ارائه یک سیستم تصمیم‌یار جهت پیش بینی خاموشی اضطراری نیروگاه‌های برق آبی با استفاده از استخراج قوانین انجمنی
مطالعه موردی: نیروگاه برق آبی مارون بهبهان

الهام پروین نیا [1*]، خسرو فرداد [2]

۱- گروه مهندسی کامپیوتر، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران
parvinnia@iaushiraz.ac.ir
۲- گروه مهندسی کامپیوتر، واحد بهبهان، دانشگاه آزاد اسلامی، بهبهان، ایران
Khossro.fardad @gmail.com

چکیده: **حسگرها وضعیت بخشهای مختلف نیروگاه برق آبی را رصد کرده و ابزارهای کنترلی دستورات لازم را جهت فعالیت نیروگاه صادر می‌کنند. همچنین کارشناسان براساس مقدار اعدادی که حسگرها و دماسنج ها نشان می‌دهند و براساس شرایط محیطی نیروگاه و تجربه در شرایط خاص تصمیم‌های لازم را جهت خاموشی اضطراری نیروگاه اتخاذ می‌کنند. در یک نیروگاه برق آبی عوامل متعددی مانند: تعمیرات، علائم هشدار در حسگرها، آسیب های فیزیکی تجهیزات و یا ارتفاع آب در پشت سد، ممکن است تولید برق را متوقف سازد. لذا تشخیص زمان مناسب فعالیت یا عدم فعالیت نیروگاه با توجه به داده های حسگرها بسیار حیاتی است. هر چند سیستمهای کنترلی موجود به نحو مطلوبی شرایط را بررسی می‌کنند ولی بعلل مختلف مانند خطای انسانی یا خطای تجهیزات ممکن است تصمیم به ادامه کار یا خاموشی اضطراری با خطا همراه باشد. در این مقاله با استفاده از تکنیک‌های داده کاوی یک سیستم تصمیم‌یار طراحی شده است که روابط معنادار بین ویژگی هایی که حسگرها همه روزه در نیروگاه برق آبی آرشیو می‌کنند را می‌یابد. این روابط مانند قوانین استنتاج می‌توانند در تصمیم گیری دقیق و سریع کارشناسان کمک شایانی نموده و از آسیب به تجهیزات در تصمیمات اشتباه یا دیرهنگام جلوگیری نماید. نتایج بدست آمده از تحلیل داده های سالهای ۹۲ الی ۹۴ نیروگاه برق آبی مارون ۴۱ قانون برای تصمیم گیری سریع بدست آمده است که با بررسی های به عمل آمده توسط متخصصان فنی این نیروگاه ۴ قانون آن جدید و سودمند ارزیابی شده است.**

واژه های کلیدی: **خاموشی اضطراری، نیروگاه برق آبی ، داده کاوی، قوانین استنتاج**

تاریخ ارسال مقاله : ۱۳۹۶/۱/۲۱

تاریخ پذیرش مقاله : ۱۳۹۷/۲/۲۲

*نام نویسنده‌ی مسئول : الهام پروین نیا

نشانی نویسنده‌ی مسئول : شیراز - کیلومتر ۵ جاده صدرا – پردیس دانشگاه آزاد شیراز – دانشکده مهندسی – گروه کامپیوتر.

۱- مقدمه

با توجه به افزایش هر روزه ارزش انرژی و تقاضا برای مصرف انرژی در سالهای اخیر، تولید انرژی اهمیت بسیار زیادی پیدا کرده است و پژوهشهای زیادی در جهت بهبود کارایی فرایند تولید انرژی انجام شده است. با توجه به ماهیت انرژی الکتریکی و استفاده‌های زیادی که از آن در مصارف خانگی و صنعتی انجام می‌شود تحقیق اخیر با هدف بهبود فرایند تولید برق انجام می‌شود. محیط نیروگاهی تجهیزات بسیار زیادی دارد که بصورت یک سیستم توزیع شده کنترل می‌شود. برای تولید برق در نیروگاه برق آبی نیاز به احداث سد بر روی رودخانه می-باشد که در کشورها با توجه به وجود رودخانه های با دبی آب مختلف امکان نصب این نوع نیروگاهها بسیار زیاد است. نیروگاههای آبی را بر اساس شیوه ساختمانشان می‌توان به دو صورت زیر تعریف کرد:

- نیروگاهی که جزیی از ساختمان سد بوده و با سد دارای ساختمان واحد است و آب را در پشت خود نگه می دارد که در اصطلاح به آن نیروگاه مسیر جریان(In-Stream) گویند.
- نیروگاهی که ساختمان آن با آب بندها یکی نیست و به آن اصطلاحا نیروگاه خارج از مسیر جریان (Off- Stream) گویند. این نیروگاه ممکن است که در پشت سد یا جای دیگر روی زمین و یا بصورت زیرزمینی ساخته شود.

عمر مفید نیروگاه برق آبی بیش از ۵۰ سال بوده و به ۱۰۰ سال هم می‌رسد که در مقایسه با عمر نیروگاه بخاری (حدود ۲۵ تا ۳۰ سال) بسیار زیاد است. حدود ۱۹ درصد از کل انرژی الکتریکی تولیدی جهان بصورت برق آبی تولید می‌شود[۱]. به منظور تولید انرژی چرخشی توربین نیازی به سوخت نمی‌باشد. در نتیجه هزینه عملکرد نیروگاه بسیار کم می‌شود. بعلاوه با توجه به عدم نیاز سوخت به مخزن‌های ذخیره سوخت هم نیازی نیست. در این نیروگاهها هیچگونه آلودگی ناشی از گازهای حاصل از احتراق وجود ندارد و از این نظر مشکلی برای محیط زیست ایجاد نمی‌کند[۲]. زمان راه‌اندازی نیروگاه برق آبی بسیار کوتاه است و در زمان کمی(حدود چند دقیقه) قادر به هماهنگی با شبکه و اتصال به آن می‌باشد. هزینه نگهداری این نیروگاهها بسیار پایین است. جلوگیری از سیل و سیلابهای فصلی (با ایجاد سد) از دیگر مزایای نیروگاه برق آبی است[۱].

یکی از مسایل قابل تامل در بحث تولید برق اینست که نمی توان برق را تولید کرده و برای استفاده‌های آتی ذخیره کرد. پس تولید بدون وقفه و براساس نیاز صنعت برق کشور امری حیاتی است ولی براساس عوامل متعددی مانند تعمیرات، علائم هشدار در حسگرها، آسیب های فیزیکی تجهیزات و یا ارتفاع آب در پشت سد ممکن است نیروگاه فرایند تولید برق را متوقف سازد. لذا تشخیص زمان مناسب فعالیت یا عدم فعالیت نیروگاه بسیار حیاتی است و با کمک سیستمهای کنترلی تعبیه شده در نیروگاه مانند گاورنر[1] تلاش می‌شود بهترین تصمیم در جهت تولید یا عدم تولید برق و مراقبت از تجهیزات بعمل بیاید. فعالیت تولیدی نیروگاه به داده‌هایی که از طریق سیستم کنترل ثبت شده بستگی داشته و اتاق کنترل براساس آنها تصمیم نهایی را اتخاذ می-نماید. بعلت ماهیت عملیاتی نیروگاه و وجود تجهیزات مختلف، داده-های ثبت شده از سیستم کنترل که خروجی حسگرهای تعبیه شده روی سیستم می‌باشد معمولا دارای نویز بوده و دقیق نمی‌باشند. داده-های پالایش نشده زیادی لحظه به لحظه از طریق حسگرها در نیروگاه تولید و آرشیو می‌شوند. استفاده از تکنیک های داده‌کاوی می‌تواند در تصمیم گیری مناسب و سریع با توجه به حجم زیاد داده‌ها بسیار مفید واقع شود.

ادامه مقاله بگونه‌ای تدوین شده است که در بخش(۲) نحوه فعالیت نیروگاه برق آبی بطور مختصر تشریح می‌گردد. سپس در بخش(۳) مسئله مورد بررسی در این مقاله تبیین می‌شود. تحقیقات گذشته در این زمینه کاری در بخش (۴) معرفی شده اند تا اهمیت و جایگاه این مسئله نزد دیگر متخصصان نیز مشخص شود. روش کار و مراحل داده-کاوی نیز در بخش (۵) بیان شده است. در بخش پایانی نیز جمع بندی و نتیجه‌گیری از این تحقیق بیان شده است.

۲- نحوه فعالیت یک نیروگاه آبی

نیروگاه برق آبی براساس تبدیل انرژیها کار می‌کند یعنی انرژی آب روان (انرژی جنبشی) پس از ذخیره به انرژی پتانسیل تبدیل شده و سپس با تنظیم جریان بشکل انرژی هیدرولیک ظاهر شده و وارد توربین و سپس ژنراتور می‌شود که حاصل کار انرژی الکتریکی است[۸]. تجهیزات نیروگاه آبی بر حسب کاری که انجام می دهند به چند گروه عمده تجهیزات تقسیم می‌گردند: تجهیزات تولید، تجهیزات کمکی، تجهیزات مکانیکی، تجهیزات برق که در شکل(۱) بصورت شماتیک یک نیروگاه برق آبی نشان داده شده است.

با توجه به اینکه تجهیزات نیروگاه گرانقیمت و بسیار حساس می‌باشند باید مکانیزمهای کنترلی مناسبی وجود داشته باشد تا از آنها محافظت بعمل بیاورد. زیرا در صورتیکه شرایط کارکرد مطلوب فراهم نباشد فرایند تولید برق به تجهیزات صدمه خواهد زد. دستگاه کنترل جریان توربین با یک تنظیم کننده گاورنر خودکار مجهز شده که کار آن ابقا سرعت چرخشی در یک سطح ثابت، بدون توجه به بار روی توربین می‌باشد. همچنین این تجهیز برای راه انداختن و توقف توربین و کنترل بار روی توربوژنراتور بکار گرفته می‌شود. گاورنر فرکانس (قدرت اکتیو-حقیقی) را کنترل نموده و یک سری اطلاعات را دریافت و در کنترل خود بررسی و بعد دستور می‌دهد. گاورنر از سه مجموعه مهم تشکیل می‌شود: حسگرها، مقایسه‌کننده و کنترل‌کننده، TC100. حسگرها شرایط مختلفی مانند: سرعت واقعی محور توربین، قدرت خروجی از

ترمینالهای استاتور ژنراتور، وضعیت گشودگی دریچه‌ها، هد خالص آب بالا دست توربین، ارتفاع خالص آب پائین دست توربین را رصد کرده و به مقایسه کننده‌ها منتقل می‌کنند[۳].

در مقایسه کننده و کنترل کننده اطلاعات ، مقدار مطلوبی که به گاورنر داده ایم با مقادیر واقعی مقایسه می‌شود و اگر برابر نباشد (که به آن سیگنال خطا می‌گوئیم) به قسمت اصلی اطلاع می‌دهد. سیگنال خطا می‌تواند تناسبی که از یک محاسبه کننده می‌گذرد باشد یا مشتق و یا انتگرالی محاسبه کند و سپس یک سیگنال کنترلی خارج شود.

تمام برنامه های گاورنر در مجموعه الکترونیکی بنام TC100واقع در پانل کنترل توربین می‌باشد. حسگرهای اطراف توربین که مقادیر مطلوب را به آنها داده‌ایم و یک سری الگوریتم‌ها و معادلاتی که برای آنها تعریف شده وارد گاورنر و شیر کنترل تناسبی که شامل دو قسمت شیر پایلوت (راهنما یا هادی) مدار فرمان و شیر کنترل اصلی (مدار قدرت) است وارد می‌شود. علت دو قسمت بودن این شیر این است که سیگنال اولیه ضعیف و بعد از تقویت وارد شیر کنترل می‌شود. در مدار پاسخ، سرعتی ایجاد می شود که حرکت اسپول هر دو شیر را کنترل کرده و مقایسه می‌کند و بطور تناسبی (اگر صفر شود همانجا عمل می-کند) اجرا می نماید[۳].

تقویت‌کنندها[۲] از سیلندری که بوسیله کپسولهای گاز نیتروژن روغن را به سیستم گاورنر اعمال می‌کند تشکیل شده است. فشار درون کپسولهای نیتروژن بین خودشان و تقویت‌کننده تقسیم شده و پیستون درون تقویت‌کننده را تحت فشار قرار داده و طبیعتا روغن را تحت فشار قرار می‌دهد و به سیستم گاورنر هدایت می‌کند. اگر در حین کار، الکتروموتورهای گاورنر از کار بیفتند تقویت کننده‌ها می‌توانند با کمک فشار گاز نیتروژن برای دستور دادن (حدود سه مرتبه) مورد استفاده قرار گیرند و همچنین می‌توانند ضربه‌های وارده از طرف پمپها در هنگام خاموشی اضطراری، باردهی و ضربه قوچ در اثر باز و بسته کردن سریع لوله های سیستم را خنثی کنند. بخش الکترونیک گاورنر شامل کنترل کننده حلقه بسته قابل برنامه ریزی و مجهز به پردازنده است که شامل سیگنالهای ورودی است و یک سیگنال خروجی دارد که نتیجه پردازش کنترل کننده را نشان می‌دهد. عملکرد این سیستم خاموش نمودن اضطراری گاورنر و همچنین بستن شیر پروانه‌ای می-باشد. تنظیم سیستم بر اساس یک سرعت ایمن انجام می شود و وقتی که سرعت از مقدار نامی بالاتر رود یک سیگنال ممتد ارسال می نماید و واحد را متوقف می نماید[۳].

۳- شرح مسئله

حسگرها وضعیت بخشهای مختلف نیروگاه را رصد کرده و ابزارهای کنترلی مانند گاورنر دستورات لازم را جهت فعالیت های کنترل نیروگاه صادر می‌کنند. همچنین کارشناسان براساس مقدار اعدادی که حسگرها و دماسنج ها نشان می‌دهند و براساس شرایط محیطی نیروگاه و تجربه در شرایط خاص تصمیم های لازم را جهت خاموشی اضطراری نیروگاه اتخاذ می‌کنند. پس براساس عوامل متعددی مانند تعمیرات، علائم هشدار در حسگرها، آسیب های فیزیکی تجهیزات و یا ارتفاع آب در پشت سد ممکن است فرایند تولید برق متوقف شود. لذا تشخیص زمان مناسب فعالیت یا عدم فعالیت نیروگاه بسیار حیاتی است.

فعالیت تولیدی نیروگاه بستگی به داده هایی دارد که از طریق سیستم کنترل ثبت می شود و اتاق کنترل براساس آنها تصمیم نهایی را اتخاذ می‌نماید. با توجه به ماهیت عملیاتی نیروگاه و وجود تجهیزات مختلف، داده های ثبت شده که خروجی حسگرهای تعبیه شده روی سیستم می‌باشد معمولا دارای نویز بوده و دقیق نمی‌باشند. هر چند سیستمهای کنترلی موجود به نحو مطلوبی شرایط عملکرد را بررسی می‌کنند ولی بعلل مختلف مانند خطای انسانی یا خطای تجهیزات ممکن است تصمیم به ادامه کار یا خاموشی اضطراری با درصدی از خطا همراه باشد. اهمیت این موضوع باعث می‌شود روشهای موازی با روشهای موجود فعلی برای تصمیم گیری و تشخیص خطاها وجود داشته باشد. با توجه به حجم بسیار زیاد داده‌های تولید شده از جانب حسگرهای موجود، بررسی سنتی و نرم افزاری متداول نمی‌تواند ایده مناسبی باشد لذا روشهای داده‌کاوی که برای بررسی حجم بالایی از داده ها توسعه یافته اند می توانند راهکارهای سریع و دقیقی در کنار روشهای موجود در نیروگاه باشند.

در واقع روش پیشنهادی این تحقیق هیچگاه حذف کارشناسان نیروگاه را مد نظر ندارد و هدف آن معرفی یک سیستم کمکی برای تصمیم-گیری سریع تر و دقیق‌تر می‌باشد.

۴-تحقیقات پیشین

تاکنون پژوهشهای زیادی در حوزه انرژی و به طور خاص در زمینه برق با استفاده از داده‌کاوی انجام شده است. درمرجع [۵و۶] شرکت برق کیوشو، داده‌های حسگرها و اطلاعات آب و هواهای مختلف را برای حفظ ایمنی نیروگاه‌های برق آبی را جمع‌آوری کرده است و علائم غیرطبیعی که ممکن است شرایط غیر معمول را بوجود آورد بررسی می‌کند. سپس نتایج که نشانگر الگوهای غیر معمول ارتعاش بلبرینگ و اطلاعات آب و هوا است با استفاده از روش ماشین بردار پشتیبان شناسایی می‌شود که این الگوها در تحمل لرزش و حفاظت از نیروگاه-های برق آبی در شرایط غیر معمول، مفید خواهد بود.

در پژوهش دیگری [۷] تجهیزات خودکاری برای بهینه‌سازی حفاظت در شرایط احتمالی خطا نصب شده است. سیستم اتوماتیک قطع و بارگذاری ژنراتور در نیروگاه برق آبی چرچیل آبشار در لابرادور نصب شده است. تکنیک‌های داده‌کاوی مانند درخت تصمیم و رگرسیون استفاده شده است. عکسهای بلادرنگ از سیستم قدرت هیدروکبک[۳] در یک دوره ۵ ساله جمع آوری شده و نتایج زیادی از شبیه‌سازی حاصل می‌شود. پردازش این داده ها با استفاده از نرم افزار توسعه یافته دانشگاه لیژ انجام شده است. این رویکرد مناسب‌ترین پارامترها و تنظیمات بهینه برای سیستم اتوماتیک قطع و بارگذاری ژنراتور را

مشخص می‌کند، به حداقل رساندن تعداد قطع ژنراتور با حفظ کارایی یک پوشش ایمن برای نیروگاه می‌باشد. قوانین جدید بر پایه دانش حاصل از این رویکرد می‌تواند تبیین شود.

در[۴] افزایش راندمان بویلرها با کمک داده‌کاوی در نیروگاه حرارتی باعث بهینه سازی بازده احتراق می‌گردد. مسئله بهینه سازی کارکرد سیستم احتراق در یک نیروگاه حرارتی ایجاد یک توازن بین افزایش بازده احتراق و کمینه سازی انتشارگازهای نامطلوب می‌باشد. فرایند احتراق در سیستم بویلر یک فرایند چندورودی چندخروجی، متغیر با زمان و به شدت غیرخطی می‌باشد. به همین خاطر تنظیم پارامترهای کنترل‌پذیر اصلی آن با استفاده از روشهای کلاسیک مبتنی بر مدلسازی مشکل می‌باشد. از آنجایی که سیستمهای کنترل توزیع شده به طور گسترده در نیروگاهها مورد استفاده قرار می‌گیرد، داده‌های مربوط به پارامترهای بویلر جمع آوری و ثبت می‌شوند. به همین خاطر استفاده از تکنیکهای داده‌کاوی یک روش بسیار موثر به منظور تنظیم بهینه پارامترهای احتراق می‌باشد. با توجه به اینکه داده‌های موجود، خروجی حسگرهای نصب شده روی سیستم است اغلب، داده ها دارای دقت مناسب نبوده و در بعضی مواقع نویز نیز روی آنها تاثیر می‌گذارد، به همین دلیل استفاده از مجموعه های فازی در تحلیل این داده‌ها بسیار مفید می‌باشد. بنابراین یک روش مبتنی بر تکنیک‌های داده‌کاوی فازی به منظور افزایش راندمان بویلر ارائه شده و سپس کارایی آن با استفاده از یک تست مجازی روی سیستم بویلر نیروگاه حرارتی ذوب آهن اصفهان مورد ارزیابی قرار می‌گیرد و میانگین کمترین افزایش راندمان طی انجام پنج آزمایش برابر با۰٫۶۵۶ درصد است.

در [۸] یک سیستم عیب یابی و نظارت برای نیروگاه برق آبی معرفی شده است که وظیفه جمع‌آوری سیگنالها از قسمتهای مختلف نیروگاه و انتقال داده‌ها به مرکز عیب‌یابی مجازی را انجام می‌دهد. یک دسته بند با کمک ماشین بردار پشتیبان[۴] برای طبقه بندی عیب‌ها مورد استفاده قرار گرفته است. روش ماشین بردار پشتیبان مورد استفاده دقت یادگیری بالایی دارد. بنابراین برای عیب مشخص شده راه‌حل مناسب را مشخص می‌کند. همچنین داده‌ها در مرکز عیب یابی و نظارت بصورت لحظه ای آنالیز شده و در اختیار کارشناسان قرار می گیرد تا کارشناسان نیز بتوانند در جریان تصمیم سیستم قرار بگیرند و حمایتهای عملیاتی لازم را مهیا سازند.

در پژوهش دیگری [۹] روشی برای پیش‌بینی خطای تاثیرگذار روی تجهیزات نیروگاهی بر اساس رگرسیون بردار پشتیبان معرفی شده است که ابتدا ضریب همبستگی برای انتخاب ویژگیهای مناسب به صورت یک بردار ویژگی محاسبه می‌شود و سپس با استفاده از روش جستجوی شبکه‌ای پارامترهای رگرسیون بردار پشتیبان بهینه شده و در انتها مدل پیش بینی با بردار ویژگی و پارامترهای بهینه‌سازی شده ساخته می‌شود. نتیجه این پژوهش نشان می‌دهد که مدل ایجاد شده توسط نویسندگان آن قادر به پیش بینی نقاط باارزش با دقت بالایی می‌باشد.

با توجه به تحقیقات اخیر استفاده از تکنیک‌های داده‌کاوی در تصمیم گیری مناسب و سریع با توجه به حجم زیاد داده‌ها امری ضروری است. همچنین داده‌های پالایش نشده زیادی لحظه به لحظه در نیروگاه تولید و آرشیو می‌شود که برای تشخیص پارامترهای تاثیرگذار قوی می‌توان از فرایند داده کاوی بهره گرفت.

همانگونه که در [۱۰] بیان شده است حالت load-following در نیروگاه های هسته ای روش پیچیده‌ای است که در اغلب نیروگاههای هسته ای که در سالهای اخیر در کشورهای مختلف راه اندازی می شود مورد استفاده می باشد(روشهای و حالتهای دیگر به نوعی منسوخ شده هستند). با توجه به اینکه همواره تغییرات قابل توجهی در بسیاری از فرآیندهای مرتبط با نیروگاه هسته‌ای وجود دارد. به منظور داشتن یک برنامه کنترل جامع که بتواند همه جوانب را در کنترل بهتر نیروگاه لحاظ کند استفاده از تکنیک های داده کاوی معرفی شده است . این مطالعه رویکردی را برای استفاده از مدل های رگرسیون پیشنهاد می دهد. مجموعه داده ها برای داده کاوی با شبیه سازی دو برنامه کنترل ایستا از VVER-1000 NPP در نرم افزار Simulink از بسته نرم افزاری متلب بدست می‌آید.

در [۱۱] مطالعات جامعی در مورد روشهای داده کاوی برای شناسایی رویدادهای ایسلندینگ (islanding) در یک سیستم توزیع قدرت با نسل های توزیع مبتنی بر اینورتر ارائه شده است. ویژگی های مهم برای هر فاز در طرح تشخیص ایسلندینگ در جزئیات مورد بررسی قرار گرفته است. این ویژگیها از زمان اندازه گیری ولتاژ، فرکانس و تثبیت هارمونیک کل جریان و ولتاژ در نقطه اتصال مشترک است. مطالعات عددی روی سیستم IEEE 34-bus برای بررسی سناریوهای گوناگون از شرایط ایسلندینگ و غیر ایسلندینگ بوده است. سپس ویژگی های به دست آمده برای آموزش چندین تکنیک داده کاوی مانند درخت تصمیم گیری، ماشین بردار پشتیبانی، شبکه عصبی، و جنگل تصادفی استفاده می شود. نتایج شبیه سازی نشان می‌دهد که پارامترهای مهم ویژگی می تواند بر اساس رابطه بین ویژگی های استخراج شده مورد ارزیابی قرار گیرد. چهار ویژگی مهم که پس از شبیه سازی تشخیص دقیق ایسلندینگ را ممکن می کند ولتاژ اصلی THD، THD جریان اصلی، سرعت تغییر ولتاژ و انحراف ولتاژ است. مقایسه روشهای داده کاوی انجام شده نشان می دهد روش جنگل تصادفی نسبت به بقیه روشهای مورد مطالعه دقت بالایی در شناسایی ایسلندینگ دارد.

پژوهشگران در [۱۲] برای گذر از شبکه های برق قدیمی و منسوخ شده به یک شبکه هوشمند با حسگرهای جاسازی شده هم مساله اقتصادی و هم مساله زیست محیطی را برای نظارت بر پایداری سیستم مورد بررسی همزمان قرار داده اند. با گسترش ابزارهای اندازه‌گیری هوشمند، رشد چشمگیر داده ها این گذر با سرعت و دقت بیشتری انجام می‌شود. [۱۲] خلاصه ای جامع در مورد وضعیت کنونی تکنولوژی داده بزرگ در ادغام شبکه هوشمند ارائه می دهد. ویژگیهای

شبکه هوشمند و چند منبع داده انرژی مورد بحث قرار می گیرد. سپس، این مقاله جامع برنامه ها را با استفاده از داده های بزرگ در شبکه هوشمند خلاصه می کند، که همچنین شامل برخی از برنامه های کاربردی با نام تجاری جدید با آخرین فن آوری های بزرگ داده است. علاوه بر این، برخی از سیستم عامل های اصلی و تکنیک های استخراج دانش به منظور ترویج بینش های داده های بزرگ مورد بررسی قرار می گیرند. در نهایت، چالش ها و فرصت ها در این مقاله ذکر شده است.

یکی از موضوع های دیگری که در سال های اخیر در صنعت برق از داده کاوی و الگوریتمهای تکاملی برای یافتن راه حل مناسب تر استفاده کرده است، تحقیقات در مورد انتخاب محل مناسب محل نصب نیروگاه می باشد[۱۳]. ارزیابی تاثیرات زیست محیطی اغلب به عنوان یک الزام قانونی در انتخاب سایت برای چندین دهه مورد استفاده قرار می گیرد. هدف از این کار فعلی، توسعه یک مدل برای تصمیم گیرندگان برای رتبه بندی یا طبقه بندی پروژه های مختلف نیروگاه با توجه به ویژگی های چند معیار مانند کیفیت هوا، کیفیت آب، هزینه تحویل انرژی، تاثیر محیط زیست، خطر طبیعی و مدت زمان پروژه است. مطالعه موردی در [۱۳] مربوط به استفاده از الگوریتم ژنتیک برای پیش بینی نسبی چند لایه برای رتبه بندی موقعیت های مختلف نیروگاه در هند است.

[۱۴] یک چارچوب تصمیم گیری چند معیاره نرو-فازی ارائه می دهد که از مزایای هر دو شبکه عصبی و رویکرد فازی استفاده می کند. هدف از این کار این است که سیاست گذاران را قادر سازیم تا بهترین محل برای نصب نیروگاه برق آبی را با در نظر گرفتن معیارهای مهم مانند اثرات زیست محیطی و محیطی، خصومت، تاثیر اجتماعی اقتصادی، هزینه تحویل انرژی، کیفیت آب و کیفیت هوا انتخاب کنند. در مطالعه موردی، شبکه عصبی چندبخشی پرسپترون آموزش دیده با الگوریتم ژنتیکی برای طبقه بندی هر یک از ویژگی های مهم سایت نیروگاه به کلاس های مختلف مورد استفاده قرار می گیرد. در مرحله نهایی، با استفاده از استدلال فازی، امتیاز به رتبه بندی و یا اولویت بندی مختلف موجود و همچنین آینده سایت برای نصب نیروگاه برق آبی در هند اختصاص داده است.

مطالعه دیگری [۱۵] به بررسی تولید برق مطلوب از مخزن منگل در پاکستان تمرکز دارد. در این مقاله یک مدل ریاضی برای نیروگاه برق آبی Mangla و تکنولوژی بهینه سازی ذرات و الگوریتم ژنتیک در این مدل برای تولید برق بهینه استفاده شده است. نتایج نشان داد که تولید برق با استفاده از تکنیک های بهینه سازی در مدل ریاضی پیشنهاد شده افزایش می یابد. الگوریتم ژنتیک می‌تواند حداکثر الکتریسیته را از بهینه سازی ذرات ذره تولید کند، اما زمان اجرای ذرات بهینه سازی ذرات بسیار کمتر از الگوریتم ژنتیک است. نیروگاه برق آبی Mangla می تواند تولید برق تا ۵۹ * ۱۰۹ کیلووات ساعت را با استفاده از جریان های بهینه از تولید ۴۷ * ۱۰۸ کیلووات ساعت از روش های سنتی تولید کند.

۵- روش کار و مراحل داده کاوی

از دیدگاه محققین فرایند داده‌کاوی در جهت کشف دانش طی مراحل زیر انجام می‌شود :

۱) جمع آوری داده ها
۲) پیش پردازش داده ها، انتخاب و مرتب کردن آنها
۳) اتخاذ بهترین روش داده کاوی برای دستیابی به هدف
۴)تحلیل داده ها بر اساس روش تعیین شده
۵) ارزیابی و اعتبارسنجی یافته ها
۶) استفاده از نتایج و تثبیت و تحکیم دانش کشف شده
۷)تصمیم گیری بر اساس دانش کشف شده

۵-۱- جمع آوری داده ها

در نیروگاه های آبی بر اساس سیاست های موجود، داده‌ها در بازه‌های زمانی مختلفی بررسی و ذخیره می‌شود. به عنوان مثال بازه ممکن است یک ساعته یا سه ساعته یا روزانه یا هفتگی و ... باشد. در این تحقیق از داده های ثبت شده بصورت سه ساعته از سال ۱۳۹۲ تا ۱۳۹۴ در نیروگاه برق آبی مارون بهبهان استفاده شده است.

در این نیروگاه داده‌ها توسط یک کارشناس با مراجعه به تک تک حسگرها یا دماسنج های تعبیه شده و یادداشت آنها جمع آوری می-شود. در گذشته فرم جمع آوری داده ها که لاگ شیت نامیده می شود بصورت کاغذی در زونکن های مخصوصی نگهداری می‌شد. ولی امروزه کارشناس هنگام جمع‌آوری داده‌ها بجای نوشتن در کاغذ آنها را در دستگاهی بنام PDL وارد می‌کند و دستگاه از طریق پورت USB یا کابل RS-232 به کامپیوتر متصل شده و داده های موجود در آنها به سیستم منتقل می‌شود.

در اغلب نیروگاههای امروزی علاوه بر روش مراجعه حضوری کارشناس یک شبکه ارتباطی مکانیزه نیز وجود دارد که داده ها را به اتاق کنترل منتقل می‌کند ولی با توجه به اینکه اغلب حسگرها مکانیکی بوده و تابع شرایط محیطی هستند ممکن است گیر کرده و نتوانند داده واقعی را ارسال کنند. بنابراین مراجعه کارشناس ضروری است تا حسگر را از نزدیک ببیند تا در صورت لزوم با ضربه یا تکان دادن، حسگر داده واقعی را نشان بدهد. خطای انسانی در یادداشت کردن داده ها امری اجتناب ناپذیر است. بعنوان مثال براساس درجه اهمیتی که اپراتور برای یک حسگر خاص قائل است عدد خوانده شده را ممکن است گرد کند.

داده‌های جمع‌آوری‌شده در هر روز از منابع مختلفی که بعضی از آنها در جدول (۱) نشان داده شده اند جمع‌آوری و بصورت فایلهای اکسل ذخیره می‌گردند. هر فایل دارای ویژگیهای مختلفی است که ویژگی واحد متبوع خود را بیان می‌کند. بعنوان نمونه در گاورنر فقط دماسنج روغن تانک ثبت می‌شود و در فایل توربین دماسنج روغن تانک و دماسنج پد ثبت می‌شود.

جدول (۱) : اسامی بعضی از فایلهای گزارش گیری بصورت اکسل

نام فایل
ارتعاشات واحد
پارامترهای تولید واحد نیروگاه
ترانس اصلی
ترانس کمکی نیروگاه
توربین
ژنراتور
گاورنر
کولینگ نیروگاه

۵-۲- پیش پردازش داده ها

برای اینکه از فایلهای جمع آوری شده در مرحله قبل اطلاعات ارزشمندی بدست آوریم که قابل تحلیل و آنالیز باشند بایستی این فایلها را با یکدیگر تجمیع کنیم که این مرحله جزو پیش پردازش داده‌ها محسوب می‌شود.

هریک از مقادیر اندازه گیری شده توسط حسگرها و دماسنج ها به عنوان یک ویژگی در نظر گرفته می شود. مجموعه ویژگیها که به صورت همزمان اندازه گیری می شوند یک رکورد را تشکیل می دهند.

پس از تجمیع فایلها، عملیات پاکسازی داده ها آغاز می گردد. پاکسازی داده ها فایل اطلاعات را به گونه ای آماده می کند که انجام الگوریتم های داده کاوی میسر گردد و شامل مراحل زیر است:

ساماندهی ویژگیهای خالی : پس از تجمیع بسیاری از ویژگیهایی که مقدار آنها خالی بود (بعلت خطای اندازه‌گیری یا سهل انگاری اپراتورهای ثبت مقادیر، خالی مانده اند) حذف می‌شوند تا کار با رکوردهایی انجام شود که همه ویژگیهای آن مشخص باشد.

حذف نویز و داده های پرت : یکی از موارد مهم دیگر حذف مقادیر نویز و داده‌های پرت در میان ویژگیهاست که در پیش پردازش انجام گرفت. برای حذف نویز از روش خوشه‌بندی استفاده گردید و رکورد مربوط به خوشه‌های تک عضوی (داده پرت) حذف شدند.

گسسته سازی : تبدیل مقادیر پیوسته به گسسته برای مقادیر ویژگیهایی که پیوسته هستند، گسسته سازی نامیده می شود. گسسته‌سازی به این منظور صورت می‌پذیرد که در مسائل پیچیده داده‌کاوی با حجم بالای رکوردها و ویژگیها، سختی مسئله کاهش یابد و سرعت کار بالاتر رود. روش مورد استفاده در این تحقیق گسسته-سازی بخشی است یعنی کل محدوده مقادیر پیوسته به n بخش بصورت بازه های[a,b] تقسیم می‌شوند. a وb مقادیر ابتدا و انتهای هر بخش می‌باشند. به هر بازه یک مقدار عددی از ۱ تا n انتساب داده می شود. سپس مقدار ویژگی بر حسب اینکه در کدام بازه قرارگیرد با عدد بازه مربوطه جایگزین می گردد.

باینری سازی : از آنجا که عملیات داده کاوی مورد نظر ما استخراج قوانین انجمنی بود، رکوردها باید به صورت تراکنشی باشند. و حضور و عدم حضور یک ویژگی در رکورد مطرح می شود. برای این منظور هر ویژگی گسسته با مقادیر ۱ تا n را به n ویژگی باینری با مقادیر صفر و یک تبدیل می کنیم.برای مثال چنانچه ویژگی A دارای ۴ بازه باشد که با مقادیر ۱ تا ۴ مقدار دهی میگردد، و رکوردی را در نظر بگیریم که مقدار ویژگی A آن ۳ باشد، در آنصورت ویژگی A به ۴ ویژگی A1, A2, A3, A4 بسط داده می شود که مقدار ویژگی A3 آن برابر با ۱ و بقیه آنها صفر می شوند.

تعیین دسته رکورد: برای تعیین دسته یک رکورد ، براساس نظر کارشناسی متخصصین نیروگاهها یک نیروگاه برق آبی در سه حالت کاری فعالیت می‌کند:

۱- معمولی (زمانی که نیروگاه در حال تولید بوده و مشکلی مشاهده نمی شود)
۲- هشدار یا Alarm (برای عملکرد هر بخش یک آستانه هشدار وجود دارد که در این شرایط کارشناسان هشدار داده شده را بررسی کرده و تصمیم مناسب اتخاذ می‌کنند)
۳- خاموشی اضطراری یا Trip (زمانی که نیروگاه از مدار تولید برق خارج می باشد)

با کمک ابزارهای داده‌کاوی داده ها را براساس سه دسته، طبقه بندی می‌کنیم. مقدار ۱ برای معمولی مقدار ۲ برای هشدار و مقدار ۳ برای خاموشی اضطراری در نظر گرفته شد.

پس از مرحله پیش پردازش فایل حاصل دارای ۳۹۵۷ رکورد می باشد.

۵-۳- انتخاب ویژگیهای تاثیر گذار و کاهش ابعاد داده ها

تعداد ویژگیهایی که در یک لاگ شیت یادداشت می شود بسیار زیاد است و ازطرف دیگر واحدهای مختلف لاگ شیت مخصوص به خود داشته و همانگونه که در جدول (۱) بیان شد فایلهای زیادی وجود دارند که با تجمیع در یک فایل، حجم فایل نهایی بسیار بالا رفته و پردازش بسیار زمان‌بر خواهد بود. از طرفی درجه اهمیت همه ویژگیها با هم برابر نیست و همچنین بعضی از ویژگیها در بازه مورد بررسی بدون تغییر هستند(ثابت). پس به راحتی می‌توان ویژگیهای ثابت را حذف کرد و در مورد حذف ویژگیهای کم اهمیت تر تصمیم‌گیری کرد.

برای کاهش ابعاد داده‌ها و حذف فیلدهای کم اهمیت تر روشهای مختلفی مانند استفاده از درخت تصمیم گیری یا تحلیل مولفه های اصلی وجود دارند[۱۶]. هدف مشترک همه روشها، کاهش حجم داده‌ها جهت افزایش سرعت عملیات می‌باشد بشرطی که دقت کار کاهش نیابد. در بین ویژگیهای بسیار زیادی که در لاگ شیت نیروگاه وجود دارد همگی ویژگیها هم اولویت و دارای ارزش برابر نیستند. در این تحقیق به دو روش ویژگیهای تاثیرگذار را انتخاب کردیم.

روش اول: استفاده از روش PCA[۵].

روش PCA بهترین روش برای کاهش ابعاد داده هاست. در این روش که مبتنی بر جبر خطی و مقادیر ویژه است، برای هر ویژگی ضریب تاثیر محاسبه می گردد. با حذف ویژگیهای با ضرایب کم‌اهمیت بدست آمده از این تبدیل، اطلاعات از دست رفته نسبت به روشهای دیگر کمتر است. در این روش محورهای مختصات جدیدی برای داده‌ها تعریف شده و داده‌ها براساس این محورهای مختصات جدید بیان می-شوند. اولین محور باید در جهتی قرار گیرد که واریانس داده‌ها بیشینه شود (یعنی در جهتی که پراکندگی داده‌ها بیشتر است). دومین محور باید عمود بر محور اول به گونه‌ای قرار گیرد که واریانس داده‌ها بیشینه شود. به همین ترتیب محورهای بعدی عمود بر تمامی محورهای قبلی به گونه‌ای قرار می‌گیرند که داده‌ها در آن جهت دارای بیشترین پراکندگی باشند[۱۶].

در جدول (۲) نمایی از ویژگیهای تاثیر گذاری که با کمک وزن دهی و نظارت کارشناسان جمع آوری شده است نشان داده شده است.

جدول (۲): ویژگیهای تاثیرگذار در داده های جمع آوری شده

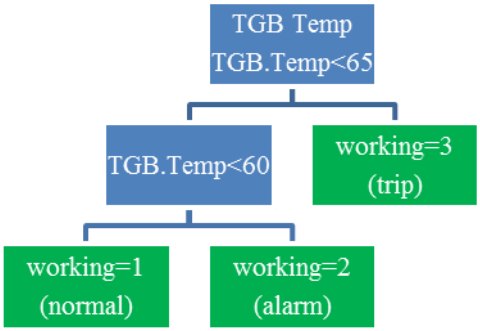
نام ویژگی	توضیح	نام بخش
TGB Temp	دماسنج TGB	توربین
Winding Temp	دماسنج سیم پیچ	ترانس اصلی
CB.A	یاتاقان کمباین	ارتعاشات یاتاقانها
CB.B	یاتاقان کمباین	ارتعاشات یاتاقانها
UGB.A	یاتاقان راهنمای بالایی	ارتعاشات یاتاقانها
UGB.B	یاتاقان راهنمای بالایی	ارتعاشات یاتاقانها
TGB.A	یاتاقان توربین	ارتعاشات یاتاقانها
TGB.B	یاتاقان توربین	ارتعاشات یاتاقانها
Working	کارکرد نیروگاه (۱ =عادی ، ۲= هشدار ، ۳=تریپ)	کارکرد کلی نیروگاه

روش دوم: برای تعیین ویژگیهای تاثیرگذار از درخت تصمیم گیری استفاده می کنیم. درخت تصمیم گیری فرایند طبقه بندی داده ها براساس مقادیر ویژگیها می باشد. در این روش ویژگیهای جدا کننده قوی در ریشه و شاخه های بالاتر قرار گرفته و ویژگیها با مقدار جدا کننده ضعیف نزدیک به برگها قرار می گیرند. براساس تعریف سه دسته معمولی و هشدار و خاموشی اضطراری درخت تصمیم گیری ویژگی دماسنج TGP را به عنوان ویژگیهای با قدرت تفکیک پذیری بالا معرفی می‌نماید(شکل۲).

حد آستانه برای ورود به یک دسته خاص در جدول ۳ نشان داده شده است.

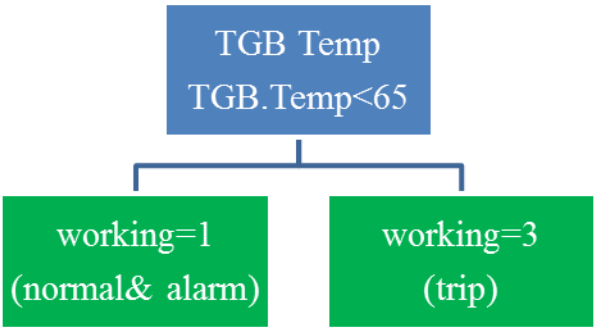
جدول(۳): آستانه تریپ و هشدار در نیروگاه برق آبی مارون بهبهان

ویژگی	Alarm		Trip	
	کانال A	کانال B	کانال A	کانال B
CB	175𝜇m	175𝜇m	200𝜇m	200𝜇m
UGB	175𝜇m	175𝜇m	200𝜇m	200𝜇m
TGB	310𝜇m	310𝜇m	350𝜇m	350𝜇m
دماسنج TGB	60°		65°	
دماسنج سیم پیچ	89°		99°	



شکل(۲): درخت تصمیم سه دسته جهت یافتن ویژگیهای جداکننده

در یک آزمایش دیگر با ادغام دسته های معمولی و هشدار در یکدیگر بعنوان یک دسته (دسته۱) و خاموشی اضطراری به عنوان دسته۲ مجدد طبقه بندی را اعمال می کنیم. طبقه بندی با دو دسته، تصمیم گیری را ساده تر می‌کند و با یک بررسی شرط ساده ما را به دسته مورد نظر می‌رساند. در این حالت نیز مانند سه دسته، ویژگی دماسنج TGP تعیین کننده و قدرت تفکیک سازی خوبی دارد(شکل۳).



شکل(۳): درخت تصمیم دو دسته جهت یافتن ویژگیهای جداکننده

در حالت کلی ویژگی دماسنج TGP قابلیت تفکیک بسیار خوبی دارد و از بررسی درخت تصمم شکل (۱) و (۲) نیز این موضوع مشخص است اما اگر این ویژگی را در ساخت درخت تصمیم دخالت ندهیم(فرض کنیم در یک شرایط خاص و عملیاتی نتوانیم دمایTGP را اندازه بگیریم) و بخواهیم با بقیه ویژگیها درخت را بسازیم با کمک درخت تصمیم مشابه شکل(۲) یا (۳) مشخص می شود که ویژگی دمای windingدر ریشه قرار می گیرد و قابلیت جداکنندگی بهتری نسبت به بقیه دارد. پس قوانینی که ویژگیهای دمایTGP و دمای winding را بررسی می‌کنند اهمیت بیشتری برای طبقه بندی دارند.

۵-۴- قواعد انجمنی

قوانین انجمنی به دنبال کشف مشارکت ها و همبستگی ها و روابط جالب بین داده های موجود در یک پایگاه داده می باشد. کشف الگوهای جالب از میان مقدار متنابهی از تراکنش های تجاری یا صنعتی می تواند در بسیاری از فرایندهای تصمیم گیری راهگشا باشد. در مسأله جاری، کشف الگوهای جالب در میان اطلاعاتی که از حسگرها و دماسنجها جمع آوری شده می تواند کارشناسان فنی را در اتخاذ تصمیمات صحیح تر یاری دهد.

فرض کنید $I=\{I_1,I_2,I_3,\ldots I_m\}$ مجموعه اقلام باشد و D پایگاه داده ای است که شامل مجموعه‌ای از تراکنش‌ها است که هر تراکنش دارای یک شناسه و همچنین مجموعه اقلامی است که زیر مجموعه ای از I باشد. اگر A و B از اعضای مجموعه اقلام I باشد. یک قانون انجمنی بصورت A➔B بیان می شود. این قانون بیان می کند که وقوع ویژگی B تحت تاثیر وقوع ویژگی A است. سمت چپ و راست قانون می تواند ترکیبی از ویژگیها باشد.

۵-۵- تولید قوانین انجمنی وارزیابی آنها

بطور کلی کشف قوانین انجمنی یک فرایند دو مرحله ای است:

مرحله اول : یافتن کلیه مجموعه های داده های مکرر(پرتکرار) .

مرحله دوم : تولید قوانین انجمنی قوی از مجموعه داده های مکرر.

برای یافتن مجموعه های پرتکرار و همچنین برای ارزیابی قوانین استخراج شده نیاز به معرفی دو معیار کمی مهم به نام پشتیبان [۶] و اطمینان [۷] می باشد.

میزان پشتیبان و اطمینان قانون A=>B در مجموعه تراکنشی D بصورت زیر تعریف می شود.

$$\text{Support}(A=>B)=P(A\cup B) = P(A,B) / \text{no. of transactions} \quad (۱)$$

$$\text{Confidence}(A=>B)=P(B|A) =P(A , B) / P(A) \quad (۲)$$

در مرحله اول با استفاده از الگوریتم معروف Apiriori مجموعه ویژگیهای پرتکرار را بدست می آوریم. این الگوریتم مجموعه هایی از ویژگیها را معرفی می کند که تکرار آنها حداقل برابر یا بزرگتر از حد آستانه پشتیبان [۸] باشد. پشتیبان برای یک مجموعه ویژگی A از رابطه زیر محاسبه می گردد.

$$A= \{a1, a2 , \ldots ak\}$$
$$\text{Support } (A) = P(a1,a2\ldots.,ak) /\text{no. of transactions} \quad (۳)$$

در مرحله دوم کلیه ترکیبات ممکن از ویژگیهای درون یک مجموعه پرتکرار مانند A برای ساخت قوانین به صورت P ➔Q استفاده می شود. P و Q هر دو زیر مجموعه های غیر تهی از A هستند که اجتماع آنها برابر با خود A می شود.

تعداد قوانین تولید شده ممکن است بسیار زیاد باشد. اما تنها قوانینی به عنوان یک قانون تاثیر گذار مورد قبول واقع می شوند که دارای حداقل مقدار سطح آستانه پشتیبانی و سطح آستانه اطمینان[۹] باشند.

چالش عمده در کاوش مجموعه قوانین انجمنی این است که اغلب تعداد بسیار زیادی از مجموعه ویژگیها هستند که پشتیبان آنها بزرگتر یا مساوی حداقل آستانه پشتیبان است. به خصوص هنگامی که آستانه پشتیبان با مقدار کوچکی تنظیم شده باشد. البته تنظیم مقدار کوچک برای آن این امتیاز را دارد که قوانین بیشتری که شاید از دید ما پنهان مانده باشند را مشخص می کند که ممکن است مورد توجه کارشناس فنی نیروگاه قرار بگیرد.

معیارهای پشتیبان و اطمینان برای انتخاب قوانین انجمنی جالب و قوانین ناجالب کافی نیستند. در واقع پشتیبان و اطمینان کمک می-کنند تراکنش‌های بهتر انتخاب شوند ولی بررسی اینکه در تراکنش انتخابی چه میزان همبستگی بین مقدم و تالی قانون وجود دارد با کمک معیارهای پشتیبان و اطمینان امکانپذیر نیست.

جهت تکمیل و تقویت دقت انتخاب می توان از معیار همبستگی نیز استفاده کرد. یکی از معیارهای همبستگی، لیفت[۱۰] می باشد[۱۶]. براساس تعاریف صورت گرفته در روابط (۱) و (۲) معیار لیفت بصورت رابطه (۴) تعریف می شود.

$$\text{Lift}(A,B) = P(A\cup B)/P(A)P(B) \quad (۴)$$

چنانچه نتیجه رابطه (۴) کوچکتر از یک باشد، مجموعه های A و B همبستگی منفی دارند. بدین معنی که محتمل است حضور و پیشامد یکی باعث نبودن دیگری شود. اگر نتیجه رابطه (۳) مقداری بزرگتر از یک باشد، پیشامدهای A و B دارای همبستگی مثبت هستند و حضور یکی دلالت بر بودن دیگری خواهد داشت. اگر نتیجه رابطه (۴) برابر با مقدار یک باشد، A و B مستقل هستند و میان آنها وابستگی وجود ندارد. در واقع معیارلیفت نشان می‌دهد که چگونه درجه یکی از پیشامدها باعث افزایش رخداد و پیشامد دیگر است.

۵-۶- نتایج بدست آمده

داده های پاکسازی شده نیروگاه برق آبی مارون که شامل ۳۹۵۷ رکورد و ۸ ویژگی انتخاب شده بودند در تولید قوانین انجمنی مورد استفاده قرار گرفتند. این داده ها دارای یک ویژگی اضافه دیگر تحت عنوان دسته نیز می باشند. دسته ها شامل معمولی و هشدار و خاموشی اضطراری می باشد. هدف از انجام این پروژه بررسی حالت هشدار است. چون حالت معمولی که عملکرد عادی است و ما به دنبال پیدا کردن روابط بین ویژگیها در حالت هشدار هستیم به این هدف که به حالت خاموشی اضطراری نرسیم. بنابراین قوانینی که به حالت هشدار منتج میشود را بیشتر مورد توجه قرار می دهیم.

به کمک الگوریتمApiriori و استفاده از پکیج Rattle قوانینی که دارای حداقل پشتیبان باشند استخراج شده اند. حداقل سطح پشتیبان ۰,۱ در نظر گرفته شده است. علت کوچک در نظر گرفتن آن ایجاد شانس برای تولید قوانین بیشتر است. تعداد این قانونها ۴۱ عدد می باشد که در جدول (۴) همراه با میزان سطح اطمینان و مقدار لیفت قانون آورده شده اند.جدول براساس مقدار بالاترین سطح اطمینان مرتب شده است. بسته به آنکه حداقل آستانه سطح اطمینان چقدر در نظر گرفته شود ، تعدادی از قوانین انتخاب می شوند. لازم به ذکر است که در صورتی بین مقدم و تالی یک قانون ارتباط مستقیم وجود دارد که مقدار لیفت آن قانون بزرگتر از یک باشد.

از آنجا که هدف از انجام این پژوهش تحلیل داده ها در حالت هشدار بود، تنها قوانینی که تالی آنها به حالت هشدار ختم می شود و دارای سطح آستانه قابل قبول و لیفت بزرگتر از یک باشند در نظر گرفته شده اند.

قوانین استخراج شده باید توسط کارشناسان فنی مورد ارزیابی قرار گیرند. بعضی از قوانین حقایق گذشته را مجددا اثبات می‌کنند و بعضی از آنها دانش نوینی است که از میان داده ها استخراج می‌گردند. ۴ قانون زیر دانش نوینی است که توسط کارشناسان نیروگاه برق مارون شناسایی شده است.

1) if TGB.temp<=62 class= between (normal,alarm]
2) if UGB.B<=85 and TGB.temp<=63 then
class= between (normal,alarm]

3) if UGB.B<= 85 then class=between (normal,alarm]
4) if UGB.A<=93 then class= between (normal,alarm]

این قوانین به زبان ساده بصورت زیر نوشته می‌شوند:

۱- اگر شرایط بگونه‌ای مهیا شود که دمای یاتاقان توربین کمتر از ۶۲ درجه سانتیگراد باشد نیروگاه نهایتا هشدار می‌دهد ولی هیچگاه تریپ بوجود نمی‌آید.

۲- اگر دمای یاتاقان توربین کمتر از ۶۳ درجه و لرزش یاتاقان بالای ژنراتور کمتر از ۸۵ میکرومتر باشد نیروگاه نهایتا هشدار می‌دهد ولی هیچگاه تریپ بوجود نمی‌آید.

۳- اگر لرزش یاتاقان بالای کانال B ژنراتور کمتر از ۸۵ میکرومتر باشد وارد تریپ نمی‌شویم.

۴- اگر لرزش یاتاقان بالای کانال A ژنراتور کمتر از ۹۳ میکرومتر باشد وارد تریپ نمی شویم.

۶- نتیجه‌گیری

در این مقاله با بکارگیری کشف قوانین انجمنی تلاش شد روابطی معنادار از ویژگی هایی که حسگرها در نیروگاه برق آبی مارون بهبهان همه روزه آرشیو می‌کنند یافته شود تا بکمک آنها بتوان مکانیزمی کمکی برای تصمیم گیری در مورد حالت خاموشی اضطراری یا فعالیت عادی نیروگاه ایجاد کرد. این مکانیزم با استفاده از تکنیک داده‌کاوی دقت تصمیم گیری را افزایش داده و از آسیب به تجهیزات در تصمیمات اشتباه یا دیرهنگام کمک می‌کند.

سپاسگزاری

در پایان از مدیریت و کارشناسان شرکت بهره برداری از سد و نیروگاه برق آبی مارون بهبهان که ما را در انجام این تحقیق یاری کردند تقدیر و تشکر می نماییم.

مراجع

[۱] م. خواجه پور، م.عارف پور، م. غیاثی، "نیروگاههای برق آبی: مزایا،معایب و مقایسه با دیگر روشهای تولید انرژی"، اولین کنفرانس بین المللی و سومین کنفرانس ملی سد و نیروگاههای برق آبی ، تهران، ایران، ۱۳۹۰.

[۲] م.عباسپور، ع.کرباسی، ع.پهلوان، ح.رحیمی پور انارکی، س.مطهری، "مطالعات پایه و بررسی اثرات زیست محیطی نیروگاههای برق آبی(نمونه موردی نیروگاه سردآبرود)"، فصلنامه علمی پژوهشی علوم و تکنولوژی محیط زیست، دوره دوازدهم، شماره دو، ۱۳۸۹ .

[۳] ح.فیروزی: آشنایی با بهره برداری از نیروگاه های برق ابی، چاپ اول ، نشر فدک ایساتیس، تهران، ۱۳۹۲.

[۴] م.پارسا، ع.وحیدیان کامیاد، م.نقیبی نیستانی، "افزایش راندمان بویلر نیروگاه حرارتی ذوب آهن اصفهان با استفاده از داده کاوی فازی" ، بیست و ششمین کنفرانس بین المللی برق ،تهران ، ایران ، ۱۳۹۰.

[5] T. Onoda, N. Ito, H. Yamasaki, "Unusual Condition Mining for Risk Management of Hydroelectric Power Plants", Proc. of ICDM-06 Workshops Workshop on Risk Mining, 2006.Plamondon, R., Lorette, G., *"Automatic Signature Verification and Writer Identification - The State of the Art"*, Pattern Recognition, Vol. 22, pp. 107-131, 1989.

7. confidence
8. min_support
9. min_confidence
10. lift

[6] T. Onoda, N. Ito,H.Koreeda,"Interactive trouble condition sign mining for hydroelectric power plants", World Automation Congress (WAC), 2010.
[7] J.A. Huang, S. Harrison, G. Vanier, A. Valette, L. Wehenkel, "Application of data mining to optimize settings for generator tripping and load shedding system in emergency control at Hydro-Québec", COMPEL - The international journal for computation and mathematics in electrical and electronic engineering, Vol. 23 Iss: 1, pp.21 – 34,2004.
[8] Selak, L., Butala, P., Sluga, A., "Condition monitoring and fault diagnostics for hydropower plants", Comput. Ind. 65(6), 924–936 (2014),2014.
[9] Jiang Liu, Guangzhen Geng, "Fault Prediction for Power Plant Equipment Based on Support Vector Regression", 2015 8th International Symposium on Computational Intelligence and Design,2015.
[10] T. Foshch, J. Machado, F. Portela, M. Maksimov, O. Maksimova, 2017, "Comparison of Two Control Programs of the VVER-1000 Nuclear Power Unit Using Regression Data Mining Models", ISSN 2073-6231. Ядерна та радіаційна безпека 3(75).
[11] Aziah Khamis, Yan Xu, Azah Mohamed,2017," Comparative Study in Determining Features Extraction for Islanding Detection using Data Mining Technique: Correlation and Coefficient Analysis", International Journal of Electrical and Computer Engineering (IJECE) , ISSN: 2088-8708,Vol. 7, No. 3, June 2017, pp. 1112 – 1124.
[12] Chunming Tu, Xi He, Zhikang Shuai, Fei Jiang,2017," Big data issues in smart grid – A review" ," Renewable and Sustainable Energy Reviews, 79 (2017) 1099–1107.
[13] Benjamin A. Shimray, Kh. Manglem Singh,Thongam Khelchandra,R. K.Mehta,2017," Ranking of Sites for Installation of Hydropower Plant Using MLP Neural Network Trained with GA: A MADM Approach" , Hindawi Computational Intelligence and Neuroscience Volume 2017, Article ID 4152140, 8 pages ,https://doi.org/10.1155/2017/4152140.
[14] Benjamin A Shimray, Kh. Manglem Singh, Thongam Khelchandra,R. K Mehta,2017," A New MLP–GA–Fuzzy Decision Support System for Hydro Power Plant Site Selection", Arab J Sci Eng, DOI:10.1007/s13369-017-2885-4
[15] Muhammad Zaman et al.,2017," Optimization of Mangala Hydropower Station, Pakistan, using Optimization Techniques", MATEC Web of Conferences 136, 02010 (2017), DOI: 10.1051/matecconf/201713602010D2ME 2017.
[16] Jiawei Han,Micheline Kamber ,Jian Pei,Data Mining:Concepts and Techniques(3rd Edition), Morgan Kaufmann Press, 2011.

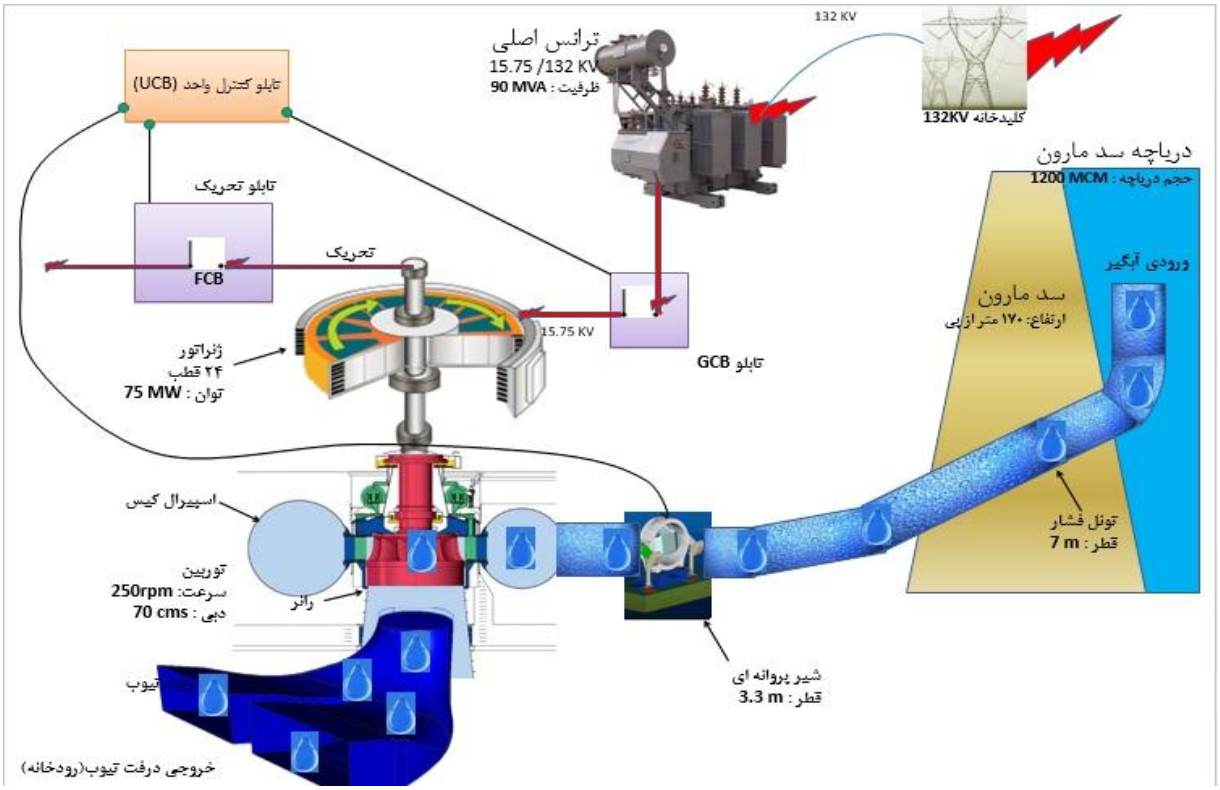
زیرنویس‌ها

1.Governer
2.Accumulator
3.Hydro-Québec
4.Support Vector Machine
5.Principal Component Analysis
6.support

جدول (۴) : تعدادی از قوانین انجمنی کشف شده

شماره قانون	قانون یافته شده	پشتیبان	اطمینان	لیفت
.۱	{TGB.temp=(64.8,65]} => {working=(2,3]}	0.1036475	1.0000000	3.8140496
.۲	{ TGB.temp=(62.8,63]} => { working=(1,2]}	0.2354641	1.0000000	1.3907584
.۳	{TGB.temp=(62.2,62.5]} => {working=(1,2]}	0.1227880	1.0000000	1.3907584
.۴	{UGB.B=(84.5,85],TGB.temp=(62.8,63]} => {working=(1,2]}	0.1209823	1.0000000	1.3907584
.۵	{TGB.temp=(61.5,62]} => {working=(1,2]}	0.1112315	1.0000000	1.3907584
.۶	{CB.A=(81.5,82]}=> {working=(1,2]}	0.1072589	0.9054878	1.2593148
.۷	{UGB.B=(84.5,85]} => { working=(1,2]}	0.3174431	0.8634578	1.2008611
.۸	{UGB.A=(92.6,93], UGB.B=(84.5,85]} => { working=(1,2]}	0.1426508	0.8494624	1.1813969
.۹	{UGB.A=(91.5,92],UGB.B=(84.5,85]} => {working=(1,2]}	0.1097869	0.8467967	1.1776896
.۱۰	{ UGB.A=(92.6,93]} => { working=(1,2]}	0.2986638	0.7573260	1.0532575
.۱۱	{UGB.B=(83.5,84]} => {working=(1,2]}	0.2184904	0.6477516	0.9008660
.۱۲	{ UGB.A=(91.5,92]} => { working=(1,2]}	0.2531600	0.6253345	0.8696892
.۱۳	{working=(2,3]} => {UGB.A=(91.5,92]}	0.1455399	0.5550964	1.3711525
.۱۴	{TGB.temp=(62.8,63]} => {UGB.B=(84.5,85]}	0.1209823	0.5138037	1.3975662
.۱۵	{TGB.temp=(62.8,63],working=(1,2]} => {UGB.B=(84.5,85]}	0.1209823	0.5138037	1.3975662
.۱۶	{ UGB.A=(92.6,93], working=(1,2]} => {UGB.B=(84.5,85]}	0.1426508	0.4776300	1.2991723
.۱۷	{UGB.B=(84.5,85]} => {UGB.A=(92.6,93]}	0.1679307	0.4567780	1.1582585
.۱۸	{ UGB.B=(84.5,85],working=(1,2]} => {UGB.A=(92.6,93]}	0.1426508	0.4493743	1.1394848
.۱۹	{ working=(1,2]} => { UGB.B=(84.5,85]}	0.3174431	0.4414867	1.2008611
.۲۰	{UGB.B=(83.5,84]} => {UGB.A=(91.5,92]}	0.1480679	0.4389722	1.0843121
.۲۱	{UGB.A=(91.5,92],working=(1,2]} => {UGB.B=(84.5,85]}	0.1097869	0.4336662	1.1795891
.۲۲	{working=(2,3]} => {UGB.B=(83.5,84]}	0.1126761	0.4297521	1.2740722
.۲۳	{UGB.A=(92.6,93]} => {UGB.B=(84.5,85]}	0.1679307	0.4258242	1.1582585
.۲۴	{ working=(1,2]} => { UGB.A=(92.6,93]}	0.2986638	0.4153692	1.0532575
.۲۵	{working=(2,3]}=> {TGB.temp=(64.8,65]}	0.1036475	0.3953168	3.8140496

۲۶.	{UGB.B=(84.5,85],working=(1,2]} => {TGB.temp=(62.8,63]}	0.1209823	0.3811149	1.6185693
۲۷.	{UGB.B=(83.5,84]} => {UGB.A=(92.6,93]}	0.1235103	0.3661670	0.9284950
۲۸.	{UGB.A=(91.5,92]} => {UGB.B=(83.5,84]}	0.1480679	0.3657449	1.0843121
۲۹.	{UGB.A=(91.5,92]} => {working=(2,3]}	0.1455399	0.3595004	1.3711525
۳۰.	{UGB.B=(84.5,85]} => {UGB.A=(91.5,92]}	0.1296497	0.3526523	0.8710920
۳۱.	{ working=(1,2]} => { UGB.A=(91.5,92]}	0.2531600	0.3520844	0.8696892
۳۲.	{UGB.B=(84.5,85],working=(1,2]} => {UGB.A=(91.5,92]}	0.1097869	0.3458476	0.8542836
۳۳.	{UGB.B=(83.5,84]} => {working=(2,3]}	0.1126761	0.3340471	1.2740722
۳۴.	{UGB.B=(84.5,85]}=> {TGB.temp=(62.8,63]}	0.1209823	0.3290766	1.3975662
۳۵.	{ working=(1,2]} => { TGB.temp=(62.8,63]}	0.2354641	0.3274736	1.3907584
۳۶.	{UGB.A=(91.5,92]} => {UGB.B=(84.5,85]}	0.1296497	0.3202498	0.8710920
۳۷.	{UGB.A=(92.6,93]} => {UGB.B=(83.5,84]}	0.1235103	0.3131868	0.9284950
۳۸.	{working=(1,2]} => {UGB.B=(83.5,84]}	0.2184904	0.3038674	0.9008660
۳۹.	{working=(1,2]} => {TGB.temp=(62.2,62.5]}	0.1227880	0.1707685	1.3907584
۴۰.	{working=(1,2]} => {TGB.temp=(61.5,62]}	0.1112315	0.1546961	1.3907584
۴۱.	{working=(1,2]} => {CB.A=(81.5,82]}	0.1072589	0.1491713	1.2593148



شکل(۱): فرایند کار نیروگاه برق آبی مارون بهبهان بصورت نمادین