Pro-active Risk Management Model of gas transmission network Using data mining and Markov decision process

Mohammad Sadegh Behrouz[1] - Mohammad Ali AfsharKazemi[2*]
Adel Azar[3]- Ezzatollah Asgharizadeh[4]

Abstract

The aim of the research is predicting critical risks of the gas transmision network and choosing optimal corrective action in optimal time and cost. Risks were predicted with data-mining algorithms based on the CRISP methodology. K-Means, Kohnen, Two Step algorithm and Neural Network Algorithms, C.5 Tree, Nearest Neighbor and Support Vector have been used for clsutering and classification. Markov decision process is also used to select the optimal control action. Decision making problem is based on back down induction in stochastic dynamic programming in the limited time of modeling and simulation and sensitivity analysis and model validation. Based on results, in 97.56% of the agreed data, learning was created and the accuracy and validity of the data mining model was estimated at 86.92%. Also, 13 risks have been identified as critical, and the simulation results show a 92% improvement rate in the cost and 77% in the control action implementation time.

Key words:
Risk Assessment, Mathematical Modeling, Data Mining, Markov Decision Process, Gas Transmission Network.

1. Ph.D. student of Industrial Management, Faculty of Management and Economics, Branch of Science and Research, Islamic Azad University, Tehran, Iran. sadeqbehruz@ut.ac.ir
2. Associate Professor, Faculty of Management, Branch of central Tehran, Islamic Azad University, Tehran, Iran. dr.mafshar@gmail.com(Responsible author)
3. Professor, Management and Accounting Faculty, Tarbiat Modares University, Tehran, Iran.azara@modares.ac.ir
4. Associate professor of management faculty of Tehran University, Tehran, Iran. asghari@ut.ac.ir

مقاله علمی ـ پژوهشی

تاریخ دریافت ۱۴۰۲/۰۳/۲۲ **تاریخ پذیرش ۱۴۰۲ /۰۵/۳۰**

مدیریت کنش‌گرایانه ریسک شبکه انتقال گاز با داده‌کاوی و فرایند تصمیم مارکوف

محمدصادق بهروز[۱] ـ محمدعلی افشارکاظمی[۲*] ـ عادل آذر[۳] ـ عزت‌اله اصغری‌زاده[۴]

چکیده

هدف پژوهش، ارائه مدل پیش‌بینی ریسک‌های بحرانی و اولویت‌دار شبکه انتقال گاز و انتخاب اقدام کنترلی بهینه از نظر زمان و هزینه است. در این پژوهش، پیش‌بینی ریسک‌ها با الگوریتم‌های داده‌کاوی بر اساس متدلوژی CRISP انجام شده و برای خوشه‌بندی از الگوریتم K-Means, Kohnen, Two Step و در ادامه، از الگوریتم‌های شبکه عصبی، درخت C.5، نزدیک‌ترین همسایگی و بردار پشتیبان برای طبقه‌بندی بهره‌گیری شده است. اقدام کنترلی بهینه با فرایند تصمیم مارکوف انتخاب و مسئله تصمیم‌گیری بر اساس استقرای بازگشتی برنامه‌ریزی پویای احتمالی زمان محدود مدل‌سازی و شبیه‌سازی شده و تحلیل حساسیت و اعتبارسنجی مدل صورت پذیرفته است. بر اساس نتایج، در ۹۷/۵۶ درصد از داده‌های موردتوافق، یادگیری ایجاد شده و صحت و اعتبار مدل داده‌کاوی، ۸۶/۹۲ درصد برآورد شده است. همچنین ۱۳ ریسک، بحرانی تشخیص داده شده‌اند و نتایج شبیه‌سازی مدل ریاضی نیز، نرخ بهبود را ۹۲ درصد در هزینه و ۷۷ درصد در زمان اجرای اقدام کنترلی نشان می‌دهد.

واژگان کلیدی: ارزیابی ریسک، مدلسازی ریاضی، داده‌کاوی، فرایند تصمیم مارکوف، شبکه انتقال گاز.

۱. دانشجوی دکتری مدیریت صنعتی گرایش تحقیق در عملیات، دانشکده مدیریت و اقتصاد، دانشگاه علوم و تحقیقات تهران، دانشگاه آزاد اسلامی، تهران، ایران. sadeqbehruz@ut.ac.ir

۲. دانشیار، گروه مدیریت صنعتی، دانشکده مدیریت، واحد تهران مرکز، دانشگاه آزاد اسلامی، تهران، ایران. (نویسنده مسئول).dr.mafshar@gmail.com

۳. استاد، گروه مدیریت صنعتی، دانشکده مدیریت و اقتصاد، دانشگاه تربیت مدرس، تهران، ایران azara@modares.ac.ir

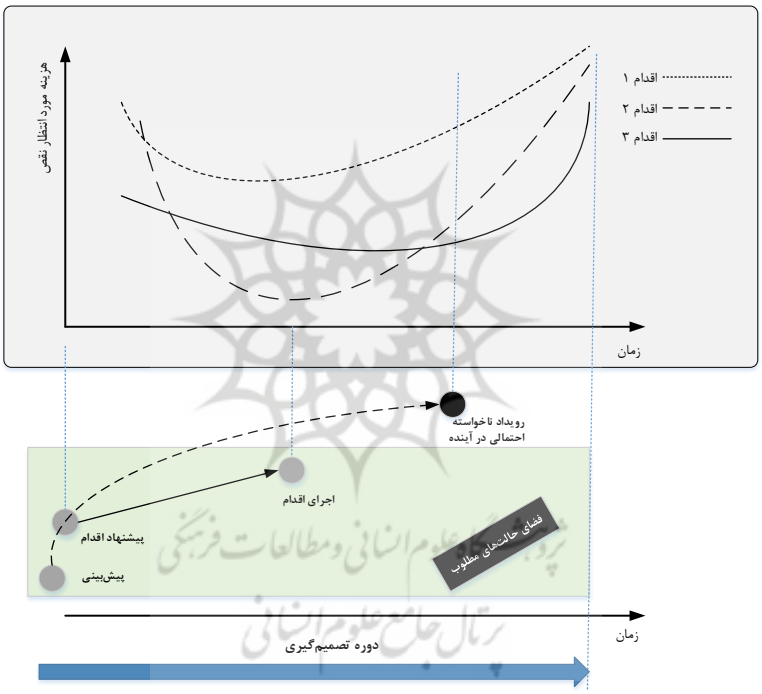
۴. دانشیار، گروه مدیریت صنعتی، دانشکده مدیریت، دانشگاه تهران، تهران، ایران. asghari@ut.ac.ir

مقدمه

یکی از صنایع پرخطر با ریسک‌فاکتورهای متنوع، شبکه‌های انتقال گاز هستند و استفاده از تکنولوژی‌های روز دنیا در این صنعت از اهمیت بالایی برخوردار است. از مهمترین اقدامات برای کاهش مخاطرات، پیاده‌سازی برنامه‌های مدیریت و ارزیابی ریسک است. مدیریت ریسک فرایند نظام‌یافته برای شناسایی، تحلیل و واکنش به ریسک است که متضمن بیشینه کردن احتمال و پیامدهای رویدادهای مثبت و کمینه کردن احتمال و پیامدهای رویدادهای منفی است (قدوسی، میرسعیدی و حسنی، ۱۳۹۹). «داده‌کاوی»، برای تصمیمات صحیح در مواقع عادی و کنترل و مدیریت جهت‌دار امور برای مدیران سازمان‌ها از ابزارهای اساسی محسوب می‌شود ضرورت بهره‌مندی از روش‌های پیش‌نگر و پیش‌بینانه در شناسایی ریسک‌های بحرانی و دارای اولویت حائز اهمیت است (مروتی شریف‌آبادی و عباس‌آبادی، ۱۴۰۱)؛ مسئله‌ای که کمتر در برنامه‌های مدیریت ریسک به‌صورت داده‌محور به آن توجه شده است. ازاین‌رو استفاده از ابزار داده‌کاوی و ارائه مدل پیش‌بینی ریسک‌های بحرانی به‌صورت داده‌محور با بهره‌گیری از الگوریتم‌های داده‌کاوی و یادگیری ماشین در این پژوهش موردتوجه قرار گرفته است؛ به صورتی که همزمان آموزش داده‌ها و افزایش اعتبار نتایج از طریق ترکیب الگوریتم‌ها به‌صورت یکپارچه (و نه مقایسه تکنیک‌ها با یکدیگر) در راستای هدف پژوهش حاصل شود. جمع‌آوری و پردازش داده با سامانه‌های هوشمند، کنترل فرایندهای کاری را تسهیل نموده و واژه «فعالانه» در سیستم‌ها و سامانه‌های اطلاعاتی داده‌محور، اشاره به کشف و تشخیص شرایط بالقوه‌ای که می‌تواند منجر به بروز رویداد تصادفی و ناخواسته شود داشته و باید جلوگیری از گرفتار شدن سیستم تولید در این شرایط را به همراه داشته باشد. (Bousdekis et al, 2018) رویکرد تصمیم‌گیری فعالانه پیش‌نگر یا کنش‌گر به ادبیات موضوع «درک و عکس‌العمل» یا «کشف و اقدام» در یک شرایط آگاهانه اشاره دارد که مبتنی بر مدل «کشف ـ پیش‌بینی ـ تصمیم‌گیری ـ اقدام» است. (Bousdekis and Mentzas, 2019). هر یک از این مراحل به‌عنوان یک عامل پردازش رویداد پیاده‌سازی و اجرا می‌شوند و برای توسعه یک سامانه اطلاعاتی یکپارچه که ویژگی‌های محاسبات رویدادمحور فعالانه را داراست، به‌کار می‌روند. مفهوم تصمیم‌گیری رویدادمحور در شکل ۱ نشان داده شده است. این رویکرد عملیاتی، بر ارائه پیشنهادها و توصیه‌ها برای انجام اقدام بهینه در زمان واقعی موردنیاز متمرکز است. در این مدل هزینه نقص یا ریسک به‌صورت تابعی از زمان نشان داده شده است. در زمان $T=0$ پیش‌بینی صورت گرفته و پیشنهاد ارائه می‌شود. تابع هزینه نقص در بازه زمانی

بین T=0 و پایان زمان تصمیم‌گیری بهینه می‌شود. (پایان زمان تصمیم‌گیری زمانی است که اقدام کنترلی بعدی تدوین می‌شود). به‌عنوان مثال برای سه اقدام پیشنهادی مشخص شده روی شکل، اقدام بهینه اقدام ۲ است و زمان بهینه برای پیاده‌سازی و اجرای اقدام زمانی است که تابع هزینه نقص مینیمم است. در این روش، تصمیم‌گیری به‌جای اینکه بلافاصله بعد از وقوع نقص باشد، بر اساس هشدارها در یک بازه زمانی بهینه و مطمئن صورت می‌گیرد.

شکل ۱. مفهوم تصمیم‌گیری رویداد محور فعالانه (Bousdekis et al, 2018)



نقش کاوش داده‌ها، در تحقق رویکرد تصمیم‌گیری فعالانه پیش‌نگر یا کنش‌گر، پررنگ است. داده‌کاوی فرایند کاملی از اعمال متدلوژی مبتنی بر کامپیوتر است که شامل تکنیک‌های جدید برای اکتشاف دانش از داده‌ها است. (علیخانزاده، ۱۳۹۲). الگوریتم‌های داده‌کاوی به دو دسته کلی نظارتی و غیرنظارتی یا پیش‌بینی و توصیفی تقسیم شده است. در الگوریتم‌های پیش‌بینی، هدف، پیش‌بینی یک ویژگی خاص بر مبنای ویژگی‌های دیگر است. ویژگی پیش‌بینی‌شونده متغیر وابسته و بقیه متغیرها متغیر مستقل نامیده می‌شوند. در الگوریتم‌های توصیفی هدف استخراج الگو از داده‌ها است که نیاز به تحلیل نتایج دارد

و به دنبال کشف راهی برای آگاهی از خصوصیات داده است (درمحمدی و همکاران، ۱۳۹۳). روش‌های مختلفی برای پیاده‌سازی و اجرای پروژه‌های داده‌کاوی وجود دارد. یکی از این موارد متدلوژی CRISP-DM[1] است. موفق‌ترین پروژه‌های داده‌کاوی، در چارچوب این فرایند استاندارد اجرا می‌شود و به دلیل گستردگی کاربرد در نمونه مطالعات مشابه، در این پژوهش از این متد بهره‌گیری شده است. در این روش، نتیجه هر مرحله به نتیجه‌ی مراحل قبلی وابستگی دارد. نکته‌ی قابل‌توجه دیگر در این مدل، امکان بازگشت به هریک از مراحل قبل و اعمال اصلاحات موردنیاز است. این متدلوژی از گام‌های درک کسب‌وکار[2]، درک داده[3]، آماده‌سازی داده[4]، مدل‌سازی[5]، ارزیابی[6] و پیاده‌سازی[7]، تشکیل شده است که هرکدام شامل زیربخش‌های مربوط به خود است. (جعفری نعیمی،۱۳۹۰)، (غضنفری و همکاران، ۱۳۹۵). در این پژوهش برخلاف سایر تحقیقات انجام شده، الگوریتم ترکیبی پیش‌بینی برای بهینه‌سازی طبقه‌بندی داده‌ها، به‌صورت تکاملی به‌کارگیری شده و در هر مرحله، هدف تقویت میزان صحت و اعتبار مدل طبقه‌بندی و افزایش میزان یادگیری داده‌ها است. درحالی‌که سایر پژوهش‌های مشابه، بر رویکرد مقایسه‌ای در به‌کارگیری تکنیک‌های داده‌کاوی متمرکز است و کارایی و صحت این تکنیک‌ها در مقایسه با یکدیگر مورد قضاوت و ارزیابی قرار می‌گیرد. استفاده از قواعد برنامه‌ریزی پویای احتمالی برای تصمیم‌گیری رویدادمحور و انتخاب اقدام اصلاحی مناسب، بر اساس حداکثر مقدار تابع مطلوبیت با توجه به زمان و هزینه بهینه و رویکرد فعالانه و کنش‌گرایانه در فرایند ارزیابی ریسک، از وجوه دیگر تمایز پژوهش است که موردتوجه پژوهشگران قرار گرفته است.

در جدول ۱ تعدادی از پژوهش‌های مشابه قبلی صورت گرفته، ذکر شده است.

روش‌شناسی

این پژوهش بر اساس هدف یک پژوهش توسعه‌ای و از نظر ماهیت پژوهشی کمّی است که برای بیان رفتار بهینه مقوله‌ی موردمطالعه، از ابزار شبیه‌سازی در آن استفاده شده است. در این پژوهش برای داده‌کاوی، از تکنیک‌های یادگیری ماشین در دو دسته داده‌کاوی «پیش‌بینی» و «توصیفی» استفاده شده است. برای خوشه‌بندی داده‌ها از

1. Cross Industry Standard Process For Data Mining
2. Business understanding
3. Data Understanding
4. Data Preparation
5. Modeling
6. Evaluation
7. Deployment

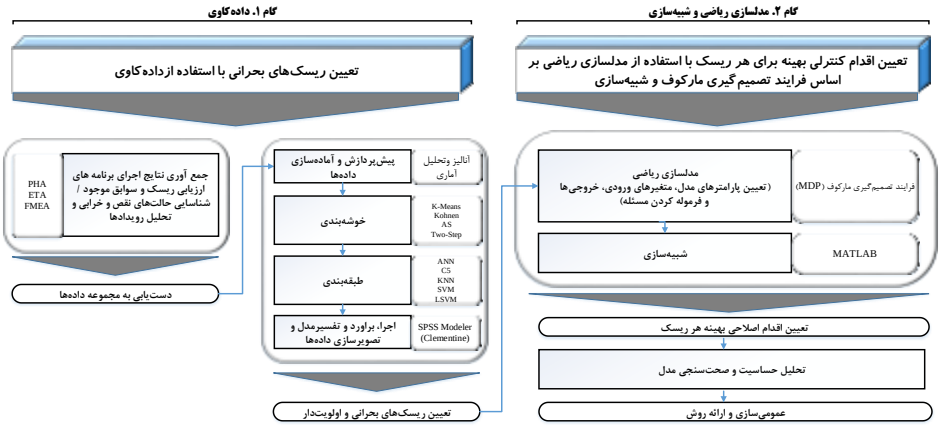
جدول ۱. مروری بر پژوهش‌های قبلی

ردیف	موضوع پژوهش	منابع
۱	ارائه مدل تشخیص ناهنجاری‌ها در جریان داده در حسگرها با بهره‌گیری از یادگیری ماشین	استین وینکل و همکاران (۲۰۲۱)
۲	ارائه مدل فعالانه بهینه‌سازی نگهداری تعمیرات و لجستیک در محیط اینترنت اشیا صنعتی تصمیم‌گیری بر اساس زنجیره مارکوف مبتنی بر پردازش رویدادها	الکساندر بوسدکیس و منتزاس (۲۰۱۹)
۳	فعال‌سازی تصمیمات در نگهداری تعمیرات مبتنی بر شرایط بر اساس پردازش‌ها و محاسبات فعالانه‌ی رویدادمحور	الکساندر بوسدکیس و منتزاس (۲۰۱۸)
۴	تصمیم‌گیری هوشمند با حسگرها برای کاهش هزینه و خطا در پایش سازه‌های پایا و نیمه پایا نسبت با فرایند تصمیم‌گیری مارکوف	نقیب‌هاشمی و همکاران (۱۴۰۰)
۵	مدل بهینه‌سازی درمان دارویی دیابت نوع ۲ بر اساس زمانبندی و توالی درمان‌های دارویی کنترل قند خون با رویکرد فرایند تصمیم‌گیری مارکوف	اقبالی زارچ و همکاران (۱۳۹۶)
۶	بهینه‌سازی یکپارچه نگهداری تعمیرات و سفارش قطعات یدکی در سیستم‌های چندبخشی با تمرکز بر خرابی‌ها و نقص‌های جزئی و ارائه مدل تصمیم‌گیری مبتنی بر زنجیره مارکوف	اکتای کارباگ و همکاران (۲۰۲۰)

الگوریتم‌های K-Means، Kohnen و Two-step و برای طبقه‌بندی از الگوریتم‌های شبکه عصبی، درخت C.5، نزدیک‌ترین همسایگی و بردار پشتیبان استفاده شده است. در این پژوهش هدف از مدل‌سازی دستیابی به مقادیر بهینه‌ی زمان و هزینه برای انتخاب و اجرای بهترین اقدام اصلاحی است. با توجه به ویژگی رویدادمحور بودن فرایند تصمیم‌گیری و شاخص‌های موردنظر مانند تعیین کمترین میزان هزینه و انتخاب اقدام بهینه و همچنین تصمیم‌گیری در شرایط تصادفی بر اساس جستجوی غیرقطعی، از فرایند تصمیم مارکوف در زمان محدود مبتنی بر استقرای بازگشتی برنامه‌ریزی پویای احتمالی، بر اساس معادله بلمن[1] برای شبیه‌سازی رویدادهای گسسته[2]، بهره‌گیری شده است. برای آنالیز هر نقص (رویداد)، حالت‌های مختلف سیستم در نظر گرفته شده است که عبارت‌اند از: ۱. وضعیت سیستم در حالت خطر ۲. وضعیت سیستم در حالتی که یک رویداد تصادفی یا نقص اتفاق بیفتد ۳. وضعیت سیستم با در نظر گرفتن اجرای اقدام ۴. وضعیت سیستم در حالت تأثیرگذاری اقدام. در شکل ۲ مراحل انجام پژوهش نمایش داده شده است.

1. Bellman
2. Discrete Events Simulation

شکل ۲. مراحل انجام پژوهش



فرایند تصمیم‌گیری مارکوف مبتنی بر برنامه‌ریزی پویای احتمالی: فرایند تصمیم‌گیری مارکوف یک مدل بهینه‌سازی برای تصمیم‌گیری در شرایط عدم قطعیت است در این مدل خروجی فرایند مارکوف از جنس سیاستی است برای بهینه کردن زمان اقدام. این سیاست بر اساس حداکثر مطلوبیت مورد انتظار یا حداقل شکست و خرابی مورد انتظار ارزیابی شده است. (نقیب‌هاشمی، توچائی و مروستی، ۱۴۰۰). مدل مارکوف با رویکرد فعالانه و پیش‌نگر در سامانه‌های اطلاعاتی به‌کار گرفته می‌شود و هدف اصلی آن ارائه اقدامات بهینه‌ی جایگزین و پیشنهادات و توصیه‌های عملیاتی در زمان مناسب و بهینه است. بر اساس حالت‌های انتقال و تغییر در مراحل زمانی ثابت، مدل کلاسیک فرایند تصمیم‌گیری مارکوف، یک مدل زمانی گسسته فرض شده است. این مدل با تعریف اقدامات جایگزین بیشتر نیز قابلیت کاربرد دارد. جدول ۲ متغیرهای مدل تصمیم‌گیری را نشان می‌دهد.

جدول ۲. معرفی پارامترها و متغیرهای مدل ریاضی

متغیرهای مدل	تعریف
$P^f(t_1, t_2)$	تابع توزیع احتمال وقوع نقص f در بازه زمانی (t_1, t_2) در شرایطی که تا زمان t_1 نقصی اتفاق نیفتد.
$P^f_{a_i}(t_1, t_2)$	تابع توزیع احتمال وقوع نقص f در بازه زمانی (t_1, t_2) در شرایطی که تا زمان t_1 نقصی اتفاق نیفتد و فرض می‌شود اقدام i دقیقاً در زمان t_1 اجرا شود.
$P(S_d, S_f)$	تابع توزیع احتمال وضعیت سیستم در حالت خطر و وضعیت سیستم در حالتی که یک رویداد تصادفی اتفاق افتاده است

متغیرهای مدل	تعریف
$P(S_d, S_{a_i})$	تابع توزیع احتمال وضعیت سیستم در حالت خطر و وضعیت سیستم با در نظر گرفتن اجرای اقدام i
$P(S_{a_i}, S_{e_i})$	تابع توزیع احتمال وضعیت سیستم با در نظر گرفتن اجرای اقدام i و وضعیت سیستم در حالت تأثیرگذاری اقدام i
$P(S_{e_i}, S_f)$	تابع توزیع احتمال وضعیت سیستم در حالت تأثیرگذاری اقدام i و وضعیت سیستم در حالتی که یک رویداد تصادفی اتفاق بیفتد
$P(S_{a_i}, S_f)$	تابع توزیع احتمال وضعیت سیستم با در نظر گرفتن اجرای اقدام i و وضعیت سیستم با یک رویداد تصادفی
C_{a_i}	هزینه اجرای اقدامi
C_f	هزینه ریسک یا نقص
S_{a_i}	وضعیت سیستم با در نظر گرفتن اجرای اقدام i
S_{e_i}	وضعیت سیستم در حالت تأثیرگذاری اقدام i
S_d	وضعیت سیستم در حالت خطر
S_f	وضعیت سیستم در حالتی که یک رویداد تصادفی اتفاق بیفتد
t_d	زمان وقوع خطر
t_{e_i}	زمان تأثیرگذاری اقدام i
t_{a_i}	زمان اجرای اقدام i
L	زمان بین وقوع نقص و اجرای اقدم
T	بازه زمانی تصمیم
U^{a_i}	تابع مطلوبیت اقدام i

در فرایند تصمیم‌گیری مارکوف با استفاده از الگوریتم بازگشتی استقرایی در افق زمانی محدود، مطلوبیت مورد انتظار برای هر اقدام را با رابطه زیر می‌توان نشان داد:

$$U^{a_i} = -C_{ue}(P^{\varepsilon}(t_0, t_{a_i}) + P^{\varepsilon}(t_{a_i}, t_{e_i}) + (1 - P^{\varepsilon}(t_{a_i}, t_{e_i}))P^{\varepsilon}_{a_i}(t_{a_i}, \acute{T})) - C_{a_i}(t_{a_i})(1 - P^{\varepsilon}(t_0, t_{a_i}))$$

که در آن t_0 زمان تشخیص الگو، t_{a_i} زمان انجام اقدام t_{e_i} زمان تأثیرگذاری اقدام i است.

$(t_{e_i} = t_{a_i} + \Delta t_{a_i})$.

T پایان زمان تصمیم‌گیری است. بدین معنی که بعد از این زمان دلیلی برای تصمیم‌گیری وجود ندارد. C_{ue} هزینه یک رویداد ناخواسته و $C_{a_i}(t_{a_i})$ تابع هزینه اقدام یا عملکرد a_i است. در این مدل تابع توزیع احتمال وقوع یک رویداد ناخواسته، بر اساس ویژگی‌ها و خصوصیت‌های مارکوفی در گروه توابع نمایی قرار دارد. (Bousdekis et al, 2018). بر اساس مدل و متغیرهای معرفی شده روابط زیر جهت محاسبه توابع مربوط به احتمال وقوع نقص و هزینه انجام اقدام اصلاحی تعریف شده است که برای رسیدن به تصمیم بهینه در زمان مناسب و با هزینه مینیمم، استفاده می‌شوند.

$$U^{a_i} = -C_{ue}(P^{\varepsilon}(t_0, t_{a_i}) + P^{\varepsilon}(t_{a_i}, \boldsymbol{t_{e_i}}) + (1 - P^{\varepsilon}(t_{a_i}, \boldsymbol{t_{e_i}}))P^{\varepsilon}_{a_i}(t_{a_i}, \acute{T})) - C_{a_i}(t_{a_i})(1 - P^{\varepsilon}(t_0, t_{a_i}))$$

$$P^{f}(t_1, t_2) = \frac{G^{f}(t_2) - G^{f}(t_1)}{1 - G^{\varepsilon}(t_1)}$$

$$P^{f}_{a_i}(t_1, t_2) = \frac{G^{f}_{a_i}(t_2) - G^{f}_{a_i}(t_1)}{1 - G^{f}(t_1)}$$

$$P(S_d, S_f) = P^{f}(t_0, t_{a_i})$$

$$P(S_d, S_{a_i}) = 1 - P^{f}(t_0, t_{a_i})$$

$$P(S_{a_i}, S_f) = P^{f}(t_{a_i}, t_{e_i})$$

$$P(S_{a_i}, S_{e_i}) = 1 - P^{f}(t_{a_i}, t_{e_i})$$

$$P(S_{e_i}, S_f) = P^{f}_{e_i}(t_{e_i}, T)$$

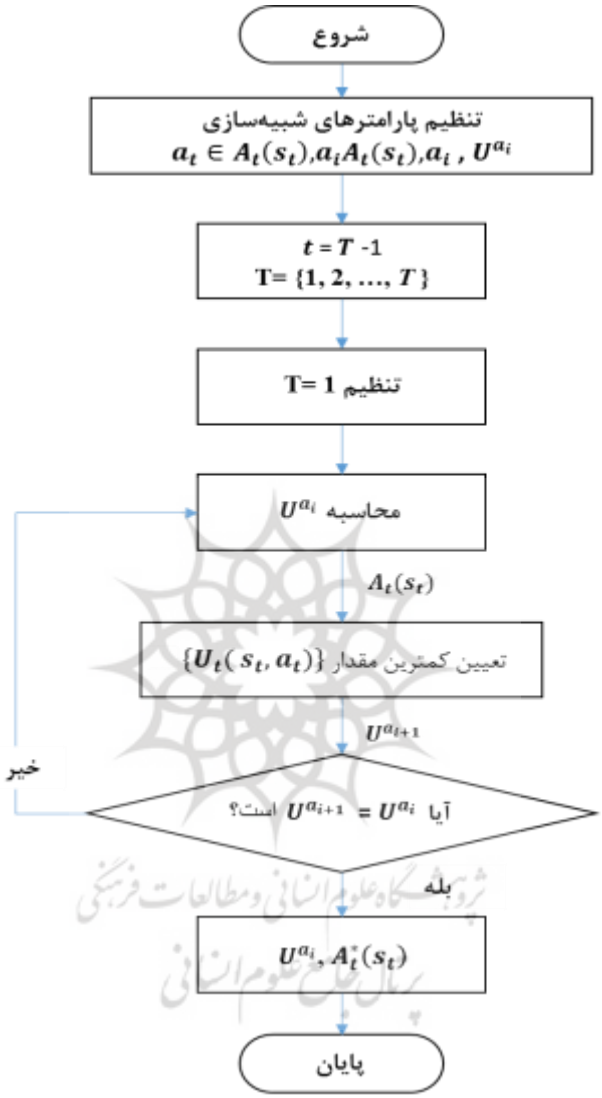
$$P^{0}(S_d, S_f) = P^{f}(t_0, T)$$

مبتنی بر مدل ریاضی و متغیرهای تعریف شده، برای تخصیص اقدام بهینه به هر رویداد در زمان مناسب و با کمترین هزینه بر اساس الگوریتم (جدول ۳) و فلوچارت (شکل ۳) زیر، مدل در نرم‌افزار MATLAB مدلسازی و اجرا شده و تصمیم بهینه برای هر رویداد یا نقص تعیین شده است.

جدول ۳. شبه کد الگوریتم فرایند تصمیم‌گیری مارکوف بازه زمانی محدود براساس استقرای بازگشتی در برنامه‌ریزی پویا

Initialize the MDP model components: T= {1, 2, …, T }, S,$A_t(s_t)$, U^{a_i}, C_f, P^{ε}.
For $s_t \in S$ do
  $Z_t(s_t)$
End
For $t = T$ -1→ 0 do
  For $a_t \in A_t(s_t)$ do
    Calculate $U^{a_i} = -C_f(P^{\varepsilon}(t_0, t_{a_i}) + P^{\varepsilon}(t_{a_i}, t_{e_i}) + (1\ P^{\varepsilon}(t_{a_i}, t_{e_i}))P^{\varepsilon}_{a_i}(t_{a_i}, T) - C_{a_i}(t_{a_i})(1 - P^{\varepsilon}(t_0, t_{a_i}))$
  End
    Calculate $Z^*_t(s_t) \leftarrow \min\{U_t(s_t, a_t)\}$
    Calculate $A^*_t(s_t) \leftarrow \text{argmin}\{U_t(s_t, a_t)\}$
End
Return Z^*, A^*

شکل ۳. فلوچارت فرایند تصمیم‌گیری مارکوف



یافته‌های پژوهش

شناسایی ریسک‌ها و خطرات، بر اساس برنامه‌های ارزیابی ریسک صورت گرفته قبلی توسط متخصصان در بازه زمانی سه‌ساله در صنعت موردمطالعه است. این داده‌ها مربوط به ۲۵۰ کیلومتر از خط لوله گاز اهواز به سنندج است و مجموعه داده اولیه را تشکیل می‌دهد. پایگاه داده ایجاد شده، متشکل از ۲۴۰ نمونه (رکورد) و ۱۵ ویژگی است که ۱۱ منطقه شامل ۴۰ ناحیه را در محیط موردمطالعه شامل می‌شود. در پایگاه داده، هر یک از

نقص‌ها و ریسک‌ها که اعضای مجموعه داده هستند برای رعایت اختصار کدگذاری و برچسب‌زنی شده است. به‌عنوان مثال کد B2F4 مربوط به نقص ۴ از ناحیه B منطقه ۲ شبکه انتقال خط لوله گاز است. برای انجام فرایند داده‌کاوی از نرم‌افزار SPSS Modeler (Clementine)**18** بهره‌گیری شده است. برای ارزیابی صحت و دقت نتایج شاخص‌های حساسیت، شفافیت، دقت و صحت مورد ارزیابی قرار گرفته است. (درمحمدی و همکاران، ۱۳۹۳) مبتنی بر نتایج به‌دست آمده از خوشه‌بندیِ رویدادها، خوشه بحرانی انتخاب و در ادامه پس از تعیین شاخص هدف به‌عنوان خروجی و تعیین ویژگی‌های ورودی، رویدادها (ریسک‌ها یا نقص‌ها) طبقه‌بندی شده‌اند. در آغاز فرایند داده‌کاوی، ابتدا ایجاد جریان داده[1] و انجام عملیات پیش‌پردازش و آماده‌سازی داده‌ها صورت گرفته است. برای این منظور آزمون نرمال بودن داده‌ها، تشخیص داده‌های مفقوده، تشخیص داده‌های پرت، انتخاب ویژگی‌های اصلی، تلفیق داده‌ها و کاهش ابعاد و نهایتاً پاکسازی داده‌ها و ایجاد مجموعه داده تمیز برای خوشه‌بندی و طبقه‌بندی، انجام شده است. برای خوشه‌بندی داده‌ها نیز از الگوریتم‌های K-Means, Kohnen, Two-Step استفاده شده است. در این پژوهش مقادیر خارج از محدوده، بدون اثر[2] در نظر گرفته شده است. با توجه به ماهیت هر ویژگی، مواردی از قبیل شماره سیستم، اجزای سیستم و حالت نقص که صرفاً جنبه برچسب‌گذاری داشته حذف شده‌اند. همچنین پس از محاسبه عدد اولویت ریسک ویژگی‌های نرخ فراوانی، شدت نقص و نرخ کشف نیز فیلتر شده‌اند. در ادامه با استفاده از گره Feature Selection ورودی‌های با ارزشِ بیش از ۹۵ درصد و با اهمیت بالا تعیین شده است. بر این اساس «هزینه نگهداری و تعمیرات»، «برآورد کلی اهمیت نقص»، «احتمال وقوع نقص»، «عدد اولویت ریسک»، «تعداد خرابی» و «میانگین زمانی وقوع نقص»، برای تجزیه‌وتحلیل آماری و ارزش‌گذاری با استفاده از این گره انتخاب شده‌اند و تست مربوط به نرمال بودن داده‌ها در مورد آن‌ها انجام شده است. برای این منظور شاخص‌های «چولگی» و «کشیدگی» محاسبه و مورد ارزیابی قرار گرفته است. مبنای نرمال بودن داده‌ها بر اساس این دو شاخص قرار گرفتن اعضای مجموعه داده در بازه (۲+ و ۲-) است. بر این اساس تمامی داده‌ها، مورد ارزیابی قرار گرفته و دو مورد داده پرت در ویژگی «میانگین زمانی وقوع نقص» تشخیص داده شده است که اقدام بی‌اثرسازی در مورد آن انجام شده و با استفاده از متد Fix که در آن مقادیر ثابت مانند میانگین و میانه یا هر مقدار ثابت دیگر جایگزین می‌شود، نسبت به آن رفتار شده است. در ادامه عملیات

1. Data Stream
2. Nullify

پیش‌پردازش داده‌ها با استفاده از گره Anomaly داده‌های پرت یا مقادیر غیرمعمول مورد واکاوی قرار گرفته است. این گره قسمتی از مدل کشف مغایرت است که اطلاعاتی راجع به ویژگی‌ها، تنظیمات و فرایند تخمین در مدل را نمایش می‌دهد. مدل کشف مغایرت یک روش غیرنظارتی است و نیازی به مجموعه داده برای آموزش و یادگیری ندارد. در این مدل به‌منظور تعیین مغایرت‌های احتمالی هر رکورد با سایر رکوردها مقایسه و در گروه‌های همتا قرار می‌گیرد. روش مورداستفاده برای پذیرش یا عدم پذیرش مغایرت‌ها حدآستانه[1] است. میزان برش و حدآستانه به‌صورت خودکار تعیین و برحسب درصد به‌عنوان یک پارامتر در مدل محسوب می‌شود. همچنین تعداد فیلدهایی که به‌عنوان شاخص مغایرت باید مدنظر قرار گیرد، در تنظیمات اولیه مشخص می‌شود (علیخانزاده، ۱۳۹۲). بر اساس آن کشف مغایرت در سه گروه همتا با تعداد رکوردهای ۵۴ و ۸۵ و ۱۰۱ صورت گرفته است که در گروه سوم ۲ رکورد به‌عنوان مقادیر مغایر تشخیص داده شده و شاخص میانگین با مقادیر مشخص جایگزین شده است. درنهایت نتایج نشان‌دهنده ۲ داده مغایر است که ۰/۸۳ درصد از کل رکوردها را شامل شده و ۹۹/۱۷ درصد داده‌ها معادل ۲۳۸ رکورد، عدم مغایرت داشته است. جدول ۴.

جدول ۴. جدول نتایج اجرای الگوریتم Anomaly

ارزش	تعداد	درصد
مغایرت	۲	۰/۸۳
عدم مغایرت	۲۳۸	۹۹/۱۷

در این پژوهش «شاخص سیلوئیت» برای انجام خوشه‌بندی به‌عنوان مبنا موردنظر قرار گرفته است و از سه الگوریتم K-Means, Kohnen, Two Step استفاده شده است که با تغییر مقادیر مربوط به تعداد خوشه (K)، بهترین مقدار برای این شاخص، مبتنی بر الگوریتم K-Means، ۰/۶۴۴۶ بدست آمده است. بر اساس نتایج بدست آمده پیاده‌سازی الگوریتم‌ها، روش K-Means به‌عنوان روش برتر انتخاب و بر این اساس داده‌ها در ۵ خوشه تقسیم‌بندی شده که تعداد اعضای هر یک از خوشه‌ها بر اساس نوع داده و درصد فراوانی آن‌ها در جدول ۵ مشخص شده است. همچنین جدول ۶ نشان‌دهنده محدوده قرارگیری شاخص سیلوئیت و مقدار آن است.

1. Cutoff

جدول ۵. فراوانی داده‌ها در هر خوشه

خوشه	تعداد اعضا	درصد
cluster-1	۶۵	۲۷/۰۸
cluster-4	۷۳	۳۰/۴۲
cluster-5	۳۲	۱۳/۳۳
cluster-3	۱۲	۵
cluster-2	۵۸	۲۴/۱۷

جدول ۶. مقدار شاخص سیلوئیت

مقدار شاخص سیلوئیت	محدوده
۰/۶۴۴۶	Good

برای تشخیص خوشه بحرانی، ویژگی‌های اصلی و دارای اولویت با استفاده از روش K-Means شناسایی و مورد تجزیه‌وتحلیل آماری قرار گرفته است. بر این اساس با استفاده از گره Analysis تعداد اعضای هر خوشه، انحراف استاندارد، خطای استاندارد و میانگین هر یک از ویژگی‌ها در خوشه‌های پنج‌گانه بررسی شده است. با توجه به تعداد اعضای بیشتر، انحراف استاندارد و درصد خطای کمتر در اغلب شاخص‌ها، میانگین عدد اولویت ریسک و تعداد خرابی نسبتاً بالا، میانگین وقوع نقص پایین، میانگین احتمال وقوع نقص نسبی تقریباً بالا به نسبت اعضای خوشه، خوشه ۴ به‌عنوان خوشه بحرانی، انتخاب شده است.

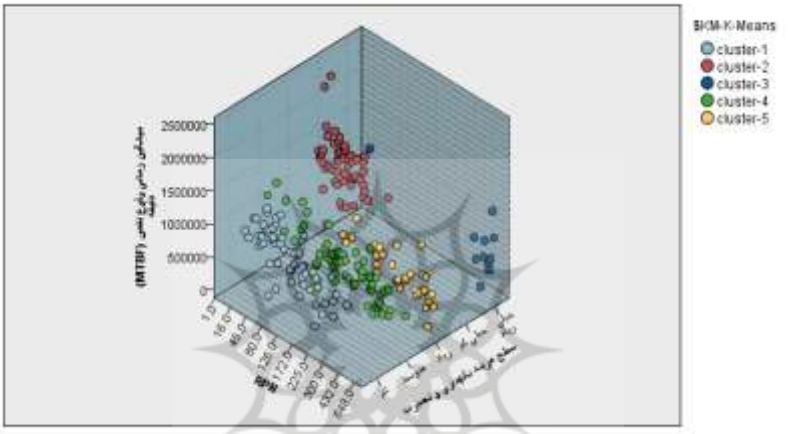
جدول ۷. اطلاعات مربوط به خوشه‌ها بر اساس اهمیت نقص

خوشه	۱			۲	۳			۴			۵	
تعداد	۲	۲۹	۳۴	۵۸	۱۰	۱	۱	۱۷	۴۸	۸	۲۴	۸
اهمیت	H	M	L	L	H	M	L	H	M	L	H	M

شکل ۴ نحوه خوشه‌بندی ریسک‌ها و نقص‌ها را با الگوریتم K-Means نشان می‌دهد. بر اساس سه ویژگی «عدد اولویت ریسک»، «سطح هزینه نگهداری و تعمیرات» و «میانگین زمانی وقوع نقص»، با توجه به شکل، خوشه ۴ دارای فراوانی بیشتر، سطح هزینه نت بین سطح زیاد و متوسط، RPN در مقادیر بالا دارای تراکم بیشتر و میانگین وقوع نقص نیز در بیشتر ریسک‌ها مقادیر کم را شامل می‌شود. درحالی‌که سطح هزینه

نت در خوشه‌های دیگر نسبت به خوشه ۴ کمتر است؛ بجز خوشه ۳ که تعداد اعضای آن به نسبت سایر خوشه‌ها دارای فراوانی کمتر است و عناوین ریسک‌ها از جامعیت کافی برای بررسی برخوردار نیستند. در خوشه ۱ بیشتر اعضا دارای عدد اولویت ریسک و سطح هزینه پایین هستند. در خوشه ۲ تقریباً همه اعضا هیچ‌یک از شرایط بحرانی بودن را ندارند.

شکل ۴. نتایج خوشه‌بندی با روش K-Means



در فرایند داده‌کاوی پژوهش، از الگوریتم‌های شبکه عصبی، درخت C.5، نزدیکترین همسایگی و بردار پشتیبان برای طبقه‌بندی بهره‌گیری شده است. برای این منظور سه ویژگی «عدد اولویت ریسک»،«احتمال وقوع نقص» و «هزینه نگهداری و تعمیرات» به‌عنوان ویژگی‌های مهم و با اهمیت بالای ۹۵ درصد برگزیده شده است. با استفاده از گره Data Analysis آزمون نرمال بودن داده‌ها مانند مرحله خوشه‌بندی صورت گرفته است. در مدل طبقه داده‌های پژوهش، ۷۰ درصد داده‌ها برای آموزش، ۲۰ درصد داده‌ها برای تست و ۱۰ درصد داده‌ها برای اعتبارسنجی در نظر گرفته شده است. در گام ابتدایی اجرای الگوریتم‌های طبقه‌بندی، از الگوریتم ماشین بردار پشتیبان خطی و غیرخطی استفاده شده است که در مرحله اول صحت طبقه‌بندی ۷۶/۸ درصد و پس از استفاده از مدل غیرخطی این الگوریتم این مقدار ۸۵/۷۱ درصد بوده است. برای الگوریتم SVM از تابع خطی نوع دوم و برای L-SVM، تابع سیگموئید با لاندای ۰/۱ و دقت رگرسیون ۰/۱ استفاده شده است. در ادامه اجرای مدل طبقه‌بندی، از شبکه عصبی پرسپترون چندلایه

استفاده شده و تنظیمات اولیه شامل تعداد لایه‌ها و نورون‌ها، قوانین توقف الگوریتم، درصد داده‌هایی که در هر تکرار برای آموزش اختصاص داده می‌شوند و تصمیم در مورد داده‌های پرت، مشخص و مدل اجرا شده است و صحت طبقه‌بندی برابر ۸۹/۳ درصد است و بنابراین استفاده از این الگوریتم در ادامه ماشین بردار پشتیبان، سبب ارتقاء میزان صحت طبقه-بندی داده‌ها شده است. میزان صحت پیش‌بینی مدل نیز بر اساس ماتریس درهم‌ریختگی در الگوریتم اجرا شده ۹۶/۲ درصد است. شبکه عصبی استفاده شده از نوع پرسپترون دولایه است که در لایه اول دو نورون و در لایه دوم یک نورون را داراست. در ساختار شبکه عصبی مدل، ۵ متغیر و Bias به‌عنوان ورودی و سطح کلی اهمیت نقص به‌عنوان خروجی در نظر گرفته شده است. نتایج بدست آمده، به‌عنوان ورودی برای الگوریتم K نزدیک‌ترین همسایگی استفاده شده است. خروجی‌های حاصل از اجرای الگوریتم و محاسبه فواصل بر اساس فاصله اقلیدسی، بهبود نتایج در دقت و صحت طبقه‌بندی را نشان می‌دهد؛ بدین ترتیب که پس از اجرای مدل، درصد داده‌های صحیح طبقه‌بندی‌شده ۹۱/۰۷ درصد است. در این الگوریتم حداقل k=3 و حداکثر K=5 در نظر گرفته شده است. برای افزایش اعتبار مدل طبقه‌بندی داده‌های پژوهش، الگوریتم درخت تصمیم C.5 برای آموزش داده‌ها به مدل اضافه شده و نتایج نشان می‌دهد که صحت مدل طبقه‌بندی بر اساس درصد داده‌هایی که آموزش داده شده است، ارتقاء یافته است. این میزان پس از استفاده از الگوریتم درخت تصمیم C.5، به ۹۲/۸۶ درصد افزایش یافته که میزان داده‌های صحیح طبقه‌بندی شده و صحت مدل را نشان می‌دهد. جدول ۸ نتایج آنالیز مدل‌ها را به‌صورت یکپارچه نشان می‌دهد.

نتایج بدست آمده از الگوریتم‌های پیاده‌سازی شده با استفاده از «گره Ensemble» نرم‌افزار ادغام شده و خروجی مدل نشان‌دهنده افزایش صحت مدل است و این میزان برابر ۹۷/۵ درصد است. در این مدل بر اساس آنالیزهای انجام‌شده، متد بالاترین درجه اطمینان برای ویژگی هدف (سطح کلی اهمیت نقص) بالاترین میزان کارایی را داشته است. مبتنی بر نتایج طبقه‌بندی داده‌های پژوهش، ۱۳ نقص دارای اهمیت بالا بوده و ۱۷/۸۱ درصد داده‌ها را شامل می‌شود. ۶۰ نقص معادل ۸۲/۱۹ درصد داده‌ها دارای اهمیت متوسط هستند. در جدول ۹، نتایج آنالیز مدل‌ها را پس از ادغام با استفاده از متد بالاترین درجه اطمینان برای داده‌های موردتوافق الگوریتم نشان داده شده است.

جدول ۸. نتایج آنالیز یکپارچه مدل ترکیبی طبقه‌بندی

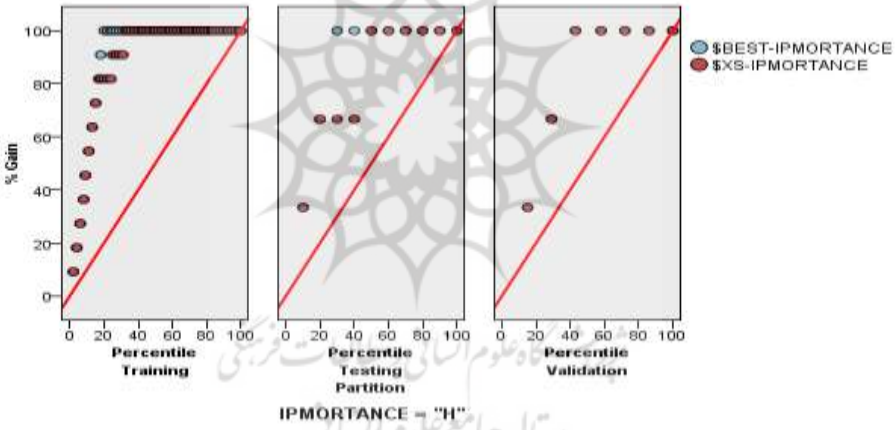
نتایج طبقه‌بندی با الگوریتم SVM						
اعتبارسنجی		آموزش		تست		Partition
۱۴٪/۲۹	۱	۶۰٪	۶	۷۶٪/۷۹	۴۳	صحیح
۸۵٪/۷۱	۶	۴۰٪	۴	۲۳٪/۲۱	۱۳	اشتباه
	۷		۱۰		۵۶	مجموع
نتایج طبقه‌بندی با الگوریتم ترکیبی SVM و LSVM						
اعتبارسنجی		آموزش		تست		Partition
۷۱٪/۴۳	۵	۷۰٪	۷	۸۵٪/۷۱	۴۸	صحیح
۲۸٪/۵۷	۲	۳۰٪	۳	۱۴٪/۲۹	۸	اشتباه
	۷		۱۰		۵۶	مجموع
نتایج طبقه‌بندی با الگوریتم ترکیبی SVM، LSVM، ANN						
اعتبارسنجی		آموزش		تست		Partition
۷۱٪/۴۳	۵	۷۰٪	۷	۸۹٪/۲۹	۵۰	صحیح
۲۸٪/۵۷	۲	۳۰٪	۳	۱۰٪/۷۱	۶	اشتباه
	۷		۱۰		۵۶	مجموع
نتایج طبقه‌بندی با الگوریتم ترکیبی SVM، LSVM، ANN، KNN						
اعتبارسنجی		آموزش		تست		Partition
۸۵٪/۷۱	۶	۷۰٪	۷	۹۱٪/۰۷	۵۱	صحیح
۱۴٪/۲۹	۱	۳۰٪	۳	۸٪/۹۳	۵	اشتباه
	۷		۱۰		۵۶	مجموع
نتایج طبقه‌بندی با الگوریتم ترکیبی SVM، LSVM، ANN، KNN، C.5						
اعتبارسنجی		آموزش		تست		Partition
۵۷٪/۱۴	۴	۵۰٪	۵	۹۲/۸۶٪	۵۲	صحیح
۴۲٪/۸۶	۳	۵۰٪	۵	۷/۱۴٪	۴	اشتباه
	۷		۱۰		۵۶	مجموع

نمودار ۵ نشان‌دهنده روند یادگیری در داده‌های موردتوافق نهایی (۴۹ رکورد) است؛ به شکلی که یادگیری در ۹۷/۵۶ درصد از داده‌های آموزش (۴۱ رکورد) محقق شده است و این میزان در رابطه با داده‌های «اعتبارسنجی» و «تست» نیز قابل‌مشاهده است.

جدول ۹. نتایج آنالیز صحت و کارایی مدل نهایی طبقه‌بندی

داده‌های موردتوافق						
Partition	تست		آموزش		اعتبارسنجی	
توافق	۱۴	۷۳٪/۲۱	۶	۶۰٪	۵	۲۸٪/۵۷
عدم توافق	۱۵	۲۶٪/۷۹	۴	۴۰٪	۲	۷۱٪/۴۳
مجموع	۵۶		۱۰		۷	
نتایج نهایی طبقه‌بندی با الگوریتم ترکیبی و Ensemble						
Partition	تست		آموزش		اعتبارسنجی	
صحیح	۴۰	۹۷٪/۵۶	۴	۶۶٪/۶۷	۱	۵۰٪
اشتباه	۱	۲٪/۴۴	۲	۳۳٪/۳۳	۱	۵۰٪
مجموع	۴۱		۶		۲	

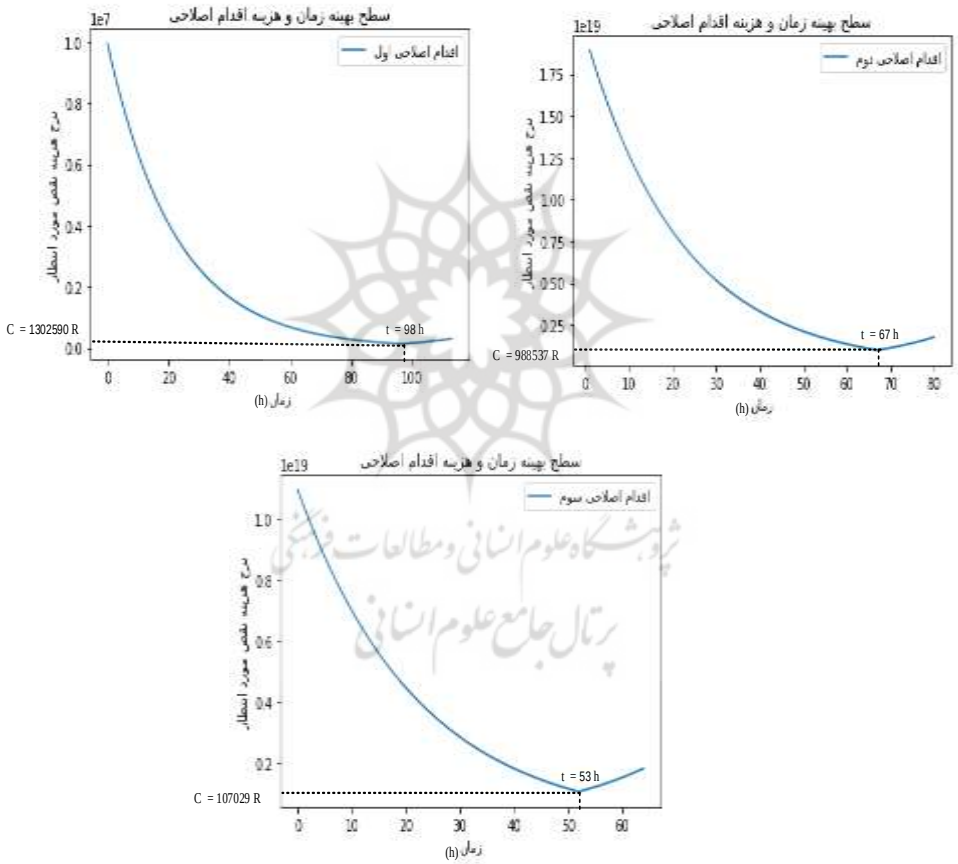
شکل ۵. روند یادگیری در داده‌های آموزش، تست و اعتبارسنجی



بر اساس نتایج شبیه‌سازی، ۱۳ ریسک فاکتور جزء نقص‌های بحرانی محسوب شده و به‌عنوان ریسک‌های دارای اولویت پیش‌بینی شده‌اند. با توجه به کلی بودن و قابلیت تعمیم اقدامات کنترلی به همه نقص‌ها، برای اختصاص هر اقدام اصلاحی به هر ریسک یا نقص، از فرایند تصمیم‌گیری مارکوف که یک مدل ریاضی تصمیم‌گیری تصادفی است، استفاده شده است. این مدل قابلیت تعیین سیاست بهینه برای هر نقص را با توجه به حداکثر میزان مطلوبیت در زمان و هزینه بهینه داراست. برای رعایت اختصار از ارائه جداول داده‌های مربوط به ماتریس اولیه ورودی و ماتریس احتمال انتقال اولیه خودداری شده است. در تمامی محاسبات توزیع مبنا، نمایی و نرخ وقوع نقص ۰/۰۴۵ در نظر گرفته

شده است. (λ=۰/۰۴۵). در شکل ۶ نتایج شبیه‌سازی سه اقدام اصلاحی به‌عنوان نمونه نشان داده شده است. اقدام اصلاحی بهینه برای هر نقص بر اساس زمان انجام اقدام اصلاحی یا اجرای اقدام که کمترین میزان هزینه نقص در آن به وقوع می‌پیوندد، انتخاب شده است. با توجه به اینکه هر چه ریسک هزینه کمتری داشته باشد اقدام اصلاحی نیز کم‌هزینه‌تر خواهد بود، در نمودارهای ۱ تا ۳ زمان اجرای یک اقدام زمانی است که ریسک یا نقص موردنظر کمترین میزان هزینه را به همراه داشته باشد.

شکل ۶. نمودارهای ۱۳ گانه اقدام اصلاحی یا سیاست بهینه مرتبط با هر نقص



جدول ۱۰ زمان اجرا و عنوان اقدام موردنظر برای هر نقص را به همراه هزینه کمینه نقص نشان می‌دهد.

جدول ۱۰. نتایج شبیه‌سازی فرایند تصمیم‌گیری مارکوف

عنوان نقص	اقدام اصلاحی/ سیاست بهینه	زمان	هزینه
ناکارآمدی سامانه حفاظت کاتدیک خطوط	استفاده از سامانه‌ها و نصب حسگرهای هشداردهنده در همه قسمت‌های شبکه	۹۸	۱۳۰۲۵۹۰۰
ناکارآمدی تجهیزات کنترل فشار و نرخ جریان در شبکه انتقال	ایجاد مرکز پایش خطرات و کنترل بحران	۶۷	۹۸۸۵۳۷۰۰
انتشار گازهای آلاینده و مواد سمی شیمیایی(مرکاپتان، روغن، گریس و	آموزش کارگران و پیمانکاران برای شناسایی و پیشگیری از ریسک‌ها و	۵۳	۱۰۷۰۲۹۰۰
عدم آموزش مناسب و توجیه نبودن پیمانکاران نسبت به موقعیت محیطی	بازنگری ضوابط و الزامات برون‌سپاری امور نت	۴۶	۱۲۲۳۲۹۰۰
نشت میعانات گازی	ارزیابی ریسک دوره‌ای و به‌روزرسانی، بازنگری و بهبود دستورالعمل‌ها، روش‌های اجرایی و مستندات	۵۷	۶۳۰۰۳۲۰۰
آسیب مکانیکی (برخورد ماشین‌آلات مانند اسکریپر، پیکور، بلدوزر و ...) با	نصب علائم، نشانگرها و تابلوهای ایمنی	۶۳	۴۵۳۸۴۵۰۰
عدم وجود فشارسنج حساس در طول	بازطراحی و ایمن‌سازی لوله‌ها، اتصالات	۵۱	۷۰۲۰۹۵۰۰
نقص در حفاری‌های ضروری	اخذ مجوزهای کاری یا گواهی‌نامه‌های	۸۲	۲۶۶۰۲۷۰۰
استفاده از مواد اولیه بی‌کیفیت در تجهیزات و قطعات	بازبینی و به‌روزرسانی تولیدکنندگان و عرضه‌کنندگان قطعات و لوازم یدکی	۵۴	۱۱۴۲۴۷۰۰
انبارش و نگهداری بیش از حد موردنیاز تجهیزات، قطعات، مواد اولیه قابل اشتعال و کپسول گازهای	بازطراحی سیستم برنامه‌ریزی و کنترل موجودی قطعات و لوازم یدکی	۷۲	۸۰۸۷۴۷۰۰
نقص شیرهای اتوماتیک قطع جریان	بازطراحی سیستم هزینه‌یابی و قیمت	۶۵	۶۹۸۲۴۲۰۰
نقص در عایق‌کاری شبکه انتقال	تغییر در روش‌های تجزیه‌وتحلیل و دوره‌های زمانی آنالیز اطلاعات نت	۷۲	۵۴۲۹۷۱۰۰
خوردگی لوله	بازرسی‌های دوره‌ای و انجام تست‌های	۷۷	۴۹۳۹۴۳۰۰

با توجه به نتایج شبیه‌سازی، کمترین میزان هزینه نقص مربوط به «انتشار گازهای آلاینده و مواد سمی شیمیایی (مرکاپتان، روغن، گریس، دوده و مواد سمی)» و اقدام اصلاحی مرتبط آن عبارت است از «آموزش کارگران و پیمانکاران در راستای شناسایی و پیشگیری از ریسک‌ها و خطرات و استفاده از تجهیزات حفاظت فردی» که در زمان $t = 53\ h$ انجام می‌شود. در جدول ۱۱ مقادیر اولیه و مقادیر بهینه به‌دست‌آمده از اجرای شبیه‌سازی در مورد «هزینه اجرای اقدام» و «زمان اجرای اقدام» ارائه شده است.

جدول ۱۱. مقادیر اولیه و بهینه هزینه و زمان اقدام

$ta_1(h)$	$ta_2(h)$	Cf_1	Cf_2
۱۸۸	۹۸	۵۰۰۸۰۰۴۹۰	۱۳۰۲۵۹۰۰
۴۲۱	۶۷	۹۵۰۸۱۷۶۵۰	۹۸۸۵۳۷۰۰
۱۹۶	۵۳	۵۵۰۶۱۳۱۲۰	۱۰۷۰۲۹۰۰
۳۱۵	۴۶	۴۶۰۲۵۰۸۴۰	۱۲۲۳۲۹۰۰
۳۷۰	۵۷	۳۸۷۵۷۳۰۴۰	۶۳۰۰۳۲۰۰
۱۳۷	۶۳	۳۶۵۰۶۲۰۷۰	۴۵۳۸۴۵۰۰
۳۵۱	۵۱	۳۳۰۳۰۷۶۸۰	۷۰۲۰۹۵۰۰
۳۴۸	۸۲	۵۰۰۲۶۶۵۵۰	۲۶۶۰۲۷۰۰
۲۴۴	۵۴	۶۱۴۶۱۴۰۷۰	۱۱۴۲۴۷۰۰
۲۱۰	۷۲	۹۷۲۶۸۸۶۲۰	۸۰۸۷۴۷۰۰
۲۶۵	۶۵	۶۱۴۱۶۸۵۱۰	۶۹۸۲۴۲۰۰
۲۳۲	۷۲	۶۵۳۰۳۷۱۹۰	۵۴۲۹۷۱۰۰
۳۸۳	۷۷	۷۴۲۸۳۹۶۴۰	۴۹۳۹۴۳۰۰

با توجه به نتایج فوق، در جدول ۱۲ درصد بهینگی حاصل‌شده از اجرای مدل با توجه به میانگین مقادیر ارائه شده است.

جدول ۱۲. میزان بهبود بر اساس میانگین زمان و هزینه اقدام

میانگین Cf_2	میانگین Cf_1	میزان بهبود	$ta_2(h)$ میانگین	$ta_1(h)$ میانگین	میزان بهبود
۴۶۶۰۲۳۳۰	۵۸۷۹۲۶۱۱۶	٪۹۲	۶۵	۲۸۱	٪۷۷

اجرای اقدام اثربخش، به صورتی که از یک طرف در بازه زمانی کوتاه‌تر پس از وقوع نقص باشد و از طرف دیگر با کمترین میزان هزینه همراه باشد، از موارد مهم و اجتناب‌ناپذیر است؛ ازاین‌رو نرخ بهبود در زمان و هزینه حائز اهمیت است. برای این منظور بکارگیری ابزار مؤثر برای بهینه‌سازی این مؤلفه‌ها نقش تعیین‌کننده و بسزایی را ایفا می‌نماید. با توجه به جدول ۹، نرخ بالای بهینگی حاصل‌شده در میانگین زمان و هزینه اجرای اقدام، مؤید کارایی و مطلوبیت مدل تصمیم‌گیری استفاده شده در این پژوهش

است. برای سنجش اعتبار و تحلیل حساسیت، مدل پژوهش با نرخ وقوع ۰/۰۲۵ نیز مورداستفاده قرار گرفته است که نتایج تفصیلی در جدول ۱۳ ارائه شده است.

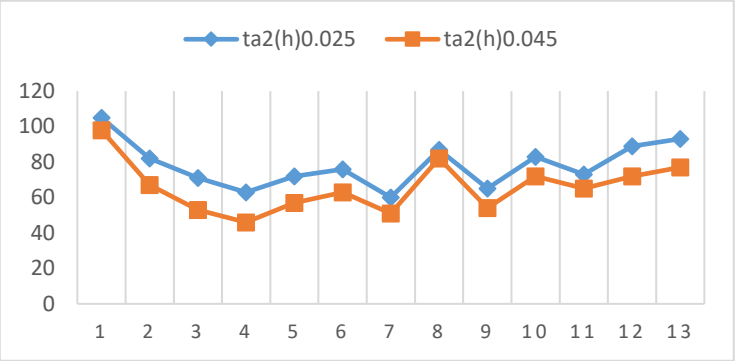
جدول ۱۳. مقادیر اولیه و بهینه هزینه و زمان اقدام

$ta_1(h)$	$ta_2(h)$	Cf_1	Cf_2
۱۸۸	۱۰۵	۵۰۰۸۰۰۴۹۰	۵۵۴۰۸۴۴۰
۴۲۱	۸۲	۹۵۰۸۱۷۶۵۰	۱۸۶۲۵۶۰۰
۱۹۶	۷۱	۵۵۰۶۱۳۱۲۰	۱۴۱۷۴۳۰۰
۳۱۵	۶۳	۴۶۰۲۵۰۸۴۰	۱۴۴۵۲۱۰۰
۳۷۰	۷۲	۳۸۷۵۷۳۰۴۰	۹۷۳۲۴۸۰۰
۱۳۷	۷۶	۳۶۵۰۶۲۰۷۰	۸۳۰۰۳۲۰۰
۳۵۱	۶۰	۳۳۰۳۰۷۶۸۰	۱۱۱۷۴۰۰۰
۳۴۸	۸۷	۵۰۰۲۶۶۵۵۰	۸۶۵۵۳۴۰۰
۲۴۴	۶۵	۶۱۴۶۱۴۰۷۰	۱۸۳۶۴۱۰۰
۲۱۰	۸۳	۹۷۲۶۸۸۶۲۰	۱۸۵۸۶۶۰۰
۲۶۵	۷۳	۶۱۴۱۶۸۵۱۰	۱۵۰۴۴۳۰۰
۲۳۲	۸۹	۶۵۳۰۳۷۱۹۰	۱۰۷۵۱۰۰۰
۳۸۳	۹۳	۷۴۲۸۳۹۶۴۰	۱۱۰۷۲۸۰۰

بر اساس نتایج بدست آمده از اجرای شبیه‌سازی با نرخ وقوع نقص ۰/۰۲۵ و با توجه به مقادیر جدول فوق، میزان بهبود، در میانگین هر دو شاخص زمان و هزینه بالای ۹۰ درصد برآورد شده است. با افزایش نرخ وقوع نقص اهمیت تصمیم‌گیری در زمان کوتاه‌تر بیشتر می‌شود؛ به عبارت دیگر هر چه نرخ وقوع نقص بیشتر باشد به‌طور طبیعی تصمیمات سریع‌تر و در بازه زمانی کوتاه‌تر اتخاذ می‌شود. این مطلب در نمودار زیر بر اساس دو سناریوی شبیه‌سازی‌شده نشان داده شده است.

همان‌طور که در شکل ۷ مشاهده می‌شود، مقادیر بهینه مربوط به میانگین نرخ وقوع ۰/۰۴۵ در تمامی سیاست‌های نت، دارای مقادیر کمتر هستند که با توجه به نرخ وقوع نقص بیشتر، تصمیم‌گیری و انتخاب سیاست بهینه بایستی در بازه زمانی کوتاه‌تر صورت پذیرد و بر این اساس نتایج به‌دست‌آمده، بیانگر صحت مدل است.

شکل ۷. نمودار تحلیل حساسیت بر اساس زمان



بحث و نتیجه‌گیری

در این پژوهش با بهره‌گیری از این رویکرد ارزیابی ریسک فعالانه، متناسب با نقص‌ها و رویدادهای شناسایی‌شده، اقدام کنترلی در سطح بهینه‌ای از زمان و هزینه نقص، تعیین شده است. برای این منظور با توجه به ماهیت مسئله و اهداف موردنظر و همچنین ویژگی تصمیم‌گیری در شرایط تصادفی بر اساس جستجوی غیرقطعی، از فرایند تصمیم مارکوف در زمان محدود مبتنی بر استقراء بازگشتی برنامه‌ریزی پویای احتمالی بهره‌گیری شده است. همچنین برای تعیین هزینه نقص و هزینه اجرای هر یک از اقدامات اصلاحی و مطلوبیت مورد انتظار برای هریک از نقص‌ها و اقدام اصلاحی مربوطه، از شبیه‌سازی رویداد گسسته استفاده شده است و کمترین هزینه و زمان بهینه، سیستم در حالت‌های مختلف محاسبه و تحلیل شده است. پس از اجرای شبیه‌سازی، علاوه بر تخصیص اقدام بهینه به هر رویداد یا ریسک، میزان بهبود در هزینه نقص‌ها و زمان اجرای یک اقدام اصلاحی برآورد و با شرایط قبل از اجرای مدل مقایسه شده است. نتایج پژوهش حاکی از آن است که نرخ بهینگی در میانگین هزینه اجرای اقدام کنترلی ۹۲ درصد و این میزان در میانگین زمان اجرا، ۷۷ درصد است. برای تحلیل حساسیت و اعتبارسنجی، مدل موردنظر با میانگین نرخ وقوع نقص ۰/۰۲۵ نیز موردسنجش قرار گرفته است. آنچه در این پژوهش برخلاف سایر پژوهش‌ها موردتوجه قرار گرفته است در فاز داده‌کاوی، آموزش داده‌ها با یادگیری ماشین و در فاز مدلسازی، ارائه مدل تصمیم‌گیری برای انتخاب راهکار و اقدام اصلاحی بهینه (علاوه بر تشخیص نقص) است؛ به‌روز نبودن داده‌های مربوط به سوابق ارزیابی ریسک‌های صورت‌گرفته و همچنین عدم امکان دسترسی به داده‌های جدید، از محدودیت‌های اصلی پژوهش است. برای مطالعات آتی پیشنهاد می‌شود نقص‌ها و

رویدادهای شناسایی‌شده در این پژوهش با استفاده از روش‌های بهینه‌سازی ترکیبی مبتنی بر رویداد و داده‌محور و بهره‌گیری از روش‌های قطعی و غیرقطعی بهینه‌سازی مانند الگوریتم‌های متاهیوریستیک، برنامه‌ریزی عدد صحیح، برنامه‌ریزی آرمانی و تئوری بازی‌ها، مورد تجزیه تحلیل قرار گرفته و مدلسازی ریاضی و بهینه‌سازی‌ها در این زمینه با تمرکز بر شاخص‌های نت و مدیریت ریسک انجام شود. استخراج قوانین انجمنی با استفاده از داده‌کاوی در زمینه مدیریت ریسک می‌تواند از پژوهش‌های آتی در این زمینه باشد.

منابع

آذر، عادل. جهانیان، سعید. (۱۳۹۲). *روش‌شناسی علم مدیریت*. مرکز نشر دانشگاهی، چاپ اول، تهران، ایران.

اقبالی زارچ، مریم. توکلی مقدم، رضا. سپهری، محمد مهدی. اصفهانیان، فاطمه. آذرون، امیر. (۱۳۹۶). بهینه‌سازی و تحلیل هزینه‌ای درمان دارویی دیابت نوع ۲ برای پیشگیری از هایپوگلیسمی با رویکرد فرآیند تصمیم مارکوف، *مجله انفورماتیک سلامت و زیست پزشکی*، دوره چهارم، شماره دوم،۸۴-۹۷.

عامری، حکیمه. علیزاده، سمیه. برزگری، اکبر. (۱۳۹۲). استخراج دانش از داده‌های بیماران دیابتی با استفاده از روش درخت تصمیم C5. *مدیریت سلامت*، سال شانزدهم، شماره ۵۳، ۵۸-۷۲.

علیخانزاده، امیر. (۱۳۹۲). *داده‌کاوی*. بابل: نشر علوم رایانه.

علیزاده، سمیه. ملک‌محمدی، سمیرا. (۱۳۹۳). *داده‌کاوی و کشف دانش گام‌به‌گام با نرم-افزار Clementine*. تهران: انتشارات دانشگاه صنعتی خواجه نصیرالدین طوسی.

غضنفری، مهدی. علیزاده، سمیه. تیمورپور، بابک. (۱۳۹۵). *داده‌کاوی و کشف دانش*. تهران: انتشارات دانشگاه علم و صنعت ایران.

قدوسی، محمد. میرسعیدی، فاطمه. حسنی، علی‌اکبر. (۱۳۹۹). ارائه مدل تحلیل ریسک در پروژه‌های شهرسازی مبتنی بر داده‌کاوی با مطالعه موردی. *چشم‌انداز مدیریت صنعتی*، سال دهم، شماره ۳۸، ۱۳۷-۱۵۹.

مروتی شریف‌آبادی، علی. زنجیرچی، سید محمود. عباس‌آبادی، ام‌البنین. (۱۴۰۱). مدیریت فرایندهای نگهداری و تعمیرات با رویکرد ترکیبی چارچوب طبقه‌بندی فرایندها و داده‌کاوی. *چشم‌انداز مدیریت صنعتی*، سال دوازدهم، شماره ۴۶، ۱۷۵-۱۹۸.

نقیب هاشمی، سیدسهند. اصغری توچائی، سیدامیر. بینش مروستی، محمدرضا. (۱۴۰۰). تصمیم‌گیری منفعلانه هوشمند برای حسگرهای بیدارشونده در پایش سازه‌ای. *نشریه مهندسی برق و مهندسی کامپیوتر ایران*، سال ۱۹، شماره ۳، ۱۷۰-۱۸۲.

Bousdekis, A., Mentzas, G. (2019). A Proactive Model for Joint Maintenance and Logistics Optimization in the Frame of Industrial Internet of Things. *Operational Research in the Digital Era – ICT Challenges*, Springer Proceedings in Business and Economics.21-45

Bousdekis, A., Papageorgiou, N., Magoutas, B., Apostolou, D. (2018). Enabling condition-based maintenance decisions with proactive eventdriven computing. *Computers in Industry*. Vol 100. 173–183.

Chang, W., Meng, T., & Lim, C. (2015) .Clustering and visualization of failure modes using an evolving tree. *Expert Systems with Applications, 42,* 7235–7244.

Karabağ, O., Eruguz, A., Basten, R. (2020). "Integrated optimization of maintenance interventions and spare part selection for a partially observable multi-component system". *Reliability Engineering and System Safety*. April vol200. 106955.

Steenwinckel, B., De Paepe, D., Heyvaert, P., & Moens, P. (2021). FLAGS: A methodology for adaptive anomaly detection and root cause analysis on sensor data streams by fusing expert knowledge with machine learning. *Future Generation Computer Systems, 116,* 30–48.

Yang, C., Zou, Y., Lai, P., & Jiang, N. (2015). Data mining-based methods for fault isolation with validated FMEA model ranking. *Springer*, 112, 434-447.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی