436. <https://doi.org/10.1016/j.asoc.2023.110436>
- Yuan, E., Wang, L., Cheng, S., Song, S., Fan, W., & Li, Y. (2024). Solving flexible job shop scheduling problems via deep reinforcement learning. *Expert Systems with Applications*, 245, 123019. <https://doi.org/10.1016/j.eswa.2023.123019>
- Zandieh, M., & Roumani, M. (2017). A biogeography-based optimization algorithm for order acceptance and scheduling. *Journal of Industrial and Production Engineering*, 34(4), 312-321. <https://doi.org/10.1016/j.ejor.2013.07.032>
- Zhang, H., Leng, J., Zhang, H., Ruan, G., Zhou, M., & Zhang, Y. (2021, July). A deep reinforcement learning algorithm for order acceptance decision of individualized product assembling. In *2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI)* (pp. 21-24). IEEE. DOI: [10.1109/DTPIS2967.2021.9540190](https://doi.org/10.1109/DTPIS2967.2021.9540190)
- Zhong, X., & Ou, J. (2017). Improved approximation algorithms for parallel machine scheduling with release dates and job rejection. *4OR*, 15, 387-406. <https://doi.org/10.1007/s10288-016-0339-6>
- Zhong, X., Ou, J., & Wang, G. (2014). Order acceptance and scheduling with machine availability constraints. *European journal of operational research*, 232(3), 435-441. <https://doi.org/10.1016/j.ejor.2013.07.032>