Ind. Manag. Persp. 2023, 13(3), 249-284; doi: 10.48308/JIMP.13.3.249

Industrial Management Perspective

Journal hompage: https://jimp.sbu.ac.ir/?lang=en

Original Article

Agent-Based Modeling of Pharmaceutical Distribution Online Monitoring System, with Reinforcement Learning Approach

Ramin Kangarlou Haghighi*, Abbas Toloie Eshlaghy**, Mohammad Reza Motadel***

Abstract

Increasing operating profit is a challenge faced by pharmaceutical distribution companies. Most of the researches conducted in this field have a cost reduction approach. The online monitoring system is one of the effective methods that can be used for managers' decision-making and improving the performance of the pharmaceutical distribution chain with the approach of reducing costs and increasing revenues. To create this system, conceptual, mathematical and computer modeling is needed. The aim of this research is to develop a mathematical model of the online monitoring system to improve the pharmaceutical distribution system based on a conceptual model using a reinforcement learning approach. The mathematical model was derived based on the factor-based conceptual model and using the multi-factorial reinforcement learning approach. After extracting the mathematical model, the effectiveness of the model was validated by comparing the results obtained from the output of the mathematical model and the actual results in a pharmaceutical distribution company. The findings of the research showed that the developed mathematical model has the ability to continuously improve the goals, decisions and performance of the pharmaceutical distribution chain processes, according to the interactions and changes in the behavior of agents and the state of the environment.

Keywords: Mathematical Modeling; Agent-based Modeling; Online monitoring system; pharmaceutical distribution System; Reinforcement Learning.

How to Cite: Kangarlou Haghighi, Ramin; Toloie Eshlaghy, Abbas; Motadel, Mohammad Reza (2023). Agent-Based Modeling of Pharmaceutical Distribution Online Monitoring System, with Reinforcement Learning Approach, Ind. Manag. Persp., 13(3), 249-284 (In Persian).

Received: Jan. 21, 2023; Revised: Mar. 07, 2023; Accepted: June. 7, 2023; Published Online: Jul. 02, 2023.

* Ph.D. Student, Industrial, IT and Technology Management department, Management faculty, Central Tehran branch, Islamic azad university, Tehran, Iran.

** Professor, Management and Economics faculty, Science and Research branch, Islamic azad university, Tehran, Iran. Corresponding author. Email: toloie@srbiau.ac.ir

*** Assistant Professor, Industrial, IT and Technology Management department, Management faculty, Central Tehran branch, Islamic azad university, Tehran, Iran.

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

چشم‌انداز مدیریت صنعتی، ۱۴۰۲، دوره ۱۳، شماره ۳، صص ۲۴۹ – ۲۸۴، doi: 10.48308/JIMP.13.3.249

چشم‌انداز مدیریت صنعتی

Journal hompage: https://jimp.sbu.ac.ir/?lang=fa

نوع مقاله: پژوهشی

مدل‌سازی عامل‌بنیان سیستم پایش برخط توزیع دارو با رویکرد یادگیری تقویتی

رامین کنگرلو حقیقی*، عباس طلوعی اشلقی**، محمدرضا معتدل***

چکیده

افزایش سود عملیاتی، چالشی است که شرکت‌های پخش دارو با آن مواجه هستند. اغلب پژوهش‌های انجام شده در این حوزه، دارای رویکرد کاهش هزینه می‌باشند. سامانه پایش برخط، یکی از روش‌های اثربخشی است که می‌تواند برای تصمیم‌گیری مدیران و بهبود عملکرد زنجیره توزیع دارو با رویکرد کاهش هزینه‌ها و افزایش درآمدها، مورد استفاده قرار گیرد. برای ایجاد این سامانه، نیاز به مدل‌سازی مفهومی، ریاضی و رایانه‌ای است. هدف این پژوهش، توسعه مدل ریاضی عامل‌بنیان سیستم پایش برخط برای بهبود نظام توزیع دارو بر اساس مدل مفهومی با استفاده از رویکرد یادگیری تقویتی است. مدل ریاضی بر اساس مدل مفهومی عامل‌بنیان و با استفاده از رویکرد یادگیری تقویتی چندعاملی استخراج شد. پس از استخراج مدل ریاضی، اثربخشی مدل با مقایسه نتایج حاصل از خروجی مدل ریاضی و نتایج واقعی در یک شرکت پخش دارو، صحه‌گذاری شد. یافته‌های پژوهش نشان داد که مدل ریاضی توسعه‌یافته، با توجه به تعاملات و تغییر رفتار عامل‌ها و وضعیت محیط، قابلیت بهبود مستمر اهداف، تصمیم‌گیری‌ها و عملکرد فرایندهای زنجیره توزیع دارو را دارد.

کلیدواژه‌ها: مدل‌سازی ریاضی؛ مدل‌سازی عامل‌بنیان؛ سیستم پایش برخط؛ نظام توزیع دارو؛ یادگیری تقویتی.

استناددهی: کنگرلو حقیقی، رامین؛ طلوعی اشلقی، عباس؛ معتدل، محمدرضا (۱۴۰۲). مدل‌سازی عامل‌بنیان سیستم پایش برخط توزیع دارو با رویکرد یادگیری تقویتی. چشم‌انداز مدیریت صنعتی، ۱۳(۳)، ۲۴۹ – ۲۸۴.

تاریخ دریافت: ۱۴۰۱/۱۱/۰۱، تاریخ بازنگری: ۱۴۰۱/۱۲/۱۶، تاریخ پذیرش: ۱۴۰۲/۰۳/۱۷، تاریخ اولین انتشار: ۱۴۰۲/۰۴/۱۱.

* دانشجوی دکتری مدیریت فناوری اطلاعات، گروه مدیریت صنعتی، فناوری اطلاعات و تکنولوژی، دانشکده مدیریت، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران.

** استاد، گروه مدیریت و اقتصاد، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.

نویسنده مسئول: Email: toloie@srbiau.ac.ir

*** استادیار، گروه مدیریت صنعتی، فناوری اطلاعات و تکنولوژی، واحد تهران مرکزی، دانشگاه آزاد اسلامی، تهران، ایران.

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

۱. مقدمه

زنجیره توزیع دارو باید دارو را در زمان و مکان مناسب، به‌میزان موردنیاز و با کیفیت و قیمت تعیین‌شده در اختیار مشتریان قرار دهد؛ به‌نحوی‌که برای ذی‌نفعان اصلی، ایجاد ارزش کند و سود عملیاتی را افزایش دهد. مأموریت اصلی شرکت‌های پخش، تأمین داروها از تأمین‌کنندگان، نگهداری مناسب آن‌ها در انبارهای پخش و توزیع‌کنندگان استانی و منطقه‌ای و توزیع به‌موقع و کافی آن‌ها به‌منظور برآورده‌سازی نیازهای داروخانه‌ها، بیمارستان‌ها و مراکز درمانی مطابق با الزامات قانونی است.

در بسیاری از مواقع، نیاز مشتریان به‌دلایلی مانند شیوع بیماری‌های واگیردار، وقوع حوادث غیرمترقبه، تغییرات فصلی، شرایط اقتصادی و غیره بسیار متغیر است و به‌راحتی قابل‌پیش‌بینی نیست؛ بنابراین ضروری است که سطح موجودی افزایش یابد. با توجه به تنوع زیاد اقلام دارویی و تعدد مراکز جغرافیایی مشتریان، می‌توان به اهمیت هوشمندسازی فرایند پخش دارو پی برد [۴].

با توجه به پیچیدگی نظام توزیع دارو، بهره‌گیری از روش‌هایی مانند شبیه‌سازی عامل‌بنیان می‌تواند بسیار سودمند باشد. در شبیه‌سازی عامل‌بنیان، هر عامل می‌تواند به‌طور مجزا و مستقل عمل متقابل خود با دیگر عامل‌ها و محیط را تعیین کند؛ بنابراین پیچیدگی‌هایی که در سیستم توزیع دارو وجود دارد را می‌توان با تقریب خوبی در این نوع مدل شبیه‌سازی کرد. در صنعت پخش دارو، با توجه به اینکه محیط و قوانین رفتاری عامل‌ها به‌طور مداوم در حال تغییر است، مدل‌سازی عامل‌بنیان می‌تواند تغییرات در سطح کل سیستم را پس از وقوع تغییرات محیطی و تغییرات رفتاری عامل‌ها، مدل‌سازی و تحلیل حساسیت کند. با استفاده از نتایج تحلیل حساسیت، سازمان می‌تواند بهترین تصمیم را برای دستیابی به اهداف خود اتخاذ کند [۱۶].

ایجاد بهبود از طریق هوشمندسازی فرایندها با استفاده از یک سیستم پایش برخط، رویکردی است که در صنعت توزیع دارو کمتر موردتوجه قرار گرفته است [۲۴].

با توجه به اینکه در صنعت پخش دارو، شرایط محیطی به‌طور دائم در حال تغییر است و روابط بین شرکت پخش با تأمین‌کنندگان، مشتریان و نهادهای قانونی به‌دلیل تغییر سیاست‌ها، دست‌خوش تغییرات می‌شود، باید از یک روش مدل‌سازی استفاده شود که بتواند پیچیدگی‌های محیطی و تغییرات را پشتیبانی کند [۴]؛ همچنین با افزایش تجربه عامل شرکت پخش، بتوان تصمیمات دقیق‌تر و کاراتری اتخاذ کرد. یکی از روش‌های مدل‌سازی که می‌تواند نیازهای ما را برطرف کند، روش شبیه‌سازی عامل‌بنیان است. این روش می‌تواند روابط بین عامل‌ها را مورد تجزیه‌وتحلیل قرار دهد و پیچیدگی‌های محیطی را نیز شبیه‌سازی کند [۳].

چنانچه روش شبیه‌سازی عامل‌بنیان با رویکرد یادگیری تقویتی که یکی از روش‌های هوش مصنوعی است با هم ترکیب شوند، امکان اتخاذ تصمیمات بهینه‌تر با توجه به یادگیری‌های قبلی را نیز فراهم می‌آورد [۸، ۲۸].

در سایر پژوهش‌ها از روش‌هایی مانند شبکه‌های عصبی هیبریدی، پژوهش عملیاتی با تابع هدف‌ها و محدودیت‌های مختلف، هوش تجاری، استفاده از فناوری‌های اینترنت اشیا و بلاک‌چین و روش‌های دیگر استفاده شده است [۴].

مزیت روش انتخابی نسبت به سایر روش‌ها این است که یک روش ترکیبی است و با ترکیب مدل‌سازی عامل‌بنیان، یادگیری تقویتی و پایش برخط، می‌توان؛ ابعاد مختلف عامل‌ها و فرایندهای آن‌ها را تجزیه‌وتحلیل کرد و تحلیل حساسیت انجام داد. همچنین عملکرد شرکت پخش، به‌عنوان یک عامل، می‌تواند از نظر شاخص‌های مختلف درآمدی و هزینه‌ای موردبررسی قرار گیرد و در وضعیت‌های مختلف و با توجه به نتایج پایش، اقدام‌های مؤثرتری انجام شود که به افزایش سود عملیاتی شرکت پخش منجر شود [۲۰، ۳۹].

هدف اصلی این پژوهش، توسعه مدل سیستم پایش برخط برای بهبود نظام توزیع محصولات دارویی بر اساس روش شبیه‌سازی عامل‌بنیان است و «مدل مفهومی عامل‌بنیان» که در پژوهش کنگرلو و همکاران (۱۴۰۰)، تدوین و صحه‌گذاری شده است [۲۷] به «مدل ریاضی عامل‌بنیان»، توسعه می‌یابد.

در این پژوهش تلاش شده است که پاسخ مناسبی برای سؤال زیر ارائه شود.

ـ چگونه می‌توان مدل مفهومی عامل‌بنیان سیستم پایش برخط نظام توزیع دارو را به مدل ریاضی عامل‌بنیان تبدیل کرد؛ به‌گونه‌ای که در پژوهش‌های آتی، مدل ریاضی عامل‌بنیان توسعه‌یافته، قابلیت تبدیل به مدل رایانه‌ای عامل‌بنیان را داشته باشد؟

این پژوهش، در شش بخش اصلی مقدمه، پیشینه پژوهش، روش، یافته‌ها، بحث و نتیجه‌گیری و منابع، ارائه شده است.

۲. مبانی نظری و پیشینه پژوهش

مبانی نظری و پیشینه پژوهش شامل سه قسمت اصلی، مروری بر سیستم پایش برخط، مدل‌سازی عامل‌بنیان و مبانی یادگیری تقویتی است.

مروری بر پیشینه سیستم پایش برخط. یک سیستم پایش برخط، یک فرایند یا مجموعه‌ای از فرایندهای توزیعی است که وظیفه آن‌ها جمع‌آوری و تفسیر پویای اطلاعات مربوط به یک برنامه کاربردی است. فعالیت‌های پایش شامل راه‌اندازی حسگرها، فعال‌سازی حسگرها، سنجش،

تفسیر رویداد و انجام اقدام است. فعالیت‌های اصلی سیستم پایش برخط نیز شامل جمع‌آوری اطلاعات، تفسیر و واکنش مناسب یا اقدام می‌شود [۳۷].

پایش به معنای نظارت است و منظور از آن هوشیاری در مورد وضعیت یک پدیده از راه مشاهده تغییراتی است که ممکن است با گذر زمان در آن سامانه یا پدیده رخ دهد [۳۸].

در پژوهش‌های مختلف، کاربردهای متعددی برای سیستم پایش برخط پیشنهاد شده است که اندازه‌گیری ضربان قلب و تنفس بیماران، با استفاده از حسگر شتاب سه‌محوری به‌صورت بی‌سیم، تشخیص خطای مکانیکی سیم‌پیچ، پایش آلاینده‌های محیط‌زیستی در حوزه آب و پساب و پایش وضعیت فرایندها، سیستم‌ها و تجهیزات نیروگاه‌های هسته‌ای، نمونه‌هایی از این موارد هستند که در ادامه به تعدادی از آن‌ها اشاره می‌شود [۲۵، ۳۵، ۳۸، ۱۷].

سپهرنیا و همکاران (۱۳۹۵)، روش صحیح دستیابی به نمونه و نحوه سنجش آب و پساب را به‌منظور به‌کارگیری در پایش برخط محیط‌زیستی پیشنهاد دادند. در این روش، نمونه‌برداری پیوسته و تحلیل مداوم و پیوسته نمونه‌ها بدون نیاز و دخالت عامل انسانی، میسر شده و دقت اندازه‌گیری به لحاظ عدنبودم وجود خطاهای اندازه‌گیری ناشی از خطاهای اپراتور و نمونه‌برداری‌های دستی، ارتقا یافته است [۳۸].

ربیعی و همکاران (۱۳۹۳)، روشی را برای پایش برخط خطاهای مکانیکی سیم‌پیچ ترانسفورماتور قدرت ارائه دادند. در این روش امکان پایش شرایط سیم‌پیچ و تشخیص خطاهای مکانیکی در حین بهره‌برداری از ترانسفورماتور فراهم می‌شود. این روش مبتنی بر ایجاد یک مکان هندسی از اختلاف ولتاژ ورودی و خروجی از یک سیم‌پیچ خاص روی محور Y با جریان ورودی سیم‌پیچ روی محور X است [۳۵].

جوادپور و محمدی، روشی را برای ایجاد سامانه‌های پایش برخط بی‌سیم به‌منظور پایش سیگنال‌های حیاتی بدن مانند، دمای بدن، ضربان قلب، ریتم تنفس، سیگنال‌های مغزی و غیره ارائه دادند. در این روش داده‌ها با استفاده از حسگرهای بی‌سیم، جمع‌آوری شده و در پایگاه داده‌ای بر روی سرور ذخیره‌سازی می‌شود. تحلیل داده‌های ذخیره‌شده توسط پزشک می‌تواند به تشخیص بیماری و شناسایی وضعیت بیمار کمک کند. این‌گونه سامانه‌ها می‌توانند در بخش‌هایی مانند مراقبت‌های ویژه و مراقبت قلبی، نقش بسیار مؤثری داشته باشند [۲۵].

رضا و امین فاضلی (۱۳۹۵)، پایش برخط تجهیزات، سیستم‌ها و فرایندهای نیروگاه‌های هسته‌ای را موردتوجه قرار دادند که مبتنی بر شناسایی موارد غیرعادی از طریق سیگنال‌های فرایند در زمان فعال‌بودن نیروگاه بود. در این رابطه مواردی مانند بررسی ارتعاش، بررسی صوتی، بررسی بخش‌های رهاشده و نویز رآکتور بررسی شد. مدل سیستم پایش برخط ارائه‌شده، امکان ایجاد هشدارهای اولیه برای شکست‌های نزدیک و ایجاد تغییر در عملکرد و کارایی فرایندهای موردنظر را دارد [۱۷].

ایمک [1]و همکاران (۲۰۱۶)، مدلی را برای پایش برخط خطا در ماشین‌های ابزار، بر اساس میزان مصرف انرژی، پیشنهاد دادند که اثربخشی کلی تجهیزات را بهبود داده و کارایی منابع و بهره‌وری در تولید را افزایش می‌دهد. برای این منظور از حسگرهایی برای پایش میزان ارتعاش، میزان دردسترس‌بودن ماشین، کیفیت فرایندهای ماشین‌کاری و غیره استفاده شد. شرایط غیرعادی ماشین با الگوی خطا مشخص می‌شود که می‌تواند نشان‌دهنده خرابی و افت کیفیت باشد. داده‌های جمع‌آوری‌شده با استفاده از تکنیک‌های داده‌کاوی تحلیل می‌شود و اطلاعات موردنیاز برای تصمیم‌گیری در مورد برنامه‌ریزی نگهداری و تعمیرات را فراهم می‌کند [۱۴].

مروری بر مدل‌سازی عامل‌بنیان. در شبیه‌سازی عامل‌بنیان، تمرکز بر روی عامل‌های موجود در سیستم و نحوه رفتار آن‌ها است؛ بنابراین مدل‌سازی در این روش از پایین به بالا با شناسایی عامل‌ها و رفتار تک‌تک آن‌ها صورت می‌گیرد. در این روش برای توصیف تعامل، عوامل را به هم متصل می‌کنند و آن را در یک محیط قرار می‌دهند. در این صورت رفتار کلی سیستم، برآیند رفتار تعداد زیادی از رفتارهای فردی عامل‌ها است [۳۶].

گیلبرت و همکاران[2] (۲۰۰۷)، مدل‌سازی عامل‌بنیان را در حوزه‌هایی که پژوهشگر به‌دنبال بررسی الگوهای رفتاری در سطح خُرد و کنش و برهم‌کنش عوامل با یکدیگر است و محیط‌های پیچیده و غیرخطی، مناسب می‌دانند [۳].

یک مدل عامل‌بنیان از سه قسمت عامل‌ها، محیط و تعامل‌های بین عامل‌ها و محیط تشکیل شده است که در ادامه تعاریفی از آن‌ها ارائه شده است.

عامل:[3] عامل، شخص، موجود، شیء یا نهادی خودمختار (متمایز) با ویژگی‌ها، اقدامات و اهداف خاص است.

محیط:[4] چشم‌اندازی است که عوامل در آن برهم‌کنش دارند. محیط شامل تمام اجزای سیستم شبیه‌سازی‌شده، به‌استثنای عامل‌ها است.

تعامل‌ها: شامل عمل و عکس‌العمل میان عامل‌ها و نیز میان عامل‌ها و محیط است. این تعامل‌ها غیرخطی هستند و بنابراین رفتار کلی برابر با مجموع رفتار همه عامل‌ها نیست. تعامل‌ها از طریق تبادل اطلاعات شکل می‌گیرند [۳].

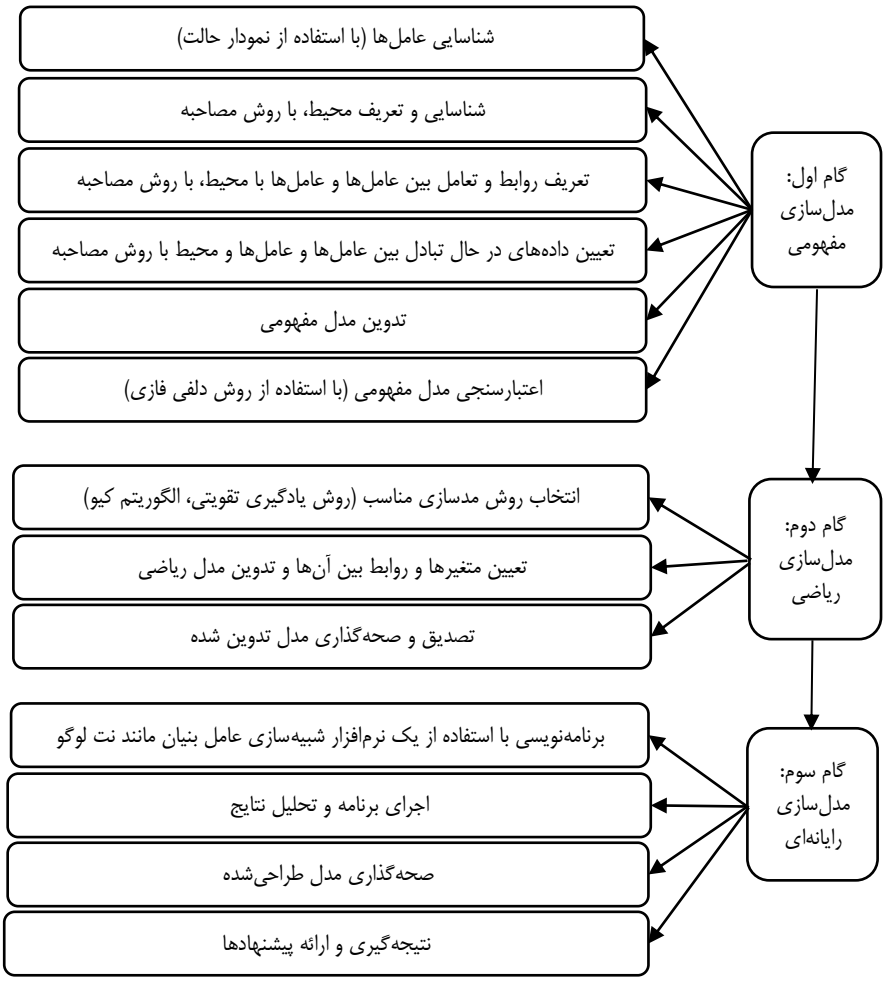
در پژوهش‌های انجام‌شده، مدل‌سازی عامل‌بنیان با گام‌های مختلفی ارائه شده است. به‌عنوان نمونه در یک پژوهش، شش گام برای مدل‌سازی عامل‌بنیان ذکر شده است. این گام‌ها شامل ۱. محدوده مدل، ۲. تطبیق نظریه در سطح خرد، ۳. تعیین ورودی ـ خروجی ـ تعیین

1. Emec, et al.
2. Gilbert, et al.
3. Agent
4. Environment

مقادیر اولیه، ۴. عامل‌ها و ویژگی‌ها، ۵. رفتار و ۶. گام زمانی است و سنجش اعتبار مدل نیز به دو قسمت اعتباربخشی و اعتبارسنجی تقسیم‌بندی شده است [۳].

در پژوهش کنگرلو و همکاران (۱۴۰۰)، چهار گام برای توسعه یک مدل عامل‌بنیان عنوان شده است. این گام‌ها عبارت‌اند از: ۱. تعریف مجموعه عامل‌ها و محیطی که عامل‌ها در آن با یکدیگر در تعامل هستند؛ ۲. طراحی مدل مفهومی عامل‌بنیان، مجموعه روابط و نحوه تراکنش میان عامل‌ها؛ ۳. برنامه‌نویسی و ساخت مدل عامل‌بنیان مدل‌سازی در محیط نرم‌افزاری؛ ۴. اجرای مدل و تحلیل نتایج [۳۴].

با جمع‌بندی دیدگاه‌های پژوهشگران در مورد گام‌های مدل‌سازی عامل‌بنیان، مراحل مدل‌سازی عامل‌بنیان پژوهش موردنظر را می‌توان در قالب شکل ۱، ارائه کرد.



شکل ۱. مراحل مدل‌سازی عامل‌بنیان سیستم پایش‌برخط توزیع دارو

پژوهش‌های متعددی با استفاده از روش عامل‌بنیان انجام شده است که در ادامه به تعدادی از آن‌ها اشاره می‌شود.

ابوالفتحی و همکاران (۱۳۹۹)، مدلی عامل‌بنیان در حوزه بازاریابی مربوط به انتشار نوآوری (انتشار تلویزیون در ایران) ارائه دادند. بر اساس راهنمای عملیاتی تدوین‌شده برای اجرای مدل‌های مبتنی بر عامل، مدل اجرا شد و سپس با درنظرگرفتن سناریوهای مختلف برای شبکه ارتباطاتی، اثر افزایش تعداد ارتباطات و افزودن شبکه‌های ارتباطاتی در بازار بالقوه برای پذیرش نوآوری (محصول و خدمت جدید) موردبررسی قرار گرفت. نتایج نشان داد که افزایش تعداد ارتباطات و افزودن شبکه‌های ارتباطاتی در پذیرش و انتشار نوآوری، تأثیری زیادی ایجاد نمی‌کند؛ ولی می‌تواند تا حدودی تأثیری مناسب فراهم آورد [۳].

خانی و همکاران (۱۳۹۸)، یک مدل عامل‌بنیان ارائه کردند که توان پوشش عامل‌های مختلف و در حال تغییر بازار را دارا باشد و با افزایش توان پیش‌بینی، نتیجه‌های ورود برند جدید به بازار، به تصمیم‌گیری بهتر برای انتخاب راهبردهای نفوذ و توسعه بازار کمک کند.

بر اساس نتایج به‌دست‌آمده، تغییر متغیرهای مؤثر بر معرفی محصول یا برند جدید می‌تواند تغییرهای اساسی در نتیجه‌ها به‌وجود آورد [۳۰].

اسکندر و همکاران (۱۳۹۸)، یک مدل عامل‌بنیان برای شبیه‌سازی و پیش‌بینی عملکرد گروهی در گروه مدیریت صنعتی، برای شناسایی عملکرد منابع انسانی و بهبود راهبردهای مدیریت منابع انسانی ارائه کردند. با توجه به رویکرد شبیه‌سازی عامل‌بنیان، شاخص‌های تأثیرگذار شناسایی شدند؛ سپس از میان این شاخص‌ها، شاخص‌هایی که بر عملکرد گروهی در گروه مدیریت صنعتی بر اساس فرایندهای مؤثر بر عملکرد گروهی تأثیرگذار است، استخراج شد. مدل مفهومی پژوهش با استفاده از مشاهدات پژوهشگر، مصاحبه با افراد مؤثر بر عملکرد گروهی و رویه‌های موجود در سایت دانشگاه استخراج شد. با استفاده از مدل شبیه‌سازی می‌توان تأثیر تغییر ورودی‌ها بر خروجی‌ها در طول زمان را مشاهده کرد. نتایج شبیه‌سازی نشان داد که مطلوبیت مدیر گروه و کارشناس آموزش برای چند سناریوی مختلف درنظرگرفته‌شده دارای بیشترین مقدار است؛ همچنین تعداد وظایفی که منتظر اجر هستند نیز در همه سناریوها از مقدار زیادی برخوردار است که نیاز است تا گروه برای بهبود راهبردهای مدیریت منابع انسانی سیاست‌های جدیدی را اتخاذ کند [۱۵].

فرح‌بخش و همکاران (۱۴۰۱)، مدلی برای شبیه‌سازی چرخه عمر صنعت برق با استفاده از شبیه‌سازی عامل‌بنیان ارائه کردند. در ابتدا عامل‌ها شناسایی و تعاملات بین عامل‌ها و عامل‌ها با محیط استخراج شد. برای بهینه‌سازی مدل، چهار سناریو با نظر خبرگان تعیین شد و شبیه‌سازی

به کمک نرم‌افزار انی‌لاجیک[1] صورت گرفت. نتایج نشان داد که بهتر است دولت برای مدیریت کمبود برق و اثرات جبران‌ناپذیر آن بر اقتصاد، تغییراتی در سیاست‌های خود به‌وجود آورد تا سرمایه‌گذاران تمایل بیشتری به سرمایه‌گذاری و ورود به بازار صنعت برق داشته باشند [۱۶].

سنانی[2] (۲۰۲۱)، در پژوهشی به بررسی چهار ابزار مدل‌سازی عامل‌بنیان شامل مت سیم، انی‌لاجیک، نت لوگو و ری‌سپت پرداخت. در این بررسی، پژوهش‌های پیشرفته انجام‌شده با این ابزارها بررسی شد و مرجعی از ابزارهای مدل‌سازی عامل‌بنیان برای پژوهشگرانی که درحال‌توسعه سیستم‌های پشتیبان تصمیم هستند، ارائه شده است. نتایج بررسی‌ها نشان داد که سیستم عامل‌بنیان به‌جای اینکه ابزاری برای حل مسئله باشد، ابزاری برای توسعه و آزمایش راه‌حل‌های جایگزین با توجه به سناریوهای مختلف است؛ هچنین مدل‌سازی عامل‌بنیان یک روش بسیار مفید برای حمایت از تصمیم‌گیرندگان در موقعیت‌های با پیچیدگی و عدم‌قطعیت بالا است. از میان این چهار ابزار، نت لوگو، به‌دلیل ویژگی‌ها و قابلیت‌های خود، انتخاب مناسبی است [۱۲].

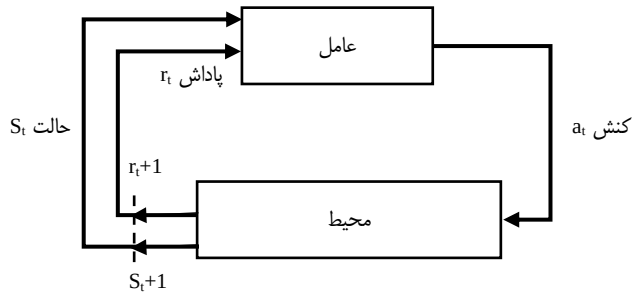
بونابو[3] (۲۰۱۲)، پس از معرفی مختصر اصول اولیه شبیه‌سازی عامل‌بنیان، چهار حوزه کاربردی شبیه‌سازی روند، شبیه‌سازی سازمانی، شبیه‌سازی بازار و شبیه‌سازی انتشار را بررسی کرد. بر اساس بررسی‌های انجام شده، نتیجه‌گیری شد که میزان دقت در تعیین ورودی‌ها و کامل‌بودن ورودی‌های مدل عامل‌بنیان که ناشی از عوامل متعددی مانند تخصص، داده‌ها و غیره است، می‌تواند خروجی متفاوتی را به‌وجود آورد؛ همچنین علی‌رغم افزایش چشمگیر توان محاسباتی در سال‌های اخیر، نیازهای محاسباتی بالای سیستم عامل‌بنیان، در مدل‌سازی سیستم‌های بزرگ، همچنان به‌صورت یک مشکل باقی می‌ماند [۱۱].

مروری بر مبانی یادگیری تقویتی. به‌طورکلی در یادگیری تقویتی، با توجه به شناخت محیط و بر اساس نتایج تعامل‌هایی که با محیط وجود داشته و سود و زیان‌هایی که بر اثر انجام کنش‌های مختلف به‌دست آمده است، سیاستی اتخاذ می‌شود که با عمل به آن در بلندمدت، مطلوبیت بیشینه شود [۳۹].

اجزای اصلی مسائل یادگیری تقویتی شامل پنج عنصر عامل[4]، محیط[5] سیاست[6]، تابع پاداش[7] و مدل هستند. عامل همان یادگیرنده است که از طریق مجموعه‌ای از حسگرها و محرک‌های

1. Anylogic
2. Cenani
3. Bonabeau
4. Agent
5. Environment
6. Policy
7. Reward Function

خود می‌تواند با محیط در تعامل باشد. محیط فضایی است که عامل کنش‌های[1] مربوطه را در آن انجام داده و روی آن تأثیر می‌گذارد. سیاست، هسته اصلی عامل یادگیرنده تقویتی است که مشخص می‌کند که عامل چه کنشی را در هر حالت باید انجام دهد. با توجه به میزان مطلوبیت کنش انجام‌شده، پاداشی به آن داده می‌شود که مقدار آن از طریق تابع پاداش مشخص می‌شود. مدل، رفتار و محیط را شبیه‌سازی کرده و حالت بعدی سیستم را پیش‌بینی می‌کند. یادگیری تقویتی جزو مجموعه هوش مصنوعی و زیرمجموعه یادگیری ماشین است. فرایند یادگیری تقویتی در شکل ۲، نشان داده شده است.



شکل ۲. فرایند یادگیری تقویتی

این مراحل به‌ترتیب شامل مشاهده حالت (S_t)، تصمیم‌گیری عامل برای انجام کنش (a_t)، انجام کنش (a_t)، شبیه‌سازی سیستم و مشاهده حالت جدید(S_t+1)، اختصاص پاداش (R_t+1)، برای کنش (a_t)، و یادگیری از تجربیات است که تا رسیدن به‌شرط پایداری، این روند تکرار می‌شود [۳۹].

در مسئله یادگیری تقویتی، عاملی وجود دارد که با سعی و خطا و تجربه، سعی بر یافتن بهترین جواب برای مسئله را دارد؛ سپس با انجام آن عمل، شرایط تغییر کرده و در قبال آن یک میزان پاداش دریافت می‌کند [۱۳].

روش یادگیری تقویتی محدودیت ناشی از مدل‌سازی را با تکیه بر یک مدل شبیه‌ساز، مرتفع می‌کند. مزیت دیگر استفاده از شبیه‌ساز آن است که در روند حل مسئله، تکیه بر حالت‌های محتمل است و به‌صورت خودکار وقت کمتری صرف انجام محاسبات برای حالت‌های غیرمحتمل می‌شود. این مزیت می‌تواند محدودیت ناشی از ابعاد مسئله را در روش یادگیری تقویتی تا حدودی تخفیف دهد [۲۲].

1. Actions

اصغری اسکویی و همکاران (۱۳۹۶)، بر روی روش‌های یادگیری تقویتی در مدل چندعاملی بازار برق ایران مطالعه کرده و آن‌ها را با دو استراتژی تصادفی و حریصانه مقایسه کردند. در این پژوهش، میزان سود واحدها و زمان رسیدن به حالت تعادل به‌عنوان ملاک ارزیابی در نظر گرفته شد. نتایج شبیه‌سازی نشان داد که استراتژی یادگیرنده، سود عامل‌ها را به‌طور معناداری افزایش می‌دهد و سرعت همگرایی به حالت تعادل را بیشتر می‌کند [۷].

نیک‌انجام و همکاران (۱۳۹۹)، با معرفی یک چارچوب ارتباطی برای سیستم‌های چندعاملی مشارکتی، تلاش کردند که در رابطه با چالش یادگیری در محیط‌های چندعاملی مشارکتی پویا، غیرقطعی و با اندازه فضای بزرگ در برنامه‌های کاربردی، پاسخی ارائه کنند. در راستای حل مشکلات مربوط به همگرایی، انتقال دانش به کار برده شده است که می‌تواند به میزان زیادی در افزایش الگوریتم‌های یادگیری تقویتی مؤثر واقع شود. همکاری میان عامل‌ها با استفاده از عامل سرگروه و هماهنگی میان آن‌ها توسط یک عامل هماهنگ‌کننده صورت می‌گیرد. چارچوب پیشنهادی برای حل مسئله گله‌داری به‌کار رفته است و نتایج تجربی، افزایش کارایی عامل‌ها را نشان می‌دهد [۳۳].

خاکسار و همکاران (۱۳۹۷)، روشی را برای بهبود عملکرد حمله در تیم شبیه‌سازی فوتبال ربات‌ها با استفاده از روش یادگیری تقویتی پیشنهاد دادند. در این خصوص الگوریتم یادگیری کیو- وی روی تیم ربات‌های فوتبالیست اجرا شده و به آموزش عامل‌های حمله پرداخته شد.

برای آموزش عامل‌ها از دو روش مختلف برای تقسیم سیگنال تقویتی بین عامل‌ها استفاده شده است. در روش اول، سیگنال تقویتی به‌طور یکسان بین عامل‌ها توزیع شده و در روش دوم، از ایده تقسیم سیگنال تقویتی بر حسب میزان خبرگی عامل‌ها، بهره گرفته شده است.

بر اساس این ایده، عامل با دانش کمتر جریمه بیشتری نسبت به عامل با دانش بیشتر دریافت می‌کند. نتایج پژوهش نشان داد که هر دو روش پیشنهادی، نسبت به حالتی که عامل‌ها آموزش ندیدند، نتیجه بهتری دارد و روش دوم نیز از روش اول بهتر است [۲۸].

آرل و همکاران کاربرد جدیدی از یک سیستم چندعاملی در چارچوب یادگیری تقویتی برای به‌دست‌آوردن یک سیاست کنترل سیگنال ترافیک کارآمد معرفی کردند. هدف از ارائه این روش، به‌حداقل‌رساندن میانگین تأخیر، ازدحام و احتمال انسداد تقاطع است.

برای این منظور، شبکه ترافیک پنج تقاطع موردمطالعه قرار گرفته که هر تقاطع توسط یک عامل هوشمند مستقل کنترل می‌شده است. یک عامل مرکزی و یک عامل خروجی به‌کار گرفته شده و عوامل خروجی سیگنال‌های ترافیکی بر اساس الگوریتم طولانی‌ترین صف اول، برنامه‌ریزی شده است. در این روش از الگوریتم ترکیبی یادگیری کیو با یک شبکه عصبی پیش‌خور استفاده شده است. نتایج تجربی مزایای کنترل مبتنی بر یادگیری تقویتی چندعاملی را نسبت به کنترل تک تقاطع ایزوله نشان می‌دهد [۶].

ژانگ و وانگ[1] (۲۰۱۸)، الگوریتمی به‌منظور یادگیری تقویتی چندعاملی برای وظایف مشارکتی طراحی کردند که در آن چندین عامل باید رفتار خود را برای دستیابی به بهترین عملکرد سیستم، هماهنگ کنند. در این رابطه یک الگوریتم همکاری چندعاملی یادگیری کیو، با عنوان «کاوش اقدام‌ها با توجه به نسبت‌های ارزش کیو[2]» پیشنهاد شده و پویایی آن بررسی و تحلیل شده است. در دو بازی تصادفی انجام‌شده، یکی با چهار عامل و دیگری با هشت عامل، الگوریتم عملکرد عالی از خود نشان داده است. نتایج تجربی نشان داد که این الگوریتم در مسئله شبکه حسگر توزیع‌شده نسبت به سایر الگوریتم‌ها عملکرد بهتری دارد [۴۲].

خلاصه‌ای از پژوهش‌های انجام‌شده مرتبط در ایران و جهان در جدول ۱، ارائه شده است.

جدول۱. پژوهش‌های انجام‌شده در ایران و جهان (مطالعات کتابخانه‌ای)

منبع	پژوهشگر (سال)	مهم‌ترین موضوعات موردمطالعه
		دسته‌بندی پژوهش: یادگیری تقویتی در ایران
[۲۱]	حصاری و اربابی، (۱۳۹۵)	با استفاده از روش یادگیری تقویتی و الگوریتم یادگیری کیو، تلفات موتورهای القایی به مقدار زیادی کاهش یافته است.
[۷]	اصغری اسکوئی و همکاران، (۱۳۹۶)	با استفاده از روش یادگیری تقویتی توزیع‌شده، مدلی چندعاملی از بازار برق ایران، با هدف حداکثرکردن سود بلندمدت ارائه شده است؛ سپس استراتژی‌های مناسب برای ارائه پیشنهاد قیمت در مکانیزم حراج انتخاب شده است. نتایج نشان داد که سود حاصل از روش یادگیری تقویتی به‌طور معناداری، بیش از استراتژی حریصانه و تصادفی است.
[۲۸]	خاکسار و همکاران، (۱۳۹۷)	روشی برای بهبود عملکرد حمله در تیم شبیه‌ساز فوتبال ربات‌ها با استفاده از یادگیری تقویتی ارائه شده است. برای آموزش عامل‌ها از دو روش ۱. توزیع سیگنال تقویتی به‌طور یکسان بین عامل‌ها و ۲. تقسیم سیگنال تقویتی بر حسب میزان خبرگی عامل‌ها، بهره گرفته شده است. طبق نتایج، روش دوم عملکرد بهتری نسبت به روش اول دارد؛ همچنین هر دو روش نسبت به حالتی که عامل‌ها آموزش ندیده‌اند، نتیجه بهتری دارد.
[۳۳]	نیک انجام و همکاران، (۱۳۹۹)	با استفاده از روش یادگیری تقویتی، چارچوبی برای یادگیری مشارکتی عامل‌ها در محیط‌های پویا برای مسئله گله‌داری ارائه شده است. نتایج نشان داد که بهبود عملکرد حاصل از روش پیشنهادی قابل‌قبول است.
		دسته‌بندی پژوهش: یادگیری تقویتی در جهان
[۶]	آرل و همکاران، (۲۰۱۰)	با استفاده از روش یادگیری تقویتی، الگوریتم یادگیری کیو، یک سیستم چندعاملی برای دستیابی به یک راهبرد کنترل سیگنال ترافیک کارآمد ارائه شده است. برنامه‌ریزی با استفاده از الگوریتم طولانی‌ترین صف اول انجام شده است. نتایج تجربی مزایای کنترل مبتنی بر یادگیری تقویتی چندعاملی را نسبت به کنترل تک–تقاطع ایزوله نشان می‌دهد.

[1]. Zhang, Z.& Wang, D
[2]. EAQR (Exploring Actions according to Q-value Ratios)

منبع	پژوهشگر (سال)	مهم‌ترین موضوعات موردمطالعه
[۴۱]	وانو و همکاران، (۲۰۱۸)	یک روش یادگیری تقویتی با پاداش‌دهی داخلی برای همکاری‌های چندعاملی، معرفی شده است. نتایج نشان داد که این روش عامل‌ها را قادر می‌سازد که رفتارهای مشارکتی مناسبی را اتخاذ کنند.
[۴۲]	ژانگ و وانگ، (۲۰۱۸)	یک الگوریتم یادگیری چندعاملی برای وظایف مشارکتی که در آن برای دستیابی به بهترین عملکرد سیستم، چند عامل نیاز به هماهنگی دارند، طراحی شده است. نتایج تجربی نشان داد که این الگوریتم در مسئله شبکه حسگر توزیع‌شده نسبت به سایر الگوریتم‌ها عملکرد بهتری دارد.
[۱۸]	گاسپروف و همکاران، (۲۰۲۱)	یک نمای کلی جامع و به‌روز از کاربردهای پیشرفته فعلی یادگیری تقویتی متمرکز بر بازارسازی بهینه ارائه شده است. نتایج نشان داد که تکنیک‌های یادگیری تقویتی عملکرد بهتری را از نظر بازار تعدیل‌شده با ریسک، نسبت به استراتژی‌های استاندارد بازار سازی که معمولاً از مدل‌های تحلیلی مشتق می‌شوند، ارائه می‌کنند.
		دسته‌بندی پژوهش: سیستم پایش برخط در ایران
[۲۵]	جوادپور و محمدی، (۱۳۹۳)	یک سامانه بهبودیافته برای اندازه‌گیری ضربان قلب و تنفس با استفاده از حسگر شتاب سه‌محوری به‌صورت بی‌سیم پیشنهاد شده است.
[۳۵]	ربیعی و همکاران، (۱۳۹۳)	روشی برای پایش برخط برای تشخیص خطای مکانیکی سیم‌پیچ از طریق تشکیل مکان هندسی ولتاژ، جریان با مقایسه نمودار مکان هندسی ترانسفورماتور ارائه شده است.
[۳۸]	سپهرنیا و همکاران، (۱۳۹۵)	سیستم پایش‌برخط برای پایش آلاینده‌های محیط‌زیستی در حوزه آب و پساب ارائه شده است. عوامل مؤثر بر اعتبار داده‌های پایش شناسایی شده است.
[۱۷]	فاضلی و همکاران، (۱۳۹۵)	پایش برخط وضعیت فرایندها، سیستم‌ها و تجهیزات نیروگاه شامل تشخیص و شناسایی موارد غیرعادی از طریق بررسی سیگنال‌های فرایند در نیروگاه‌های هسته‌ای فعال بررسی شده است.
		دسته‌بندی پژوهش: سیستم پایش برخط در جهان
[۱۰]	باهونگ[1] و همکاران، (۲۰۱۲)	کاربرد استاندارد آی‌ای‌سی۶۱۸۵۰ در سیستم پایش برخط پست فرعی هوشمند، شامل بررسی کارکردها و الزامات عملکرد سیستم پایش برخط و ایجاد مدل کارکردی و مدل رابط پایش برخط مبتنی بر آن
[۱۹]	جیوئوزا[2] و همکاران، (۲۰۱۴)	ارائه سیستمی برای شناسایی دقیق خطاهای عملکرد و ارائه ساختار نرم‌افزاری برای سیستم پیشنهادی برای بررسی عملکرد برنامه‌های کاربردی موازی و سلامت سیستم
[۲۳]	ایشاباکی و کایجج[3]، (۲۰۱۵)	استفاده از برچسب‌های رادیوشناسه با فرکانس فوق‌العاده بالا برای برچسب‌زدن روی داروها برای پایش آن‌ها، ارائه یک معماری برای سیستم پایش و مدیریت داروها
[۱۴]	ایمک [4]و همکاران، (۲۰۱۶)	پیشنهاد یک چارچوب پایش عیب بر اساس تکنیک‌های داده‌کاوی در قالب نگهداری و تعمیر مبتنی بر انرژی به‌منظور بهبود کارایی منابع
		دسته‌بندی پژوهش: مدل‌سازی عامل بنیان در ایران
[۳۰]	خانی و همکاران،	مدلی با استفاده از روش عامل‌بنیان ارائه شده است که توان پوشش عامل‌های مختلف

1. Baohong, et al
2. Gioiosa, et al.
3. Ishabakaki, & Kaijage
4. Emec, et al.

منبع	پژوهشگر (سال)	مهم‌ترین موضوعات موردمطالعه
	(۱۳۹۸)	و درحال‌تغییر بازار را دارا است و با افزایش توان پیش‌بینی نتایج ورود برند جدید به بازار به تصمیم‌گیری بهتر برای انتخاب راهبردهای نفوذ و توسعه بازار کمک کند. بر اساس نتایج، با تغییر متغیرهای مؤثر بر معرفی محصول یا برند جدید می‌توان به تغییرات اساسی در نتایج دست یافت.
[۱۵]	اسکندر و همکاران، (۱۳۹۸)	مدلی برای شبیه‌سازی و پیش‌بینی عملکرد تیمی در گروه مدیریت صنعتی برای شناسایی عملکرد منابع انسانی و بهبود راهبردهای مدیریت منابع انسانی ارائه شده است. برای شبیه‌سازی از نرم‌افزار انی‌لاجیک، استفاده شده است. نتایج نشان داد که مطلوبیت مدیر گروه و کارشناس آموزش برای چند سناریوی مختلف درنظرگرفته‌شده دارای بیشترین مقدار است.
[۳]	ابوالفتحی و همکاران، (۱۳۹۹)	با استفاده از مدل‌سازی عامل‌بنیان، روشی برای بهبود انتشار نوآوری ارائه شده است. مدل عامل‌محور در ترکیب با مدل انتشار باس ایجاد شده است. روشی که برای بهبود انتشار نوآوری معرفی شده است در قالب ۱۰ سناریو، مورد بررسی و آزمون قرار گرفته و پس از تحلیل خروجی‌ها، روش پیشنهادی در بهبود انتشار نوآوری، به‌ویژه برای غلبه بر تأخیر اولیه در پذیرش نوآوری، مؤثر تشخیص داده شده است.
[۱]	عباسی سیر و همکاران، (۱۴۰۰)	شبیه‌سازی عامل‌بنیان در مورد رفتار خرید آنی مصرف‌کنندگان با درنظرگیری عامل‌های تخفیف، زمان مطلوب حضور مشتری در خرید از طرف مدیر فروشگاه و یادگیری مصرف‌کننده از خریدهای قبلی صورت گرفته است. شبیه‌سازی با استفاده از مدل عامل‌بنیان و نرم‌افزار نت لوگو انجام شده است. نتایج نشان می‌دهد که عامل‌های درنظرگرفته‌شده در این پژوهش رفتار خرید آنی مصرف‌کنندگان را به‌عنوان یک تحلیل اقتصادی و بر پایه ارتباط بین مشتری ـ محصول به‌خوبی توصیف می‌کند.
[۸]	آذر و همکاران، (۱۴۰۰)	مدلی عامل‌بنیان برای شبیه‌سازی زنجیره تأمین فولاد به‌منظور برآورد میزان تولید و مصرف ارائه شده است. در بخش کیفی از مصاحبه با خبرگان و روش تحلیل تم بهره گرفته شده است. در بخش کمّی برای درک و بررسی روابط علّی و معلولی از پرسشنامه استفاده شده است. آزمون مدل به روش دیمتل انجام شده است. درنهایت مدل با استفاده از نرم‌افزار انی‌لاجیک و برنامه‌نویسی جاوا طراحی شده است. تأیید مدل رایانه‌ای نیز با استفاده از نظر خبرگان انجام شده است.
[۲]	عباسی سیر و همکاران، (۱۴۰۱)	مدلی پویا برای تحلیل رفتار سرمایه‌گذاران با درنظر گرفتن مدل‌سازی عامل‌بنیان، رفتار سهامداران و درنظرگرفتن ویژگی‌های شخصیتی متنوع آن‌ها و تحلیل حساسیت رفتار معامله‌گری، تحت محیط نرم‌افزاری نت لوگو ارائه شده است. نتایج پژوهش نشان می‌دهد که درصد سهامداران با شخصیت نوسان‌گیر و هیجانی در تعادل قیمتی به ۵۲ و ۴۸ درصد، بین سهامداران هیجانی و نوسان‌گیر توزیع شده است.
[۱۶]	فرح‌بخش و همکاران، (۱۴۰۱)	مدلی برای شبیه‌سازی چرخه عمر صنعت برق با استفاده از شبیه‌سازی عامل‌بنیان ارائه شده است. شبیه‌سازی رایانه‌ای با استفاده از نرم‌افزار انی‌لاجیک انجام شده است. برای بهینه‌سازی مدل چهار سناریو با نظر خبرگان ارائه شده است. در سناریوی چهارم که بهترین سناریو است، تشویق دولت برای ایجاد نیروگاه‌های خصوصی موردتوجه قرار گرفته است.
		دسته‌بندی پژوهش : مدل‌سازی عامل‌بنیان در جهان
[۵]	فیلیپ آرایا، (۲۰۲۰)	با بررسی مزایا و محدودیت‌های مدل‌سازی عامل‌بنیان، خطوط راهنمایی برای اجرای

منبع	پژوهشگر (سال)	مهم‌ترین موضوعات موردمطالعه
		مدل‌سازی عامل بنیان در بخش ساخت‌وساز ارائه شده است. نتایج نشان می‌دهد که مدل‌سازی عامل‌بنیان این امکان را فراهم می‌کند که پیچیدگی‌های نامشخص مرتبط با فرایندها و پروژه‌های ساخت‌وساز را در هم ادغام کند.
[۱۲]	شهناز سنانی، (۲۰۲۱)	بررسی چهار ابزار مدل‌سازی مبتنی بر عامل، شامل نت لوگو متسیم، انی‌لاجیک، ری‌پست[1] صورت گرفته است و همچنین تعیین اینکه مدل‌های عامل‌محور، کجا، چرا و چگونه مورداستفاده قرار می‌گیرند. یافته‌ها نشان داد که نرم‌افزارهای نت لوگو و انی‌لاجیک در زمینه‌های مختلف پژوهشی مثل طراحی معماری، برنامه‌ریزی شهری و حمل‌ونقل، بیشتر موردتوجه پژوهشگران قرار گرفته‌اند؛ درحالی‌که متسیم و ری‌پست کاربرد نسبتاً محدودی دارند.
[۳۲]	مونتی و همکاران، (۲۰۲۲)	یک پروتکل برای یادگیری ریز متغیرهای پنهان در یک مدل‌سازی عامل‌بنیان پیشنهاد شده است. پروتکل پیشنهادی در یک مدل‌سازی عامل‌بنیان بازار مسکن نشان داده شده که در آن نمایندگان با درآمدهای متفاوت قیمت‌های بالاتری را برای زندگی در محله‌های پردرآمد پیشنهاد می‌کنند. نتایج نشان داد که مدل به‌دست‌آمده امکان برآورد دقیق متغیرهای پنهان را فراهم می‌کند؛ درحالی‌که رفتار کلی مدل عامل‌بنیان را حفظ می‌کند.
[۱۱]	بونابو، (۲۰۲۲)	اصول اولیه شبیه‌سازی عامل‌بنیان و چهار حوزه کاربردی آن شامل شبیه‌سازی جریان، شبیه‌سازی سازمانی، شبیه‌سازی بازار و شبیه‌سازی انتشار با استفاده از برنامه‌های کاربردی دنیای واقعی موردبحث قرار گرفته است.
	دسته‌بندی پژوهش: بهینه‌سازی زنجیره تأمین دارو در ایران	
[۲۹]	خلیل‌زاده و همکاران، (۱۳۹۳)	ارائه یک روش جدید و دقیق برای پیش‌بینی فروش شرکت‌های توزیع دارو، استفاده از مدل آریما و شبکه‌های عصبی هیبریدی و آزمون مدل در «شرکت پخش هجرت»
[۲۶]	جویبان و همکاران، (۱۳۹۷)	توسعه مدلی برای زنجیره تأمین چنددوره‌ای و چندمحصولی که در آن تصمیمات استراتژیک و تاکتیکی و زمان‌های حمل‌ونقل بهینه شوند.
[۴]	علی‌دوست و همکاران، (۱۳۹۹)	یک مدل ریاضی چندهدفه برای زنجیره تأمین دارو، در حوادث غیرمترقبه، با اهداف کمّینه‌سازی هزینه‌ها، کمینه‌سازی درصد کمبود دارو و بیشینه‌سازی پراکنش مراکز توزیع با رویکرد امدادرسانی بهتر در زمان وقوع حادثه ارائه شده است. برای حل مسئله از روش ترابی– حصینی و رابطه توابع هدف با یکدیگر استفاده شده است. نتایج نشان داد که بیشترین تقاضا در حالت فعال‌شدن گسل ری و کمترین تقاضا مربوط به فعال‌شدن گسل مشا است؛ همچنین افزایش مطلوبیت توابع هدف کمّینه‌سازی، درصد کمبود دارو و بیشینه‌سازی پراکنش مراکز توزیع به‌ترتیب هزینه بیشتری را به سیستم تحمیل خواهند کرد.
[۲۴]	جعفری و همکاران، (۱۴۰۱)	چگونگی تأثیر هوشمندی کسب‌وکار بر یکپارچگی، چابکی و عملکرد زنجیره تأمین بررسی شده است. ابتدا با مرور پیشینه پژوهش، یک مدل مفهومی ارائه شده است. برای آزمون فرضیه‌ها نیز از مدل‌سازی معادلات ساختاری استفاده شده است. نتایج نشان داد که هوشمندی کسب‌وکار، یکپارچگی و چابکی، نقش مهمی در دستیابی به عملکرد بهتر زنجیره تأمین دارند؛ همچنین هوشمندی کسب‌وکار بیشترین تأثیر را بر عملکرد زنجیره تأمین دارد. یکپارچگی نیز تأثیر مستقیمی بر چابکی زنجیره تأمین

1. Repast

منبع	پژوهشگر (سال)	مهم‌ترین موضوعات موردمطالعه
		می‌گذارد.
دسته‌بندی پژوهش: بهینه‌سازی زنجیره تأمین دارو در جهان		
[۴۰]	لمی وانگ و یون وانگ، (۲۰۲۲)	تجزیه‌وتحلیل مدیریت خدمات مالی زنجیره تأمین مبتنی بر به‌اشتراک‌گذاری داده‌های اینترنت اشیا در فناوری بلاک‌چین انجام شده است. نتایج نشان داد که هزینه‌های تجاری کاهش یافته و کارایی شرکت بهبود می‌یابد و مخاطرات تأمین‌کنندگان مالی زنجیره تأمین کاهش می‌یابد.
[۹]	باها، (۲۰۲۳)	بررسی و تحلیل کاربردهای هوش مصنوعی در مدیریت زنجیره تأمین، با هدف تعیین شکاف پژوهشی تأثیر هوش مصنوعی بر عملکرد زنجیره تأمین و شناسایی تکنیک‌های هوش مصنوعی که می‌توانند عملکرد مدیریت زنجیره تأمین را افزایش دهند، صورت گرفته است. نتایج نشان داد که شکاف پژوهشی شامل عدم‌وجود استانداردسازی در اجرای هوش مصنوعی در مدیریت زنجیره تأمین و فقدان داده‌های مشخص در سازمان‌ها است. بیشتر راه‌حل‌های مدیریت زنجیره تأمین مبتنی بر هوش مصنوعی بر بهینه‌سازی فرایندها و صرفه‌جویی در هزینه‌ها، بدون درنظرگرفتن عامل انسانی و نیاز به آموزش و توسعه نیروی کار تمرکز دارد؛ همچنین با یکپارچه‌سازی هوش مصنوعی و فناوری بلاک‌چین، می‌توان مدیریت زنجیره تأمین ایمن‌تر و شفاف‌تر ارائه داد.

با بررسی پیشینه پژوهش، پژوهشی که با استفاده از رویکرد عامل‌بنیان و روش یادگیری تقویتی، مدلی برای پایش برخط به‌منظور بهبود نظام توزیع دارو ارائه کرده باشد، یافت نشد. در هیچ‌یک از پژوهش‌های پیشین، پایش برخط فرایندها یا سیستم‌ها با دید مدیریتی موردتوجه قرار نگرفته است و بیشتر به محصولات و خدمات پرداخته شده است.

پژوهش حاضر از این نظر دارای نوآوری است که در راستای حل مسائل مدیریتی شرکت‌های پخش دارو و بهبود عملکرد فرایندها و درنهایت افزایش سود عملیاتی از یک مکانیزم ترکیبی شامل مدل‌سازی عامل‌بنیان، روش یادگیری تقویتی (روش کیو) و مکانیزم پایش برخط، به‌صورت یکپارچه استفاده کرده است.

طبق بررسی‌های انجام‌شده و مرور پژوهش‌های پیشین در ایران و جهان، مدل‌سازی عامل‌بنیان، روش یادگیری تقویتی و مکانیزم پایش برخط به‌صورت جداگانه یا به‌صورت ترکیبی با سایر روش‌ها مانند شبکه‌های عصبی، معادلات ساختاری، پژوهش عملیاتی و غیره مورداستفاده قرار گرفته است.

به‌عنوان نمونه شبیه‌سازی عامل‌بنیان برای پیش‌بینی عملکرد گروهی، بهبود انتشار نوآوری، شبیه‌سازی رفتار خرید آنی مصرف‌کنندگان، شبیه‌سازی زنجیره تأمین فولاد، شبیه‌سازی چرخه عمر صنعت برق، تحلیل رفتار سرمایه‌گذاران، برنامه‌ریزی شهری، برنامه‌ریزی حمل‌ونقل و غیره

مورداستفاده قرار گرفته است؛ همچنین روش یادگیری تقویتی برای مواردی مانند کاهش تلفات موتورهای القایی، حداکثرکردن سود بلندمدت در بازار برق ایران، شبیه‌سازی در فوتبال ربات‌ها، گله‌داری، کنترل سیگنال ترافیک و غیره به‌کار گرفته شده است [۲، ۳، ۷، ۸، ۱۶، ۲۱، ۲۸، ۳۳].
سیستم پایش برخط برای مواردی مانند پایش علائم حیاتی بیماران، پایش عیب در قالب نگهداری و تعمیرات مبتنی بر انرژی، تشخیص خطای مکانیکی سیم‌پیچ‌ها، پایش آلاینده‌های محیط‌زیستی، پایش وضعیت فرایندها، سیستم‌ها و تجهیزات نیروگاه‌ها و غیره موردتوجه قرار گرفته شده است [۱۷، ۲۵، ۳۵، ۳۸].

ترکیب هدفمند مدل‌سازی عامل‌بنیان، روش یادگیری تقویتی و مکانیزم پایش برخط با رویکرد مدیریتی در پژوهش‌های جست‌جوشده و بررسی‌شده یافت نشد.

۳. روش‌شناسی پژوهش

با توجه به اینکه هدف اصلی پژوهش حاضر، افزایش سود عملیاتی در شرکت‌های پخش دارو با ارائه مدل سیستم پایش برخط است، از نظر هدف، کاربردی به‌شمار می‌رود.

در این پژوهش نیاز به گردآوری نتایج پژوهش‌های پیشین در خصوص موضوع موردنظر است که از این نظر روش گردآوری داده‌ها و اطلاعات موردنیاز، کتابخانه‌ای است؛ همچنین قسمتی از داده‌های موردنیاز نیز باید از شرکت‌های تولیدکننده دارو، پخش دارو و داروخانه‌ها جمع‌آوری شود که روش گردآوری داده‌ها در این حالت میدانی است.

ازآنجاکه در نظر است با استفاده از مدل مفهومی عامل‌بنیان تدوین‌شده در شکل ۳، نتایج حاصل از مرحله تحلیل کیفی به یک مدل ریاضی تبدیل شود، رویکرد این پژوهش از نوع کمّی است.

از مراحل مختلف مدل‌سازی عامل‌بنیان ارائه‌شده در شکل ۱، در این پژوهش مرحله مدل‌سازی ریاضی اجرا شده است.

برای جمع‌آوری داده‌های موردنیاز برای مدل‌سازی مفهومی از نمونه‌گیری به روش گلوله‌برفی استفاده شده است. در مدل‌سازی مفهومی عامل‌بنیان حجم نمونه توسط داده‌های جمع‌آوری‌شده و تحلیل آن‌ها مشخص می‌شود. نمونه‌گیری بدون هیچ محدودیتی در تعداد مشارکت‌کننده‌ها تا اشباع نظری که ۸ مصاحبه بود، ادامه یافت. برای مدل‌سازی مفهومی عامل‌بنیان شکل ۳، از روش مطالعه کتابخانه‌ای برای بررسی مبانی نظری پژوهش استفاده شده است؛ همچنین برای جمع‌آوری داده‌های موردنیاز برای شناسایی فرایند پخش دارو، عامل‌ها، محیط و تعامل بین آن‌ها و همچنین مؤلفه‌های پژوهش، از روش مطالعات میدانی شامل بررسی مستندات و مصاحبه‌های نیمه‌ساختاریافته، با متخصصان و خبرگان بهره گرفته شد.

مصاحبه‌شوندگان از میان مدیران شرکت‌های پخش، تولیدکنندگان محصولات دارویی، داروخانه‌داران، مشاوران و ارزیاب‌های صنعت پخش دارو انتخاب شدند.

برای بررسی روایی مصاحبه‌ها، پس از طراحی چارچوب مصاحبه، از سه نفر متخصص و خبره در زمینه پژوهش، نظرخواهی و تغییرات لازم مطابق نظرات آن‌ها اعمال شد. با توجه به اینکه مصاحبه به‌صورت باز انجام شده است، در طول مصاحبه سؤال‌های دیگری که به روشن‌شدن بیشتر ابعاد پژوهش کمک می‌کرد به چارچوب مصاحبه اضافه شد.

طی فرایند مصاحبه، دیدگاه‌های افراد از طریق پرسش‌های باز دریافت شده و سپس پاسخ‌ها یادداشت و کدگذاری شدند. برای محاسبه پایایی مصاحبه‌ها در این پژوهش از روش توافق موضوعی دو کدگذار استفاده شد. بدین منظور سه مصاحبه به‌صورت تصادفی توسط پژوهشگر و همکار کدگذار، انتخاب و کدگذاری شد. درنهایت پایایی بین دو کُدگذار(شاخص تکرارپذیری) که به‌عنوان شاخص پایایی تحلیل به‌کار می‌رود؛ محاسبه شد. درصد پایایی بین باز آزمون ۷۲/۱۵ درصد به‌دست آمد که بزرگ‌تر از۶۰ درصد است؛ بنابراین پایایی تأیید می‌شود [۲۷].

برای مدل‌سازی عامل‌بنیان در صنعت پخش دارو به‌طورکلی مراحل زیر باید انجام شود.

الف) مدل‌سازی مفهومی عامل‌بنیان؛ ب) تبدیل مدل مفهومی عامل‌بنیان به مدل ریاضی عامل‌بنیان؛ ج) تبدیل مدل ریاضی عامل‌بنیان به مدل رایانه‌ای عامل‌بنیان.

برای مدل‌سازی مفهومی عامل‌بنیان که در پژوهش کنگرلو و همکاران (۱۴۰۰)، جزئیات آن ارائه شده است [۲۶]، مراحل زیر انجام شد. عامل‌ها که شامل مشتریان، تأمین‌کنندگان، شرکت پخش دارو، نهادهای قانون‌گذار و ناظر و زیرساخت فناوری اطلاعات و ارتباطات هستند، شناسایی شدند؛ سپس با استفاده از روش داده‌بنیاد، خصوصیات عامل‌ها و شناسایی محیط (بازار دارویی در ایران)، تعیین و روابط و تعامل بین عامل‌ها و عامل‌ها و محیط شناسایی شد؛ همچنین قوانین رفتاری عامل‌ها نیز از طریق مصاحبه تعیین شد. برای مثال، مشتریان شرکت پخش در زمان شیوع بیماری‌های فصلی، موجودی انبار خود را برای داروهای مرتبط با بیماری‌های فصلی افزایش می‌دهند؛ یا وقتی دوره وصول مطالبات شرکت‌های پخش دارو افزایش می‌یابد، آن‌ها اقداماتی را انجام می‌دهند که دوره وصول مطالبات را کاهش و میزان نقدینگی خود را افزایش دهند.

برای این منظور اقداماتی را در دستور کار خود قرار می‌دهند؛ مانند افزایش زمان تحویل و کاهش اولویت تحویل سفارش‌های مشتریان بدحساب که دوره وصول مطالبات آن‌ها بالاست. در مرحله بعد، با توجه به داده‌های جمع‌آوری‌شده، مدل مفهومی عامل بنیان تدوین و با استفاده از روش دلفی فازی صحه‌گذاری شد.

برای مدل‌سازی ریاضی عامل‌بنیان که موضوع این پژوهش است، ابتدا مدل مفهومی عامل‌بنیان موردبررسی قرار گرفت و تعاملات بین عامل‌ها و عامل‌ها و محیط به تعدادی شاخص

عملکردی قابل‌اندازه‌گیری تبدیل شد؛ سپس این شاخص‌ها با استفاده از روش مقایسات زوجی اولویت‌بندی شدند. سپس با توجه به اینکه از روش یادگیری تقویتی بهره گرفته شده است، برای هر یک از شاخص‌ها یک جدول کیو تشکیل شد. در این جدول حالت‌ها یا وضعیت‌های مختلفی که احتمال وقوع دارند، فهرست می‌شوند و اقدام یا اقدام‌های عملی برای مواجه‌شدن با این وضعیت تعیین می‌شود؛ بنابراین بر اساس روش یادگیری تقویتی، باید اقدامی در هر وضعیت انتخاب شود که بیشترین پاداش (سود عملیاتی بالاتر) را به همراه داشته باشد.

چون شاخص‌های عملکردی متعددی مشخص شده است که به‌صورت برخط پایش می‌شوند، باید تغییراتی در فرمول محاسبه یادگیری تقویتی (الگوریتم کیو)، به‌وجود آید که پاداش بلندمدت بیشتری که همان سود عملیاتی است را برای شرکت پخش دارو به همراه داشته باشد. از روش یادگیری تقویتی (الگوریتم کیو)، نیز برای محاسبه میزان پاداش هر اقدام در حالت‌های وقوع یافته استفاده شد؛ همچنین برای کلیه شاخص‌ها جدول کیو ایجاد شده است.

جزئیات انجام کار در بخش تحلیل داده و یافته‌های پژوهش تشریح شده است. درنهایت برای تبدیل مدل ریاضی عامل‌بنیان به‌دست‌آمده به مدل رایانه‌ای عامل‌بنیان که امکان محاسبات سریع و تحلیل حساسیت را فراهم می‌آورد، باید از یک نرم‌افزار عامل‌بنیان مانند نت لوگو و انی‌لاجیک استفاده شود و در محیط آن کدنویسی صورت گیرد. تبدیل مدل ریاضی عامل‌بنیان به مدل رایانه‌ای عامل‌بنیان، موضوع پژوهش دیگری است.

۴. تحلیل داده‌ها و یافته‌های پژوهش

این پژوهش کوشیده شد تا با تحلیل و تدوین مدل مفهومی عامل‌بنیان در سیستم پایش برخط و تبدیل آن به یک مدل ریاضی عامل‌بنیان، در نظام پخش و توزیع محصولات دارویی سود عملیاتی را افزایش داد. برای مدل‌سازی مفهومی عامل‌بنیان ابتدا باید عامل‌ها و محیط شناسایی می‌شدند. برای شناسایی عامل‌ها از رسم نمودار حالت[1] استفاده شد. با استفاده از نتایج جمع‌بندی مصاحبه‌ها و همچنین نمودار حالت، عامل‌های سیستم پایش برخط نظام پخش محصولات دارویی به همراه نوع، اهداف و مشخصه‌های آن‌ها تعیین شد.

بر اساس نتایج مصاحبه‌ها، بازار محصولات دارویی در ایران به‌عنوان محیط سیستم تعیین شد. قوانین رفتاری عامل‌ها نیز با استفاده از روش اگر ـ آنگاه و نتایج مصاحبه‌ها استخراج شد. چنانچه عاملی، کنشی را از یک عامل دیگر یا محیط دریافت کند، بر اساس قوانین رفتاری حاکم، واکنشی از خود نشان می‌دهد. با یادگیری از محیط، قوانین رفتاری و رفتار عامل‌ها تغییر می‌کند. مدل مفهومی عامل‌بنیان سیستم پایش برخط نظام پخش محصولات دارویی که شامل عامل‌ها و

1.State diagram

ارتباط آن‌ها با یکدیگر و با محیط، ورودی‌ها[1]، خروجی‌ها[2] و دوره زمانی تعامل‌ها است، در شکل ۳، ارائه شده است [۲۶].

برای بررسی اعتبار مدل مفهومی ارائه‌شده، مدل دلفی فازی با عدد فازی مثلثی به‌کار رفت. در این پژوهش برای تجمیع نظر افراد خبره از روش میانگین فازی استفاده شد.

اگر دیدگاه هر فرد خبره به‌صورت عدد فازی مثلثی (l, m, u) نمایش داده شود، میانگین فازی n عدد فازی مثلثی به‌صورت رابطه ۱، محاسبه خواهد شد.

رابطه (۱) $F_{AVE} = \frac{\sum l}{n}, \frac{\sum m}{n}, \frac{\sum u}{n}$

پس از تجمیع فازی دیدگاه افراد خبره، فازی‌زدایی مقادیر به‌دست‌آمده مطابق روابط ۲ و ۳، صورت گرفت.

رابطه (۲) $F = (l, m, n)$

رابطه (۳) $X = \frac{l+m+u}{3}$

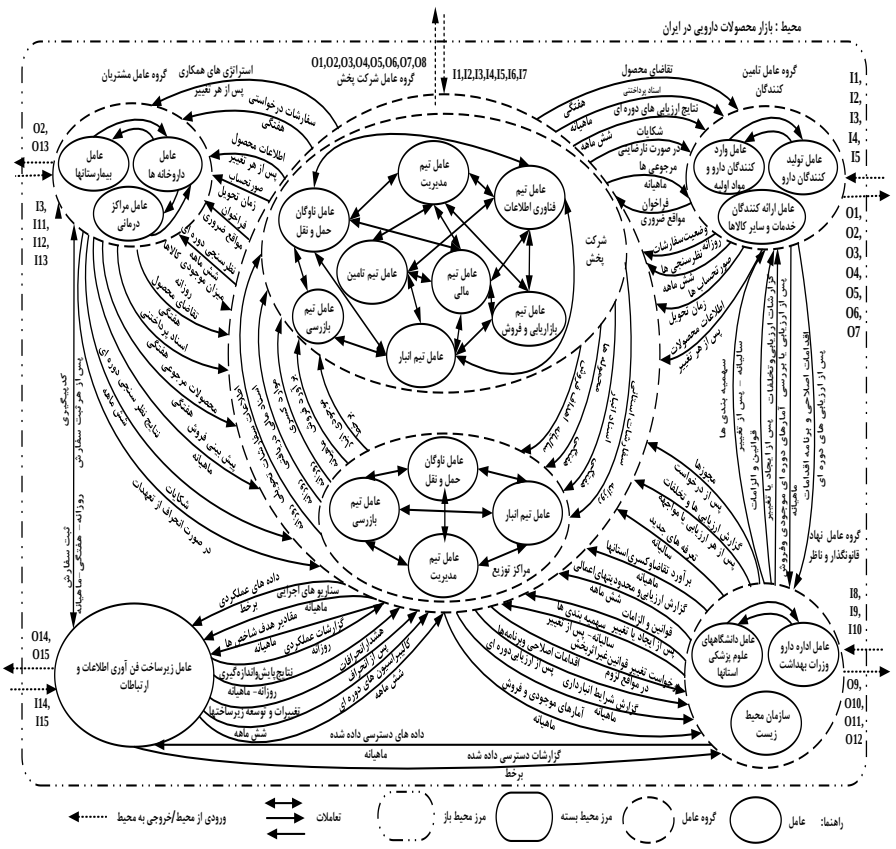
l: کمینه مقادیر عدد فازی M

m: محتمل‌ترین مقدار یک عدد فازی

u: بیشینه مقادیر عدد فازی M

X: مقدار فازی‌زدایی شده میانگین عدد فازی

1. I (input)
2. O (Out Put)



شکل ۳. مدل مفهومی عامل‌بنیان سیستم پایش برخط توزیع دارو [۲۶]

برای انجام مراحل دلفی فازی، ابتدا یک گروه پنج نفره تشکیل شد؛ سپس پرسشنامه‌ای بر اساس ساختار مدل مفهومی عامل‌بنیان تدوین‌شده، طراحی و برای اعضا ارسال و از آن‌ها درخواست شد تا با استفاده از متغیرهای کلامی، میزان اهمیت عامل‌ها، روابط بین عامل‌ها و روابط بین عامل‌ها و محیط را که شامل ۸۴ مورد بود؛ مشخص کنند؛ همچنین خواسته شد که اگر مورد جدیدی مدنظر آن‌ها بود، در پرسشنامه اضافه کنند.

در مرحله بعد، متغیرهای کلامی به اعداد فازی مثلثی تبدیل شدند؛ سپس میانگین فازی ۵ عدد فازی مثلثی طبق رابطه ۱، محاسبه شد. پس از تجمیع فازی، بر اساس روابط ۲ و ۳، فازی‌زدایی صورت گرفت. در مرحله اول، میانگین فازی‌زدایی‌شده چهار مورد، کمتر از آستانه تحمل (۰/۷)، بود که در مرحله دوم، مجدداً نظرسنجی شد. با توجه به اینکه میانگین فازی‌زدایی‌شده، موارد مرحله دوم، حداقل ۰/۷ بود، تمامی موارد موردتأیید قرار گرفت. با توجه به

اینکه اختلاف‌نظر افراد خبره در دو مرحله، از حد آستانه (۰/۲)، کمتر بود، فرایند متوقف شد؛ بنابراین صحت اعتبار مدل مفهومی تأیید شد [۲۷]. نتایج بررسی‌های مرحله دوم نظرسنجی در جدول ۲ و مقایسه نتایج مرحله اول و دوم فازی در جدول ۳، ارائه شده است.

جدول ۲. نتایج بررسی‌های مرحله دوم نظرسنجی [۲۵]

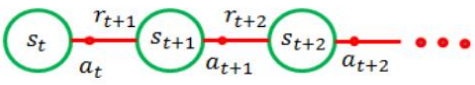
ردیف	موارد موردبررسی	u	m	l	X	نتیجه
۲	تعاملات بین عامل شرکت پخش و مشتریان					
۱-۲	استراتژی‌های همکاری	۱/۰۰	۰/۸۵	۰/۶۰	۰/۸۱۷	تأیید
۶	تعاملات بین شرکت پخش و نهاد قانون‌گذار و ناظر					
۱-۶	درخواست تغییر قوانین غیر اثربخش	۰/۹۵	۰/۷۰	۰/۴۵	۰/۷۰۰	تأیید
۸	تعاملات بین شرکت پخش و زیرساخت فناوری اطلاعات و ارتباطات					
۵-۸	کالیبراسیون‌های دوره‌ای	۰/۹۵	۰/۷۵	۰/۵	۰/۷۳۳	تأیید
۱۵	تعامل عامل‌ها با محیط (به نهادهای متولی امنیت سایبری)					
۱۴-۱۵	گزارش امنیتی	۰/۹۵	۰/۷۰	۰/۴۵	۰/۷۰۰	تأیید

جدول ۳. مقایسه نتایج مرحله اول و دوم دلفی فازی [۲۵]

موارد موردبررسی	X_1	X_2	(X_1-X_2)	نتیجه
۱-۲	۰/۶۵	۰/۸۱۷	۰/۱۶۷	<۰/۲
۱-۶	۰/۶۸۳	۰/۷۰۰	۰/۰۱۷	<۰/۲
۵-۸	۰/۶۵	۰/۷۳۳	۰/۰۸۳	<۰/۲
۱۴-۱۵	۰/۶۵	۰/۷۰۰	۰/۰۵۰	<۰/۲

با توجه به اینکه برای مدل‌سازی ریاضی عامل‌بنیان از مدل مفهومی عامل‌بنیان به‌دست‌آمده استفاده می‌شود، پژوهش از حالت کیفی به کمّی تبدیل می‌شود.

برای مدل‌سازی باید از روشی استفاده شود که رویکرد عامل‌بنیان داشته باشد. یکی از این روش‌های عامل‌بنیان، روش یادگیری تقویتی است. یادگیری تقویتی یک نگاشت از موقعیت به عمل است؛ یعنی نشان می‌دهد در هر موقعیت یا حالتی چه عملی باید انجام شود تا به هدف رسید. زنجیره حالت و عمل در شکل ۴، ارائه شده است.



شکل۴. زنجیره حالت و عمل

برای حل مسئله‌ای که در این پژوهش تعریف شده است، با توجه به اینکه مطابق مدل مفهومی، عامل‌های مختلفی با هم در تعامل بوده و هر یک از عامل‌ها به دنبال دستیابی به اهداف خود هستند، از سیستم‌های چندعاملی غیرهمکارانه استفاده می‌شود که در آن گروهی از عامل‌ها برای رسیدن به اهداف خود تلاش می‌کنند. برای آموزش عامل‌ها برای رسیدن به یک سیاست بهینه از یادگیری تقویتی، روش یادگیری تفاوت زمانی و الگوریتم یادگیری کیو استفاده شده است. در این روش ارزش زوج (حالت ـ عمل) در همان لحظه انجام می‌شود و نیازی به اتمام فرایند برای ارزش‌گذاری زوج‌های (حالت ـ عمل) نیست؛ بنابراین این روش، امکان پایش برخط را نیز فراهم می‌آورد. در الگوریتم کیو، به‌روزرسانی ارزش زوج (حالت ـ عمل) (s_t , a_t) طبق رابطه ۴ که به «رابطه بلمن»[1] معروف است، انجام می‌شود.

رابطه (۴)

$$Qnew(s_t , a_t) = Qcurrent(s_t , a_t) + \alpha[r_{t+1} + \gamma\ max\ Q(s_{t+1} , a_t) - Q(s_t , a_t)]$$

در رابطه ۴، α نرخ یادگیری[2] است که عددی بین صفر و یک $\alpha \in [0,1]$ است.

هرچه آلفا بیشتر باشد، سرعت یادگیری هم بالاتر می‌رود؛ ولی باید توجه داشت که مقادیر بزرگ آلفا، یادگیری را ناپایدار می‌کند.

نرخ یادگیری زیاد و نزدیک به یک، باعث کُندشدن یادگیری عامل می‌شود؛ زیرا عامل به یادگیری‌هایی که از هر مرحله به‌دست آورده است، توجهی نکرده و فقط به یادگیری‌های جدید توجه می‌کند؛ بنابراین همگرایی دیرتر شکل می‌گیرد.

γ نیز ضریب تنزیل یا تخفیف[3] است که عددی بین صفر و یک $\gamma \in [0,1]$ است. اگر پارامتر گاما نزدیک به صفر باشد، عامل تمایل دارد فقط پاداش فوری را در نظر بگیرد. اگر گاما نزدیک به یک باشد، عامل به پاداش‌های آینده توجه بیشتری خواهد کرد و تمایل دارد پاداش را به تأخیر بیندازد.

$max\ Q(s_{t+1} , a_t)$، بیشترین مقدار ارزش کیو است، وقتی اقدام a_t در وضعیت s_{t+1} انجام شود.

$Q(s_t , a_t)$، سود بلندمدت ناشی از اقدام a تحت سیاست π در وضعیت کنونی s است (مجموع پاداش‌های دریافت‌شده وقتی که از حالت s شروع و عمل a انجام شود و به دنبال آن خط‌مشی موجود دنبال شده باشد) [۳۴].

1. Bellman
2. Learning Rate
3. Discount Factor

برای محاسبه میزان ارزش کیو باید زوج‌های «حالت ـ عمل» تعیین شود. بدین منظور باید عامل اصلی که شرکت پخش محصولات دارویی است، موردبررسی قرار گیرد. نخستین موردی که در رابطه با عامل باید مشخص شود، هدف اصلی است. در این پژوهش هدف اصلی، افزایش سود عملیاتی است. سود عملیاتی بر اساس رابطه ۵، محاسبه می‌شود.

رابطه (۵)

سود عملیاتی= درآمد عملیاتی– هزینه کالاهای فروخته‌شده– (هزینه‌های عملیاتی– استهلاک)

با توجه به رابطه ۵، دو دسته کلی درآمدها و هزینه‌های عملیاتی باید در این رابطه موردتوجه قرار گیرند؛ بنابراین باید فرایندها و شاخص‌هایی که روی آن‌ها تأثیر زیادی دارند، در نظر گرفته شوند. فرایندهای تأثیرگذار شامل بازاریابی و فروش، مدیریت تأمین، انبارش، بارگیری، حمل و تحویل هستند. مهم‌ترین شاخص‌های عملکردی باید استخراج و وزن‌دهی شوند و در دوره‌های تعیین‌شده مورد پایش و اندازه‌گیری قرار گیرند.

زوج‌های «حالت ـ عمل» نیز بر اساس شاخص‌های دارای بیشترین تأثیرگذاری روی سود عملیاتی شناسایی می‌شوند و جدول کیو بر اساس آن‌ها ایجاد می‌شود. شاخص‌ها از مصاحبه‌های انجام‌شده استخراج شده و وزن‌دهی آن‌ها بر اساس روش مقایسات زوجی صورت می‌گیرد. برای وزن‌دهی شاخص‌ها گام‌های زیر انجام شده است:

گام اول: ایجاد ماتریس مقایسات زوجی و تکمیل ماتریس توسط سه نفر خبره؛ به این صورت که شاخص‌های هر سطر با ستون‌ها مقایسه و اولویت آن در محل تلاقی سطر و ستون نوشته می‌شود. برای امتیازدهی از مقیاس ۹ درجه ساعتی[1] مطابق جدول ۴، استفاده می‌شود.

جدول ۴. جدول امتیازدهی در مقایسات زوجی

ارزش	j نسبت به i وضعیت مقایسه	توضیح
۱	ترجیح یکسان[2]	شاخص i نسبت به j اهمیت برابر دارد یا نسبت به هم ارجح نیستند.
۳	کمی ارجح[3]	شاخص i نسبت به j کمی مهم‌تر است.
۵	خیلی ارجح[4]	شاخص i نسبت به j مهم‌تر است.
۷	خیلی زیاد ارجح[5]	شاخص i نسبت به j دارای ارجحیت خیلی بیشتری است.
۹	کاملاً ارجح[6]	شاخص i نسبت به j خیلی مهم‌تر و قابل‌مقایسه با j نیست.
۶-۴-۲	بینابین	ارزش‌های بینابین را نشان می‌دهد. برای مثال ۸ نمایانگر اهمیتی بیشتر از ۷ و کمتر از ۹ برای i است.

1.Saaty
2. Equally Preferred
3. Moderately Preferred
4. Strongly Preferred
5. Very Strongly Preferred
6. Extrmely Preferred

اعداد روی قطر ماتریس، همگی ۱ و خانه‌های پایین قطر ماتریس، معکوس اعداد بالای قطر ماتریس هستند.

گام دوم: در این گام نرمال‌سازی ماتریس صورت می‌گیرد.

گام سوم: متوسط اعداد هر سطر محاسبه و روبه‌روی آن نوشته می‌شود که وزن متغیر همان سطر خواهد بود.

گام چهارم: متوسط‌گیری از اوزان ارائه‌شده توسط خبرگان انجام می‌شود. درنهایت بر اساس نتایج نهایی، اولویت‌بندی شاخص‌ها صورت می‌گیرد.

در مرحله بعد باید برای هر یک از شاخص‌های تعیین‌شده، ماتریس زوج «حالت ـ عمل» ایجاد شود. عملکرد فرایندها باید بر اساس شاخص‌ها به‌صورت مستمر مورد پایش و اندازه‌گیری قرار گیرند و نتایج پایش و اندازه‌گیری برای تعیین انحراف از اهداف مورداستفاده قرار گیرند. پس از وقوع هر حالت، بهترین اقدام تعیین و اثرات آن موردبررسی قرار می‌گیرد. فرایند الگوریتم یادگیری Q، مطابق گام‌های تعیین‌شده زیر است:

گام اول: مقداردهی اولیه

گام دوم: انتخاب عمل a

گام سوم: انجام عمل

گام چهارم: سنجش پاداش

گام پنجم: به‌روزرسانی جدول و سپس برگشت به گام دوم

حالت‌ها، موارد دریافتی از محیط هستند که تحت کنترل عامل نیستند. اقدام‌ها کنش‌هایی هستند که عامل در قبال حالت‌های مختلف از خود نشان می‌دهد. در جدول ۵، نمونه‌ای از جدول کیو ارائه شده است.

جدول ۵. نمونه جدول کیو

عمل (a)	**a_1**	**a_2**	**…**	**a_m**	**Q_{max}**	**پاداش r_t**
حالت (s)						
s_1	Q_{11}	Q_{12}	…	Q_{1m}	Q_{1x}	r_1
s_2	Q_{21}	Q_{22}	…	Q_{2m}	Q_{2x}	r_2
s_3	Q_{31}	Q_{32}	…	Q_{3m}	Q_{3x}	r_3
…	…	…	…	…	…	…
s_n	Q_{n1}	Q_{n2}	…	Q_{nm}	Q_{nx}	r_n
امید ریاضی پاداش‌ها						$E\{r_t\}$

مقادیر Q برای هر زوج «حالت ـ عمل»، مطابق رابطه ۶ که به رابطه بلمن معروف است، محاسبه می‌شود.

رابطه (۶)

$$Q_{\text{جدید}}(s_t , a_t) = Q_{\text{فعلی}}(s_t , a_t) + \alpha[r_{t+1} + \gamma\, max\, Q(s_{t+1} , a_t) - Q(s_t , a_t)]$$

با توجه به توضیحات ارائه‌شده در بالا برای آلفا و گاما، مقدار ۰/۲= α و ۰/۸ =γ مناسب است. با این مقادیر، یادگیری پایدارتر و سود بلندمدت موردتوجه قرار می‌گیرد؛ سپس از میان مقادیر Q محاسبه‌شده، بزرگ‌ترین Q به‌عنوان Q_{max} انتخاب می‌شود.

بر اساس نتایج حاصل از اجرای هر اقدام در حالت موردنظر، میزان پاداش هر حالت ـ اقدام، مطابق جدول ۶ تعیین شده و محاسبات پاداش انجام می‌شود؛ سپس بر اساس آن دوباره جدول کیو به‌روزآوری می‌شود. مراحل محاسبه جدول پاداش و به‌روزآوری جدول کیو تا زمان دستیابی به هدف ادامه می‌یابد. پاداش هر وضعیت یا حالت، معادل پاداش Q_{max} خواهد بود.

جدول ۶. تعیین میزان پاداش بر اساس نتایج اقدام‌ها

نتیجه انجام اقدام‌ها	میزان پاداش
بیش از ۲۰ درصد انحراف مثبت از هدف تعیین‌شده	+۶۰۰
معادل ۱۵ تا ۲۰ درصد انحراف مثبت از هدف تعیین‌شده	+۵۰۰
معادل ۱۰ تا ۱۵ درصد انحراف مثبت از هدف تعیین‌شده	+۴۰۰
معادل ۵ تا ۱۰ درصد انحراف مثبت از هدف تعیین‌شده	+۳۰۰
معادل ۵ صفر تا ۵ درصد انحراف مثبت از هدف تعیین‌شده	+۲۰۰
بدون انحراف از هدف تعیین‌شده (تحقق هدف)	+۱۰۰
معادل ۵ صفر تا ۵ درصد انحراف منفی از هدف تعیین‌شده	−۱۰۰
معادل ۵ تا ۱۰ درصد انحراف منفی از هدف تعیین‌شده	−۲۰۰
معادل ۱۰ تا ۱۵ درصد انحراف منفی از هدف تعیین‌شده	−۳۰۰
معادل ۱۵ تا ۲۰ درصد انحراف منفی از هدف تعیین‌شده	−۴۰۰
بیش از ۲۰ درصد انحراف منفی از هدف تعیین‌شده	−۵۰۰

محاسبه پاداش کل هر شاخص بر اساس رابطه ۷، انجام می‌شود که امید ریاضی پاداش‌های هر شاخص است.

رابطه (۷)

$$r_i = \frac{\sum_{t=1}^{n} r_t}{n} = E\{r_t\}$$

هدف عامل این است که میزان پاداش هر شاخص را حداکثر کند. علت میانگین‌گرفتن از پاداش‌ها به‌صورت زیر اثبات می‌شود.

$r \sim N(\mu, \sigma^2)$

$[r_1, r_2, r_3, \dots r_n]$ مستقل

تخمین $\hat{\mu} = E(r) \cong \frac{1}{n} \sum_{i=1}^{n} r_i$

$$E(\hat{r}) = E(\frac{1}{n}\sum_{i=1}^{n} r_i) = \frac{1}{n}\sum_{i=1}^{n} E(r_i)$$

$$E(r_i) = \mu$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n} * n\mu = \mu$$

$$var(\hat{\mu}) = var(\frac{1}{n}\sum_{i=1}^{n} r_i)$$

$$var(r_i) = \sigma^2$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} var(r_i) = \frac{n * \sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$$\lim_{n\to\infty} var(\hat{\mu}) = 0$$

پس از محاسبه مقدار پاداش کل برای هر شاخص، به‌دلیل اینکه تمامی ۲۰ شاخص اولویت‌بندی‌شده در جدول ۷، بر روی سود عملیاتی تأثیرگذار هستند، پاداش کل شاخص‌ها از رابطه ۸، محاسبه می‌شود. چون میزان تأثیرگذاری شاخص‌ها بر روی هدف یکسان نیست، بنابراین ضریب وزنی نیز باید در مدل ریاضی در نظر گرفته شود. هدف عامل این است که مقدار پاداش کل تمامی شاخص‌ها را حداکثر کند.

رابطه (۸) $\mathrm{R_t} = \mathrm{R_{t+1}} + \mathrm{R_{t+2}} + \mathrm{R_{t+3}} + \cdots = \sum_{k=0}^{n} \mathrm{R_{t+k+1}}$

با درنظرگرفتن ضریب وزنی برای شاخص‌ها، رابطه ۹، ایجاد می‌شود.

رابطه (۹) $\mathrm{R_t} = \mathrm{w_1} * \mathrm{R_{t+1}} + \mathrm{w_2} * \mathrm{R_{t+2}} + \mathrm{w_3} * \mathrm{R_{t+3}} + \cdots = \sum_{k=0}^{n} \mathrm{w_i} * \mathrm{R_{t+k+1}}$

در سیر مراحل یادگیری عامل، هر چه به جلو حرکت کرد، همگرایی بیشتر می‌شود؛ بنابراین باید اثرات پاداش‌های جدید نسبت به پاداش‌های قبلی افزایش یابد. برای افزایش اثر پاداش‌های جدید باید ضریب تنزیل یا تخفیف را نیز به رابطه ۹، اضافه کرد. با اعمال ضریب تنزیل در رابطه ۹، رابطه ۱۰، حاصل می‌شود.

$$R_t = w_1 * R_{t+1} + w_2 * \gamma * R_{t+2} + w_3 * \gamma^2 * R_{t+3} + \cdots = \sum_{k=0}^{n} w_i * \gamma^k * R_{t+k+1} \quad \text{رابطه(۹)}$$

در رابطه ۱۰، R_t پاداش کل شاخص‌های عملکردی تأثیرگذار بر روی هدف، w_i ضریب وزنی شاخص‌های عملکردی، γ ضریب تنزیل یا تخفیف $\gamma\epsilon[0,1]$ که عددی بین صفر و یک است و n، تعداد شاخص‌های عملکردی است؛ بنابراین طبق رابطه‌های به‌دست‌آمده برای افزایش سود عملیاتی شرکت پخش که هدف اصلی عامل موردنظر است، اقدامات زیر انجام شد:

۱. شاخص‌های مهم درآمدی و هزینه‌ای که بر روی سود عملیاتی مؤثر هستند، با استفاده از روش مصاحبه، تعیین و بر اساس روش مقایسات زوجی، ضریب وزنی آن‌ها تعیین و اولویت‌بندی شد؛

۲. برای هر یک از شاخص‌های تعیین‌شده، حالت‌ها یا وضعیت‌های مختلف محیطی تعیین و برای هر یک از وضعیت‌ها، اقدام‌های مؤثر سازمان تعیین شدند؛

۳. مقدار ارزش هر عمل (Q)، طبق رابطه یادگیری تقویتی بلمن محاسبه شد؛

۴. مقدار پاداش حداکثری بر اساس حداکثر مقدار ارزش محاسبه‌شده (Q_{max}) برای اقدام‌های مختلف، تعیین و به‌عنوان پاداش حالت موردنظر در نظر گرفته شد؛

۵. مقدار پاداش برای حالت‌های مختلف هر شاخص محاسبه شد؛

۶. امید ریاضی پاداش‌های به‌دست‌آمده برای هر شاخص محاسبه شد؛

۷. پاداش کل نیز بر اساس رابطه ۱۰، محاسبه شد. به این ترتیب که حاصل‌ضرب ضریب وزنی هر شاخص در نرخ تنزیل و امید ریاضی پاداش‌های آن شاخص محاسبه و درنهایت نتیجه حاصل‌ضرب‌ها برای شاخص‌های تعیین‌شده با یکدیگر جمع شد. هدف که دستیابی به بالاترین میزان پاداش کل است و به افزایش سود عملیاتی منجر می‌شود، حاصل شد.

با روش پیشنهادی، برای حالت‌های مختلف محیطی که وقوع می‌یابند، بهترین اقدام‌های سازمان تعیین می‌شود. اقداماتی که با انجام آن‌ها دستیابی به پاداش حداکثری محقق می‌شود و درنهایت به افزایش سود عملیاتی شرکت پخش منجر می‌شود.

اولویت‌بندی شاخص‌ها بر اساس ضریب وزنی در جدول ۷، ارائه شده است.

جدول ۷. جدول اولویت‌بندی شاخص‌ها

اولویت	عنوان شاخص	ضریب وزنی
۱	درصد افزایش سود عملیاتی	۰/۱۷۴
۲	درصد تغییر هزینه‌ها	۰/۰۹۹
۳	میانگین درصد رضایت مشتریان	۰/۰۷۲
۴	دوره وصول مطالبات	۰/۰۷۱
۵	میانگین زمان تحویل سفارش‌ها	۰/۰۷۰
۶	سطح خدمات	۰/۰۶۴
۷	درصد اقلام راکد در انبارهای توزیع	۰/۰۵۹
۸	درصد اقلام راکد در انبارهای پخش	۰/۰۵۸
۹	میانگین دوره گردش کالاها در انبار	۰/۰۴۹
۱۰	میانگین ارزش ریالی تحویلی در هر بار حمل	۰/۰۴۵
۱۱	نسبت مسافت طی‌شده به حجم ریالی تحویلی	۰/۰۴۳
۱۲	درصد سفارش‌های مواجه‌شده با کسری	۰/۰۴۱
۱۳	سرانه فروش مراکز توزیع	۰/۰۳۹
۱۴	درصد محموله‌های مرجوعی	۰/۰۲۲
۱۵	میانگین درصد رشد قیمت خرید محصولات (پس از حذف اثر تورم)	۰/۰۲۰
۱۶	درصد داروهای منقضی‌شده	۰/۰۱۸
۱۷	درصد رشد تنوع سبد کالا	۰/۰۱۷
۱۸	درصد محصولات آسیب‌دیده در بارگیری، حمل و تحویل	۰/۰۱۶
۱۹	درصد محصولات آسیب‌دیده در انبارداری	۰/۰۱۵
۲۰	درصد محموله‌های فراخوان شده	۰/۰۱۰

نمونه‌ای از حالت‌ها یا وضعیت‌های مختلفی که ممکن است برای هر شاخص وقوع یابد و همچنین اقدام‌های مرتبط با آن‌ها در جدول ۸، با عنوان جدول ارتباط بین حالت‌ها و اقدام‌ها، ارائه شده است. درمجموع ۴۰ حالت و ۸۳ اقدام در این رابطه شناسایی شد. برای تصدیق مدل ریاضی عامل‌بنیان به‌دست‌آمده از نظرهای سه نفر از افراد خبره بهره گرفته شد و نظرهای آن‌ها در تکمیل مدل مورداستفاده قرار گرفت. برای صحه‌گذاری مدل ریاضی، داده‌های موردنیاز مربوطه به سه شاخص نمونه دوره وصول مطالبات، میانگین زمان تحویل سفارش‌ها و سطح خدمات، شامل وضعیت‌های وقوع‌یافته، اقدام‌های انجام‌شده و نتایج و میزان تأثیر آن بر روی سود عملیاتی از یکی از شرکت‌های بزرگ پخش دارو که از نظر فروش و سهم بازار در بین سه شرکت برتر ایران قرار داشت، جمع‌آوری شد؛ سپس بر اساس وضعیت‌های وقوع‌یافته و اقدام‌های انجام‌شده، میزان پاداش کل بر اساس مدل ریاضی حاصل از این پژوهش محاسبه و نتایج حاصل از محاسبات با نتایج واقعی مقایسه شد. انحراف معناداری در این رابطه مشاهده نشد. نتایج نشان

داد که با پایش مستمر شاخص‌های عملکردی شرکت پخش دارو و تعیین میزان انحراف از اهداف فرایندی تعیین‌شده و انجام اقدام‌های مناسب در وضعیت‌های مختلف وقوع یافته، هدف کل که افزایش سود عملیاتی است، حاصل می‌شود.

با یادگیری مستمری که در این مدل صورت می‌گیرد، مقادیر شاخص‌های فرایندی به‌طور مستمر در حال بهبود است و این فرایند تا دستیابی به میزان بهینه ادامه می‌یابد.

برای تحلیل حساسیت، میزان پاداش برای یک حالت یکسان و با اقدام‌های مختلف محاسبه می‌شود و بر اساس رابطه ۱۰، میزان پاداش کل محاسبه می‌شود و نتایج موردمقایسه قرار می‌گیرد. به‌طور خلاصه می‌توان نتایج نهایی حاصل از این پژوهش را به شرح زیر عنوان کرد:

شرکت پخش دارو باید بر اساس بیست شاخص ارائه‌شده در جدول ۷، عملکرد خود را به‌صورت برخط پایش کند. با هر تغییر حالت یا وضعیتی، مقادیر شاخص‌های عملکردی تغییر می‌کند و چنانچه این مقادیر خارج از محدوده هدف‌گذاری‌شده باشد، انجام اقدام مؤثری برای برگرداندن مقدار شاخص به محدوده هدف‌گذاری‌شده ضرورت می‌یابد. با توجه به اینکه اقدامات مختلفی قابل‌انجام است، باید اقدامی صورت گیرد که میزان پاداش کل آن بالاتر از سایر اقدام‌ها باشد و به افزایش سود عملیاتی شرکت منجر شود.

برای محاسبه پاداش کل از رابطه ۱۰ که یکی از دستاوردهای پژوهش است، استفاده می‌شود. به‌عنوان نمونه، چنانچه شاخص میانگین زمان تحویل سفارش‌ها به میزان ۳۰ درصد افزایش یابد و میزان شاخص را از میزان هدف‌گذاری شده ۳۰ ساعت به ۴۲ ساعت افزایش دهد و این افزایش به‌دلیل وقوع حالت افزایش غیرمترقبه تقاضا در بازار دارو باشد، باید اقدامی انتخاب و انجام شود که بیشترین تأثیر را در بهبود مقدار شاخص داشته باشد و موجب قرارگیری آن در محدوده هدف‌گذاری شده شود؛ همچنین سود عملیاتی شرکت را نیز افزایش دهد. دو مورد از اقدام‌هایی که در این زمینه می‌توان انجام داد، توسعه ناوگان حمل‌ونقل و افزایش زمان آماده‌به‌کاری خودروهای حمل‌ونقل با نگهداری و تعمیرات پیشگیرانه است. با محاسبه میزان پاداش هر یک از اقدام‌ها به روش یادگیری کیو، توسعه ناوگان حمل‌ونقل اقدام مؤثرتری است و پاداش بیشتری را به خود اختصاص می‌دهد؛ ولی باید مشخص شود که توسعه ناوگان حمل‌ونقل به چه میزانی انجام شود. در این رابطه سناریوهای درصد توسعه ناوگان حمل‌ونقل به میزان ۳، ۵، ۷ و ۱۰ درصد تعیین شد. میزان پاداش کل بر اساس رابطه ۱۰، به‌دست آمد. نتایج نشان داد که با هفت درصد توسعه ناوگان حمل‌ونقل، بیشترین پاداش کل به‌دست می‌آید و به میزان ۱۲/۰۴ ساعت میانگین زمان تحویل سفارش‌ها را کاهش می‌دهد و شاخص به محدوده هدف‌گذاری‌شده می‌رسد. کاهش میانگین زمان تحویل سفارش‌ها به کاهش دوره وصول مطالبات، افزایش نقدینگی و کاهش هزینه‌های مالی شرکت منجر می‌شود و درنتیجه سود عملیاتی را افزایش می‌دهد که هدف اصلی پژوهش است.

شرکت‌های پخش دارو می‌توانند برای بهبود فرایندها، افزایش سود عملیاتی و تحقق اهداف تعیین‌شده، یک سیستم پایش برخط را در زیرمجموعه‌ها و مراکز توزیع خود ایجاد کرده و بر اساس شاخص‌های کلیدی یادشده، داشبورد مدیریتی خود را ایجاد کنند. داده‌های موردنیاز برای پایش برخط را بر اساس فناوری و زیرساخت‌های اینترنت اشیا، جمع‌آوری، انتقال و مورد تفسیر قرار دهند؛ همچنین شرکت‌های پخش می‌توانند فرایندهای پخش و توزیع دارو را بر اساس ارتباط بین عامل‌ها و قوانین رفتاری حاکم بر آن‌ها و همچنین نمودار مفهومی و ریاضی عامل‌بنیان که از یافته‌های این پژوهش است، با استفاده از یک نرم‌افزار شبیه‌ساز عامل‌بنیان، مانند نت لوگو، انی‌لاجیک یا متلب، شبیه‌سازی و تحلیل حساسیت کنند و بر اساس نتایج، تصمیمات صحیح و به‌موقع را اتخاذ کنند.

جدول ۸. جدول ارتباط بین حالت‌ها و اقدام‌ها

حالت‌ها		اقدام‌های قابل‌انجام
S1		بالاتررفتن هزینه‌های پخش نسبت به تعرفه‌های پخش به میزان x درصد
	A1	کاهش هزینه‌های عملیاتی به میزان y درصد
	A2	افزایش درآمد عملیاتی به میزان y درصد
	A3	کاهش قیمت خرید محصولات به میزان y درصد (پس از حذف اثر تورم)
S2		نایاب‌شدن داروهای موجود در سبد کالا به میزان x درصد به‌دلیل تحریم‌ها
	A4	کاهش دوره وصول مطالبات به میزان y درصد
	A6	افزایش سبد کالا برای داروهای تولید داخل به‌میزان y درصد
S3		حذف ارز دولتی برای x درصد از داروهای موجود در سبد کالا
	A8	افزایش میانگین سرانه فروش مراکز توزیع و نمایندگی‌ها به‌میزان y درصد
	A9	افزایش تعداد ناوگان حمل‌ونقل به میزان y درصد
S4		افزایش میانگین دوره وصول مطالبات در سطح صنعت توزیع دارو به‌میزان x درصد
	A10	حذف مشتریانی که دارای بالاترین دوره وصول مطالبات هستند، به‌میزان y درصد
	A11	تأمین منابع مالی موردنیاز از طریق وام بانکی به‌میزان y درصد
	A12	کاهش میانگین زمان وصول چک‌های برگشتی به‌میزان y درصد
S5		افزایش تورم به میزان x درصد
	A13	کاهش هزینه‌های سربار به‌میزان y درصد
	A14	کاهش هزینه‌های انرژی به‌میزان y درصد
	A15	کاهش هزینه‌های پرسنلی به‌میزان y درصد
...	...	...
S40		افزایش محموله‌های آزمایش نشده تحویلی به شرکت پخش به میزان x درصد
	A80	بررسی تصادفی گزارش‌های آزمایش y درصد از محموله‌های دریافتی از تأمین‌کنندگان داخلی
	A82	قطع همکاری با تولیدکنندگانی که بیش از y درصد از محموله‌های آن‌ها در یک سال گذشته فراخوان شده است.
	A83	افزایش بازدیدها و ارزیابی‌های دوره‌ای شرکت‌های تولیدی طرف قرارداد به‌میزان y درصد

۵. نتیجه‌گیری و پیشنهادها

صنعت داروسازی و توزیع دارو، به‌عنوان یک صنعت استراتژیک، همواره موردتوجه بوده است. توزیع عادلانه و متناسب با نیاز دارو، همواره یکی از دغدغه‌های اصلی در این حوزه است. بهبود هزینه‌های پخش، میانگین زمان تحویل سفارش‌ها و میزان سود عملیاتی، همواره موردتوجه پژوهشگران بوده و در پژوهش‌های مختلف با رویکردهای مختلفی به آن پرداخته شده است.

تاکنون پژوهش‌های متعددی در مورد سیستم پایش برخط، مدل‌سازی عامل‌بنیان و یادگیری تقویتی توسط پژوهشگران داخلی و خارجی انجام شده است. بهبود نظام توزیع دارو با استفاده از سیستم پایش برخط با رویکرد مدل‌سازی عامل‌بنیان و روش یادگیری تقویتی می‌تواند رویکرد جدیدی را در صنعت پخش و توزیع دارو مطرح کند. با توجه به اینکه نظام توزیع دارو از پیچیدگی بالایی برخوردار است و عوامل مختلفی در آن تأثیرگذار هستند، از مدل‌سازی عامل‌بنیان که دارای ویژگی‌های مناسبی برای شبیه‌سازی پدیده‌های اجتماعی پیچیده است، استفاده شد. اطلاعات موردنیاز برای ارائه مدل مفهومی عامل‌بنیان از طریق ابزار مصاحبه نیمه‌ساختاریافته به‌دست آمد. مدل‌سازی عامل‌بنیان از چهار بخش انتخاب روش مدل‌سازی، مدل‌سازی و پیاده‌سازی مدل، سنجش اعتبار مدل و نتیجه‌گیری و ارائه پیشنهادها تشکیل می‌شود. در این پژوهش، به‌دلیل گستردگی فعالیت‌ها، به مدل‌سازی مفهومی که در پژوهش دیگری موردبررسی قرار گرفته، به‌صورت گذرا اشاره شده است [۲۷].

در ادامه مدل‌سازی ریاضی عامل‌بنیان در دستور کار قرار گرفت که قسمت تبدیل مدل کیفی به مدل کمّی را پوشش داد. در مدل‌سازی مفهومی عامل‌بنیان با استفاده از نمودار حالت، عامل‌ها شناسایی شد و سپس تصدیق عامل‌ها و خصوصیات آن‌ها، روابط و تعاملات میان عامل‌ها، شناسایی محیط، تعیین تعاملات میان عامل‌ها و محیط و تعیین قوانین رفتاری عامل‌ها با استفاده از ابزار مصاحبه صورت گرفت. جامعه آماری انتخاب‌شده برای مصاحبه‌ها، افراد خبره در صنعت تولید، پخش عمده و مویرگی دارو و همچنین افراد صاحب‌نظر در حوزه پایش برخط بودند. نمونه‌گیری به روش گلوله برفی صورت گرفت و تا دستیابی به اشباع نظری، مصاحبه‌ها ادامه یافت. برای بهبود سیستم پخش و توزیع محصولات دارویی، با استفاده از سیستم پایش برخط، چهار مرحله جمع‌آوری، ذخیره، تفسیر داده‌ها و واکنش مناسب صورت گرفت. برای جمع‌آوری داده‌های باکیفیت از ابزار اینترنت اشیاء و فناوری اطلاعات و ارتباطات بهره‌گیری شد. برای ذخیره‌سازی و طبقه‌بندی داده‌ها نیز زیرساخت فناوری اطلاعات و ارتباطات موردبهره‌برداری قرار گرفت. درنهایت تفسیر داده‌ها انجام شد. عملکرد سیستم از طریق پایش و اندازه‌گیری شاخص‌های عملکردی فرایندها، اندازه‌گیری شده و میزان انحراف آن‌ها از اهداف و الزامات تعیین‌شده مشخص شد. در صورت شناسایی انحراف عملکرد از اهداف و الزامات، اقدام‌های مختلفی تعیین و اجرا شدند و اثربخشی اقدامات انجام‌شده، موردسنجش قرار گرفت. در صورت

اثربخش‌نبودن اقدامات، اقدام‌های اثربخش جدیدی بر اساس روش یادگیری تقویتی، جایگزین شده و اجرا شد و مجدداً اثربخشی اقدامات موردسنجش قرار گرفت. با توجه به نتایج، در نظام توزیع دارو، پنج عامل اصلی سازمان، نهاد قانون‌گذار و ناظر، تأمین‌کننده‌ها، مشتریان و زیرساخت فناوری اطلاعات و ارتباطات، شناسایی شد که تعاملات آن‌ها با سایر عامل‌ها و محیط، در مدل مفهومی عامل بنیان ارائه شده است. اعتبار مدل مفهومی عامل‌بنیان نیز با استفاده از روش دلفی فازی در دو مرحله و با استفاده از دیدگاه‌های پنج خبره موردتأیید قرار گرفت [۲۷]. پس از صحه‌گذاری مدل مفهومی عامل‌بنیان، با استفاده از روش یادگیری تقویتی، الگوریتم یادگیری کیو، مدل مفهومی به یک مدل ریاضی عامل‌بنیان تبدیل شد. به این منظور ابتدا شاخص‌های استخراج‌شده از مصاحبه‌ها، وزن‌دهی و اولویت‌بندی شدند. در مرحله بعد بر اساس شاخص‌های مختلف، زوج‌های «حالت ـ اقدام» شناسایی شد. درمجموع برای ۲۰ شاخص تعیین‌شده، ۴۰ حالت و ۸۳ اقدام شناسایی شد که نمونه‌ای از ارتباط بین حالت‌ها و اقدام‌ها در قسمت تحلیل و یافته‌های پژوهش ارائه شده است. برای تصدیق مدل ریاضی نیز از دیدگاه‌های خبرگان در مراحل مختلف ایجاد مدل ریاضی استفاده شده و تغییرات موردنیاز اعمال شد. اثبات ریاضی مدل نیز صورت گرفت. برای صحه‌گذاری مدل ریاضی عامل‌بنیان، خروجی محاسبات حاصل از مدل پیشنهادشده با عملکرد واقعی یک شرکت پخش مقایسه شد و کارایی مدل به اثبات رسید.

بر اساس نتایج پژوهش، کیفیت داده‌ها، زیرساخت فناوری اطلاعات و ارتباطات، سنجش و ارزشیابی خودکار و اقدام و ارزیابی خودکار از عوامل مؤثر بر سیستم پایش برخط برای بهبود نظام پخش و توزیع محصولات دارویی هستند که زیرساخت فناوری اطلاعات و ارتباطات، کیفیت داده‌ها، سنجش و ارزشیابی خودکار و اقدام و ارزیابی خودکار به‌ترتیب دارای بالاترین تأثیرگذاری هستند؛ همچنین بر اساس نتایج و انجام مقایسات زوجی، شاخص‌های درصد افزایش سود عملیاتی، درصد تغییر هزینه‌ها، میانگین درصد رضایت مشتریان، دوره وصول مطالبات، میانگین زمان تحویل سفارش‌ها، سطح خدمات، درصد اقلام راکد در انبارهای توزیع و پخش، میانگین دوره گردش کالاها در انبار، میانگین ارزش ریالی تحویلی در هر بار حمل، نسبت مسافت طی‌شده به حجم ریالی تحویلی، درصد سفارش‌های مواجه‌شده با کسری موجودی، سرانه فروش مراکز توزیع، درصد محموله‌های مرجوعی، میانگین درصد رشد قیمت خرید محصولات، درصد داروهای منقضی‌شده، درصد رشد تنوع سبد کالا، درصد محصولات آسیب‌دیده در بارگیری، حمل و تحویل و انبارداری، درصد محموله‌های فراخوان‌شده، به‌ترتیب دارای بیشترین اولویت هستند. ضریب‌های وزنی شاخص‌ها در قسمت تحلیل و یافته‌های پژوهش ارائه شده است؛ بنابراین شاخص‌های کلیدی دارای تأثیرگذاری متفاوتی بر روی سیستم پایش برخط و دستیابی به اهداف هستند؛ همچنین می‌توان میزان سود عملیاتی و شاخص‌های عملکردی شرکت‌های پخش و توزیع دارو را با استفاده از این رویکرد بهبود داد و مدل مفهومی و ریاضی عامل‌بنیان ارائه‌شده از

اعتبار لازم برخوردار هستند. وجه تمایز این پژوهش در مقایسه با پژوهش‌های پیشین، استفاده از رویکرد پایش برخط برای بهبود سیستم‌ها و فرایندهای یک شرکت پخش دارو به‌منظور افزایش سود عملیاتی است. رویکرد مورداستفاده در این پژوهش، جامع‌نگر است و کلیه فرایندهای یک شرکت پخش را موردتوجه قرار می‌دهد. درصورتی‌که در پژوهش‌های پیشین، نگاه جزءنگر وجود داشته و فقط بر روی یک یا تعدادی از فرایندهای سازمان تمرکز شده است. دیدگاه جزءنگر، برخلاف دیدگاه کل‌نگر، نمی‌تواند سود عملیاتی سازمان را از ابعاد مختلف درآمدی و هزینه‌ای افزایش دهد؛ همچنین با توجه به اینکه از روش عامل‌بنیان برای پژوهش استفاده شده است، می‌توان تغییرات عوامل محیطی را نیز شناسایی و شبیه‌سازی کرد.

مهم‌ترین محدودیت این پژوهش این است که در وضعیت‌ها و شرایط یکسان، ممکن است شاهد رفتارهای متفاوتی از شرکت‌های پخش دارویی مختلف بود؛ بنابراین با فرض یکسان‌بودن رفتارهای شرکت‌های پخش در وضعیت‌های مشابه، پژوهش انجام شده است.

با توجه به اینکه تمرکز پژوهش حاضر بر روی صنعت پخش دارو بوده و تعمیم نتایج آن به سایر صنایع، منوط به پژوهش جداگانه‌ای است، می‌توان در پژوهش‌های بعدی، صنعت پخش در سایر صنایع را موردبررسی قرار داد؛ همچنین می‌توان به‌جای استفاده از مدل‌سازی عامل‌بنیان، از سایر رویکردهای مدل‌سازی استفاده کرده و نتایج را با نتایج حاصل از این پژوهش مقایسه کرد. دستاوردهای این پژوهش نشان می‌دهد که سیستم پایش برخط، یکی از رویکردهای مؤثر برای بهبود سیستم‌ها و فرایندها است و می‌تواند در کانون توجه صنایع مختلفی که به‌دنبال ایجاد بهبود در سود عملیاتی، سیستم‌ها یا فرایندهای خود هستند، قرار گیرد.

تعارض منافع. برای ارائه مطالب و نگارش این مقاله هیچ‌گونه کمک مالی از هیچ فرد، نهاد و سازمانی دریافت نشده است و نتایج و دستاوردهای این مقاله به نفع یا ضرر سازمان یا فردی خاص نخواهد بود. حضور نویسندگان در این پژوهش به عنوان شاهدی بی‌طرف ولی متخصص بوده است و نویسندگان هیچ‌گونه تعارض منافعی ندارند.

منابع

1. Abbasi Siar, S., Hashemighohar, M. & Feyzi, A. (2022). Agent-based Modeling Behavior of Traders in Tehran Stock Exchange (Case Study: Mobarakeh Esfahan Steel Company). *Modern Research in Decision Making*, *7(1)*, 88-114. (In Persian)
2. Abbasi Siar, S., Keramati, M.A. & Motadel, M.R. (2021). Agent-Based Simulation of Concumer Behavior in Impulse Buying. *Journal of Industrial Management Studies*, *19(62)*, 99-138. (In Persian)
3. Abolfathi, E., Toloie Eshlaghy, A., & Hamidi zadeh, M.R. (2018). An operational guide for agent-based modeling based on the classification of research conducted in the humanities: dissemination of innovation in Iran. *New research in decision making*, *3(2)*, 1-25. (In Persian)

4. Alidoost, F., Bahrami, F., & Safari, H. (2020). Multi-Objective Pharmaceutical Supply Chain Modeling in Disaster (Case Study: Earthquake Crisis in Tehran). *The Journal of Industrial Management Perspective*, *10(39)*, 99-123. (In Persian)
5. Araya, F. (2020). Agent Based Modeling: a tool for Construction Enginnering and Management, *Research Gate*, *32(5)*, 1-18.
6. Arel, L., Urbanik, T., & Kohls, A.G. (2010). Reinforcement Learning-based Multi agent system for network traffic signal control, *IET intelligent transport system*, *4(2)*, 128-135.
7. Asghari Oskoei, M.R., Fallahi, F., Doostizadeh, M., & Moshiri, S. (2018). Reinforcement Learning Applied to Multi Agent Modeling the case of the Iranian Power Market. *Iranian Energy Economics Research*, *7(25)*, 1-40. (In Persian)
8. Azar, A., Mashayekhi, M., Amiri, M., & Safari, H. (2021). Modeling Steel Supply Chain and Estimating Its Consumption through ABM Methodology. *The Journal of Industrial Management Perspective*, *11(41)*, 33-51. (In Persian)
9. Baha, M.M. (2023). Impact of Artificial Intelligence on Supply Chain Management Performance. *Journal of Service Science and Management*, *2023(16)*, 44-58.
10. Baohong, G., Zhiming, W., Weichun, Ge., Chenggang, W., Shuyang, Li.& Qiang, G. (2012). The Application of IEC61850 Standard in the Online Monitoring System of the Smart Substation. *IEEE*, 1-6.
11. Bonabeau, E. (2023). Agent-Based Modeling: Methods and Techniques for simulation human systems, *Colloquium*, *99(3)*, 7280-7287.
12. Cenani, S. (2021). Emergency and Complexity in Agent-Based Modeling: Review of state-of-the-art tools, *Research Gate*, *2(2)*, 1-24.
13. Edrisi, A., Bagherzadeh Chehreh, K. & Nadi, A. (2017). Presenting an Innovative Model Based on Reinforcement Learning for Intelligent Routing of Vehicles in Dynamic Network (Case Study of Isfahan City). *International Conference on Civil Engineering*, Architecture and Urban Planning of Contemporary Iran, 1-13. (In Persian)
14. Emec, S., Kruger, J. & Seliger, G. (2016). Online fault-monitoring in machine tools based on energy consumption analysis and non-invasive data acquisition for improved resource-efficiency. *Elsevier*, *40*, 236-243.
15. Eskandar, F., Radfar, R. & Toloi, A. (2019). A Simulation Model to Predict and Improve the Performance of the Working Team and Achieve Better Human Resource Management Strategies. *Journal of Strategic Management Researches*, *25(74)*, 49-73. (In Persian)
16. Farah Baksh, M., Modiri, M., Khatami Firouzabadi, M.A., & Pour Ebrahimi, A.R. (2023). Power Industry's Lise Cycle Simulation using Agent Based Modeling. *The Journal of Industrial Management Perspective*, *12(48)*, 9-35. (In Persian)
17. Fazeli, R. & Fazeli, A. (2016). Study methods of online monitoring equipment, systems and processes in nuclear power plants. *3rd National Conference on Engineering Science Development*, 1-10. (In Persian)
18. Gasperov, B., Begusic, S., Simovic, Petra, P., & Kostancar, Z. (2021). Reinforcement Learning Approaches to optimal market making, *Mathematics*, *9(2689)*, 1-22.
19. Gioiosa, R., Kestor, G. & Kerbyson, D.J. (2014). Online Monitoring Systems for Performance Fault Detection. *IEEE, (1)*, 1475-1484.
20. Hajibabaie, M., & Behnamian, J. (2021). Stochastic Bilevel programing Design of A JIT Pharmaceutical Supply Chain Network: Modeling and Algorithm. *The Journal of Industrial Management Perspective, 11(4)*, 137-165. (In Persian)
21. Hesari, S., & Arbabi, S. (2016). Using the Q-Learning in Order to Improving the Efficiency in Drive of Induction Motors. *First International Comprehensive Competition Conference of Engineering Sciences in Iran*, 1-11. (In Persian)
22. Hooshyar, M., Moosavi, J., & Mahootchi, M. (2011). The application of Reinforcement Learning in Optimizing the Exploitation of Surface Reservoirs. *The 6th National Congress of Civil Engineering*, 1-8. (In Persian)
23. Ishabakaki, P. & Kaijage, S. (2015). RFID-based Drug Management and Monitoring System, Case of Public Hospitals in Tanzania, A Review Paper. *Computer Engineering and Applications*, *4(3)*, 165-172.

24. Jafari, T., Zarei, A., Azar, A. & Moghadam, A.R. (2022). Designing a Model for the Impact of Business Inteligence on Supply Chain Performance with an Emphasis on Integration and Agility. *The Journal of Industrial Management Perspective*, *12(47)*, 279-315. DOI: 10.52547/JIMP.12.3.279 (In Persian)
25. Javadpour, A., & Mohammadi, S. (2014). Implement a portable wireless system for online heart rate monitoring. *The first conference on new approaches to medical engineering in the field of cardiovascular diseases*. (In Persian)
26. Jouyban, F., Yousefi-Nejad, M., & Neyshaboori, E. (2018). Presenting a bi-objective stochastic pharmaceutical supply chain model considering time and cost. *Journal of industrial management*, *13(44)*, 15-28. (In Persian)
27. Kangarlou Haghighi, R., Toloie Eshlaghy, A., Motadel, M.R. (2022). Agent-based Conceptual Model of Online Monitoring System, To Improve Pharmaceutical Distribution System. *Journal of Business Intelligence Management Studies*, *10(38)*, 267-316. Doi:10.22054/IMS.2021.58644.1904 (In Persian)
28. Khaksar, M., Derhami, V. & Rezaeian, M. (2018). Improve Performance of Attack in the Team Robots Soccer Using Reinforcement Learning. *Tabriz Journal of Electrical Engineering*, *48(2)*, 585-594. (In Persian)
29. Khalil zadeh, N., sepehri, M.M & Farvaresh, H. (2014). Intelligent sales prediction for pharmaceutical distribution companies: A Data Mining based approach. *Mathematical problems in engineering*, 1-15. (In Persian)
30. Khani, M., Saghaei, A., & Heidarzadeh Hanzaee. (2019). Modeling New product Launch Strategies Within Agent-Based Simulation. *Journal of New Research in Decision Making*, *4(4)*, 177-204. (In Persian)
31. Mehregan, M.R, Abooyee ardakan, M., Sadeqi arani, Z. & Zoormand, O. (2016). Simulation of market buying behavior: an agent-based modeling approach. *Modiriat-E-Farda journal*, *15(48)*, 35-64. (In Persian)
32. Monti, C., Pangallo, M., Morales, G.F., & Bonchi, F. (2022). On Learning Agent-Based Models from data, *arxiv*, *2022*, 1-26.
33. Nikanjam, A., Abdoos, M., & Mahdavi Moghadam, M. (2021). Environments using Knowledge Transfer for Herding Problem. *Journal of Control*, *14(4)*, 55-66. (In Persian)
34. Niki Oskoui, K., Behboudi, D., & Asgharpoor, H. (2018). Designing an agent-based model to determine Iran's strategy in international gas trade. Quarterly *Journal of Energy Policy and Planning Research*, *4(10)*, 67-97. (In Persian)
35. Rabiee, M., razaz, M., & joorabian, M. (2014). Provide a method for monitoring and detecting the mechanical fault of the power transformer winding online. *7th Conference on Power Plants, Hormozgan*. (In Persian)
36. Salehnia, N., Moosavi, F. (2019). *Agent-based modeling using NetLogo software*. First edition, Tehran, Dibagaran Cultural and Artistic Institute, 13-14. (In Persian)
37. Schroeder, B.A. (1995). *Online-Monitoring: A tutorial computing practices*. *5(18)*, 72-77.
38. Sepehrnia, S., & Azimi, B. (2016). Investigate instant and online monitoring and suggest a method for obtaining a sample and how to measure water and wastewater from the perspective of reference standards. *First International Conference on Environmental Pollution Sampling and Monitoring*, 115-122. (In Persian)
39. Shahverdi, K., & Monem, M.J. (2015). Development of Reinforcement Learning Algorithm for Automation of slide Gate Check Structure in Canals. *Journal of Water and Soil*, *29(4)*, 828-837. (In Persian)
40. Wang, L., Wang, Y. (2022). Supply chain Financial Service Management System Based on Block Chain IOT data Sharing and Edge Computing. *Alexandria Engineering Journal*, *61(1)*, 147-158.
41. Wano, F., Tatebe, N., Tajima, Y., Nakata, M., Kovacs, T. & Takadama, K. (2018). Multi Agent Cooperation based on Reinforcement Learning with internal Reward in maze problem, *SICE Journal of control measurement integration*, *11(4)*, 321-330.
42. Zhang, Z., & Wang, D. (2018). EAQR: A Multiagent Q-learning Algorithm for coordination of Multiple Agents, *Hindawi*, *2018(7172614)*, 1-14.