

یک مدل آماری و هوشمند برای تشخیص نقاط خرابی در نگهداری و تعمیرات پیشگویانه

مبینا خوش‌سیرت (دانشجوی کارشناسی ارشد)

مهدی خاشعی* (استادیار)

ناصر ملاوردی (استادیار)

دانشکده‌ی مهندسی صنایع و سیستم‌ها، دانشگاه صنعتی اصفهان

مهندسی صنایع و مدیریت شریف، زمستان ۱۴۰۱
دوری ۱-۳۸، شماره ۲، ص. ۵۷-۴۷، (پژوهشی)

امروزه نگهداری و تعمیرات اهمیت بسیاری در صنایع تولیدی پیدا کرده است. یک راه‌حل کارا به منظور جلوگیری از توقف فرایندهای تولید، پیش‌بینی خرابی تجهیزات است. در این مقاله، به منظور تشخیص زمان انجام اقدامات نگهداری و تعمیرات پارچوبی پیشنهاد شده تا حتی‌الامکان از توقف در فرایند جلوگیری شود. لذا هدف اصلی این پژوهش، ارائه‌ی یک مدل آماری و هوشمند برای تشخیص نقاط خرابی در زمینه‌ی نگهداری و تعمیرات پیشگویانه است. در نهایت عملکرد مدل پیشنهادی با سایر شیوه‌های داده‌محور به صورت تکی و ترکیبی، شامل مدل‌های لاجبیت، ماشین بردار پشتیبان و شبکه‌های عصبی پرسپترون چندلایه است، مقایسه شده است. نتایج حاکی از آن است که مدل ماشین بردار پشتیبان در مقایسه با مدل‌های تکی و ترکیبی سری، از دقت بالاتری برخوردار است و نیز نتایج نشانگر کارآمدی و کارایی ساختار ترکیب موازی پیشنهادی در مسائل مدل‌سازی و پیش‌بینی است.

واژگان کلیدی: نگهداری و تعمیرات پیشگویانه، پیش‌بینی، یادگیری ماشین، ساختار ترکیب.

۱. مقدمه

نگهداری و تعمیرات نقش به‌سزایی در ارتقای کیفیت تولید ایفا می‌کند. یکی از خواسته‌های مدیریت هر سازمانی افزایش دوره‌ی کاربری مفید هر دستگاه است. به طور سنتی تصمیم نگهداری و تعمیرات، از خرابی سیستم یا برنامه‌های تکراری ناشی می‌شود.^[۱] پیش‌بینی سری‌های زمانی که در آن مشاهدات جمع‌آوری شده‌ی گذشته‌ی یک متغیر، به منظور به دست آوردن روابط اساسی بین مشاهدات و تعیین یک مدل توصیفی تجزیه و تحلیل می‌شود و سپس مدل به دست آمده به منظور برون‌یابی سری‌های زمانی مورد استفاده قرار می‌گیرد، یکی از مهمترین زمینه‌های پیش‌بینی است.^[۲] مدل‌های خودرگرسیون میانگین متحرک انباشته یکی از مهمترین و پرکاربردترین روش‌های پیش‌بینی سری‌های زمانی آماری هستند. با وجود انعطاف‌پذیری این‌گونه مدل‌ها، به دلیل پیش فرض خطی بودن‌شان، الگوهای غیرخطی نمی‌توانند توسط این‌گونه از مدل‌ها، مدل‌سازی شوند و به همین دلیل تخمین مدل‌های خطی برای مسائل دنیای واقعی که اغلب مدل‌ها غیرخطی هستند، همیشه رضایت‌بخش نخواهد بود. از این رو، مدل‌های هوشمند با توجه به قابلیت مدل‌سازی غیرخطی انعطاف‌پذیرشان در پیش‌بینی به کار گرفته شده‌اند.^[۳] روش‌های نگهداری و تعمیرات را می‌توان در سه دسته کلی تقسیم‌بندی کرد: نگهداری و تعمیرات

اصلاحی، نگهداری و تعمیرات پیشگیرانه و نگهداری و تعمیرات پیشگویانه^[۴]. به طور کلی، یک راهکار مناسب نگهداری و تعمیرات باید شرایط تجهیزات را بهبود دهد، نرخ خرابی تجهیزات را کاهش دهد و هزینه‌های نگهداری و تعمیرات را کمینه کند، که در نهایت منجر به بیشینه‌سازی عمر مفید تجهیزات شود.^[۴] با مرور ادبیات موضوع مرتبط با مدل‌های پیش‌بینی و مدل‌سازی، می‌توان این مدل‌ها را در دو دسته‌ی اصلی مدل‌های آماری و هوشمند جای داد. به طور کلی سه نوع الگوریتم یادگیری ماشین وجود دارد: یادگیری نظارت شده، یادگیری بدون نظارت و یادگیری تقویتی.^[۵] در سال‌های اخیر بابتیستا و همکاران^[۶] با یکپارچه‌سازی مدل‌های خودرگرسیون میانگین متحرک انباشته و شیوه‌های داده‌محور به پیش‌بینی رویدادهای خرابی، پرداخته‌اند. نتایج آنها نشان داد که رویکرد ترکیبی پیشنهادی‌شان عملکرد بهتری را در مقایسه با مدل‌های تکی نشان می‌دهد؛ آنها یک مورد صنعتی واقعی از موتور هواپیما را مطالعه کرده‌اند. از میان همه شیوه‌های داده‌محور، ماشین بردار پشتیبان عملکرد بهتری را در ارزیابی انحراف استاندارد نمونه^۲، میانگین خطا^۳، میانگین خطای مطلق^۴ و درصد خطا^۵ از خود نشان داده است. راشل و همکاران^[۷] با یکپارچه‌سازی مدل‌های پیش‌بینانه و پیشگویانه با استفاده از مدل‌های شبکه‌ی بیزی^۶ برای مدل‌های احتمالی و روش خودرگرسیون میانگین متحرک انباشته برای مدل‌های پیشگویانه و با استفاده از یک رویداد واقعی از تراش نصب شده که در یک صنعت خودروی برزیل ارائه شده است دریافتند که حدود ۱/۴ درصد از کل زمان

* نویسنده مسئول

تاریخ: دریافت ۱۴۰۰/۰۶/۲۹، اصلاحیه ۱۴۰۰/۱۱/۱۱، پذیرش ۱۴۰۱/۰۱/۳۰.

DOI:10.24200/J65.2022.57224.2189

m.khoshsirat@in.iut.ac.ir
khashei@cc.iut.ac.ir
naserm@cc.iut.ac.ir

جدول ۱. ادبیات موضوع پیش‌بینی در حوزه‌ی نگهداری و تعمیرات پیشگویانه در سال‌های اخیر.

| نویسندگان | سال انتشار | مدل‌های استفاده شده | نوع/ساختار مدل | توضیحات |
|----------------------------------|------------|----------------------------|---|--|
| چن و همکاران ^[۱۱] | ۲۰۲۱ | شبکه‌های عصبی بازگشتی عمیق | ترکیبی (هوشمند) ساختار ترکیبی مبتنی بر پیش‌پردازش داده‌ها | در این مقاله یک روش ترکیبی مبتنی بر پیش‌پردازش داده‌ها با استفاده از روش‌های تجزیه به عوامل اصلی و شبکه‌های عصبی بازگشتی عمیق ارائه شده است. |
| زوتا و همکاران ^[۱۲] | ۲۰۲۲ | شبکه‌های عصبی عمیق | تکی (هوشمند) | در این مقاله، یک شبکه عصبی عمیق به منظور تعمیرات و نگهداری پیشگیرانه ارائه شده است. |
| کیمیرا و همکاران ^[۱۳] | ۲۰۲۰ | ماشین‌های بردار پشتیبان | تکی (هوشمند) | در این مقاله، یک ماشین بردار پشتیبان به منظور تعمیرات و نگهداری پیشگیرانه ارائه شده است. |
| تسلیر و همکاران ^[۱۴] | ۲۰۲۱ | ماشین‌های یادگیر | تکی (هوشمند) | در این مقاله، ماشین‌های یادگیر به منظور تعمیرات و نگهداری پیش‌گویانه در صنعت خودرو بررسی و مقایسه شده‌اند. |
| فیلا و همکاران ^[۱۵] | ۲۰۲۰ | محاسبات و رایانش ابری | تکی (هوشمند) | در این مقاله، یک مدل مبتنی بر رایانش ابری به منظور تعمیرات و نگهداری پیشگیرانه ارائه شده است. |

حاضر به منظور دستیابی به نتایج و پیش‌بینی‌های دقیق‌تر، یک رویکرد ترکیبی موازی نامتجانس از مدل‌های کلاسیک آماری لاجیت و الگوریتم‌های یادگیری ماشین شامل شبکه‌های عصبی پرسپترون چندلایه و ماشین‌های بردار پشتیبان ارائه شده است. در روش پیشنهادی همچنین به منظور وزن‌دهی مدل‌های تشکیل‌دهنده، از یک الگوریتم حداقل مربعات استفاده شده است. در حالت کلی، مزایا و خصوصیات مثبت روش پیشنهادی را می‌توان خلاصه‌وار چنین برشمرد:

۱. مدل‌سازی جامع‌تر: روش پیشنهادی به دلیل استفاده همزمان از مدل‌های خطی و غیرخطی، هم الگوهای خطی را مدل‌سازی خواهد کرد و هم توانایی مدل‌سازی الگوهای غیرخطی را خواهد داشت. بر این اساس، نسبت به مدل‌های تکی خطی آماری و همچنین هوشمند غیرخطی، قابلیت مدل‌سازی الگوهای بیشتری را خواهد داشت و لذا نسبت به این‌گونه مدل‌ها، مدل جامع‌تری خواهد بود.

۲. رسیدن به نتایج دقیق‌تر: روش پیشنهادی به دلیل مدل‌سازی جامع‌تر الگوهای موجود در داده‌های مورد مطالعه، فارغ از نوع و خصوصیات داده‌ها، در حالت کلی، نسبت به سایر روش‌های پیش‌بینی با جامعیت مدل‌سازی پایین‌تر، قابلیت دستیابی به نتایج دقیق‌تر دارد.

۳. هزینه‌های محاسباتی پایین‌تر: الگوریتم وزن‌دهی روش پیشنهادی، یک الگوریتم غیرتکراری است. بر این اساس، روش پیشنهادی، نسبت به سایر مدل‌های ترکیبی موازی که از الگوریتم‌های ابتکاری یا فراابتکاری در تعیین اوزان استفاده می‌کنند، وزن‌ها را در مدت زمان بسیار کمتری محاسبه خواهد کرد. لذا روش پیشنهادی، از نقطه نظر هزینه‌های محاسباتی، نسبت به این‌گونه از مدل‌های ترکیبی موازی در جایگاه بالاتری قرار خواهد گرفت.

۴. وزن‌دهی کارا‌تر: الگوریتم وزن‌دهی روش پیشنهادی، یک الگوریتم بهینه‌ی جهانی است. بر این اساس، روش پیشنهادی، نسبت به سایر مدل‌های ترکیبی موازی که از الگوریتم‌های ابتکاری یا فراابتکاری در تعیین اوزان استفاده می‌کنند، وزن‌ها را

تلف شده قابل کاهش است. همچنین آنها دریافتند که ترکیب مدل احتمالی شبکه‌ی بیزی با مدل‌های پیش‌بینی تخمین افق‌های میان‌مدت و بلندمدت را که رضایت‌بخش است، تضمین می‌کند. همچنین اخیراً دویید و فیلمون^[۸] یک رویکرد نگهداری و تعمیرات پیشگویانه را نسبت به سیستم هشداردهنده‌ی خرابی با استفاده از یادگیری ماشین، در تقابل با فناوری حسگر ارائه داده‌اند. نتایج حاکی از آن بود که پمپ اسکله ممکن است بین هفته‌های هفتم و هشتم خراب شود یا نیاز به تعمیرات داشته باشد. انحراف ۱۳/۸۵٪ از زمان خرابی واقعی را می‌توان به کیفیت و حجم کم داده‌های عملیاتی ثبت شده نسبت داد. با این اوصاف، با ابهام کمتر داده‌ها، ابزار پیش‌بینی نگهداری و تعمیرات می‌تواند به عنوان مبنایی قیل از استفاده از فناوری حسگر در پمپ‌های اسکله مورد استفاده قرار گیرد. برای اولین بار بی‌تس و گرنجر^[۹] نشان دادند که ترکیب خطی از مدل‌های تکی می‌تواند عملکرد بهتری نسبت به مدل‌های منفرد داشته باشد. پس از آن بسیاری از تحقیقات، شواهدی مبنی بر تأیید این ادعا ارائه داده‌اند. همچنین نتایج مطالعه‌ی بان^[۱۰] نشان می‌دهد که ترکیب مدل‌های مختلف روشی برای بهبود دقت پیش‌بینی است که از مزایای مدل‌های پایه استفاده می‌کند. برخی از مهمترین اقدامات صورت گرفته در حوزه‌ی نگهداری و تعمیرات پیشگویانه در سال‌های اخیر در جدول ۱ خلاصه شده است.

بررسی‌های انجام شده در ادبیات موضوع روش‌های پیش‌بینی در حوزه‌ی تعمیرات پیشگویانه نشان می‌دهد که ترکیب مدل‌های نامتجانس با وجود عملکرد بالایی که از خود نشان داده‌اند، در حوزه‌ی نگهداری و تعمیرات پیشگویانه زیاد مورد استفاده قرار نگرفته است.^[۱۶] همچنین مقالات مورد بررسی نشان می‌دهند که ترکیب مدل‌های کلاسیک و هوشمند در مقایسه با کاربرد این مدل‌ها به صورت تکی در حوزه‌ی نگهداری و تعمیرات پیشگویانه از دقت بالاتری برخوردارند. در میان مدل‌های ترکیبی سری و موازی نامتجانس نیز، مدل ترکیبی موازی علی‌رغم برخورداری از مزایای بسیار، چندان مورد استفاده قرار نگرفته‌اند. بر این اساس، با توجه به مزیت‌های مدل‌های ترکیبی موازی نامتجانس در کسب نتایج دقیق‌تر و همچنین عدم استفاده از این‌گونه ساختارهای ترکیبی در حوزه‌ی نگهداری و تعمیرات پیشگویانه، در نوشتار

شبکه‌های عصبی، دقت پیش‌بینی‌ها تا حد قابل قبولی افزایش یافته است. مدل لاجیت مدل رگرسیونی است که در آن تابع یا متغیر وابسته، خود ماهیتاً بیان‌گر دو گروه است که هر یک مقادیر صفر و ۱ را اختیار می‌کنند. برای زمانی که متغیر وابسته آن متغیر موهومی است مدل ۱ مفروض است:

$$P_i = E(Y = 1 | X_i) = \frac{1}{1 + e^{-(B_1 + B_2 X_i)}} \quad (1)$$

که در آن e پایه لگاریتم طبیعی است. برای راحتی، معادله ۱ را مطابق رابطه ۲ بازنویسی می‌کنیم:

$$P_i = \frac{1}{1 + e^{-z_i}} \quad (2)$$

که در آن $z_i = B_1 + B_2 X_i$ بیان‌گر تابع توزیع تجمعی لجستیک است. اگر P_i احتمال وقوع واقعه‌یی باشد در این صورت $(1 - P_i)$ که احتمال عدم وقوع واقعه است، به صورت زیر نشان داده می‌شود:

$$1 - P_i = \frac{1}{1 + e^{z_i}} \quad (3)$$

چنانچه از معادله ۳ لگاریتم طبیعی بگیریم، نتیجه عبارت خواهد بود از:

$$L_i = Ln\left(\frac{P_i}{1 - P_i}\right) \quad (4)$$

یعنی L_i که لگاریتم نسبت برتری یا مزیت است، نه تنها برحسب X ، بلکه برحسب پارامترها نیز خطی است. در رابطه ۴ نیز L_i به نام لاجیت معروف است.^[۱۸] در این پژوهش با به کارگیری ۲۰ متغیر ورودی، خرابی به وسیله‌ی مدل لاجیت پیش‌بینی شد که نتایج در ادامه ارائه داده شده است.

۲.۲. شبکه‌های عصبی پرسپترون چندلایه

شبکه‌های عصبی پرسپترون چندلایه از جمله روش‌های مدل‌سازی و پیش‌بینی هستند که قادر به تخمین موارد غیرخطی متعدد در داده‌ها بوده و یک چارچوب محاسبه‌ی انعطاف‌پذیر برای دامنه‌ی وسیعی از مسائل غیرخطی به شمار می‌آیند.^[۱۹] یکی از مزایای برجسته‌ی این گونه از مدل‌ها در تقابل با مدل‌های غیرخطی دیگر، این است که این نوع از شبکه‌های عصبی یک تقریب زنده‌ی جهانی هستند که می‌توانند هر نوع تابعی را با دقت دلخواه تقریب بزنند. نقطه‌ی قوت شبکه‌های پرسپترون چندلایه قدرت پردازش موازی اطلاعات داده‌هاست. پرسپترون‌های چندلایه با یک لایه‌ی پنهان از کار بردی‌ترین و معروف‌ترین نوع این گونه مدل‌ها هستند. دلیل اصلی این موضوع این است که در حالت کلی یک لایه‌ی پنهان برای تقریب انواع توابع کافی است.^[۲۰] ساختار کلی پرسپترون‌های چندلایه با یک لایه‌ی پنهان در شکل ۱ قابل مشاهده است. رابطه‌ی بین خروجی (y_t) و ورودی‌ها $(X_1, \dots, X_2, \dots, X_P)$ به صورت رابطه ۵ بیان می‌شود:

$$y_t = b_o + \sum_{j=1}^Q w_{oj} \cdot g(b_{oj} + \sum_{i=1}^P w_{ij} \cdot x_{ti}) + \varepsilon_t \quad (5)$$

به طوری که w_{oj} و w_{ij} پارامترهای مدل هستند که وزن‌های اتصالاتی نام‌گذاری می‌شوند، P تعداد گره‌های ورودی و Q تعداد گره‌های مخفی است. تابع سیگموئیدی که در رابطه ۶ ارائه شده، در بیشتر مواقع به عنوان تابع فعال‌سازی لایه‌ی مخفی مورد استفاده قرار می‌گیرد.

$$g(x) = \frac{1}{1 + \exp(-x)} \quad (6)$$

به صورت دقیق‌تری محاسبه خواهد کرد. لذا روش پیشنهادی، از نقطه نظر دقت نتایج نهایی ترکیب، نیز نسبت به این‌گونه از مدل‌های ترکیبی موازی در جایگاه بالاتری قرار خواهد گرفت.

۵. دقت و عدم‌تجانس بالاتر اجزاء: شاید مهمترین ویژگی روش پیشنهادی عدم‌تجانس بالای مدل‌های تشکیل‌دهنده‌ی آن باشد. بر اساس ادبیات موضوع، دقت نهایی یک ساختار ترکیبی موازی، به دو عامل دقت منفرد اجزاء تشکیل‌دهنده و عدم‌تجانس بین اجزاء تشکیل‌دهنده بستگی دارد. در حالت کلی، هرچه دقت منفرد اجزاء تشکیل‌دهنده یا عدم‌تجانس بین این اجزاء تشکیل‌دهنده در یک ساختار ترکیبی موازی بیشتر باشد، انتظار می‌رود که دقت نهایی بالاتری حاصل شود. بر این اساس، در روش پیشنهادی سعی شده تا عدم‌تجانس و دقت اجزاء بیشینه شوند. بدین منظور، در روش پیشنهادی، اولاً از یک روش مدل‌سازی آماری خطی و دو روش هوشمند غیرخطی استفاده شده است. بدین ترتیب، دو گروه مختلف مدل‌سازی (آماري و هوشمند) که از لحاظ مدل‌سازی تفاوت‌های معنی‌داری با یکدیگر دارند، در روش پیشنهادی لحاظ شده‌اند. همچنین در گروه مدل‌های هوشمند، دو نوع مختلف مدل‌سازی هوشمند (دقت‌محور و ریسک‌محور) که از دو تفکر متفاوت مدل‌سازی پیروی می‌کنند، استفاده شده است. همچنین در روش پیشنهادی، سعی شده تا هر یک از مدل‌های انتخاب شده در هر گروه تا حد امکان دقیق باشد. بر این اساس، در روش پیشنهادی و در گروه مدل‌های آماری، از یک روش بهینه و در گروه مدل‌های هوشمند از دو تا از دقیق‌ترین روش‌های معمول استفاده شده است.

به عبارت دیگر، مراحل کلی روش پیشنهادی، شامل انتخاب ساختار ترکیب، انتخاب مدل‌های تشکیل‌دهنده‌ی ترکیب، انتخاب روش وزن‌دهی اجزاء تشکیل‌دهنده، به‌گونه‌ی طراحی شده است تا اهداف روش پیشنهادی، شامل مدل‌سازی تا حد امکان جامع، هزینه‌های محاسباتی تا حد امکان پایین و نهایتاً حصول نتایج تا حد امکان دقیق، محقق شود.

در این راستا و به منظور راستی‌آزمایی و ارزیابی مدل پیشنهادی از داده‌های محک^۷ استفاده خواهد شد. سایر قسمت‌های ارائه شده در این مقاله بدین صورت است که در بخش دوم به بررسی مدل‌های کلاسیک آماری و یادگیری ماشین پرداخته شده است. در بخش سوم به ارائه‌ی مدل ترکیبی پیشنهادی پرداخته شده است. در بخش چهارم عملکرد ساختارهای ترکیب ارزیابی می‌شوند و نهایتاً در قسمت پنجم به نتیجه‌گیری از مباحث مطرح شده در این مقاله پرداخته شده است.

۲. مدل‌های کلاسیک آماری و یادگیری ماشین

روش‌های ترکیبی یکی از شناخته شده‌ترین ابزارهای موجود برای مقابله با محدودیت‌های مدل‌های تک‌ی، استفاده همزمان از خصوصیت‌های منحصر بفرد آنها و به طور ویژه بهبود دقت پیش‌بینی‌ها هستند.^[۶] در این قسمت با توجه به رویکرد اصلی این مقاله مبتنی بر ارائه یک ساختار ترکیب موازی کارا، به تعریف مدل‌های استاندارد کلاسیک آماری و یادگیری ماشین پرداخته شده است.

۱.۲. مدل لاجیت

با گسترش شیوه‌ها و استفاده از روش‌های پیچیده و دقیق‌تر مانند رگرسیون‌های خطی، مدل‌های لاجیت و پروبیت^[۱۷] در دهه‌های اخیر و استفاده از مدل‌های مبتنی بر

۳.۲. ماشین بردار پشتیبان

ماشین بردار پشتیبان به دلیل دقت بالایی که دارد یکی از روش‌های پرکاربرد یادگیری با نظارت است که از آن برای طبقه‌بندی و رگرسیون استفاده می‌شود. [۲۳] و [۲۴] ایده اصلی استفاده از ماشین بردار پشتیبان در طبقه‌بندی الگوها را می‌توان به صورت خلاصه چنین شرح داد: [۱۶]

ابتدا بردارهای ورودی در یک فضای مشخص به صورت خطی یا غیرخطی که مربوط به انتخاب تابع کرنل است، ترسیم می‌شود. پس از مرحله اول در فضای ویژگی به دنبال یک تقسیم‌بندی خطی بهینه - یک ابرصفحه که دو کلاس (یا چند کلاس) را جدا می‌کند - است. آموزش ماشین بردار پشتیبان همیشه به دنبال یک جواب بهینه جهانی است که از برزش بیش از حد آن جلوگیری می‌کند، بنابراین توانایی مقابله با تعداد زیادی از ویژگی‌ها را دارد. [۲۵]

فرض می‌شود با توجه به یک سری از نمونه‌های داده شده، یعنی بردارهای ورودی x_i مطابق با رابطه ۸:

$$x_i \in R^d (i = 1, 2, \dots, N) \quad (8)$$

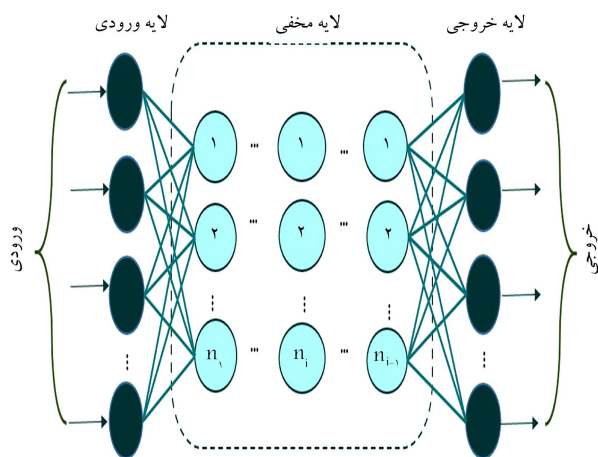
برچسب‌های مربوط به $y_i \in \{-1, +1\} (i = 1, 2, \dots, N)$ جایی که -1 و $+1$ برای دو کلاس قابل استفاده است. هدف ایجاد یک طبقه‌بندی باینری یا استخراج یک تابع تصمیم‌گیری از نمونه‌های موجود است، که احتمال کمی برای طبقه‌بندی نادرست یک نمونه در آینده دارد. با وجود کرنل‌های مختلفی مانند کرنل‌های خطی، چندجمله‌ای و غیره، مدل ماشین بردار پشتیبان با کرنل چندجمله‌ای با درجه ۳ مورد آزمایش و ارزیابی قرار گرفت که در ادامه به تشریح آن پرداخته شده است.

در ماشین بردار پشتیبان با حاشیه نرم یک پارامتر ضریب جریمه به نام C وجود دارد که میان خطای آموزشی اجازه داده شده و بیشینه‌سازی حاشیه، مصالحه ایجاد می‌کند. با افزایش مقدار C ، هزینه طبقه‌بندی اشتباه به ازای هر داده‌ی آموزشی افزایش یافته و منجر به ایجاد مدلی دقیق‌تر برای داده‌های آموزشی و در نتیجه کمتر شدن عمومیت طبقه‌بندی‌کننده برای داده‌های آزمون می‌شود. در واقع این پارامتر پیچیدگی تابع را کنترل می‌کند، به طوری که با افزایش مقدار C پیچیدگی مدل نیز افزایش پیدا می‌کند. در این پژوهش از حاشیه نرم با مقدار 0.1% استفاده شده است.

۳. مدل ترکیبی پیشنهادی بر پایه کمینه مربعات خطا

در ساختار ترکیب موازی، داده‌ها به صورت همزمان به عنوان متغیرهای ورودی برای هر یک از مدل‌های ترکیبی پیش‌بینی در نظر گرفته شده‌اند. پس از آن، کلیه ساختارهای خطی و غیرخطی موجود در داده‌ها توسط مدل‌های مورد بررسی تشخیص داده شده و در نهایت در مرحله ترکیب مقادیر پیش‌بینی شده با هم‌دیگر ترکیب می‌شوند. فرایند طراحی ساختارهای ترکیب به صورتی است که در ادامه آمده است. شکل ۲ ساختار کلی یک مدل ترکیبی موازی n جزئی را نشان می‌دهد. [۲۶]

در ترکیب موازی خطی مدل‌های ترکیبی پیش‌بینی، پس از تولید مقادیر پیش‌بینی شده توسط مدل‌های ترکیبی، به هر یک از مقادیر پیش‌بینی شده توسط مدل‌های ترکیبی وزن مشخصی اختصاص داده می‌شود. وزن‌های تخصیص داده شده سهم هر مدل را در پیش‌بینی ترکیبی مشخص می‌کنند. لذا تعیین روش وزن‌دهی مناسب، نقش مهمی را در عملکرد این دسته از ساختارهای موازی ایفا می‌کند. روش‌های وزن‌دهی را می‌توان در سه دسته میانگین ساده، الگوریتم‌های فراابتکاری^۹ و روش



شکل ۱. ساختار کلی شبکه‌های عصبی پرسپترون چندلایه. [۲۱]

همان‌طور که در رابطه ۷ نشان داده شده است، پرسپترون چندلایه به عنوان یک نگاشت غیرخطی از مقادیر ورودی به مقادیر خروجی در نظر گرفته می‌شود. پس رابطه بین ورودی و خروجی در شبکه‌های عصبی را می‌توان مطابق رابطه ۷ در نظر گرفت:

$$y_i = f(x_1, x_2, \dots, x_p, w) + \varepsilon_t \quad (9)$$

به طوری که W نشان‌دهنده بردار همه پارامترهاست و f تابعی که توسط ساختار شبکه‌های عصبی و وزن‌های اتصالی تعیین می‌شود. از این رو شبکه‌های عصبی، معادل یک مدل خودرگرسیون غیرخطی هستند. [۲]

پس از مشخص شدن یک ساختار شبکه‌ای، شبکه آماده آموزش فرایند تخمین پارامترها می‌شود. این فرایند به وسیله الگوریتم‌های بهینه‌سازی غیرخطی مانند الگوریتم پس انتشار خطا^۸ انجام می‌شود. در این مقاله بعد از سعی و خطا یک لایه پنهان برای مدل‌سازی در نظر گرفته شد چرا که از اجرای ساختارهای مختلف پرسپترون چندلایه با یک لایه مخفی، نتایج قابل قبولی حاصل می‌شود ضمن این که لایه مخفی به تعداد نرون فراوانی نیز نیاز نداشت. [۲۲] بیشتر راه‌حل‌های معمول در مشخص کردن تعداد نرون‌های لایه پنهان با روش سعی و خطا مشخص می‌شود. در این پژوهش برای بررسی کارایی، برای لایه پنهان تعداد ۱۰ نرون در نظر گرفته شده است. در حالت کلی تعداد نرون‌های لایه ورودی برابر تعداد متغیرهای موجود در بردار ورودی است. بنابراین در این پژوهش نیز تعداد نرون‌های لایه ورودی برابر تعداد متغیرهای ورودی مسئله، یعنی ۲۰ نرون در نظر گرفته شده است. همچنین از آن‌جا که در این مقاله دو کلاس برای پیش‌بینی خرابی وجود دارد (خرابی و عدم خرابی) مربوط به سیستم فشار هوای کامیون، تنها یک نرون در لایه خروجی مورد نیاز است. توابع فعال‌سازی یا توابع انتقال، مشخص‌کننده روابط بین داده‌های ورودی و خروجی به نرون مذکور هستند. در حالت کلی این گونه توابع بیان‌گر درجه‌ی غیرخطی است که در بیشتر کاربردهای شبکه‌های عصبی مصنوعی بسیار ارزشمند و تأثیرگذار است. در حالت کلی یک شبکه ممکن است توابع فعال‌سازی متفاوتی برای هر نرون در لایه‌های مختلف داشته باشد. در این پژوهش برای نرون‌های لایه پنهان و لایه خروجی از تابع فعال‌سازی سیگموئید استفاده شده است. در حالت کلی، آموزش شبکه‌های عصبی مصنوعی یک مسئله بهینه‌سازی غیرخطی محسوب می‌شود که از طریق اصلاح اوزان به ترتیبی خاص معیار سنجش را بهینه می‌سازد. در این پژوهش برای آموزش شبکه‌ی عصبی پرسپترون چندلایه از الگوریتم پس انتشار خطا استفاده شده است.

با ساده‌سازی معادلات نرمال، روابط ۱۴ تا ۱۶ حاصل می‌شود:

$$\hat{W}_{logit} \sum_{t=1}^m \hat{y}_{logit,t} + \hat{W}_{MLP} \sum_{t=1}^m \hat{y}_{logit,t} \hat{y}_{MLP,t} + \hat{W}_{SVM} \sum_{t=1}^m \hat{y}_{logit,t} \hat{y}_{SVM,t} = \sum_{t=1}^m y_t \hat{y}_{Logit,t} \quad (14)$$

$$\hat{W}_{logit} \sum_{t=1}^m \hat{y}_{logit,t} \hat{y}_{MLP,t} + \hat{W}_{MLP} \sum_{t=1}^m \hat{y}_{MLP,t} + \hat{W}_{SVM} \sum_{t=1}^m \hat{y}_{MLP,t} \hat{y}_{SVM,t} = \sum_{t=1}^m y_t \hat{y}_{MLP,t} \quad (15)$$

$$\hat{W}_{logit} \sum_{t=1}^m \hat{y}_{logit,t} \hat{y}_{SVM,t} + \hat{W}_{MLP} \sum_{t=1}^m \hat{y}_{MLP,t} \hat{y}_{SVM,t} + \hat{W}_{SVM} \sum_{t=1}^m \hat{y}_{SVM,t} = \sum_{t=1}^m y_t \hat{y}_{SVM,t} \quad (16)$$

که فرم ماتریسی این معادلات مطابق رابطه ۱۷ تعریف می‌شوند. در نهایت با حل معادلات فوق \hat{W}_{MLP} و \hat{W}_{SVM} محاسبه می‌شود.

۴. بررسی عملکرد مدل ترکیبی پیشنهادی در پیش‌بینی

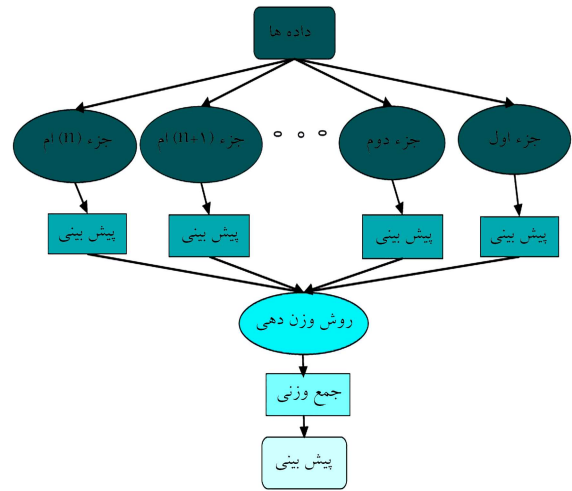
خرابی

مسئله‌ی نهایی در ارتباط با انتخاب ساختار ترکیب موازی، بررسی عملکرد ساختارهای ترکیب به منظور مدل‌سازی و پیش‌بینی سیستم‌های تحت مطالعه است. به همین دلیل ارزیابی عملکرد ساختارهای ترکیب طراحی شده، ما را به سوی دستیابی به یک ساختار کارآمد سوق می‌دهد. لذا در این قسمت، در راستای معرفی یک ساختار ترکیب موازی کارا، عملکرد ساختارهای ترکیب ارائه شده مورد تجزیه و تحلیل قرار گرفته است.

۱.۴. اندازه‌گیری عملکرد

اندازه‌گیری عملکرد در بسیاری از موارد پیش‌بینی معمولاً از یک ماتریس مقایسه 2×2 با نام ماتریس درهم‌ریختگی^{۱۱} محاسبه می‌شود؛ جایی که هر ورودی (i,j) شامل تعداد پیش‌بینی‌های درست/نادرست است.^[۱۷] در این مقاله از ماتریس درهم‌ریختگی مطابق شکل ۳ برای ارزیابی عملکرد مدل‌ها در پیش‌بینی خرابی استفاده شده، که نتایج آن در ادامه ارائه شده است.

با توجه به شکل ۳، یکی از متداول‌ترین، اساسی‌ترین و ساده‌ترین معیار اندازه‌گیری کیفیت یک «دسته‌بندی معیار حساسیت^{۱۲}» است که آن را «نرخ پاسخ‌های مثبت درست^{۱۳}» نیز می‌گویند. در واقع زمانی که از این پارامتر به عنوان پارامتر ارزیابی برای دسته‌بند استفاده می‌شود، هدف دستیابی به نهایت صحت در تشخیص نمونه‌های کلاس مثبت است. در نقطه‌ی مقابل این پارامتر، ممکن است در مواقعی صحت تشخیص کلاس منفی حائز اهمیت باشد. از متداول‌ترین پارامترها که معمولاً در کنار حساسیت بررسی می‌شود، پارامتر «خاصیت^{۱۴}» است که به آن «نرخ پاسخ‌های منفی درست^{۱۵}» نیز می‌گویند. زمانی که داده‌های مربوط به هر کلاس نسبت به هم بالانس نباشد، با توجه به اهمیت بالای ارزیابی عملکرد در هر دو کلاس، روش‌های جایگزین مورد استفاده قرار می‌گیرد. یک معیار شناخته شده، میانگین هندسی واقعی



شکل ۲. ساختار کلی یک مدل ترکیبی موازی n جزئی.

کمینه مربعات^{۱۶} جای داد. در این مقاله فرایند پیاده‌سازی مدل‌های ترکیبی موازی خطی، بر پایه‌ی روش وزن‌دهی مربعات خطا بوده است. بنابراین ترکیب خطی موازی مدل‌های لاجیت و شبکه‌های عصبی پرسپترون چندلایه و ماشین بردار پشتیبان به صورت رابطه ۹ تعریف می‌شود. به طوری که $\hat{W}_i (i = 1, 2, \dots, n)$ وزن تخصیص داده شده به نامین مدل است.

$$y_t = \hat{W}_{logit} \hat{y}_{logit,t} + \hat{W}_{MLP} \hat{y}_{MLP,t} + \hat{W}_{SVM} \hat{y}_{SVM,t} + e_t \quad (t = 1, 2, \dots, m) \quad (9)$$

یکی از معیارهای مناسب برای تخمین وزن‌های تخصیص داده شده، کمینه کردن خطای حاصله است. از آن‌جا که خطاهای مثبت و منفی همدیگر را خنثی می‌کنند، لذا مجموع خطاها یعنی $\sum e_t$ برابر صفر است. از این رو از کمینه کردن مجموع مربعات خطا طبق رابطه ۱۰ استفاده می‌شود.

بنابراین بر اساس روابط ۱۰ تا ۱۳ که به «معادلات نرمال» معروف‌اند، \hat{W}_{logit} ، \hat{W}_{MLP} و \hat{W}_{SVM} باید به‌گونه‌ی تعیین شود که مجموع مربعات خطا کمینه شود.

$$\frac{\partial \sum_{t=1}^m e_i^2}{\partial \hat{W}_{logit}} = \sum_{t=1}^m (y_t - \hat{W}_{logit} \hat{y}_{logit,t} - \hat{W}_{MLP} \hat{y}_{MLP,t} - \hat{W}_{SVM} \hat{y}_{SVM,t})^2 \quad (10)$$

$$\frac{\partial \sum_{t=1}^m e_i^2}{\partial \hat{W}_{logit}} = 2 \sum_{t=1}^m (y_t - \hat{W}_{logit} \hat{y}_{logit,t} - \hat{W}_{MLP} \hat{y}_{MLP,t} - \hat{W}_{SVM} \hat{y}_{SVM,t})(-\hat{y}_{logit,t}) = 0 \quad (11)$$

$$\frac{\partial \sum_{t=1}^m e_i^2}{\partial \hat{W}_{MLP}} = 2 \sum_{t=1}^m (y_t - \hat{W}_{logit} \hat{y}_{logit,t} - \hat{W}_{MLP} \hat{y}_{MLP,t} - \hat{W}_{SVM} \hat{y}_{SVM,t})(-\hat{y}_{MLP,t}) = 0 \quad (12)$$

$$\frac{\partial \sum_{t=1}^m e_i^2}{\partial \hat{W}_{SVM}} = 2 \sum_{t=1}^m (y_t - \hat{W}_{logit} \hat{y}_{logit,t} - \hat{W}_{MLP} \hat{y}_{MLP,t} - \hat{W}_{SVM} \hat{y}_{SVM,t})(-\hat{y}_{SVM,t}) = 0 \quad (13)$$

$$\begin{bmatrix} \sum_{t=1}^m \hat{y}_{logit,t}^V & \sum_{t=1}^m \hat{y}_{logit,t} \hat{y}_{MLP,t}^V & \sum_{t=1}^m \hat{y}_{logit,t} \hat{y}_{SVM,t}^V \\ \sum_{t=1}^m \hat{y}_{logit,t} \hat{y}_{MLP,t}^V & \sum_{t=1}^m \hat{y}_{MLP,t}^V & \sum_{t=1}^m \hat{y}_{MLP,t} \hat{y}_{SVM,t}^V \\ \sum_{t=1}^m \hat{y}_{logit,t} \hat{y}_{SVM,t}^V & \sum_{t=1}^m \hat{y}_{MLP,t} \hat{y}_{SVM,t}^V & \sum_{t=1}^m \hat{y}_{SVM,t}^V \end{bmatrix} \begin{bmatrix} \hat{W}_{logit} \\ \hat{W}_{MLP} \\ \hat{W}_{SVM} \end{bmatrix} = \begin{bmatrix} \hat{y}_{logit,1} & \hat{y}_{logit,2} & \dots & \hat{y}_{logit,m} \\ \hat{y}_{MLP,1} & \hat{y}_{MLP,2} & \dots & \hat{y}_{MLP,m} \\ \hat{y}_{SVM,1} & \hat{y}_{SVM,2} & \dots & \hat{y}_{SVM,m} \end{bmatrix} \begin{bmatrix} y_1 \\ y_1 \\ \dots \\ y_m \end{bmatrix} \quad (17)$$

هوا^{۱۹} خراب‌اند و کلاس دوم مربوط به کامیون‌هایی است که قطعات مربوط به سیستم فشار هوا در آنها خراب نیست. همان‌طور که انتظار می‌رود، در دنیای واقعی تعداد نمونه‌های متعلق به کلاس دوم بسیار بیشتر از کلاس اول خواهد بود. بنابراین، توزیع نمونه‌ها در بین کلاس‌ها به شدت نامتعادل است. همچنین در کل مجموعه داده‌ها، ۱۷۱ ویژگی موجود است که در ادامه شرح داده شده است.

۱.۲.۴. پیش‌پردازش داده‌ها

داده‌های خام اغلب برای استفاده به عنوان متغیرهای ورودی انواع مختلف مدل‌ها و به ویژه در مدل‌های شبکه‌های عصبی مصنوعی مناسب نخواهد بود. بنابراین در این پژوهش، پیش‌پردازش داده‌های ورودی قبل از مدل‌سازی توسط روش‌های یادگیری ماشین مدنظر قرار گرفته است.

الف) آنالیز مؤلفه‌های اصلی

تجزیه و تحلیل مؤلفه‌های اصلی^{۲۰} یک روش آماری است که برای تبدیل مجموعه‌ی از مشاهدات متغیرهای احتمالاً هم‌بسته، به مجموعه‌ی از مقادیر متغیرهای غیرهم‌بسته‌ی خطی به نام مؤلفه‌های اصلی استفاده می‌شود؛ یعنی روشی برای کاهش ابعاد.^[۳۲] این تحول در نوع خود خطی است و اقدامات آماری زیادی را برای دستیابی به این تحول انجام می‌دهد. کاهش ابعاد با توجه به این که اطلاعات لازم در آن مجموعه داده از بین نمی‌رود انجام می‌شود. مجموعه داده‌های جدید متعادل‌تر بعد از این که به صورت خطی پیش‌بینی شدند، به ابعاد دیگری که حاوی یک سری مؤلفه‌های اصلی است، فرستاده می‌شوند. این مؤلفه‌های اصلی بر اساس مقادیر ویژه مربوطه در ترتیب نزولی بردارهای ویژه تنظیم می‌شوند. بردار ویژه یا مؤلفه اصلی که دارای بیشترین مقدار ویژه است، به عنوان اولین مؤلفه اصلی به دنبال دومین جزء اصلی و غیره تا آخرین ستون شناخته می‌شود. بردارهای ویژه یا مؤلفه‌های اصلی که مقادیر ویژه مربوطه به آنها بسیار ناچیز است دور انداخته می‌شوند و بدین ترتیب یک ماتریس جدید به نام Y تولید می‌شود. در چنین حالتی مجموعه داده‌های اولیه X که از ابعاد r × c است در نهایت تبدیل شده و به یک ماتریس Y با بعد k × r تبدیل می‌شود، در جایی که k کمتر از c است.^[۳۳] قبل از پاکسازی داده‌ها، با استفاده از روش آنالیز مؤلفه‌های اصلی از تمام ۱۷۰ متغیر موجود که به دلایل ناشناخته فاقد نام هستند، پس از حذف بر اساس واریانس نزدیک به صفر و همچنین جایگزینی داده‌های گمشده، تعداد ۱۰۰ متغیر باقی مانده است که بعد از یکپارچه‌سازی این متغیرها در نهایت ۲۰ متغیر یکپارچه به دست آمده است.

برجسب های پیش بینی شده

| | | | | |
|--------|----------|-----------------------------------|---|---|
| | | Positive | Negative | |
| واقعیت | Positive | True Positive (TP) | False Negative (FN) | Sensitivity $\frac{TP}{TP + FN}$ |
| | Negative | False Positive (FP) | True Negative (TN) | Specificity $\frac{TN}{TN + FP}$ |
| | | Precision $\frac{TP}{TP + FP}$ | Negative Predictive Value $\frac{TN}{TN + FN}$ | Accuracy $\frac{TP + TN}{TP + TN + FP + FN}$ |

شکل ۳. ماتریس درهم‌ریختگی برای یک مسئله‌ی طبقه‌بندی باینری.

است.^[۲۸]

$$Gmean = \sqrt{TPR \times TNR} \quad (18)$$

همچنین میانگین حسابی یا متوسط حسابی^{۱۶}، نوعی سنجش‌گرایش به مرکز است و عبارت است از مجموع مقادیر موجود در یک مجموعه داده تقسیم بر تعداد آنها:

$$Amean = \frac{TPR + TNR}{2} \quad (19)$$

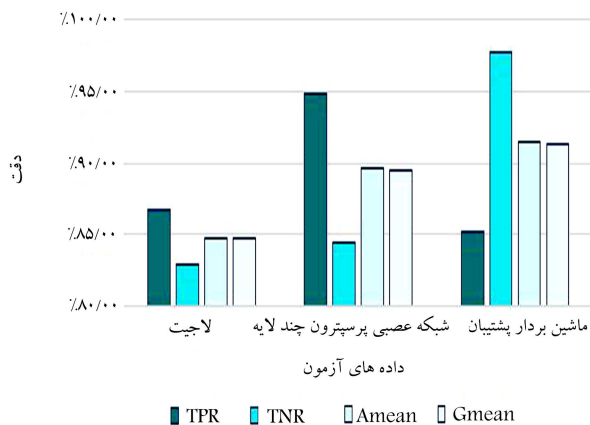
۲.۴. مجموعه داده‌ها

در دنیای واقعی، داده‌های گمشده^{۱۷} یک مسئله‌ی رایج در برنامه‌های یادگیری ماشین است. در چالش پیش‌بینی یادگیری ماشین در سال ۲۰۱۶، مجموعه‌ی از داده‌های به دست آمده از کامیون‌های اسکانیا ارائه شده است که در آن داده‌های حاکی از حسگرهای مختلف، موجود نیست. در این پژوهش این داده‌ها از داده‌های محک ثبت شده در پایگاه داده ایالت کالیفرنیا آمریکا^{۱۸}، که مؤسس آن دانشگاه کالیفرنیا در ایرواین است، ارائه داده شده است. داده‌های این پایگاه در تحقیقات بسیاری در حوزه‌های بین‌المللی برای آزمون الگوریتم‌های خود مورد استفاده قرار گرفته است. در این مجموعه داده که در مقالات متعددی به کار گرفته شده‌اند^[۲۹-۳۱]، داده‌ها از دو کلاس تشکیل شده‌اند. کلاس اول مربوط به کامیون‌هایی است که در ارتباط با یک جزء خاص از سیستم فشار

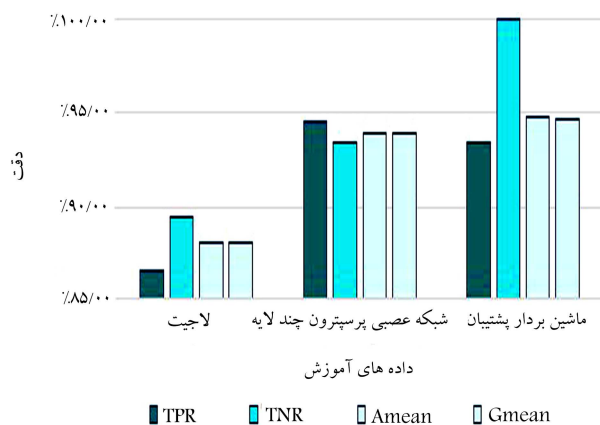
مدل‌های logit-SVM و logit-MLP، در داده‌های آزمون و آزمایش و همچنین کل داده‌ها نشان داده شده است.

۳.۳.۴. پیش‌بینی با ساختار ترکیب موازی

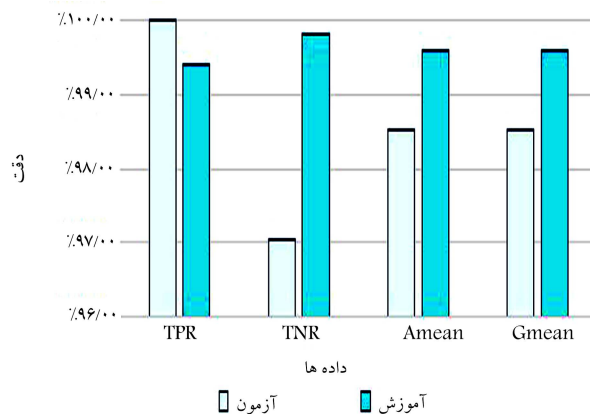
جدول ۶ و ۷ نشان‌دهنده نتایج به دست آمده برای مدل ترکیبی موازی هستند؛ شکل‌های ۴ و ۵ به ترتیب بیانگر نمودار میله‌ی روش‌های تکی در داده‌های آزمون و آموزش هستند. شکل ۶ نیز بیانگر نمودار میله‌ی روش ترکیبی برای سه مدل تکی



شکل ۴. نمودار میله‌ی معیارهای عملکرد مدل‌های تکی برای داده‌های آزمون.



شکل ۵. نمودار میله‌ی معیارهای عملکرد مدل‌های تکی برای داده‌های آموزش.



شکل ۶. نمودار میله‌ی معیارهای عملکرد مدل‌های ترکیبی موازی برای داده‌های آموزش و آزمون.

ب) پاکسازی داده‌ها^{۲۱}

در این پژوهش مجموعه داده شامل تعدادی مقادیر از دست رفته در هر ویژگی است. بنابراین برای رفع مشکل موجود، داده‌هایی که شامل بیشتر از ۵٪ مقادیر گمشده بودند، حذف و مابقی مقادیر گمشده با میانگین جایگزین شده‌اند.

ج) متعادل کردن داده‌ها

پس از جایگزینی میانگین با داده‌های گمشده، کل مجموعه داده موجود به دو قسمت داده‌های آموزش و داده‌های آزمون تقسیم شده‌اند. داده‌های آموزش برای گسترش توانایی مدل‌های پیش‌بینی و از سوی دیگر داده‌های آزمون برای ارزیابی توانایی مدل‌ها به کار گرفته می‌شوند. مجموعه‌ی آموزش شامل ۱۰۰۰ مشاهده و مجموعه آزمون شامل ۲۷۰ مشاهده است. در هر دو مجموعه داده‌های آموزش و داده‌های آزمون بعد از متعادل کردن کامل همه‌ی داده‌های موجود از مرحله‌ی پاکسازی داده‌ها، از ۱۰۰۰ داده‌ی آموزش ۵۰۰ داده مربوط به کلاس مثبت یا ۱ و ۵۰۰ داده مربوط به کلاس منفی یا صفر (از ۲۷۰ داده آزمون باقی مانده ۱۳۵ داده مربوط به کلاس مثبت یا ۱ و ۱۳۵ داده مربوط به کلاس منفی یا صفر هستند).

۳.۴. پیش‌بینی با مدل‌های تکی و ترکیبی

در این قسمت عملکرد مدل پیشنهادی در مقابل مدل‌های تکی تشکیل‌دهنده‌ی آن در پیش‌بینی خرابی سیستم فشار هوای کامیون بررسی می‌شود.

۱.۳.۴. پیش‌بینی با مدل‌های تکی

در جداول ۲ و ۳ به ترتیب عملکرد مدل‌های لاغیت، شبکه‌ی عصبی پرسپترون چندلایه و ماشین بردار پشتیبان در داده‌های آزمون و آزمایش و همچنین کل داده‌ها نشان داده شده است.

۲.۳.۴. پیش‌بینی با ساختار ترکیب سری

با توجه به فرایند خاص پیاده‌سازی ساختار ترکیب سری به صورت دو جزء خطی و غیرخطی و مدل‌سازی آنها به صورت متوالی و پشت سرهم و همچنین ترتیب به کارگیری مراحل مدل‌سازی خطی و غیرخطی، در این قسمت مراحل پیاده‌سازی مدل‌های ترکیبی سری به منظور پیش‌بینی مجموعه داده‌های محک به صورت مختصر شرح داده شده است.

الف) پیش‌بینی با مدل ترکیبی سری logit-MLP و مدل ترکیبی سری logit-SVM

- مرحله‌ی مدل‌سازی خطی: در این مرحله با در نظر گرفتن داده‌های علی و سببی به عنوان ورودی یک مدل کلاسیک آماری مانند مدل لاغیت، روابط و الگوهای خطی موجود در داده‌ها مدل‌سازی می‌شوند.
- مرحله‌ی مدل‌سازی غیرخطی: با در نظر گرفتن باقی‌مانده‌های تولید شده از مرحله‌ی مدل‌سازی خطی، به عنوان ورودی‌های یک مدل هوشمند مانند شبکه عصبی پرسپترون چندلایه در logit-MLP و ماشین بردار پشتیبان در logit-SVM، در این گام روابط غیرخطی آنالیز می‌شوند. ساختارهای نهایی مدل‌های لاغیت و شبکه‌ی عصبی پرسپترون چندلایه طراحی شده در فرایند پیاده‌سازی مدل‌های ترکیبی سری logit-SVM و logit-MLP به منظور پیش‌بینی آورده شده است.
- مرحله‌ی ترکیب: در گام آخر، حاصل جمع مقادیر پیش‌بینی شده از دو مرحله‌ی قبل پیش‌بینی نهایی ترکیبی را تولید می‌کند. در جداول ۴ و ۵ به ترتیب عملکرد

جدول ۲. نتایج مربوط به عملکرد مدل‌های تکی در پیش‌بینی (داده‌های آزمون و آموزش).

| مدل | میزان عملکرد | | | | | | | |
|--------------------------------|----------------|---------------|---------------|---------------|----------------|---------------|---------------|---------------|
| | داده‌های آموزش | | | | داده‌های آزمون | | | |
| | میانگین هندسی | میانگین حسابی | نرخ درست منفی | نرخ درست مثبت | میانگین هندسی | میانگین حسابی | نرخ درست منفی | نرخ درست مثبت |
| لاجیت | ٪۸۷/۹۹ | ٪۸۸/۰۰ | ٪۸۹/۴۰ | ٪۸۶/۶۰ | ٪۸۴/۷۹ | ٪۸۴/۸۱ | ٪۸۲/۹۶ | ٪۸۶/۶۷ |
| شبکه‌های عصبی پرسپترون چندلایه | ٪۹۴/۰۰ | ٪۹۴/۰۰ | ٪۹۳/۴۰ | ٪۹۴/۶۰ | ٪۸۹/۴۸ | ٪۸۹/۶۳ | ٪۸۴/۴۴ | ٪۹۴/۸۱ |
| ماشین بردار پشتیبان | ٪۹۴/۶۶ | ٪۹۴/۸۰ | ٪۱۰۰/۰۰ | ٪۹۳/۴۰ | ٪۹۱/۲۶ | ٪۹۱/۴۸ | ٪۹۷/۷۸ | ٪۸۵/۱۹ |
| میانگین نرخ‌های دسته‌بندی | ٪۹۲/۲۲ | ٪۹۲/۲۷ | ٪۹۴/۲۷ | ٪۹۱/۵۳ | ٪۸۸/۵۱ | ۸۸/۶۴ | ٪۸۸/۳۹ | ٪۸۸/۸۹ |

جدول ۳. نتایج مربوط به عملکرد مدل‌های تکی در پیش‌بینی (کل داده‌ها).

| مدل | میزان عملکرد کل داده‌ها | | | |
|---------------------------------|-------------------------|---------------|---------------|---------------|
| | میانگین هندسی | میانگین حسابی | نرخ درست منفی | نرخ درست مثبت |
| لاجیت | ٪۸۷/۳۲ | ٪۸۷/۳۳ | ٪۸۸/۰۳ | ٪۸۶/۶۲ |
| شبکه‌های عصبی پرسپترون چند لایه | ٪۹۳/۰۶ | ٪۹۳/۰۸ | ٪۹۱/۵۰ | ٪۹۴/۶۵ |
| ماشین بردار پشتیبان | ٪۹۵/۵۱ | ٪۹۵/۶۰ | ٪۹۹/۵۳ | ٪۹۱/۶۶ |
| میانگین نرخ‌های دسته‌بندی | ٪۹۱/۹۹ | ٪۹۲/۰۰ | ٪۹۳/۰۲ | ٪۹۰/۹۸ |

جدول ۴. نتایج مربوط به عملکرد مدل ترکیبی سری در پیش‌بینی (داده‌های آزمون و آموزش).

| مدل | میزان عملکرد | | | | | | | |
|--------------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | آموزش | | | | آزمون | | | |
| | میانگین هندسی | میانگین حسابی | نرخ درست منفی | نرخ درست مثبت | میانگین هندسی | میانگین حسابی | نرخ درست منفی | نرخ درست مثبت |
| مدل ترکیبی سری logit-MLP | ٪۹۲/۶۰ | ٪۹۲/۶۰ | ٪۹۲/۰۰ | ٪۹۳/۲۰ | ٪۸۹/۲۴ | ٪۸۹/۲۶ | ٪۸۷/۴۱ | ٪۹۱/۱۱ |
| مدل ترکیبی سری logit-SVM | ٪۹۵/۵۰ | ٪۹۵/۵۰ | ٪۹۵/۶۰ | ٪۹۵/۴۰ | ٪۹۲/۹۶ | ٪۹۲/۹۶ | ٪۹۴/۰۷ | ٪۹۱/۸۵ |
| میانگین ترکیبی سری | ٪۹۴/۰۵ | ٪۹۴/۰۵ | ٪۹۳/۸۰ | ٪۹۴/۳۰ | ٪۹۱/۱۰ | ٪۹۱/۱۰ | ٪۹۰/۷۴ | ٪۹۱/۴۸ |

جدول ۵. نتایج مربوط به عملکرد مدل‌های ترکیبی سری در پیش‌بینی (کل داده‌ها).

| مدل | میزان عملکرد کل داده‌ها | | | |
|--------------------------|-------------------------|---------------|---------------|---------------|
| | میانگین هندسی | میانگین حسابی | نرخ درست منفی | نرخ درست مثبت |
| مدل ترکیبی سری logit-MLP | ٪۹۱/۸۸ | ٪۹۱/۸۹ | ٪۹۱/۰۲ | ٪۹۲/۷۶ |
| مدل ترکیبی سری logit-SVM | ٪۹۴/۹۶ | ٪۹۴/۹۷ | ٪۹۵/۲۸ | ٪۹۴/۶۵ |
| مدل ترکیبی سری | ٪۹۳/۴۲ | ٪۹۳/۴۳ | ٪۹۳/۱۵ | ٪۹۳/۷۱ |

بالاترین نرخ دسته‌بندی در داده‌های آزمون و آموزش، عملکرد بالاتری دارد. همچنین با توجه به فرایند خاص پیاده‌سازی ساختار ترکیب سری به صورت دو جزء خطی و غیرخطی و مدل‌سازی آنها به صورت متوالی و پشت سرهم و همچنین ترتیب به کارگیری مراحل مدل‌سازی خطی و غیرخطی با توجه به جداول ۴ و ۵، مدل ترکیبی سری logit-SVM در معیارهای نرخ درست مثبت، نرخ درست منفی، میانگین حسابی و میانگین هندسی در مقایسه با مدل ترکیبی سری logit-MLP از عملکرد بهتری برخوردار بوده است.

آنهاست. همچنین در ادامه، نرخ درست مثبت و نرخ درست منفی مدل‌های تکی و ترکیبی موازی داده‌های آزمون و آموزش مقایسه شده است (جداول ۸ و ۹).

۴.۴. نتایج

با توجه به نتایج به دست آمده از پیش‌بینی با مدل‌های تکی در جداول ۲ و ۳، یادآور می‌شود که در میان سه مدل، مدل ماشین بردار پشتیبان به طور میانگین با

جدول ۶. نتایج مربوط به عملکرد مدل ترکیبی در پیش‌بینی (داده‌های آزمون و آموزش).

| مدل | میزان عملکرد | | | | | | | | |
|---------------------------|---------------|----------|----------|---------------|---------|----------|---------------|---------|--------|
| | آموزش | | | | آزمون | | | | |
| | میانگین | نرخ درست | نرخ درست | میانگین | میانگین | نرخ درست | نرخ درست | میانگین | |
| مدل ترکیبی موازی پیشنهادی | میانگین هندسی | مثبت | منفی | میانگین هندسی | مثبت | منفی | میانگین حسابی | مثبت | منفی |
| | ٪۹۹/۶۰ | ٪۹۹/۶۰ | ٪۹۹/۸۰ | ٪۹۸/۵۱ | ۹۸/۵۲ | ٪۹۷/۰۴ | ٪۱۰۰/۰۰ | ٪۹۹/۶۰ | ٪۹۹/۸۰ |

جدول ۷. نتایج مربوط به عملکرد مدل ترکیبی در پیش‌بینی (کل داده‌ها).

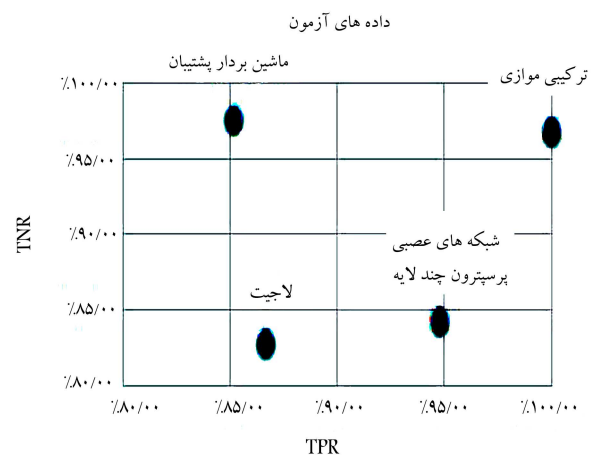
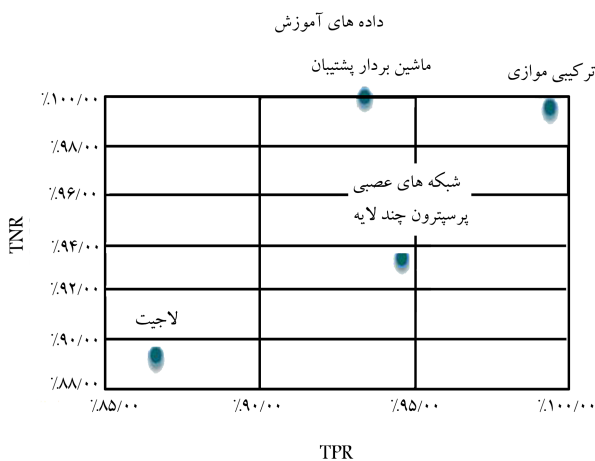
| مدل | میزان عملکرد کل داده‌ها | | | |
|---------------------------|-------------------------|---------------|---------------|---------------|
| | میانگین هندسی | نرخ درست مثبت | نرخ درست منفی | میانگین حسابی |
| مدل ترکیبی موازی پیشنهادی | ٪۹۹/۳۶ | ٪۹۹/۵۳ | ٪۹۹/۲۱ | ٪۹۹/۳۷ |

جدول ۸. مقایسه‌ی نرخ درست مثبت و نرخ درست منفی مدل‌های تکی و ترکیبی موازی (داده‌های آزمون و آموزش).

| | داده‌های آموزش | | | | داده‌های آزمون | | | |
|---------------|----------------|---------------------|------------------|--------|----------------|---------------------|------------------|--------|
| | مدل پیشنهادی | ماشین بردار پشتیبان | پرسپترون چندلایه | لاجیت | مدل پیشنهادی | ماشین بردار پشتیبان | پرسپترون چندلایه | لاجیت |
| نرخ درست مثبت | ٪۹۹/۴۰ | ٪۹۳/۴۰ | ٪۹۴/۶۰ | ٪۸۶/۶۰ | ٪۱۰۰/۰۰ | ٪۸۵/۱۹ | ٪۹۴/۸۱ | ٪۸۶/۶۷ |
| نرخ درست منفی | ٪۹۹/۶۰ | ٪۱۰۰/۰۰ | ٪۹۳/۴۰ | ٪۸۹/۴۰ | ٪۹۷/۰۴ | ٪۹۷/۷۸ | ٪۸۴/۴۴ | ٪۸۲/۹۶ |

جدول ۹. مقایسه‌ی نرخ درست مثبت و نرخ درست منفی مدل‌های تکی و ترکیبی موازی (کل داده‌ها).

| | کل داده‌ها | | | |
|---------------|--------------|---------------------|------------------|--------|
| | مدل پیشنهادی | ماشین بردار پشتیبان | پرسپترون چندلایه | لاجیت |
| نرخ درست مثبت | ٪۹۹/۵۳ | ٪۹۱/۶۶ | ٪۹۴/۶۵ | ٪۸۶/۶۲ |
| نرخ درست منفی | ٪۹۹/۲۱ | ٪۹۹/۵۳ | ٪۹۱/۵۰ | ٪۸۸/۰۳ |



شکل ۸. نرخ درست مثبت در مقابل نرخ درست منفی در مدل‌های تکی و ترکیبی موازی داده‌های آموزش.

شکل ۷. نرخ درست مثبت در مقابل نرخ درست منفی در مدل‌های تکی و ترکیبی موازی داده‌های آزمون.

روابط مدل ترکیبی موازی پیشنهادی این مقاله جداول ۶ و ۷، بیان‌گر آن است که ساختار پیشنهادی ترکیبی موازی، در قیاس با مدل‌های تکی و مدل ترکیبی سری، به مراتب عملکرد بهتری برای پیش‌بینی خرابی از خود نشان می‌دهد و منجر به بهبود نتایج شده است. همچنین در شکل‌های ۷ و ۸ مشاهده می‌شود که مدل ترکیبی موازی پیشنهادی عملکرد بالاتری نسبت به مدل‌های تکی از خود به نمایش

همچنین با توجه به نتایج به دست آمده از مدل‌های تکی و ترکیبی سری، مدل ترکیبی سری دقت پیش‌بینی‌ها را به طور میانگین به میزان ۱٪ در داده‌های آموزش و ۲٪ در داده‌های آزمون بهبود بخشیده است. لذا می‌توان نتیجه گرفت که مدل ترکیبی سری در مقایسه با مدل‌های تکی به کارگرفته شده از دقت پیش‌بینی بالاتری برخوردار است. همچنین ارزیابی ساختارهای طراحی شده با بهره‌گیری از تحلیل

۵. نتیجه‌گیری

دقت پیش‌بینی‌ها یکی از مهمترین عوامل مؤثر بر کیفیت تصمیمات اتخاذی در زمینه نگهداری و تعمیرات پیشگویانه است. مدل‌های پیش‌بینی خطی همان‌گونه که از نامشان پیداست، تنها قادر به مدل‌سازی الگوها و ساختارهای خطی موجود در داده‌ها هستند. به همین ترتیب مدل‌های پیش‌بینی قطعی، تنها قادر به مدل‌سازی الگوهای قطعی و غیرفازی موجود در داده‌ها هستند. این موضوع برای مدل‌های فصلی/غیر فصلی نیز صادق بوده و ساختارهای خودهم‌بسته فصلی موجود در داده‌ها در مدل‌سازی غیرفصلی لحاظ نمی‌شود. در این مقاله یک روش ترکیبی به منظور پیش‌بینی خرابی هم‌زمان الگوهای خطی و غیرخطی ارائه شده است. ایده‌ی روش پیشنهادی ترکیب مدل‌های کلاسیک آماری با مدل‌های یادگیری ماشین به منظور بهره‌گیری از مزایای منحصر به فرد هر یک از این مدل‌هاست. به منظور مقایسه‌ی عملکرد مدل پیشنهادی با سایر مدل‌های تکی تشکیل‌دهنده‌ی آن، داده‌های معیار خرابی سیستم فشار هوای کامیون در نظر گرفته شده است. نتایج نشانگر آن است که دقت پیش‌بینی مدل پیشنهادی در معیارهای ارزیابی از سایر مدل‌های تکی و ترکیبی سری بررسی شده بهتر بوده است. لذا برای مطالعات آتی، با توجه به دقت بیشتر ترکیب مدل‌های کلاسیک آماری و یادگیری ماشین به صورت موازی، این ساختار برای به کارگیری در مسائل نگهداری و تعمیرات پیشگویانه به عنوان ابزاری نسبتاً دقیق و کارا توصیه می‌شود. در آخر محققان می‌توانند با استفاده از ساختار ترکیبی سری - موازی در رویکرد پیشنهادی و بررسی نتایج حاصل از آن و همچنین استفاده از سایر مدل‌های آماری و هوشمند دیگر برای ساختن مدل LPHM در مدل‌سازی و پیش‌بینی را مورد بررسی قرار دهند.

گذاشته است؛ یعنی دقت پیش‌بینی‌های مدل پیشنهادی در تمام معیارهای ارزیابی از مدل‌های تکی و ترکیبی سری بهتر بوده است. بدین ترتیب پس از ارزیابی و سنجش عملکرد ساختارهای ترکیب و پس از کسب اطمینان از کارایی و کارآمدی ساختار ترکیب موازی در مدل‌سازی و پیش‌بینی، مدل ترکیبی موازی در مقایسه با مدل‌های تکی دقت پیش‌بینی‌ها را به طور میانگین به میزان ۱۱٪ در داده‌های آزمون و ۷٪ در داده‌های آموزش بهبود بخشیده است.

خلاصه‌ی نتایج مقایسه‌ی مدل‌های تکی و ترکیبی سری و ترکیبی موازی پیشنهادی:

- عملکرد نسبتاً بالای مدل ماشین بردار پشتیبان در مقایسه با مدل‌های لاجیت و شبکه‌های عصبی پرسپترون چندلایه؛
- برتری مدل‌های یادگیری ماشین در معیارهای نرخ درست مثبت و نرخ درست منفی، در مقایسه با مدل کلاسیک آماری؛
- بهبود عملکرد مدل ترکیبی موازی پیشنهادی در مقایسه با مدل‌های تکی و ترکیبی سری؛
- مدل ترکیبی موازی پیشنهادی، در مقایسه با مدل‌های تکی، دقت پیش‌بینی‌ها را به طور میانگین به میزان ۱۱٪ در داده‌های آزمون و ۷٪ در داده‌های آموزش بهبود بخشیده است؛
- مدل ترکیبی سری logit-SVM عملکرد بهتری در مقایسه با مدل ترکیبی سری logit-MLP از خود نشان داده است؛
- مدل ترکیبی سری، در مقایسه با مدل‌های تکی، به طور میانگین دقت پیش‌بینی‌ها را در داده‌های آموزش و آزمون بهبود بخشیده است.

پانویس‌ها

1. predictive maintenance
2. standard deviation
3. median error (ME)
4. mean absolute error (MAE)
5. percentage error
6. bayesian network (BN)
7. benchmark
8. back propagation
9. meta heuristic algorithm
10. least squares
11. confusion matrix
12. sensitivity
13. true positive rate
14. specificity
15. true negative rate
16. arithmetic mean (Amean)
17. missing data
18. University of California at Irvine (UCI)
19. air pressure system (APS)
20. principal component analysis (PCA)
21. data cleaning
22. median
23. balancing

منابع (References)

1. Márquez, A.C. *The maintenance management framework: models and methods for complex systems maintenance*, Springer Series in Reliability Engineering, Softcover reprint of hardcover 1st ed. (2007).
2. Khashei, M. and Bijari, M. "Improving the performance of financial forecasts by combining linear and nonlinear models of stacked moving average self-regression and artificial neural networks", *Sustainable Growth and Development Research (Economic Research)*, **8**, pp. 83-100, (In Persian) (2008).
3. Khashei, M., Hejazi, S.R. and Bijari, M. "A new hybrid artificial neural networks and fuzzy regression model for time series forecasting", *Fuzzy Sets and Systems*, **159**, pp. 769-786 (2008).
4. Khashei, M. and Bijari, M. "Artificial neural networks for predicting and analyzing the prices of products made by Isfahan Steel Company in Tehran Metal Exchange (Part 2)", *Foolad Monthly*, **134**, pp. 31-28, (In Persian) (1385).

5. Bukhsh, Z.A., Saeed, A., Stipanovic, I. and et al. "Predictive maintenance using tree-based classification techniques: A case of railway switches", *Transportation Research Part C: Emerging Technologies*, **101**, pp. 35-54 (2019).
6. Baptista, M., Sankararaman, S., de Medeiros, I.P. and et al. "Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling", *Computers & Industrial Engineering*, **115**, pp. 41-53 (2018).
7. Ruschel, E., Santos, E.A.P. and Loures, E.d.F.R. "Mining shop-floor data for preventive maintenance management: integrating probabilistic and predictive models", *Procedia Manufacturing*, **11**, pp. 1127-1134 (2017).
8. Kimera, D. and Nangolo, F.N. "Predictive maintenance for ballast pumps on ship repair yards via machine learning", *Transportation Engineering*, 100020 (2020).
9. Bates, J.M. and Granger, C.W. "The combination of forecasts", *Journal of the Operational Research Society*, **20**, pp. 451-468 (1969).
10. Bunn, D. "Forecasting with more than one model", *Journal of forecasting*, **8**, pp. 161-166 (1989).
11. Chen, j., Chen, T., Liua, W. and et al. "Combining empirical mode decomposition and deep recurrent neural networks for predictive maintenance of lithium-ion battery", *Advanced Engineering Informatics*, **50**, 101405 (2021).
12. Zonta, T. Costa, C. Zeiser, F. and et al. "A predictive maintenance model for optimizing production schedule using deep neural networks", *Journal of Manufacturing Systems*, **62**, pp. 450-462 (2022).
13. Kimera, D. and Nangolo, F. "Improving ship yard ballast pumps' operations: A PCA approach to predictive maintenance", *Maritime Transport Research*, **1**, 100003 (2020).
14. Theissler, A., velázquez, P., Kettelgerdes, M. and et al. "Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry", *Reliability Engineering & System Safety*, **215**, 107864 (2021).
15. FILA, R., EL KHAILI, M. and MESTARI, M., "Cloud computing for industrial predictive maintenance based on prognostics and health management", *Procedia Computer Science*, **177**, pp. 631-638 (2020).
16. Cai, Y.-D., Liu, X.-J., Xu, X.-b. and et al. "Support vector machines for prediction of protein subcellular location", *Molecular Cell Biology Research Communications*, **4**, pp. 230-233 (2000).
17. Cramer, J.S. "An introduction to the logit model for economists paperback", *Timberlake Consultants* (2001).
18. Cramer, J.S., "The origins and development of the logit model", *Logit models from economics and other fields*, **2003**, pp. 1-19 (2003).
19. Khashei, M. and Bijari, M. "An artificial neural network (p, d, q) model for timeseries forecasting", *Expert Systems with applications*, **37**, pp. 479-489 (2010).
20. Khashei, M. "Soft smart decision making", *Faculty of Industrial and Systems Engineering, Isfahan University of Technology*, (In Persian) (2013).
21. MR, G.R., Somu, N., and Mathur, A. "A multilayer perceptron model for anomaly detection in water treatment plants", *International Journal of Critical Infrastructure Protection*, pp. 100393 (2020).
22. Chauvin, Y. and Rumelhart, D.E. "Backpropagation: theory, architectures, and applications", *Psychology press*, 1st Edition (1995).
23. Sexton, T., Brundage, M.P., Hoffman, M. and et al. "Hybrid datafication of maintenance logs from ai-assisted human tags," *Proceeding of. 2017 IEEE International Conference on Big Data (Big Data)*, IEEE, pp. 1769-1777 (2017).
24. Chang, C.-C. and Lin, C.-J. "LIBSVM: A library for support vector machines", *ACM transactions on intelligent systems and technology (TIST)*, **2**, pp. 27 (2011).
25. Vapnik, V.N. "The nature of statistical learning", *Theory*, (1995).
26. Rahimi, Z.H. and Khashei, M. "A least squares-based parallel hybridization of statistical and intelligent models for time series forecasting", *Computers & Industrial Engineering*, **118**, pp. 44-53 (2018).
27. Carvalho, T.P., Soares, F.A. Vita, R. and et al. "A systematic literature review of machine learning methods applied to predictive maintenance", *Computers & Industrial Engineering*, **137**, pp. 106024 (2019).
28. Cleofas-Sánchez, L., Garcia, V., Marqués, A. and et al. "Financial distress prediction using the hybrid associative memory with translation", *Applied Soft Computing*, **44**, pp. 144-152 (2016).
29. Cerqueira, V., Pinto, F., Sá, C. and et al. "Combining boosted trees with metafeature engineering for predictive maintenance," *Proceeding of. International Symposium on Intelligent Data Analysis*, Springer, pp. 393-397 (2016).
30. Gondek, C., Hafner, D. and Sampson, O.R. "Prediction of failures in the air pressure system of scania trucks using a random forest and feature engineering," *Proceeding of. International Symposium on Intelligent Data Analysis*, Springer, pp. 398-402 (2016).
31. Ozan, E.C., Riabchenko, E., Kiranyaz, S. and et al. "An optimized k-nn approach for classification on imbalanced datasets with missing data," *Proceeding of. International Symposium on Intelligent Data Analysis*, Springer, pp. 387-392 (2016).
32. Jolliffe, I.T. "Springer series in statistics", *Principal component analysis*, Springer Series in Statistics, 2nd Edition, **29** (2002).
33. Prusty, M.R. Jayanthi, T. Chakraborty, J. and and et al. "Feasibility of ANFIS towards multiclass event classification in PFBR considering dimensionality reduction using PCA", *Annals of Nuclear Energy*, **99**, pp. 311-320 (2017).