فصلنامه‌ی مدیریت مهندسی نوین
دوره نهم، شماره سوم، پاییز ۱۴۰۲

توسعه و پیاده‌سازی روش تحلیل مؤلفه‌های اصلی جهت پایش وضعیت توربین گازی

سمیرا پیری نیارق[۱]، الهام قنبری[۲]

چکیده

توربین‌های گازی ماشین‌های پیچیده و گران‌قیمتی هستند که هزینه‌های تعمیر خرابی‌های پیش‌بینی‌نشده‌ی آن‌ها بالاست. حس‌گرهای سامانه‌ی کنترلی توربین‌های گازی به‌طور پیوسته داده‌های زیادی را گزارش می‌کنند که ثبت و جمع‌آوری مقادیر آن‌ها، تشکیل داده‌های حجیم را می‌دهند. با داده‌کاویِ داده‌هایِ حجیمِ توربین گازی، می‌توان عیب و خرابی توربین را قبل از رخداد پیش‌بینی نمود. مجموعه‌ی ماتریسِ داده‌های تحقیق حاضر، حاصل ثبت و تجمیع مقادیر برخی از حس‌گرهای نصب‌شده بر روی توربین گاز فریم ۹ یکی از نیروگاه‌های کشور هست. در این پژوهش پس از نرمالیزه کردن داده‌های ماتریس، سطرهای ماتریس بر حسب سالم یا معیوب بودن برچسب-گذاری شدند. سپس با بهره‌گیری از تکنیک مؤلفه‌های اصلی، ابعاد ماتریسِ داده، از هفت بُعد به چهار بُعد کاهش داده شد و ویژگی‌های اصلی آن استخراج گردید. پس از یادگیریِ ماشین که بر روی ۸۰٪ داده‌ها انجام گرفت معیار دقت، میزان خطا و همگرایی مدل برای کاهش‌های ابعادی مدل از دو تا شش بُعد بر روی ۲۰٪ داده‌ها مورد مطالعه قرار گرفت که در نهایت مشخص گردید علاوه بر اینکه مدل ساخته شده به‌خوبی از پس کاهش ابعاد ماتریس، استخراج ویژگی‌ها با کمک تکنیک «تحلیل مؤلفه‌های اصلی» برآمده همچنین توانسته با تکنیک «شبکه عصبی مصنوعی» وقوع عیب را شناسایی و همچنین کلاس عیب را تشخیص دهد. مدل ساخته شده با تلفیق تکنیک «تحلیل مؤلفه‌های اصلی» با تکنیک «شبکه مصنوعی عصبی» توانسته دقتی بیش از ۹۰٪ و با دقت خوب و بیشترین میزان همگرایی ماتریس داده‌ها را نمایش دهد و کلاس عیب توربین را مشخص نماید.

کلید واژه‌ها: توربین گاز، پایش وضعیت نیروگاه، تحلیل مؤلفه‌های اصلی، شبکه عصبی مصنوعی

---

۱. کارشناس ارشد گروه کامپیوتر، واحد یادگار امام خمینی (ره) شهر ری، دانشگاه آزاد اسلامی، تهران، ایران

۲. استادیار گروه کامپیوتر، واحد یادگار امام خمینی (ره) شهر ری، دانشگاه آزاد اسلامی، تهران، ایران (نویسنده مسئول)

el.ghanbari@iau.ir

تاریخ وصول:۱۴۰۲/۴/۱۳ **تاریخ پذیرش:۱۴۰۲/۷/۱۶**

مقدمه

شناسایی و تشخیص عیب، مهم‌ترین اصل در پایش وضعیت توربین گاز[1] است که در سال‌های اخیر توجه زیادی را به خود جلب کرده‌است. توربین‌های گازی نقشی بسیار مهم و حیاتی در صنایع نفت، گاز و نیرو ایفا می‌کنند. مدیریت رفتارهای دینامیکی و پیچیده توربین‌های گازی، نیازمند سامانه‌های کنترل[2]، حفاظت و پایش وضعیت[3] توربین گاز با قابلیت اطمینان، انعطاف در برنامه‌ریزی و دسترسی‌پذیری بالاست. از آنجا که هرگونه خطا[4] و عیب[5] در توربین گاز و در عملیات و سامانه‌های کنترل و حفاظت آن، هزینه‌های سنگینی را در پی دارد. تشخیص و شناسایی عیب، یک مسئله بسیار مهم اقتصادی و ایمنی در توربین‌های گازی است و با شناسایی و تشخیص به‌موقع عیب و رفع آن قبل از رخداد خرابی، می‌توان از خسارات و هزینه‌های بالای آن جلوگیری نمود.

در سامانه‌های پایش توربین‌های گازی، حجم عظیمی از داده‌های ارزشمند توسط حس‌گرهای نصب شده بر روی توربین تولید و ثبت می‌شوند و بدون آنکه پردازش شوند بلا استفاده رها می‌شوند. درحالی‌که با توسعه ابزارها و تکنیک‌های داده‌کاوی می‌توان الگوها و روابط پنهان در لابلای داده‌ها را استخراج و در جهت شناسایی به موقع عیب و پیشگیری از خرابی و خاموشی توربین، همچنین اصلاح و ارتقا طراحی توربین‌های جدید اقدام نمود. مسئله اصلی در این پژوهش بررسی اثربخشی کاربرد تکنیک‌های داده‌کاوی و تلفیق آن‌ها جهت پردازش داده‌های حجیم و بررسی امکان‌پذیریِ پیش‌بینی عیوب توربین گاز قبل از رخداد خرابی می‌باشد. کاربرد فناوری داده‌کاوی می‌تواند منجر به یافتن الگوهای پنهان در داده‌های حجیم توربین گاز و استخراج ویژگی‌های[6] مفید، جهت شناسایی عیوب و سپس تفکیک و طبقه‌بندی آن‌ها شود. پایش وضعیت توربین به منظور نظارت بر عملکرد توربین گاز و پیشگیری از خرابی‌های ناگهانی در برخی از نیروگاه‌ها استفاده می‌شود. بر اساس گزارش شرکت آمریکایی بنتل[7]، هزینه‌های افت عملکرد و افزایش سوخت مصرفی، ناشی از رسوب‌گرفتی و کثیفی پره‌های کمپرسور، به ازای هر واحد توربین گازی V94.2 در هر سال، حدود ۳.۳۷ میلیون دلار می‌شود که با اجرای آنالیز عملکردی توربین و شستشوی کمپرسور[8]، این هزینه قابل بازگشت می‌باشد (راهمون[9] و همکاران، ۲۰۲۱).

پس رویکرد پژوهش حاضر استفاده از تکنیک‌های داده‌کاوی در حوزه سامانه‌های پایش توربین‌های گازی می‌باشد. حوزه پایش توربین‌های گازی و نیروگاهی، حوزه‌ای نو در صنعت نیروگاهی جهان می‌باشد که هنوز ابتدای راه پژوهشی خود بوده و نیاز به بهبود و توسعه ابزارهای پایش و بیشتر از منظر نرم‌افزاری می‌باشد. در سامانه‌های پایش توربین‌های گازی داده‌های حجیمی توسط حس‌گرهای نصب شده بر روی توربین تولید و

1 Gas turbine
2 Control System (CS)
3 Condition Monitoring (CM)
4 Fault
5 Failure
6 Feature
7 Bechtel
8 Compressor Washing
9 Rahmoune

بدون پردازش ثبت می‌شوند. با توسعه ابزارها و تکنیک‌های داده‌کاوی می‌توان الگوها و روابط پنهان در لابلای داده‌ها را استخراج و در جهت شناسایی به موقع عیب و پیشگیری از خرابی و خاموشی توربین اقدام نمود.

پس در واقع هدف کاربردی طرح، استفاده از داده‌کاوی در شناسایی خطاهای حس‌گرها و تجهیزات توربین گاز جهت افزایش کارایی و بهبود عملکرد آن می‌باشد. ضرورت طرح در این است که افت کارایی توربین‌های گازی منجر به ایجاد هزینه‌های سنگین می‌شود و از طرفی حجم زیاد داده‌های سیستم کنترل و پایش وضعیت، ظرفیت پیاده‌سازی و به‌کارگیری روش‌های داده‌کاوی را فراهم می‌سازد. توربین‌های گازی همواره مواجه با توقف‌هایی هستند که ناشی از عدم امکان پیش‌بینی خطای حس‌گر و تجهیزات توربین و در نتیجه خرابی توربین می‌باشد که این اتفاقات منجر به کاهش کارایی شده و بار هزینه‌های تولید را افزایش می‌دهد. این موضوع به‌وضوح ضرورت انجام چنین پژوهش‌هایی را نشان می‌دهد.

این پژوهش در چهار مرحله کلی انجام می‌پذیرد. در مرحله اول، سیگنال‌های مهم، توسط حس‌گرهای نصب شده بر روی توربین ثبت و تمیز می‌شوند. در مرحله دوم، بعد داده‌ها کم می‌شود. کاهش بعد هم بر اساس دانش تجربی هم با کمک تکنیک داده‌کاوی تحلیل مؤلفه‌های اصلی[1] (عبدی[2] و همکاران، ۲۰۱۰) انجام انجام می‌پذیرد. در این مرحله با کاربرد «تحلیل مؤلفه‌های اصلی»، ویژگی‌های داده‌ها نیز استخراج می‌شود. در این روش محورهای مختصات جدیدی برای داده‌ها تعریف شده و داده‌ها بر اساس این محورهای مختصات جدید بیان می‌شوند و در نهایت در مرحله آخر با کمک شبکه عصبی عیوب دسته‌بندی‌شده و کلاس عیب پیش‌آمده در توربین گاز تعیین می‌شود. پس درواقع روش پیشنهادی مورد استفاده در این پژوهش به طور خلاصه در زیر بیان شده است:

۱) تهیه پایگاه داده‌های پیش‌پردازش شده: مقادیر حس‌گرهای اندازه‌گیری و ثبت شده به صورت ماتریس اولیه داده‌ها درمی‌آید که برای شناسایی و تشخیص خطا تحت عملیات داده‌کاوی قرار می‌گیرند پیش‌پردازش شده و به صورت فایل متلب ثبت و مورد استفاده پروژه قرار می‌گیرد.

۲) کاهش ابعاد داده‌ها و استخراج ویژگی‌ها: با کمک تحلیل مؤلفه‌های اصلی، ابعاد داده‌ها کاهش‌یافته و سپس ویژگی‌های اصلی استخراج می‌شوند و متعاقباً محورهای مختصات جدیدی برای داده‌ها تعریف می‌شود.

۳) شناسایی عیوب با استفاده از ویژگی‌ها و بر اساس شبکه عصبی مصنوعی: از شبکه عصبی جهت طبقه‌بندی و تفکیک و شناسایی عیوب استفاده می‌شود.

سپس با ارزیابی روش پیشنهادی بر روی داده‌های واقعی، کیفیت روش پیشنهادی مورد ارزیابی قرار می‌گیرد. مجموعه داده‌های واقعی تحقیق حاضر از سنسورهای نصب شده بر روی توربین گاز فریم ۹ استخراج شده‌است. ابتدا برچسب‌گذاری ماتریسِ داده‌ها صورت گرفت تا داده‌های سالم و معیوب در هر

[1] Principal Component Analysis (PCA)
[2] Abdi

نمونه داده‌ای مشخص شود. سپس با بهره‌گیری از تکنیک مؤلفه‌های اصلی، ابعاد ماتریس داده از هفت به چهار بعد کاهش داده شد و ویژگی‌های اصلی آن استخراج گردید. در ادامه و با اجرای تکنیک شبکه عصبی مصنوعی، مدلی ایجاد می شوذ که با کمک تکنیک‌های یادگیری ماشین می‌تواند نمونه‌های عیب در ماتریس داده‌ها را شناسایی و حالتِ (کلاس) نمونه داده‌ها از نظر سالم یا معیوب بودن را تشخیص بدهد. پس از یادگیری ماشین که از روی ۸۰٪ داده‌ها انجام گرفت دقت، صحت و همگرایی مدل برای کاهش‌های ابعادی مدل به دو تا شش بُعد بررسی و مشخص گردید که علاوه بر اینکه مدل به‌خوبی کار می‌کند بهترین کاهش ابعاد برای ماتریس داده‌های توربین گاز را که ۴ بود را نیز تعیین نمود. در نهایت مدلی به دست آمد که با دقتی بیش از ۹۰٪ و با صحت بالا و بیشترین میزان همگرایی ماتریس داده‌ها را ارزیابی و کلاس عیب توربین را مشخص نماید.

در ادامه در ابتدا مساله مورد نظر به دقت بیان می شود و فرضیه‌های موجود در این بین معرفی می‌شود، سپس کارهای پیشین در این حوزه بررسی می‌شود. در بخش بعدی روش پیشنهادی ارائه می‌شود و ارزیابی آن نیز مشخص می‌گردد و در انتها نتیجه‌گیری و کارهای آتی در این حوزه بیان می‌شود.

بیان مسأله

شناسایی و تشخیص عیب، مهم‌ترین اصل در پایش وضعیت توربین‌های گازی است که در سال‌های اخیر توجه زیادی را به خود جلب کرده‌است (راهمون و همکاران، ۲۰۲۱). توربین‌های گازی نقشی بسیار مهم و حیاتی در صنایع نفت، گاز، پتروشیمی و نیرو ایفا می‌کنند. مدیریت رفتارهای دینامیکی و پیچیده توربین‌های گازی، نیازمند سامانه‌های کنترل[1] حفاظت و پایش وضعیت[2] توربین گاز با قابلیت اطمینان، انعطاف برنامه‌ریزی و دسترسی‌پذیری بالاست. از آنجا که هرگونه عیب در توربین گاز و در عملیات و سامانه‌های کنترل و حفاظت آن، عواقب سنگینی را در پی دارد. تشخیص و شناسایی عیب، یک مسئله بسیار مهم اقتصادی و ایمنی در توربین‌های گازی است و با شناسایی و تشخیص به‌موقع عیب و رفع آن قبل از رخداد خرابی از خسارات و هزینه‌های بالای آن جلوگیری نمود (لی[3] و همکارن، ۲۰۱۹).

وقتی پایش یک توربین گاز به اجرا می‌رسد حجم زیادی از داده‌های اندازه‌گیری شده وجود دارند که با سرعت (فرکانس) زیادی تولید و در بانک اطلاعاتی سیستم کنترل ذخیره می‌شوند. انجام عملیات هم‌زمان تشخیص و شناسایی عیب با تکیه بر کار فقط اپراتور امری کاملاً غیرممکن است. روش‌های گذشته مبتنی بر مدل‌سازی توربین و کشف عیب، به دلیل حجم بالای داده‌هایی که از ثبت سیگنال حس‌گرهای متعدد نصب شده بر روی توربین گاز تولید می‌شوند عملاً غیرممکن می‌باشد. با توسعه فناوری‌های جدید نظیر داده‌کاوی، ضرورت کاربرد چنین فناوری‌هایی در صنعت پایش توربین گازی پیش از پیش احساس می‌شود.

[1] Control System (CS)
[2] Condition Monitoring (CM)
[3] Lee

در عمل تلاش جهت کاربرد فناوری‌های جدید اطلاعاتی، منجر به توسعه سیستم‌های پایش وضعیت توربین مبتنی بر فناوری داده‌کاوی شده‌است (ملانی[1] و همکاران، ۲۰۲۱). مزایایی که با کاربرد فناوری داده‌کاوی در پردازش داده‌های حس‌گرهای توربین کسب می‌شود یافتن الگوهای پنهان در داده‌های حجیم و استخراج ویژگی‌های[2] مفید، شناسایی عیوب و سپس تفکیک و طبقه‌بندی آن‌ها می‌باشد (آن[3] و همکاران، ۲۰۲۲). مسئله اصلی در این پروژه یافتن راه‌حل تکنیک‌های مفید داده‌کاوی و تلفیق آن‌ها جهت پردازش داده‌های حجیم می‌باشد. کاربرد نتایج پژوهش حاضر مستقیماً در صنعت نیروگاهی مورد استفاده قرار می‌گیرد. لذا لازم است در ابتدا به صورت اجمالی تعاریف و توضیحاتی از صنعت نیروگاهی و توربین گاز ارائه شود تا اهمیت نتایج مالی ملموس‌تر شود.

توربین گاز یک موتور درون‌سوز از نوع ماشین‌های دوار است که بر اساس انرژی گازهای ناشی از احتراق کار می‌کند. هر توربین گاز شامل یک کمپرسور برای فشرده کردن هوا، یک محفظه احتراق برای مخلوط کردن هوا با سوخت و محترق‌کردن آن و یک توربین برای تبدیل انرژی درونی گازهای داغ و فشرده به انرژی مکانیکی است. بخشی از انرژی مکانیکی تولیدشده در توربین، صرف چرخاندن کمپرسور خود توربین شده و باقی انرژی، بسته به کاربرد توربین گاز، ممکن است مولد الکتریکی را بچرخاند (توربوژنراتور). نیروگاه گازی[4] با استفاده از توربین‌های گازی و بخاری، ژنراتورهایی را به گردش در می‌آورد که تولید برق می‌کنند. پایش وضعیت توربین به منظور نظارت بر عملکرد توربین گاز و پیشگیری از خرابی‌های ناگهانی در برخی از نیروگاه‌ها استفاده می‌شود. بر اساس گزارش شرکت آمریکایی بنتل افت عملکرد و سوخت مصرفی، بیشتر ناشی از رسوب‌گرفتی کمپرسور، به ازای هر واحد توربین گازی V94.2 در هر سال، حدود ۳.۳۷ میلیون دلار می‌شود که با اجرای آنالیز عملکردی توربین و شستشوی کمپرسور[5]، این هزینه قابل بازگشت می‌باشد (ربانی[6]، ۲۰۱۶). به‌کارگیری روش پایش وضعیت توربین هزینه‌های تعمیرات توربین را تا ۲۵٪ کاهش می‌دهد که این میزان بدون در نظر گرفتن صرفه‌جویی‌های ناشی از پیشگیری از کاهش تولید می‌باشد. این مقادیر نشان از اهمیت پروژه پایش وضعیت توربین دارد.

شاید یکی از قدیمی‌ترین روش‌های تشخیص خطا، تهیه یک مدل بسیار دقیق از فرآیند و واحدهای کنترل آن و مقایسه پارامترهای اندازه‌گیری شده توسط حس‌گرها با پارامترهای مشابه در مجموعه مدل‌سازی و شبیه‌سازی‌شده و تهیه و استخراج اختلاف آن‌ها[7] می‌باشد. این مقدار اختلاف، به‌عنوان مشخصه‌ای جهت آشکارسازی خطا در نظر گرفته می‌شود. همان‌گونه که مشخص است ابتدا باید خطا در یک سیستم آشکار شود و سپس عملیات تحلیل و تشخیص آن به اجرا برسد. خطا در واقع به عنوان یک تخطی

[1] Melani
[2] Feature
[3] An
[4] Gas power plant
[5] Compressor Washing
[6] Rabbani
[7] Residual

غیرمجاز در حداقل یکی از مشخصه‌ها و پارامترهای سیستم از مقدار قابل قبول/ متعارف و استاندارد آن ارائه تعریف می‌شود. به دلیل حجم بالای آلارم‌ها و داده‌های تولیدی توسط حس‌گرهای سیستم کنترل توربین گاز، عملاً امکان ارزیابی و جداسازی خطا از عیب توسط اپراتور یا سیستم کنترل، غیر ممکن بوده و لذا هیچ‌گونه تحلیلی از زمان وقوع یک خرابی و روش‌های تعویض قطعات پیش از یک تعمیر اساسی توربین با فرایندهای مرسوم وجود ندارد. به همین دلیل در پایش وضعیت توربین هم به صورت زمان حقیقی هم آفلاین، کاربرد تکنیک‌های داده‌کاوی نظیر «تحلیل مؤلفه‌های اصلی» اجتناب‌ناپذیر است. توربین گاز را می‌توان به عنوان یک ماشین پیچیده و بسیار گران‌قیمت در نظر گرفت. به عنوان مثال، از نظر پیچیدگی، تعداد قطعات موجود در یک توربین گاز بالغ بر بیست هزار قطعه می‌باشد، قیمت توربین‌های گازی نیز در حدود چندین میلیون دلار است. چنین ماشین گران‌قیمتی می‌تواند به راحتی بر اثر به وجود آمدن یک عیب در قطعات آن از بین برود. با وجود همه مزایای بالای به‌کارگیری توربین‌های گازی، استفاده از آن‌ها در دما و گاهی دور بالا سبب پیدایش مشکلات بسیاری شده و هزینه تعمیرات زیادی را ایجاد می‌کند. فقط هزینه تعمیرات نیروگاهی ایران حدود ۲۵۰۰ میلیارد تومان می‌باشد که رقم بسیار بالایی است. علاوه بر هزینه‌های تعمیراتی هزینه‌ها و صدمات خاموشی‌های پیش‌بینی‌نشده نیز بسیار زیاد می‌باشد. به دلیل هزینه‌های بسیار بالای تعمیرات توربین‌های گازی و هزینه‌های سنگین خاموشی‌های پیش‌بینی‌نشده، ضرورت بکار گیری تکنیک‌های جدید تعمیراتی مبتنی بر فناوری اطلاعاتی و داده‌پردازی احساس می‌شود.

پیشینه پژوهش

موضوع شناسایی و طبقه‌بندی عیوب توربین گاز با تکنیک داده‌کاوی در کارهای پیشین مورد تحقیق قرار گرفته‌است (ملانی و همکاران، ۲۰۲۱)(ژای[1] و همکاران،۲۰۲۳). برای نمونه در (ربانی،۲۰۱۶) و در توربین-های گازی ساخت داخل V94.2قابلیت تشخیص و تفکیک دو کلاس خطا در این واحد گازی با کمک تکنیک استخراجِ ویژگی‌های مبتنی بر «تحلیل مؤلفه‌های اصلی» مورد ارزیابی قرار گرفته‌است.

در (خرم کشکولی[2]،۲۰۱۷) با کمک الگوریتمk-means ابعاد داده‌های اولیه کم شده و سپس با پیاده‌سازی «تحلیل مؤلفه‌های اصلی» دانشی که درون داده‌های «وضعیت نرمال توربین» پنهان بوده استخراج و با استفاده از آن به تشخیص و شناسایی عیب توربین گاز پرداخته شود.

در (توماج[3]،۲۰۱۲) از تکنیک «تحلیل مؤلفه‌های اصلی» جهت کاهش ابعاد ماتریس استفاده شده و از تکنیک «تحلیل مؤلفه‌های اصلیِ» ماتریس داده‌ها، جهت طبقه‌بندی وضعیت و عیوب موتور درونسوز استفاده شده-است. نتایج نشان می‌دهد که روش پیشنهادی موثر است.

[1] Xie
[2] Khorram Kashkooli
[3] Tomaj

در (لی و همکاران، ۲۰۱۹) توسعه یک مدل عیب‌یابی بر پایه ویژگی سیگنال‌های (داده‌های) ارتعاشی برای تشخیص عیوب محتمل در توربین گاز مورد توجه قرار گرفته است در این مقاله مدل یک شبکه عصبی در شرایط کاری متفاوت، تحت آزمایش قرار گرفته‌است .نتایج تحقیق نشان داده که روش «شبکه عصبی مصنوعی» توانسته توان خروجی توربین گاز را با دقت خوبی پیش‌بینی نماید. در این تحقیق از تکنیک «یادگیری ماشین» جهت طبقه‌بندی عیوب بر پایه ویژگی‌های سیگنال‌های (داده‌ها) ارتعاشی آن‌ها استفاده‌شده است.

در (شاکیا[1]،۲۰۲۱) تکنیک «تحلیل مؤلفه‌های اصلی» برای شناسایی عیوب با کمترین میزان خطا مورد ارزیابی قرار گرفته است. در (سویرز و مروسکوسکا[2]، ۲۰۲۰) کاربرد «تحلیل مؤلفه‌های اصلی» در تشخیص عیبِ نشتی بویلر در تأسیسات نفت خام فورت‌مک‌مری کانادا بررسی شده و مزیت این تکنیک در تشخیص عیب به همراه کاهش آلارم‌های اشتباه مورد تاکید قرار گرفته است.

در (هدروگ[3] و همکاران، ۲۰۲۲) به بررسی استراتژی تشخیص عیب فازی برای تشخیص اشکال در ارتعاشات توربین گاز می‌پردازد. هدف اصلی این مقاله، تعیین رابطه بین علائم و اشکال مربوطه در ارتعاشات توربین گاز است. در راستای این هدف، از روشهای فازی برای تشخیص اشکال در ارتعاشات توربین گاز استفاده می‌کند. رویکرد فازی به عنوان یک روش مبتنی بر قوانین و قاعده‌ها، امکان تطبیق با پیچیدگی‌ها و عدم قطعیت‌های مرتبط با توربین گاز را فراهم می‌کند.

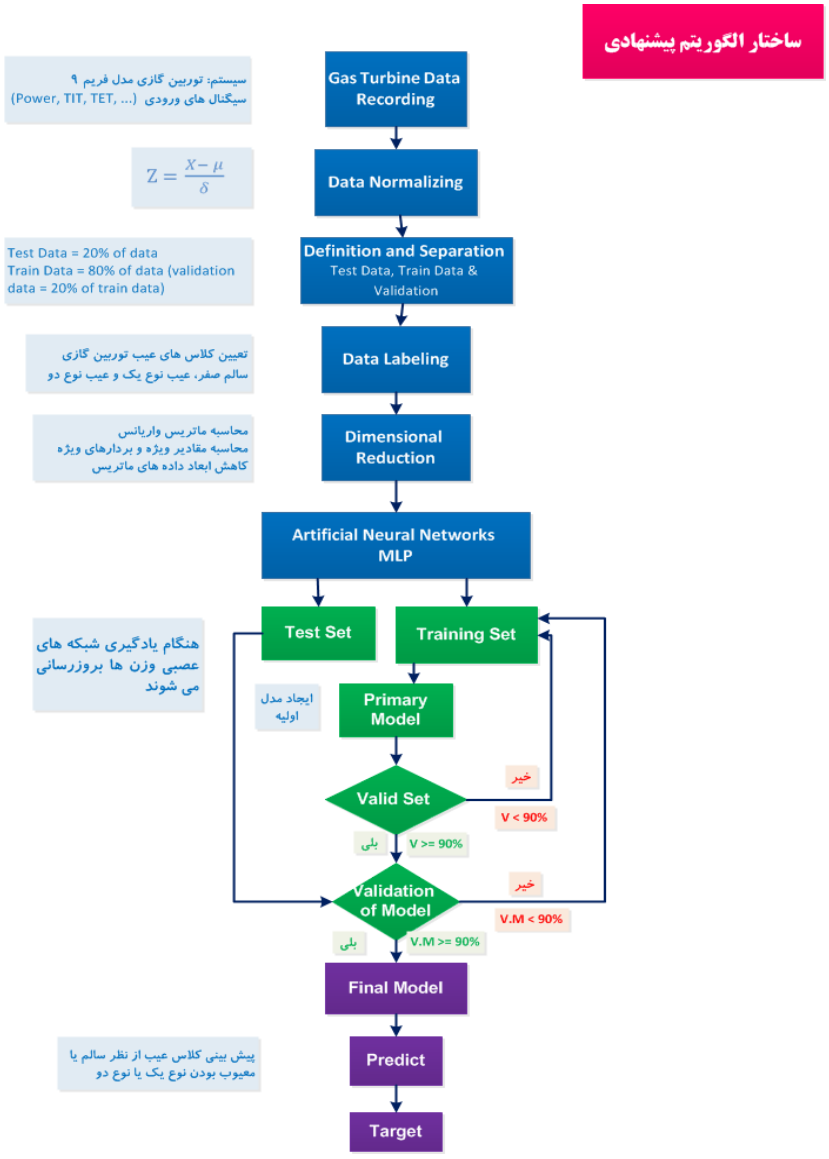
در (یو[4] و همکاران، ۲۰۲۲)، یک روش نوین برای تشخیص عیب در موتور با استفاده از تحلیل مؤلفه اصلی (PCA)و سیستم قاعده باور گسترده (DBRB) ارائه شده است PCA .یک تکنیک آماری است که معمولاً برای کاهش ابعاد و استخراج ویژگی‌ها استفاده می‌شود. این تکنیک در شناسایی اجزا یا ویژگی‌های مهم‌تر که در تغییرات داده‌ها نقش دارند، کمک می‌کند. نویسندگان پیشنهاد داده‌اند که PCA را با سیستم DBRB که یک رویکرد مبتنی بر قوانین برای مدیریت ابهام و نامعینی در فرآیند تشخیص استفاده می‌شود، ترکیب کنند. سیستم DBRB از توابع باور برای نمایش و انتشار عدم قطعیت در فرآیند تشخیص استفاده می‌کند. با ترکیب PCAو سیستمDBRB ، نویسندگان هدف خود را بهبود دقت و قابلیت اعتماد در تشخیص عیب در موتور قرار داده‌اند.

روش پژوهش و روش پیشنهادی

در این بخش به معرفی روش پیشنهادی پرداخته می‌شود. در ابتدا چارچوب کلی روش ارائه شده به صورت کامل شرح داده و سپس اجزای آن به صورت زیر توضیح داده می‌شود. این چارچوب را می توان در شکل ۱ مشاهده کرد. پیاده‌سازی الگوریتم ترکیبی «تحلیل مؤلفه‌های اصلی» و «شبکه عصبی مصنوعی» شکل (۱)

[1] Shakya
[2] Swiercz & Mroczkowska
[3] Hadroug
[4] Yu

توسط زبان برنامه‌نویسی پایتون با ورژن 3.7 انجام شده‌است. زبان برنامه‌نویسی پایتون شامل کتابخانه‌های قدرتمند جهت انجام عملیات مختلف داده‌کاوی و یادگیری ماشین می‌باشد. ماژول قدرتمند Tensorflow به‌کاررفته در این زبان برنامه‌نویسی، برای پیاده‌سازی شبکه عصبی مصنوعی می‌باشد. چارچوب Keras از توابع اصلی و بنیادین TnsorFlow استفاده می‌کند.



شکل ۱. ساختار الگوریتم پیشنهادی ارائه‌شده در این پژوهش

همان‌طور که در شکل دیده می‌شود در ابتدا باید با روش داده‌کاوی در پایش وضعیت توربین سه مرحله انجام پذیرد که در مرحله اول، سیگنال‌های مهم (مقادیر حس‌گرها) تعیین و ثبت می‌شوند. در مرحله دوم، بعد داده‌ها کم می‌شود. کاهش بعد را می‌توان بر اساس دانش تجربی و یا به صورت دلخواه تعیین نمود. مرحله سوم، تشخیص و شناسایی عیب بر اساس تکنیک‌های داده‌کاوی انجام می‌گیرد. در صورت تشخیص وقوع عیب در مرحله سوم (با توجه به نمونه‌های واردشده به الگوریتم)، مرحله چهارم وارد عمل شده و کلاس عیب پیش‌آمده را تعیین می‌کند یا به عبارتی دیگر نوع عیب و خرابی توربین را تشخیص می‌دهد.

پس بخش اول ساختار اصلی الگوریتم که در شکل آمده است عملیات مربوط به تحلیل مؤلفه‌های اصلی می‌باشد. در این بخش بردار ورودی، داده‌های اندازه‌گیری شده از هفت حس‌گر نصب‌شده بر روی توربین گاز می‌باشد که ابتدا بر روی‌داده‌ها، عمل پیش‌پردازش و نرمالیزه کردن انجام گرفته و سپس عملیات تحلیل مؤلفه‌های اصلی که یکی از تکنیک‌های پرکاربرد داده‌کاوی می‌باشد بر روی داده‌های نرمالیزه شده انجام می شود؛ که طی عملیات تحلیل مؤلفه‌های اصلی، ابعاد داده‌ها از هفت بعد به چهار بعد کاهش پیدا کرد.

تحلیل داده‌های چندگانه از نقش اساسی در تحلیل اطلاعات برخوردار است. مجموعه داده‌های چندگانه، حالت‌ها یا متغیرهای زیادی را برای هر مشاهده در بردارند. اگر در هر مجموعه داده n متغیر وجود داشته باشد، هر متغیر می‌تواند دارای چند بعد باشد. با توجه به اینکه اغلب درک و شهود فضای چندبعدی دشوار است، روش تحلیل مؤلفه‌های اصلی PCA ابعاد کلیه داده‌ها را بر اساس ویژگی‌های مشابه کاهش می‌دهد. پس یکی از کاربردهای اصلیِ تحلیل مؤلفه‌های اصلیPCA عملیاتِ کاهشِ بعد[1] است (آن و همکاران، ۲۰۲۲).

تحلیل مؤلفه‌های اصلی علاوه بر کاهش ابعاد داده‌ها می‌تواند ویژگی‌های اصلی را که نیز ارزش بیشتری دارند شناسایی و معرفی نماید. در مرحله بعدی می‌بایست به تشخیص و تعیین کلاس عیب و جداسازی آن با استفاده از شبکه عصبی مصنوعی پرداخت. در این مرحله عیب‌های واقع شده بایستی تشخیص و جداسازی شوند که این امر در پژوهش حاضر با کمک تکنیک شبکه عصبی مصنوعی انجام گرفته است. کلاس‌های عیب بر اساس دانش تجربی تعریف می‌شوند. بعد از انجام عملیات مؤلفه‌های اصلی بر روی داده‌های پژوهش حاضر، بردار ورودی داده‌های اندازه‌گیری شده از هفت حس‌گر نصب شده بر روی توربین گاز، از هفت بعد به چهار بعد کاهش پیدا کرد. داده‌ها متشکل از سه بخش مهم داده آموزشی، داده تست و داده ارزیابی می-باشد.

داده‌های آموزشی به شبکه‌های عصبی تزریق می‌شوند. این الگوریتم، وظیفه‌ی یادگیری و واکشیِ الگوهای مختلف، در داده‌ها را دارند. بعد از به دست آوردنِ الگو توسط الگوریتم شبکه عصبی، مدل ساخته می‌شود. این مدل دوباره تنظیم می‌شود تا ایرادات قبلی رفع شود. بخش نهایی الگوریتم مدل ساخته شده در حافظه ذخیره می‌شود. بعد از ذخیره‌ی مدل، با استفاده از داده‌های جدید سیستم توانایی پیش‌بینیِ رفتار را دارد. در ادامه هر بخش آن به صورت جزئی‌تر بیان می‌شود:

برچسب‌گذاری

به هر یک از این مجموعه‌ی داده‌ها یک ستون برچسب[2] اضافه کردیم که امکان تجمیع سه مجموعه‌ی داده در یک ماتریس باشد تا عملیات طبقه‌بندی را آغاز کنیم. (جدول ۱) ویژگی تکنیک طبقه‌بندی، وجود برچسب در داده است. نوع بر چسب در این مجموعه داده‌ها سالم یا معیوب بودن می‌باشد. برچسب شماره

[1] Dimensionality Reduction
[2] label

صفر به معنای داده‌های حالت سالم، برچسب شماره ۱ داده‌های حالت معیوب نوع اول و برچسب شماره ۲ حالت معیوب نوع دوم می‌باشد.

جدول ۱. ماتریس ترکیبی از مجموعه داده‌ها به همراه افزودن ستون کلاس

Index	0	1	2	3	4	5	6	labeles
0	-0.01542	0.050435	-0.05644	-0.19689	-0.43266	-0.0047	-0.01227	0
1	-0.01553	0.050341	-0.05642	-0.19986	-0.43494	-0.00471	-0.0123	0
2	-0.01545	0.050237	-0.05633	-0.19853	-0.43311	-0.0047	-0.01226	0
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
300000	-0.01499	0.04221	-0.05663	-0.21263	-0.42032	-0.00457	-0.01229	1
300001	-0.01494	0.042115	-0.05656	-0.21214	-0.41924	-0.00456	-0.01227	1
300002	-0.01509	0.042029	-0.05654	-0.21543	-0.42203	-0.00458	-0.0123	1
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
777599	-0.14053	0.080291	-0.11187	-2.48751	-2.50832	-0.02168	-0.04237	2
777600	-0.14053	0.080355	-0.11191	-2.48719	-2.50831	-0.02168	-0.04237	2

اجرای تکنیک تحلیل مؤلفه‌های اصلی PCA

هدف اصلی از اجرای تکنیک «تحلیل مؤلفه‌های اصلی» شناسایی همبستگی بین متغیرها و در نتیجه کشف الگوهای پنهان می‌باشد. در تحقیق حاضر پس از برچسب‌گذاری، داده‌ها را به دو گروه آموزشِ داده، با حجم ۸۰٪ داده‌ها و آزمون مدل، با حجم ۲۰٪ داده‌ها تفکیک کرده‌ایم. بعد از عملیات تفکیک، تکنیک PCA را روی هر دو گروه داده‌ها اعمال شد تا با شناسایی الگوهای پنهان ناشی از همبستگی، ابعاد ماتریس کاهش پیدا کند. جهت کسب بهترین کاهش ابعاد، آزمایشات متعددی صورت گرفت که طی این آزمایشات، ابعاد ماتریس در دو، سه، چهار، ... هفت بعد کاهش یافت تا ماتریس بهینه کاهش ابعاد مشخص گردد.

اجرای تکنیک شبکه عصبی مصنوعی

در تحقیق حاضر جهت طبقه‌بندی کلاسِ عیب از الگوریتم «شبکه عصبی مصنوعی» که یکی از تکنیک‌های پرکاربرد الگوریتم یادگیری ماشین محسوب می‌شود استفاده شد (دی جورجی[1]، ۲۰۲۲). طبقه‌بندی داده‌ها نتیجه یادگیری ماشین از برچسب‌گذاری نمونه‌ها توسط ناظر می‌باشد. در تحقیق حاضر، ماشین بر اساس برچسب‌هایی که به حالت سالم (برچسب ۰)، حالت معیوب یک (برچسب ۱)، حالت معیوب دو (برچسب ۲) بر روی داده‌هایی که توسط محقق به عنوان ناظر داده است برچسب‌گذاری را یاد می‌گیرد. جهت طبقه-بندی داده‌ها یک ستون به ماتریس اضافه شد تا حالت توربین (کلاس داده‌ها) را از نظر سالم و معیوب بودن مشخص نماید. هر دو نوع عیب، از دسته عیوب فولینگ[2] یا کثیفی یا جرم‌گرفتگی پره‌های کمپرسور می‌باشد

[1] De Giorgi
[2] Fouling

که عیب نوع یک "جرم‌گرفتگی با شدت کمتر" و عیب نوع دو "جرم‌گرفتگی با شدت بیشتر" می‌باشد که در حالت دوم علاوه بر کاهش دبی، موجب کاهش شدید بازده توربین می‌شود. با شناسایی به موقع عیب ۱ می‌توان از بروز عیب ۲ و تحمیل هزینه‌های سنگین‌تر تعمیر جلوگیری نمود. در الگوریتم تحقیق حاضر داده‌ها به صورت ۸۰ به ۲۰ تقسیم‌بندی شدند. مدل با کمک تکنیک «شبکه عصبی مصنوعی» از ۸۰ درصد داده‌های برچسب‌گذاری شده یاد می‌گیرد و برای ۲۰ درصد باقیمانده، عیب را پیش‌بینی می‌کند. درصد دقت و صحت پیش‌بینی مدل بررسی و در نتیجه میزان دقت مدل محاسبه می‌شود. اگر دقت مدل قابل قبول بود با همین مدل می‌توانیم عیوب را به موقع شناسایی و به موقع رفع عیب کنیم.

ایجاد مدل

ابتدا مدل شبکه عصبی از نوع پرسپترون چندلایه[1] ایجاد می‌کنیم. شبکه‌های «چند لایه‌ی پیشخور[2]» یکی از مهم‌ترین ساختارهای شبکه‌های عصبی مصنوعی می‌باشد (دی جورجی، ۲۰۲۲). به طور معمول این شبکه‌ها شامل مجموعه‌ای از واحدهای حسی (نرون‌های پایه) می‌باشند که از یک لایه‌ی خروجی، یک یا چند لایه‌ی پنهان تشکیل شده‌اند. سیگنال ورودی در خلال شبکه و در مسیری رو به جلو به صورت لایه به لایه منتشر می‌شود. ساختار الگوریتم این تحقیق به این شکل می‌باشد که لایه اول، ورودی‌ها یا همان ۴ ویژگی‌های ماتریسی هستند که توسط تکنیک تحلیل مؤلفه‌های اصلی PCA پس از کاهش ابعاد به دست آمده است. این مقادیر به عنوان ورودی‌ها، به شبکه عصبی وارد می‌شود. هر ورودی دارای یک مقدار و یک وزن مربوطه می باشد. بعد از ایجاد لایه اول به خاطر خاصیت شبکه چندلایه MLP یک لایه پنهانی به شبکه اضافه می‌کنیم. ثابت شده‌است که هر تابع می‌تواند حداکثر با سه لایه پنهان تقریب زده‌شده. اندازه‌ی لایه‌ی پنهان عموماً به صورت تجربی به دست می‌آید. در تحقیق حاضر یک لایه پنهانی ایجاد شد تا دقت مدل افزایش پیدا کند و خطا به کمترین میزان خود برسد. در آزمایش‌های مختلفی که صورت گرفت با تعداد ۱۰۰ نرون به نتیجه مطلوب رسیدیم. مرحله بعد ایجاد لایه خروجی می‌باشد. در این لایه تعداد حالت (کلاس) عیب مشخص می شود. همان‌طور که در بخش‌های قبل اشاره شده سیستم به سه حالت سالم، معیوب شماره ۱ و معیوب شماره ۲ تقسیم شده‌است که اندازه نرون در لایه خروجی عدد ۳ می‌باشد. بعد از ایجاد مدل مرحله کامپایل کردن مدل یا همان اجرای مدل بر اساس پارامترهای تعریف شده می‌باشد.

معرفی محدوده مورد مطالعه

مجموعه داده‌ها تحقیق حاضر از حس‌گرهای نصب شده بر روی توربین گاز فریم ۹ استخراج شده‌است. جمع‌آوری و ثبت داده‌ها از یک سامانه کنترلی نصب شده بر روی توربین گازی صورت می‌گیرد. در سامانه کنترلی، سیگنالِ حس‌گرهای نصب شده در قسمت‌های مختلف توربین، به یک یا چند کارت شبکه صادره

[1] Multilayer Perceptron (MLP)
[2] Feed forward Multilayer

می‌شوند. کارت‌های شبکه نیز توسط کابل‌های هم‌محور به صورت موازی به یک پنل سامانه کنترلی وصل بوده که توسط ابزارهای داده‌گیری[1] و نرم‌افزاری به صورت داده‌های فایلی به یک سامانه رابط انسان – ماشین[2] HMI ارسال می‌شوند.

داده‌ها شامل چهار مجموعه داده می‌باشد که در بردارنده‌ی سه حالت مختلف توربین گاز شامل حالت سالم و عیب‌های سیستمی با نام‌هایdata_s1 data_s2, data_s6و همچنین ترند دمایی با نام data_trend_tem است. اولین مجموعه‌ی داده، تحت عنوان data_s1، حالت سالم توربین گاز فریم ۹ را شامل می‌شود. هر یک از مجموعه‌ی داده‌ها به صورت نرمال شده با فرمت mat. بوده و شامل ۷ ستون می‌باشد. هر یک از ستون‌ها مطابق جدول (۲) داده‌های اندازه‌گیری شده توسط یک سنسور خاص را نمایش می‌دهند.

جدول ۲. نمونه جدول داده‌های نرمال‌سازی شده

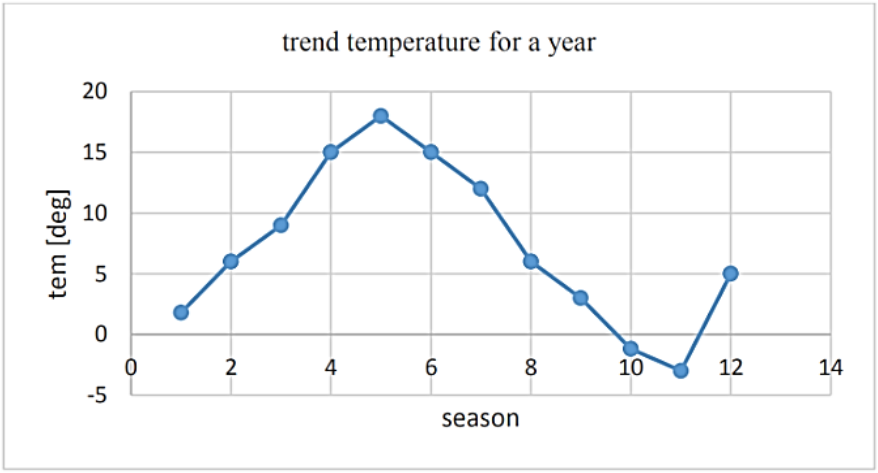
سنسور	ستون	داده اندازه گیری شده توربین فریم ۹
توان	Column 0	data_si i =1,…,4
دبی جرمی ورودی به کمپرسور	Column 1	
دمای خروجی کمپرسور	Column 2	
فشار خروجی کمپرسور	Column 3	
دبی سوخت	Column 4	
دمای ورودی توربین TIT	Column 5	
دمای خروجی اگزوز TET	Column 6	

در ماتریس داده‌ها، جدول (۲) هر سطر نمایشگرِ هفت داده اندازه‌گیری شده از هفت سنسور در یک زمان می‌باشد. گام اصلی ارزیابی عملکرد توربین گاز، تهیه یک بانک داده کامل و دقیق در تمامی ساعات کارکرد و نیز کارنکرد (خاموشی) توربین می‌باشد. همچنین توسعه مدل مربوط به توربین نیز از دیگر ضروریات است. ابتدا بایستی داده پالایش و تحلیل شوند. ماحصل تحلیل‌ها رسم نمودارهای مختلف پارامترهای عملکردی و نیز پارامترهای مربوط به فرآیند کار توربین نظیر تعداد استارت‌های موفق، ناموفق، پرج و غیره می‌باشد. همچنین فیلتر داده‌ها، رسم چندین نمودار در یک پلات برای مقایسه بهتر داده‌ها و نمایش مقدار مینیمم و ماکزیمم هر پارامتر در هر بازه زمانی دلخواه از دیگر مواردی است که می‌توان به آن‌ها اشاره نمود.

ازآنجایی‌که داده‌ها در بازه‌ی تغییرات متفاوت می‌توانند تأثیر غیر دلخواهی بر روی همدیگر و به‌تبع آن بر روی الگوریتم، قرار دهند پس داده‌ها باید در یک بازه‌ی مساوی نسبت به یکدیگر قرار بگیرند، مثلاً همه در یک بازه‌ای مانند صفر تا یک قرار داشته باشند و به این کار نرمال‌سازی داده‌ها[3] گفته می‌شود. مجموع داده های تحقیق حاضر نرمال‌سازی شده و در بازه‌ی منفی یک تا مثبت یک قرار گرفته‌اند. داده‌ها در مدت یک

[1] Data Logging
[2] Human-Machine Interface (HMI)
[3] Data Normalization

سال با روند دمایی مطابق شکل (۲) و با ثابت زمانی ۲ دقیقه نمونه‌برداری شده‌اند. در نتیجه هر یک از مجموعه داده‌ها به صورت ماتریس شامل ۲۵۹۲۰۰ سطر نمونه داده و ۷ ستون داده‌های اندازه‌گیری می‌باشد.



شکل ۲. روند دمایی یک ساله در زمان کار توربین گاز فریم ۹

در تحلیل عملکرد توربین با استفاده از پارامترهای ترمودینامیکی نظیر دما و فشار، عملکرد هر یک از اجزای واحد، بررسی می‌شود و در صورت انحراف از مقادیر مرجع گزارش ارائه می‌گردد. عملکرد توربین گاز به‌شدت تابع شرایط محیطی یعنی دما، فشار محیط و بار واحد است. بعلاوه عواملی نظیر استهلاک اجزا نیز بر عملکرد واحد اثر می‌گذارند. پس این بانک داده نه‌تنها باید دارای اطلاعات جامع از تمامی پارامترهای توربین باشد، بلکه باید دارای اطلاعات کافی از شرایط محیطی که توربین در آن کار می‌کند نیز باشد. زیر ساخت مناسب برای ذخیره اطلاعات، یکی از پیش‌نیازهای مهم در طراحی و توسعه هر نرم‌افزار محاسباتی است؛ بنابراین طراحی پایگاه داده از اولویت‌های اساسی می‌باشد. پایگاه داده وظیفه ذخیره اطلاعات موجود در فضای کاری را بر عهده دارد. برای طراحی پایگاه داده می توان از نرم‌افزار Microsoft SQL Server استفاده شود. سپس به این علت که در این پروژه از متلب استفاده می‌شود باید کل داده‌های موجود در پایگاه داده قبلی در یک فایل متنی قرار بگیرند، در نهایت خروجی داده‌ها، به صورت فایل‌های متنی CSV می‌باشد، سپس هر فیلد جدول پایگاه داده با (:) از هم جدا می‌شوند، سپس به کمک همان فایل متنی داده‌ها که حاوی داده‌های قدیمی پایگاه داده می‌باشد به وسیله‌ی (::) جدا می‌شوند که هر قسمت از (::) نمایانگر یک فیلد از جدول پایگاه داده می‌باشد، در پایگاه داده‌ی جدید قرار می‌گیرد. به عبارتی فایل متنی به عنوان مترجم یا رابط بین داده‌های پایگاه داده‌ی قدیم و جدید به کار می‌رود (در قالب فایل‌های CSV ذخیره می‌گردند). سپس این فایل‌ها با استفاده از یک کد نوشته شده در نرم‌افزار متلب، به داده‌های با پسوند mat. ذخیره می‌شوند.

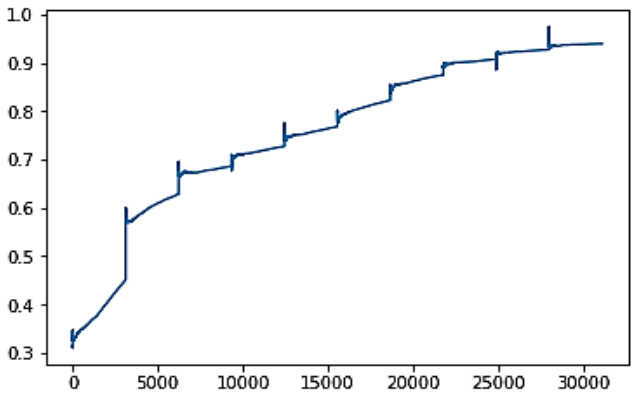
معیارهای ارزیابی

الگوریتم‌های پیشنهادی مقاله حاضر «الگوریتم‌های کاهش بعد» و «الگوریتم طبقه‌بندی» و از مباحث یادگیری ماشین می‌باشد این الگوریتم‌ها برای تشخیص و شناسایی نوع عیب استفاده می‌شود و از طریق

معیارهایی مانند صحت، میزان خطا و همگرایی الگوریتم می‌توان آن‌ها را ارزیابی نمود. صحت[1] معیار درستی پیش‌بینی مدل می‌باشد. به این معناست که مدل تا چه اندازه خروجی را درست پیش‌بینی می‌کند. میزان خطا نیز اختلاف مقدار اندازه‌گیری شده با مقدار واقعی است. همگرایی مدل، معیاری است که می‌تواند مقادیر واقعی را به درستی پیش‌بینی کند. سرعت همگرایی با بهبودی یادگیری الگوریتم رابطه مستقیم دارد.

یافته‌های پژوهش

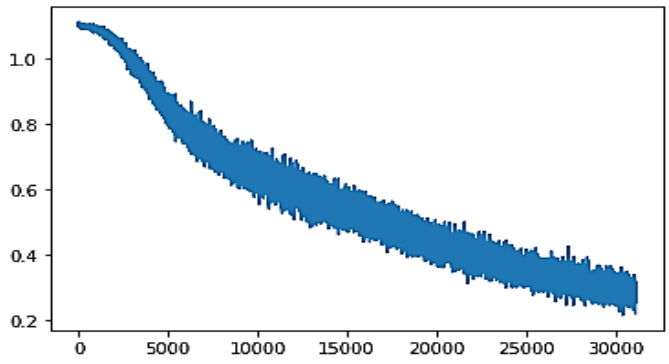
محاسبات کاهش ابعاد در ابعاد ۲ تا ۷ انجام گرفت که طی محاسبات «عملیات PCA»، «اجرای الگوریتم شبکه عصبی مصنوعی»، «صحت سنجی» و «محاسبه خطای مدل» مشخص گردید که بهترین نتایج در بعد ۴ (n_components = 4) به دست می‌آید. جهت ارزیابی صحتِ الگوریتم، محاسبات ابتدا در ۸۰٪ داده‌ها انجام شد. پس از یادگیری در ۲۰٪ باقیمانده داده‌ها، صحتِ مدل توسط شبکه عصبی مصنوعی مورد ارزیابی قرار گرفت. نتیجه محاسباتِ صحت الگوریتم بالای ۹۰٪ برای چهار بعد را نشان می‌دهد (شکل ۳). در واقع مدل توانسته به‌خوبی دقت مطلوبی برای این تعداد بعد ایجاد کند.



شکل ۳. محاسبات صحت الگوریتم برای کاهش ابعاد تا چهار بعد

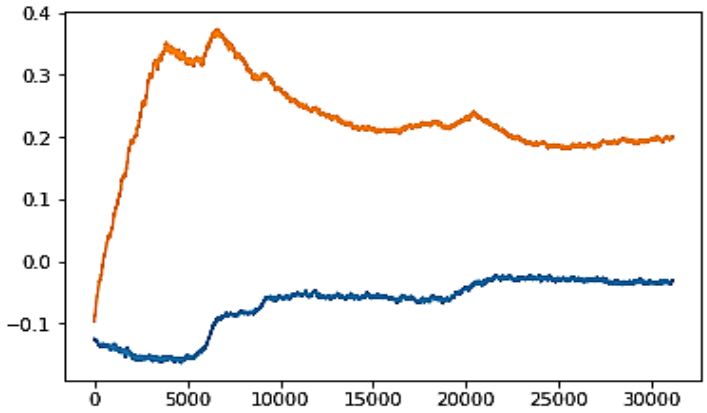
نتیجه محاسبات خطای الگوریتم در شکل (۴) به وضوح دیده می‌شود که نتیجه حاصل عدد 0.2832 می‌باشد. این نتیجه نشان‌دهنده خطای پایین می‌باشد. در واقع مدل با خطای خیلی کمتری مواجه شده‌است. نمودار نشان‌دهنده پراکندگی کم خطا و نظم در کاهش خطا دیده می‌شود.

[1] Accuracy



شکل ٤. نتیجه محاسبات خطای الگوریتم در کاهش ابعاد تا چهار بعد

شکل (۵) همگرایی وزن‌های به‌دست‌آمده از الگوریتم را نشان می‌دهد. این نمودار نسبت به ابعاد دیگر نتایج بهتری نشان می‌دهد.



شکل ۵. همگرایی وزن‌های به‌دست‌آمده از الگوریتم برای کاهش ابعاد تا چهار بعد

پیش‌بینی مدل

بعد از این مراحل، داده‌های تست را به شبکه می‌دهیم تا دقت مدل را تعیین کنیم. به این صورت که پس از یادگیری، مدل حالت عیب را بر روی ۲۰٪ مابقی داده‌ها پیش‌بینی می‌کند که با مقایسه پیش‌بینی مدل با برچسب‌های واقعی درصد دقت و خطا محاسبه می‌شود. همان‌طور که در همین فصل و در بخش اجرای تکنیکِ تحلیل مؤلفه‌های اصلی PCA شرح داده‌شده با ابعاد کاهشی متفاوت، نتایج متفاوتی از نظر دقت و خطا کسب کردیم. این نتایج به صورت نمودار دقت و خطا رسم شد که در نهایت با ابعاد n = 4 بهترین نتیجه و به دقتی معادل 0.935 و به خطای 0.2832 رسیدیم. این مقدار نشان‌دهنده کسب نتیجه بسیار مطلوبی از مدل تحقیق حاضر می‌باشد. علاوه بر دقت و خطا، همگرایی وزن‌ها نیز در n = 4 دیده‌شد. در نتیجه با اطمینان می‌شود گفت که مدل توانسته است سیستم را به خوبی پیش‌بینی کند.

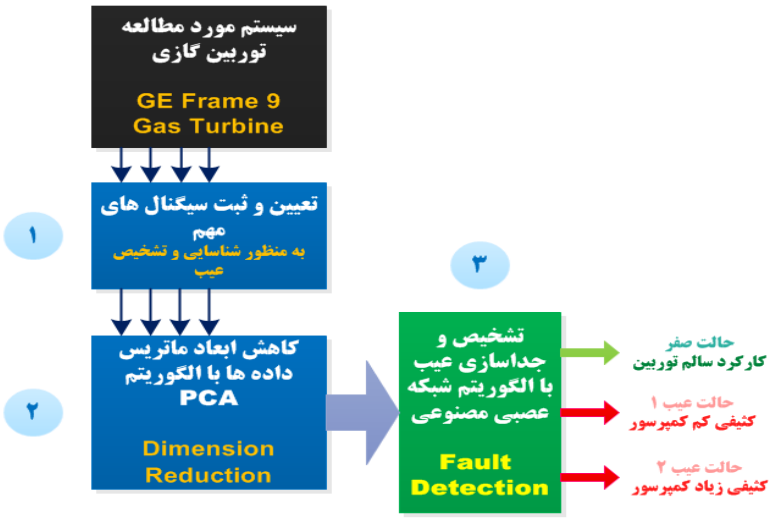
مقادیر ضرایب (وزن‌ها)

در این الگوریتم با تکرار، مجموعه‌ای از وزن‌ها آموزش می‌بینند تا برچسب (کلاس) دادهِ چندتایی‌های ورودی را پیش‌گویی کنند. در الگوریتم حاضر w1 & w2، وزن‌های الگوریتم یا همان ضرایبِ الگوریتم «شبکه عصبی مصنوعی» می‌باشند که از اجرای محاسبات «شبکه عصبی مصنوعی» به دست آمده است.

همان‌گونه که در شکل (۵) مشاهده می‌شود نحوه همگرایی وزن‌ها نشان می‌دهد دو منحنی به‌تدریج همگرا شده و مدل به‌خوبی عمل کرده و توانسته مقادیر واقعی را به‌درستی پیش‌بینی نماید. پس نتیجه گرفته می‌شود که سرعت همگرایی با بهبودی یادگیری الگوریتم رابطه‌ی مستقیم دارد

ارزیابی مدل

الگوریتم کل فرایند اجرایی پروژه در شکل (۶) به صورت خلاصه نشان داده شده است. ابتدا به ارزیابی اولین الگوریتم «تحلیل مؤلفه‌های اصلی» PCA می‌پردازیم. همان‌طور که قبلاً ذکر شد ابعاد مجموعه‌ی داده‌ها در ۷ پارامتر تعریف شدند (تعداد ستون داده‌ها) که با توجه به حجیم و پیچیده بودن داده‌ها از روش PCA برای کاهش ابعاد ماتریسِ داده‌ها استفاده شد.



شکل ٦. ساختار کلی الگوریتم مورد استفاده در روش پیشنهادی

پیچیدگی زمانی الگوریتم مورد استفاده در روش پیشنهادی که از دو بخش PCA و MLP تشکیل شده است که به صورت زیر است: در ابتدا پیچیدگی زمانی PCA که به تعداد نمونه‌ها (N) و تعداد ویژگی‌ها (D) بستگی دارد. برای محاسبه PCA، ابتدا ماتریس همبستگی بین ویژگی‌ها را محاسبه می‌کنیم که پیچیدگی زمانی آن برابر است با O(N * D^2). سپس تجزیه مقادیر ویژه را بر روی ماتریس همبستگی انجام می‌دهیم و پیچیدگی زمانی این مرحله بستگی به اندازه ماتریس همبستگی (D * D) دارد؛ بنابراین، پیچیدگی زمانی کل PCA می‌تواند تقریباً به صورت (O(N * D^2 + D^3) برآورد شود. پیچیدگی زمانی MLP که بستگی به تعداد نمونه‌ها (N)، تعداد ویژگی‌ها (D)، تعداد لایه‌ها و تعداد نورون‌ها در هر لایه دارد. پیچیدگی زمانی MLP به صورت O(N * D * L * M) برآورد می‌شود که در آن L تعداد لایه‌ها و M تعداد نورون‌ها در هر لایه است.

با کاهش ابعاد در بعدهای مختلف، نتایج متفاوتی از نظر دقت و خطا کسب شد. نتایج از نظر دقت و خطا مورد ارزیابی قرار گرفت و مشاهده شد که با کاهش ابعاد تا حد بعد n = 4 نتیجه‌ای عالی با دقتی معادل 0.935 و خطایی برابر 0.2832 به دست آمد که نشان‌دهنده کسب نتیجه بسیار مطلوب نیز از مدل می‌باشد.

علاوه بر دقت و خطا، همگرایی وزن‌ها نیز در n = 4 دیده‌شد. در نتیجه با اطمینان می‌شود گفت که ارزیابی مدل، آن را مدل موفقی نشان می‌دهد. در جدول (۴) نتیجه پیش‌بینی الگوریتم را می‌بینیم که در این جدول ستون‌های Y_test برچسب‌های داده‌های تست و ستون Output_Class برچسب داده‌های پیش‌بینی‌شده توسط الگوریتم را نشان می‌دهد. مقادیر ستون Fault نتیجه مقایسه‌ی داده‌های واقعی با داده‌های پیش‌بینی‌شده می‌باشد. عدد صفر نشان‌دهنده‌ی توربین در حالت سالم می‌باشد به این معنی که الگوریتم توانسته به درستی پیش‌بینی کند، مقادیر ۱ در واقع نشان‌دهنده‌ی عیب نوع ۱ و مقادیر ۲ نشان‌دهنده‌ی عیب نوع ۲ می‌باشد. محاسبه خطا به‌صورت میانگین قدر مطلق تفاضل بین مقادیر پیش‌بینی‌شده و مقادیر واقعی را محاسبه می‌کند که از این طریق، کارایی شبکه عصبی تولید شده را ارزیابی می‌کند.

جدول ۳. کیفیت شناسایی خطاها در الگوریتم

Index	Y_test	Output_Class	Faul
0	0	0	0
1	2	2	0
2	2	2	0
.			
.			
505	0	1	1
506	1	1	0
507	0	2	2
508	0	2	2
.			
.			
155517	1	1	0
155518	0	0	0
155519	2	2	0

FALSE	8688
total	155519
TRUE	146831
%	94%

بحث و نتیجه‌گیری

مجموعه داده‌های واقعی تحقیقِ حاضر، از حس‌گرهای نصب شده بر روی توربین گاز فریم ۹ استخراج شده‌است. ابتدا برچسب‌گذاری ماتریسِ داده‌ها صورت گرفت تا داده‌های سالم و معیوب در هر نمونه داده‌ای مشخص شود. سپس با بهره‌گیری از تکنیک مؤلفه‌های اصلی، ابعاد ماتریس داده از هفت به چهار بعد کاهش داده شد و ویژگی‌های اصلی آن استخراج گردید. در ادامه و با اجرای تکنیک شبکه عصبی مصنوعی، مدلی ایجاد شد که با کمک تکنیک «یادگیری ماشین» می‌توانست نمونه‌های خطا در ماتریس داده‌ها را شناسایی و حالتِ (کلاس) نمونه داده‌ها از نظر سالم یا معیوب بودن را تشخیص بدهد. پس از یادگیری ماشین که از روی ۸۰٪ داده‌ها انجام گرفت دقت، صحت و همگرایی مدل برای کاهش‌های ابعادی ماتریسِ داده‌ها به دو تا شش بُعد بررسی و مشخص گردید که مدل به‌خوبی کار می‌کند و با دقتی بیش از ۹۰٪ و با صحت بالا و بیشترین میزان همگرایی می‌تواند ماتریسِ داده‌ها را ارزیابی و کلاس عیب توربین را مشخص نماید. نتایج تحقیق حاضر نشان داد که تلفیق تکنیک «تحلیل مؤلفه‌های اصلی» با تکنیک «شبکه عصبی مصنوعی» توانسته به‌خوبی و با دقت بالا وقوع عیب را در توربین گازی پیش‌بینی نماید.

رویکرد پژوهش حاضر استفاده از تکنیک‌های داده‌کاوی در حوزه سامانه‌های پایش توربین‌های گازی می-

باشد. حوزه پایش توربین‌های گازی و نیروگاهی به همراه کاربرد داده‌کاوی، حوزه‌ای نو در صنعت نیروگاهی جهان می‌باشد که هنوز در ابتدای راه پژوهشی خود بوده و نیاز به بهبود و توسعه سخت‌افزارهای پایشی توربین دارد. همچنین لزوم استفاده از تکنیک‌های نرم‌افزاری به‌روز که شامل روش‌های توسعه‌یافته داده‌کاوی و یادگیری ماشین می‌باشد هست. کاربرد روش داده‌کاوی در پایش وضعیت توربین گاز و استفاده از تلفیق تکنیک‌های "تحلیل مؤلفه‌های اصلی" و "شبکه عصبی مصنوعی" برای کاهش ابعاد داده‌ها و استخراج ویژگی‌ها و متعاقباً طبقه‌بندی ویژگی‌ها در صنعت نیروگاهی که در این پایان‌نامه مورد استفاده واقع شده است از کارهای جدید و نوپا در این حوزه محسوب می‌شود. داده‌های مورد استفاده در این تحقیق برای اولین بار در این دست کارهای تحقیقی بوده و در ایران برای اولین بار نیز می‌باشد.

استفاده از شبکه‌های عصبی در مسائل طبقه‌بندی و تفکیک بهترین روش در سال‌های اخیر بوده‌است. تنظیم پارامترهای یک شبکه عصبی، گرچه کمی پیچیده و زمان‌بر می‌باشد اما به خوبی از پس طبقه‌بندی ویژگی‌ها برمی‌آید. این نتیجه‌گیری با داده‌های نیروگاهی کشور در این پژوهش بررسی مجدد و تائید شده‌است. اجرای پروژه حاضر نیاز به داده‌های واقعی داشته که از حس‌گرهای نصب شده بر روی یکی از توربین‌های نیروگاه گازی کشور استخراج شده‌است. جهت توسعه کار حاضر، می‌توان با نصب حس‌گرهای بیشتر مقادیر تجربی بیشتر و در نتیجه ماتریس داده‌ای حجیم‌تر به دست آورد که می‌توان به نتایج درست‌تر و با دقت بیشتری رسید. علاوه بر آن پیشنهاد می‌شود که حس‌گرها به توربین‌هایی نصب شود که خیلی سالم نبوده و تقریباً فرسوده باشد و احتمال خرابی در آن بیشتر باشد تا با یادگیری بیشتر مدل، پایگاه داده‌ای قوی‌تری ایجاد کنیم که باعث تشخیص، شناسایی و پیش‌بینی دقیق‌تر عیوب توربین‌های گازی قبل از رخداد خرابی بشود.

در این پژوهش از دو تکنیک PCA و شبکه عصبی مصنوعی که از تکنیک‌های داده‌کاوی می‌باشد استفاده شده‌است. با توجه به احتمال افزایش تعداد عیوب توربین یا همان ابعاد ماتریس داده و همین‌طور افزایش پیچیدگی و حجم داده‌ها، پیشنهاد می‌شود در کارهای بعدی از روش شبکه عصبی عمیق استفاده شود. دیگر پیشنهاد مرتبط با پروژه حاضر می‌توان به تلفیق تکنیک‌های KPCA با مدل نروفازی اشاره نمود. داده‌های توربین گازی عمدتاً غیرخطی است لذا KPCA کاربرد بیشتری در کاهش ابعاد دقیق‌تر در ماتریس داده‌های توربین گازی دارد. مدل نروفازی کار طبقه‌بندی و تشخیص عیوب را انجام می‌دهد. مزیت مدل نروفازی نسبت به مدل تحقیق حاضر در واقع مقاوم بودن در برابر نویز داده‌های ثبت شده‌می‌باشد. ازآنجایی‌که نویز در داده‌های توربین بسیار شایع بوده و باعث عدم دقت ماتریس داده‌ها می‌شود لذا استفاده از مدل نروفازی برای کارهای بعدی پیشنهاد می‌گردد.

سپاس‌گزاری

نویسندگان این مقاله از همکاری شرکت توربوکمپرسور تک خاورمیانه (توربوتک) به‌ویژه سرکار خانم مهندس جعفری کارشناس ارشد طراحی، جناب آقای مهندس عظیمی مدیر پروژه Condition Monitoring و

دیگر همکاران که در ارائه داده‌های عددی و راهنمایی‌هایی اجرایی همکاری نمودند کمال سپاس‌گزاری را دارد.

منابع

Rabbani, A., Rahimpour Daghian, M, (2016). Monitoring the state of output power of gas turbine V94.2 by means of time series neural network. 30th International Electricity Conference. *(in Persian)*

M. Khorram Kashkooli , M. Dehghani ,(2017). Fault Detection, Identification and Isolation of South Pars Gas Turbine Using a Combined Method Based on the Data Mining Techniques, k-means, PCA and SVM. *Tabriz Journal of Electrical Engineering*, vol74, no.2, pp 501-515, 2017. (in Persian)

Tomaj, A., Ahmadi, A., Behzadi, N., Esmaili, A. and Samadi, M, (2012). Gas turbine vibration fault diagnosis using decision tree and scoring methods. *8th specialized conference on condition monitoring and fault diagnosis*, Iran, Tehran, Sharif University of Technology. (in Persian)

Lee, W. J., Mendis, G. P., Triebe, M. J., & Sutherland, J. W. (2019). Monitoring of a machining process using kernel principal component analysis and kernel density estimation. *Journal of Intelligent Manufacturing*, *31*(5), 1175–1189. https://doi.org/10.1007/s10845-019-01504-w

Shakya, S. (2021). A Self Monitoring and Analyzing System for Solar Power Station using IoT and Data Mining Algorithms. https://irojournals.com/jscp/article/view/3/2/4

Swiercz, M., & Mroczkowska, H. (2020). Multiway PCA for Early Leak Detection in a Pipeline System of a Steam Boiler—Selected Case Studies. *Sensors*, *20*(6), 1561. https://doi.org/10.3390/s20061561

De Andrade Melani, A. H., De Carvalho Michalski, M. A., Da Silva, R. A., & De Souza, G. F. M. (2021). A framework to automate fault detection and diagnosis based on moving window principal component analysis and Bayesian network. *Reliability Engineering & System Safety*, *215*, 107837. https://doi.org/10.1016/j.ress.2021.107837

Rahmoune, M. B., Hafaifa, A., Kouzou, A., Chen, X., & Chaibet, A. (2021). Gas turbine monitoring using neural network dynamic nonlinear autoregressive with external exogenous input modelling. *Mathematics and Computers in Simulation*, *179*, 23–47. https://doi.org/10.1016/j.matcom.2020.07.017

De Giorgi, M. G., Strafella, L., Menga, N., & Ficarella, A. (2022). Intelligent Combined Neural Network and Kernel Principal Component Analysis Tool for Engine Health Monitoring Purposes. *Aerospace*, *9*(3), 118. https://doi.org/10.3390/aerospace9030118

An, Z., Cheng, L., Li, K., Ren, M., Yang, Z., Feng, W., Ling, J., Chen, H., & Chen, W. (2022). *Fault diagnosis for nuclear power equipment based on a recursive principal component analysis approach*. https://doi.org/10.1117/12.2627200

Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2(4), 433-459.

Hadroug, N., Hafaifa, A., Alili, B., & et al. (2022). *Fuzzy Diagnostic Strategy Implementation for Gas Turbine Vibrations Faults Detection: Towards a*

Characterization of Symptom–fault Correlations. Journal of Vibration Engineering & Technologies, 10, 225–251.

Xie, J., Sage, M., & Zhao, Y. F. (2023). *Feature selection and feature learning in machine learning applications for gas turbines*: A review. Engineering Applications of Artificial Intelligence, 117, 105591.

Yu, H., Gao, H., He, Y., Lin, Z., & Xu, X. (2022). *A novel motor fault diagnosis method based on principal component analysis (PCA) with a discrete belief rule base (DBRB) system*. Measurement Science and Technology, 34(3), 035012.