



دانشگاه صنعتی شاهرود

DOR:

[20.1001.1.23223146.1399.7.2.1.5](https://doi.org/10.1001.1.23223146.1399.7.2.1.5)

نشریه علمی-فنی «مندی برق»

دوره ۷ - شماره ۲

پاییز و زمستان ۱۳۹۹

صفحات ۶۴ الی ۸۷

ISSN: 2322-3146

<http://journals.sut.ac.ir/jnsee>

# ارائه مدلی برای هشداردهی وقوع خطا در خطوط انتقال قدرت با استفاده از روش‌های یادگیری ماشین با در نظر گرفتن شرایط آب و هوایی

علی قائمی<sup>۱</sup> و امین صفری<sup>۲</sup>

<sup>۱</sup> دانشجوی دکتری مهندسی برق-قدرت، دانشکده فنی مهندسی، دانشگاه شهید مدنی آذربایجان [ali.ghaemy@gmail.com](mailto:ali.ghaemy@gmail.com)

<sup>۲</sup> نویسنده مسئول، دانشیار مهندسی برق-قدرت، دانشکده فنی مهندسی، دانشگاه شهید مدنی آذربایجان  
[asafari1650@yahoo.com](mailto:asafari1650@yahoo.com)

(تاریخ دریافت مقاله: ۱۳۹۸/۰۹/۲۴ تاریخ پذیرش مقاله: ۱۳۹۹/۰۹/۲۱)

## چکیده

### واژه‌های کلیدی

خطا، خط انتقال،

یادگیری ماشین،

پیش‌بینی.

توان بالای عبوری از خطوط انتقال و هزینه‌های هنگفت ناشی از بروز خطاها در این خطوط باعث توجه ویژه محققین به مسائل مربوط به حفاظت در این حوزه شده است. ضعف‌های موجود در روش‌های حفاظت سنتی و وابستگی شدید آن‌ها به شرایط بهره‌برداری سیستم اهمیت موضوع تشخیص پیش از موعد خطا و پیش‌بینی آن با روش‌های جدید را دوچندان می‌کند. تشخیص به‌موقع و صدور هشدارهای مربوط به احتمال وقوع خطا با تحلیل داده‌ها و اطلاعات به‌دست آمده از سیستم و بررسی روابط بین پارامترهای مختلف قابل انجام است. در این مقاله از روش‌های مبتنی بر یادگیری ماشین که با دقت مناسب و مستقل از ناحیه عملکردی سیستم توان پیش‌بینی وقوع خطا را دارند استفاده شده است. برای بررسی عملکرد مدل‌ها تعداد زیادی داده در شرایط بهره‌برداری متنوع تولید شده و به‌عنوان ورودی به الگوریتم‌های تحت بررسی داده شده است. همچنین تاثیر شرایط آب و هوایی مختلف به‌عنوان یکی از عوامل مهم در وقوع خطاها در خطوط انتقال در این مطالعه لحاظ شده است. به‌منظور افزایش جامعیت، بررسی صحت و مقایسه‌پذیری نتایج از سه روش KNN، SVM و درخت تصمیم در دو حالت (داده‌های نامتعادل و متعادل‌سازی شده در دسته‌بندی‌های موجود) استفاده شده و نتایج آن ارائه گردیده است. شبیه‌سازی‌ها و مدل‌سازی‌های ارائه شده در این مقاله با استفاده از نرم‌افزارهای Python و MATLAB انجام شده است.



Sahand University  
of Technology

DOR:

[20.1001.1.23223146.1399.7.2.1.5](https://doi.org/10.1001.1.23223146.1399.7.2.1.5)

Journal of Nonlinear  
Systems in Electrical  
Engineering

Vol.7, No.2

Autumn and Winter 2020

ISSN: 2322 – 3146

<http://journals.sut.ac.ir/jnsee>

# A Machine Learning Based Fault Warning Model for Transmission Lines Considering Weather Conditions

Ali Ghaemi<sup>1</sup> and Amin Safari<sup>2</sup>

<sup>1</sup>Ph.D. Candidate, Department of Electrical Engineering, Azerbaijan Shahid Madani University, Tabriz, Iran, Email: ali.ghaemy@gmail.com

<sup>2</sup>**Corresponding author**, Department of Electrical Engineering, Azerbaijan Shahid Madani University, Tabriz, Iran, Email: safari@azaruniv.ac.ir

## ABSTRACT

### Keywords

Fault,  
Transmission lines,  
Machine learning,  
Prediction.

The high power passing through transmission systems and the high costs due to the fault occurrence in these lines have encouraged researchers to pay special attention to protection issues in this area. The limitations and deficiencies of traditional protection methods and their strong dependencies on the system operating conditions doubles the importance of early fault detection and its prediction utilizing new techniques. Timely detection and warning issuance toward the possibility of fault occurrence can be accomplished by analyzing the data and information obtained from the system and examining the relationships between different parameters. In this paper, machine learning methods are used, which have the ability to predict the occurrence of faults with appropriate accuracy independent of the operating area of the system. To evaluate the performance of the models, a large amount of data has been generated in various operating conditions and applied as input to the algorithms under study. Also, the effects of different weather conditions as one of the important factors have been considered. *For the sake of greater generality, accuracy check, and comparability of the results*, three methods including KNN, SVM, and decision tree in two modes (unbalanced and balanced data in the existing classes) have been used, and the outcomes have been presented. The simulations and modeling presented in this paper have been implemented using Python and MATLAB.

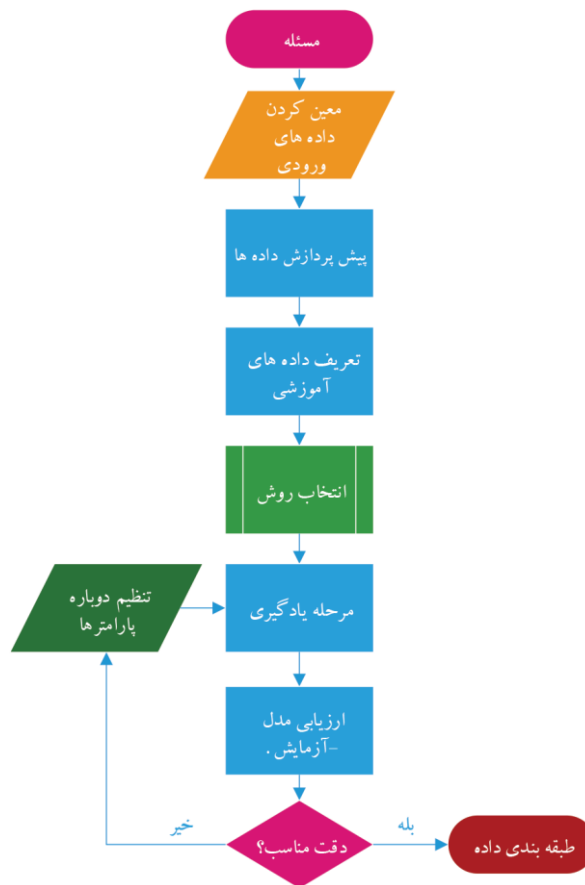
## ۱- مقدمه

خطوط انتقال هوایی یکی از مهم‌ترین عناصر سیستم‌های قدرت می‌باشند. با تجدید ساختار سیستم قدرت و ایجاد بازارهای برق رقابتی اهمیت خطوط انتقال برای غلبه بر موضوع انحصار دوچندان شده است. عملکرد سیستم قدرت در مرزهای بهره‌برداری باعث افزایش احتمال بروز خطا در المان‌های شبکه خواهد شد. از طرف دیگر ایجاد هرگونه خطا یا مشکل در خطوط انتقال می‌تواند منجر به تحمیل هزینه‌های اقتصادی و اجتماعی فراوان در سطحی وسیع شود. حفاظت خطوط انتقال یکی از مباحث مهمی است که از ابتدا مورد توجه مهندسين قدرت بوده است [۴-۱]. قواعد حفاظتی سنتی در خطوط انتقال بر پایه محاسبات فازوری و نیز مقایسه‌های فازوری با حالات از پیش تعریف شده استوار است. دقت این روش‌ها در هنگام وقوع خطا و نیز در مواقعی که شرایط بهره‌برداری سیستم متفاوت از وضعیت‌های مورد انتظار است دچار کاهش خواهد شد. از طرف دیگر نقص عملکرد رله‌ها به تنهایی عامل ۷۰٪ از اغتشاشات بزرگ و نیز خاموشی‌های آبخاری می‌باشد [۴]. افزایش ظرفیت سیستم قدرت و نیاز به عملکرد همگام تمامی بخش‌ها در مناطق جغرافیایی مختلف سبب می‌شود تا اهمیت موضوع تشخیص خطا و نیز صدور هشدار بروز خطا در کمترین زمان ممکن با توجه به شرایط بهره‌برداری به عنوان یکی از مسائل مهم جهت حفظ پایداری سیستم مطرح باشد. وجود یک سیستم هشداردهی سریع خطا می‌تواند با پیدا کردن رابطه بین داده‌های پیش از خطا، تحلیل داده‌های مربوط به خود خطا و نیز پارامترهای سیستم قبل از وقوع خطا، احتمال بروز نقص در عناصر سیستم را مشخص نماید. هنگامی که بهره‌برداران هشدارها و اطلاعات مربوط به خطا را پیش از وقوع آن دریافت نمایند می‌توانند با اتخاذ اقدامات پیشگیرانه کنترلی از بروز آن جلوگیری کنند. در سیستم‌های قدرت واقعی به صورت روزانه خطاهای مختلفی اتفاق می‌افتد که می‌توان با جمع‌آوری و تحلیل اطلاعات مربوط به آن‌ها رابطه بین پارامترها و نقایص عناصر سیستم را مشخص نمود.

تشخیص و طبقه‌بندی خطاهای سیستم به روش‌های مرسوم با استفاده از بررسی تغییرات صرف ولتاژ، جریان و امپدانس انجام می‌گیرد [۵،۶]. اما در سال‌های اخیر با توجه به پیشرفت‌های به وجود آمده در حوزه‌ی داده‌کاوی و علوم مربوط به یادگیری ماشین، روش‌های جدید و بهتری برای این منظور معرفی شده و مورد مطالعه قرار گرفته است که فارغ از ناحیه عملکردی سیستم می‌توانند با دقت مناسبی خطاها را تشخیص داده و هشدارهای لازم را صادر نمایند.

اخیراً روش‌ها و مدل‌های یادگیری ماشین به صورت گسترده برای مدل‌سازی، طراحی و پیش‌بینی موارد مختلف در سیستم‌های قدرت مورد استفاده قرار می‌گیرند. سیستم‌های قدرت ترکیبی از عناصر الکتریکی، مکانیکی، شیمیایی و غیره می‌باشند که این موضوع باعث پیچیده‌تر شدن تصمیم‌گیری‌ها به ویژه در مواقع بحرانی می‌شود [۷،۸]. همچنین افزایش استفاده از جمع‌آوری کنندگان داده در سیستم‌های قدرت سبب به وجود آمدن حجم بزرگی از اطلاعات شده است که در حالت کلی می‌توان آن‌ها را به عنوان یک امتیاز و نیز یک چالش در نظر گرفت. با استفاده از روش‌های یادگیری ماشین می‌توان به استخراج اطلاعات با ارزش کمک کرده و تصمیم‌گیری‌های مهم را ساده‌تر نمود. امروزه روش‌های یادگیری ماشین به دلیل دقت و سرعت مناسب به بخش مهم و ضروری در مدل‌های پیش‌بینی بار، پیش‌بینی مصرف و نیز تحلیل تقاضا تبدیل شده‌اند [۹،۱۰]. یادگیری ماشین در واقع شاخه‌ای از علم کامپیوتر بوده و جزء روش‌های هوش مصنوعی است که می‌تواند علاوه بر مسائل ساده، مسائلی که قابلیت توصیف صریح ریاضی ندارند را نیز حل کرده و تحلیل نماید.

حوزه کاربرد این روش‌ها با توجه به ویژگی‌های مطرح شده گستره وسیعی از قبیل تشخیص الگو، مسائل مربوط به طبقه‌بندی، فیلترینگ و جداسازی، داده‌کاوی و پیش‌بینی موارد مختلف را شامل می‌شود. ورودی روش‌های یادگیری ماشین مجموعه‌ای از داده‌هاست که بسته به نوع یادگیری می‌تواند دارای برچسب یا بدون برچسب باشد. برچسب‌ها در واقع مشخص‌کننده‌ی دسته‌ای هستند که هر داده در آن جای گرفته است. روش‌های یادگیری ماشین را می‌توان به دو بخش کلی "نظارت‌شده" و "غیر نظارت‌شده"<sup>۲</sup> تقسیم‌بندی کرد. برای آموزش الگوریتم در روش نظارت‌شده، برچسب‌گذاری داده‌ها ضروری می‌باشد [۱۱]. روش‌های مختلفی برای هر یک از موارد مذکور ارائه شده است. نحوه عملکرد کلی یادگیری ماشینی مورد استفاده در این مقاله به صورت یک فلوچارت در شکل ۱ نشان داده شده است.



شکل ۱. نحوه عملکرد کلی روش‌های یادگیری ماشینی

در مراجع [۱۶-۱۲] از روش‌های مختلف مبتنی بر شبکه‌های عصبی برای تشخیص خطاها و نقص‌ها در خطوط انتقال استفاده شده است. در مرجع [۱۲] جریان‌ها و ولتاژهای نرمالیزه شده‌ی یک سمت خط جهت تشخیص و طبقه‌بندی خطاهای احتمالی به عنوان ورودی به شبکه عصبی داده شده است. برای شبیه‌سازی خط انتقال و نیز پیاده‌سازی شبکه عصبی در این مقاله از سیمولینک MATLAB و جعبه‌ابزار شبکه‌های عصبی استفاده شده است. مدل استفاده شده شامل شبکه عصبی پیش‌خور و الگوریتم یادگیری پس‌انتشار معمول می‌باشد که شش ورودی (ولتاژها و جریان‌های سه فاز) را دریافت کرده و نوع خطا را در خروجی مشخص می‌کند.

<sup>1</sup> Supervised learning

<sup>2</sup> Un-supervised learning

نوع خطاهای تعیینی در این مرجع محدود به خطاهای خط به زمین می‌باشد. مرجع [۱۳] یک تشخیص‌دهنده‌ی خطای مبتنی بر شبکه‌های عصبی را برای یک سیستم انتقال دومداره‌ی به کار گرفته است. در وسط یکی از خطوط مورد بررسی در این مطالعه از یک جبران‌کننده‌ی استاتیک سنکرون<sup>۱</sup> استفاده شده است. وجود این جبران‌ساز باعث بروز حالات غیرخطی در وضعیت دینامیکی شده و می‌تواند باعث بروز مشکل در روش‌های حفاظتی سنتی گردد. در این مقاله نیز برای تشخیص خطا و محل آن از شبکه عصبی متداول با ورودی‌های ولتاژ و جریان یک سمت خط استفاده شده است. در مرجع [۱۴] یک روش تشخیص خطا و حفاظت سریع برای خطوط EHV ارائه شده است. اساس روش مورد استفاده در این مقاله نیز مشابه دو مرجع قبل می‌باشد با این تفاوت که در اینجا از دو دسته شبکه عصبی مجزا برای تشخیص و یافتن مکان تقریبی بروز خطا به کار گرفته شده است. ساختار شبکه‌های عصبی مورد استفاده به صورت تجربی و با استفاده از آزمون و خطا انتخاب گردیده‌اند. مرجع [۱۵] از ترکیب روش شبکه‌های عصبی فازی مبتنی بر نظریه تشدید انطباقی (ART) و نمونه‌برداری همگام شده از هر دو طرف خط برای تشخیص و طبقه‌بندی خطاهای خط استفاده نموده است. در نمونه‌برداری از یک سمت خط، حجم داده‌ها، تعداد اندازه‌گیرها و نیز پیچیدگی‌ها کمتر است اما در این حالت نیاز به فرض‌هایی برای برخی پارامترها نظیر مقاومت خط وجود دارد که در برخی موارد باعث کاهش جزئی دقت خواهد شد. در نمونه‌برداری از دو سمت خط از تجهیزاتی نظیر PMU ها استفاده شده و ولتاژها و جریان‌های دو طرف به طور همزمان در تحلیل‌ها وارد می‌شوند. خطاهای به وجود آمده در مقره‌های خط و تاثیر آن بر نقایص سیستم انتقال با استفاده از داده‌های مربوط به تصاویر مقره‌ها و تحلیل آن‌ها به وسیله شبکه عصبی پیچشی<sup>۲</sup> در مرجع [۱۶] بررسی شده‌اند. شبکه‌های عصبی پیچشی نوعی از شبکه‌های پیش‌خور هستند که اغلب در مباحث مرتبط با پردازش تصاویر مورد استفاده قرار می‌گیرند. در واقع در این مقاله با تلفیق یادگیری عمیق و داده‌های به دست آمده از تصویربرداری هوایی یک روش جدید برای بررسی خطاهای خط معرفی شده است. سیستم پردازشی ابتدا تصاویر را برای مکان‌یابی مقره‌ها بررسی کرده و پس از تشخیص و اخذ تصاویر مقره‌ها، وجود یا عدم وجود نقص در آن‌ها را به عنوان خروجی مدل مشخص می‌کند. در مرجع [۱۷] از چند روش موجود در حوزه یادگیری ماشین برای تشخیص انواع مختلف خطا در یک خط انتقال استفاده شده و سپس نتایج مربوط به هر کدام مورد مقایسه قرار گرفته‌اند. در این مرجع برای تشخیص نوع خطا و مکان آن از دامنه ولتاژها و جریان‌ها در حین بروز خطا استفاده شده است. مطالعه بر روی یک خط ۶۰۰ کیلومتری انجام شده است که برای پیش‌بینی مکان خط ابتدا خط به ۱۰ قسمت ۶۰ کیلومتری تقسیم شده و مدل معین می‌کند که خطا در کدام قسمت از خط اتفاق افتاده است. در واقع در این مرجع مسئله مکان‌یابی به صورت یک مسئله طبقه‌بندی مورد بررسی قرار گرفته و مکان دقیق خطا پیش‌بینی نمی‌شود بلکه صرفاً بخش مربوطه قابل تعیین خواهد بود. در مراجع [۱۸-۲۰] از اطلاعات به دست آمده از واحدهای اندازه‌گیری فازوری (PMU ها) برای تشخیص و مکان‌یابی خطا استفاده شده است. در مراجع [۱۸، ۱۹] با استفاده از مقادیر لحظه‌ای قرائت شده از PMU های نصب شده در دو انتهای خط، خطا و محل آن تشخیص داده شده اما فقط حالات محدودی از انواع خطا مورد بررسی قرار گرفته است. روش معرفی شده مبتنی بر محاسبه یک شاخص به صورت لحظه‌ای است که هر چند برای یک خط ساده نتایج مناسبی ارائه داده اما در حالات پیچیده‌تر بار محاسباتی بالایی را به سیستم تحمیل خواهد کرد. در مرجع [۲۰] از یک روش یادگیری نظارت‌شده مبتنی بر بررسی رفتار گذرای سیستم استفاده شده است که داده‌های مورد نیاز خود را به صورت لحظه‌ای

<sup>1</sup> STATCOM<sup>2</sup> Convolutional neural network

و زمان واقعی از میکرو-PMU ها دریافت می‌کند. برای پیاده‌سازی این روش نویسندگان ابتدا یک شاخص برحسب ولتاژها و جریان‌ها تعریف کرده و با ترکیب اطلاعات حاصل با داده‌های میکرو-PMU ها و اعمال آن‌ها به مدل یادگیری ماشین نتایج را ارائه نموده‌اند. روش ارائه شده دارای دو بخش عملکردی آنلاین و آفلاین است که به ترتیب از آن‌ها برای آموزش و پایش حالات غیرنرمال استفاده گردیده است. بخش برون خط از داده‌های ثبت شده قبلی و بخش برخط از اندازه‌گیری‌های بی‌درنگ به عنوان ورودی استفاده می‌کنند. مراجع [۲۱-۲۳] از ترکیب روش تبدیل موجک و روش‌های یادگیری ماشین بهره گرفته‌اند. در مرجع [۲۱] از روش تبدیل موجک گسسته<sup>۱</sup> و شبکه‌های عصبی مصنوعی برای تشخیص مکان خطا در خطوط انتقال استفاده شده است. تبدیل موجک در تحلیل سیگنال‌ها کاربرد دارد و برخلاف تبدیل فوریه به سبب قابلیت تنظیم باند تحت بررسی امکان اخذ مولفه‌های فرکانسی بالا و پایین در آن وجود دارد. در این روش، تبدیل موجک وظیفه پیش‌پردازش داده‌ها را داشته و خروجی آن به عنوان داده‌های آموزشی به شبکه عصبی داده می‌شود. در این مقاله در مورد نحوه تشخیص بروز خطا و نوع آن بحثی نشده و صرفاً مکان وقوع تحت بررسی قرار گرفته است. در مرجع [۲۲] نیز از تبدیل موجک گسسته و بردار ماشین پشتیبان برای تشخیص و طبقه‌بندی خطاها در یک سیستم انتقال شش فازه استفاده شده است. باید توجه داشت که هرچند توان قابل انتقال در سیستم شش فازه بالاتر است تعداد حالات خطا نیز در آن بالاتر رفته و حفاظت پیچیده‌تر خواهد شد. در این مقاله برای تولید داده‌های آموزشی از تغییر پارامترهای مختلف نظیر نوع، محل، مقاومت و زاویه وقوع خطا استفاده شده و وجود یا عدم وجود خطا روی هر فاز در خروجی برای تشخیص نوع خطا تعیین گردیده است. مرجع [۲۳] به موضوع تشخیص و طبقه‌بندی خطاهای خط انتقال در حضور جبران‌سازهای سری پرداخته است. سیستم تحت مطالعه یک خط سه فاز متصل به نیروگاه می‌باشد. خط مذکور به دو قسمت تقسیم شده و بر روی هر بخش یک جبران‌ساز سری قرار گرفته است. برای شبیه‌سازی مدل‌ها از جعبه‌ابزار شبکه‌های عصبی و تبدیل موجک در MATLAB استفاده شده است. در این مرجع نتایج مختلفی ارائه گردیده اما این نتایج از لحاظ میزان صحت و دقت مورد بررسی قرار نگرفته و معیارهای اعتبارسنجی متداول برای آن بررسی نشده است. تاثیر تغییرات در شیوه نمونه‌برداری از داده‌ها بر میزان دقت یک الگوریتم یادگیری عمیق مبتنی بر شبکه عصبی پیچشی در مرجع [۲۴] بررسی شده و از آن برای طبقه‌بندی خطاها در خط انتقال استفاده شده است. در مورد نحوه پیاده‌سازی مدل یادگیری ماشینی مورد استفاده در این مقاله توضیح خاصی ارائه نشده اما نتیجه کار نشان دهنده این موضوع است که کاهش بیش از حد در فرکانس نمونه‌برداری باعث از دست رفتن بخشی از اطلاعات شده و دقت خروجی برای مدل نهایی را کاهش می‌دهد. در مراجع [۲۵، ۲۶] یادگیری ماشین برای بهبود عملکرد رله‌های دیستانس در خط انتقال به کار گرفته شده است. روش ارائه شده در مرجع [۲۵] با بهره‌گیری از بردار پشتیبان فاز معیوب و خطای خط به زمین را در شرایط مختلف تشخیص می‌دهد. در این مرجع فرض بر این است که وقوع خطا قبلاً تشخیص داده شده و در اینجا فاز معیوب و وجود یا عدم وجود زمین در خطا تشخیص داده می‌شود. مرجع [۲۶] با استفاده از سیگنال‌های ولتاژ و جریان و نیز امپدانس ظاهری به عنوان داده‌های ورودی، الگوهای پنهان را به کمک شبکه عصبی بازگشتی برای تشخیص خط آشکار کرده و از آن‌ها برای حفاظت دیستانس استفاده می‌کند. در مرجع [۲۷] از یک سیستم دسته‌بندی چندگانه برای تشخیص خطا در خط انتقال با استفاده از اطلاعات مربوط به خط موردنظر و نیز خطوط مجاور به عنوان ورودی بهره گرفته شده است. هر یک از دسته‌بندی کننده‌های پایه در این روش برای یکی از خطوط انتقال در نظر گرفته شده‌اند که با بروز خطا هر کدام محدوده‌ای را برای وقوع آن پیش‌بینی کرده و تصمیم‌گیری نهایی با استفاده از میانگین وزنی

<sup>1</sup>Discrete wavelet transform

نتایج حاصله اتخاذ می‌شود. مرجع [۲۸] از ترکیب روش KNN و نظریه فازی برای پیش‌بینی نوع خطا و نیز محل بروز آن در سیستم انتقال استفاده کرده است. در اینجا برای ایجاد داده‌های آموزشی از نرم‌افزار MATLAB و جعبه‌ابزار PSB استفاده شده است و فرض بر این است که داده‌های استخراج شده برای استفاده در مدل فاقد برچسب می‌باشند. لذا ابتدا این داده‌ها با استفاده از یک روش یادگیری نظارت نشده تحت عنوان ISODATA خوشه‌بندی و برچسب‌گذاری شده و سپس از آن‌ها برای آموزش و پیاده‌سازی مدل اصلی استفاده شده است. جزئیات روش پیش‌بینی و یا نتایج مکان‌یابی به طور واضح ارائه نشده و همچنین در بررسی صحت و اعتبار نتایج صرفاً به معیار دقت بسنده شده است که به تنهایی ابزار مناسبی برای اطمینان از عملکرد مناسب مدل نمی‌باشد.

اغلب، اطلاعاتی که از سیستم قدرت واقعی در رابطه با وقوع خطاها در دسترس است بسیار پراکنده و محدود بوده و نمی‌توان یک مجموعه داده با حجم مناسب برای خطاها را در یک قسمت خاص از سیستم قدرت مانند خطوط انتقال با جزئیات لازم پیدا کرد. از این رو در این مقاله داده‌های مربوط به خطا با توجه به عملکرد خط در شرایط بهره‌برداری مختلف تولید گردیده و از آن‌ها برای آموزش مدل‌های یادگیری ماشین و هشداردهی بروز خطا استفاده خواهد شد. یکی دیگر از عوامل مهمی که در اغلب مقالات نادیده گرفته شده تاثیر شرایط آب و هوایی بر وقوع خطاها در خط انتقال می‌باشد. شرایط جوی مختلف هر یک به میزان متفاوتی بر وقوع خطاها تاثیر گذار هستند که این موضوع در سه وضعیت (نرمال، بد و وخیم) در این مقاله لحاظ گردیده است. در این مقاله ابتدا عوامل شبکه‌ای موثر بر وقوع خطا مورد بررسی قرار گرفته و مدل‌هایی کاربردی جهت وارد کردن تاثیر آن‌ها در مطالعات معرفی شده است، سپس اثر عوامل جوی نیز با آن‌ها تلفیق شده و نتایج کلی نزدیک به واقعیت استحصال شده است. مسئله تحت بررسی در این مقاله با استفاده از سه الگوریتم قوی در حوزه یادگیری نظارت‌شده در چندین حالت مختلف مورد تحلیل قرار گرفته است. علاوه بر موارد فوق موضوع مهم دیگری که اغلب مقالات اخیر به آن توجه نکرده‌اند لزوم پاسخ‌دهی مناسب مدل‌ها برای داده‌های منطبق با شرایط واقعی می‌باشد. اغلب مقالات نتایج را با مجموعه داده‌هایی به دست آورده‌اند که نسبت داده‌های آموزشی آن‌ها در کلاس‌های مختلف مشابه حالات عملی نمی‌باشد. روش‌های یادگیری ماشینی مانند بردار پشتیبان در مواقعی که تعادل بین تعداد داده‌ها در کلاس‌های مختلف برقرار نباشند یا تعداد داده در کلاس‌های مهم مورد نظر جهت پیش‌بینی نسبت به سایر کلاس‌ها (مانند حالات وقوع خطا نسبت به نرمال) کمتر باشد دچار کاهش دقت قابل ملاحظه‌ای خواهند شد. در این مقاله برای حل این مشکل و افزایش دقت پیش‌بینی‌ها از روش‌های متعادل‌سازی داده‌ها استفاده گردیده است. با استفاده از این روش‌ها می‌توان با اعمال یک مرحله پیش‌پردازش روی داده‌ها برای مجموعه‌هایی با نسبت‌های واقعی در کلاس‌های مختلف، نتایج مناسب به دست آورد. توضیحات جزئی‌تر در این باره در بخش‌های آتی مقاله ارائه شده است.

بخش‌های بعدی مقاله به شرح مقابل ارائه گردیده‌اند: در بخش دوم عوامل موثر بر وقوع خطا همراه با مدل‌ها و توابع احتمالی مربوط به آن‌ها ارائه گردیده و بررسی شده‌اند. مطالعات عددی شامل تولید داده‌ها، اعمال روش‌های یادگیری ماشین و استخراج معیارهای ارزیابی آن‌ها در بخش سوم ارائه شده‌اند. در پایان بخش سوم، مقایسه و تحلیلی از نتایج به دست آمده از روش‌های مختلف بیان شده و در نهایت نتیجه‌گیری کلی مقاله در بخش چهارم آورده شده است.



## ۲- عوامل موثر بر وقوع خطا در خط انتقال

افزایش توان عبوری از خط انتقال می‌تواند باعث افزایش احتمال بروز خطا گردد. این موضوع می‌تواند ناشی از افزایش تولید حرارت به دلیل عبور جریان‌های بزرگ باشد. با بیشتر شدن دمای خط استحکام مکانیکی آن به تدریج دچار تضعیف شده و نیز فلش خط بزرگ‌تر می‌گردد که این موضوع به نوبه‌ی خود می‌تواند سبب تسهیل وقوع خطاهای ناشی از برخورد با عوارضی مانند درختان بلند شود. اگر توان عبوری باز هم افزایش پیدا کند ممکن است خط به حد پایداری حرارتی خود برسد و در نهایت توسط عوامل حفاظتی از مدار خارج شود. عبور جریان بیش از ظرفیت حرارتی تنها برای مدت محدودی می‌تواند مجاز باشد. طبق تعریف ظرفیت نامی برای جریان یک خط انتقال، میزان جریانی است که می‌تواند به مدت نامحدود از آن عبور کند [۲۹]. بنابراین اضافه‌بار می‌تواند باعث بروز خطا یا عملکرد تجهیزات حفاظتی برای جلوگیری از آسیب دیدن خط شده و خاموشی‌های ناخواسته را به دنبال داشته باشد. انتقال توان در فواصل طولانی، نیاز به استفاده از ولتاژهای بالا در خطوط انتقال را اجتناب‌ناپذیر کرده است. یکی دیگر از عوامل موثر بر احتمال وقوع خطا در خط اندازه‌ی ولتاژ شین‌های متصل به آن می‌باشد. افزایش بیش‌ازحد ولتاژ می‌تواند باعث دشارژ بین شین و مقره‌ها شده و خطای گذرا یا دائمی را در پی داشته باشد [۳۰]. از طرف دیگر کاهش بیش‌ازحد ولتاژ خط نیز می‌تواند سبب افزایش جریان برای انتقال توان یکسان شده و همان مشکلات ناشی از افزایش دمای خط و عملکرد رله‌ها را در پی داشته باشد.

عامل بسیار مهم دیگری که احتمال وقوع خطا در خط انتقال را تحت تاثیر قرار می‌دهد شرایط جوی و آب و هوایی محیط می‌باشد [۳۱]. برای مثال در روزهایی که وزش باد شدید باشد خطوط به صورت مواج به سمت یک دیگر حرکت کرده و امکان وقوع خطاهای اتصال کوتاه افزایش می‌یابد. احتمال بروز خطا در شرایط طوفانی و رعد و برق نیز به شدت بیشتر می‌شود. عوامل بدتر دیگر نظیر سیل، زلزله و آتش‌سوزی می‌توانند به‌طور کامل برج‌های خط انتقال را درهم‌شکسته و خط را از مدار خارج نمایند.

### ۲-۱ مدل احتمالی خطا

با توجه به مباحث انجام شده در بخش قبل می‌توان نتیجه گرفت که خطاهای خط انتقال تا حد زیادی به توان‌های اکتیو و راکتیو عبوری از خط، ولتاژ شین‌ها و شرایط جوی وابسته می‌باشد. بنابراین در مدل احتمالی خطا در خط باید موارد مذکور لحاظ شده و مورد توجه قرار گیرند. محدوده مقادیر پارامترهای مورد نظر در خط انتقال به مقادیر نامی، نحوه انجام محاسبات و تجهیزات لحاظ شده در زمان طراحی خط بستگی دارد که در مقاله حاضر حداکثر این مقادیر برای کارکرد درست خط به ترتیب برای توان اکتیو، توان راکتیو و ولتاژ به فرم  $V_{Lmax}$ ،  $Q_m$ ،  $P_m$  و برای حداقل ولتاژ به صورت  $V_{Lmin}$  در نظر گرفته شده‌اند. بدیهی است که میزان احتمال وقوع خطا زمانی که به ظرفیت نهایی خط نزدیک می‌شویم به سبب بروز عوامل ذکر شده در بخش‌های قبلی، همچنین تنش‌های وارده به سیستم، اغتشاشات احتمالی، پایین آمدن قدرت مانور و نیز عوامل ناشی از فرسودگی تجهیزات و... دچار افزایش می‌شود. اگر مقدار پارامتر مورد نظر از محدوده‌ی مجاز و مناسب برای کارکرد سیستم (مانند حداکثر ولتاژ قابل تحمل یا حداقل ولتاژ مورد نیاز) تخطی کند مسلماً سیستم دچار نقص خواهد شد.

می‌توان گفت که الگوی بروز این اغتشاش‌ها فاقد ماهیت خطی ندارد و با نزدیک‌تر شدن به کران‌های مجاز پارامترها احتمال بروز خطا نیز افزایش خواهد یافت، فلذا استفاده از توابع درجه اول در محاسبه روابط مربوط به احتمال مناسب به نظر نمی‌رسد. از طرف دیگر توابع درجه سوم و بالاتر ضمن افزایش پیچیدگی و بار محاسباتی برای بازه مورد نظر مزیت اضافی چندانی ارائه نمی‌دهند.



بنابراین، در این مقاله از بخشی از توابع درجه دو برای نشان دادن تغییرات پارامترها در بازه بین مقادیر مجاز و حداکثری استفاده شده است. برای بدست آوردن روابط مربوط به توابع درجه دو، از سه نقطه شامل نقاط مربوط به ناحیه مجاز پارامترها و نیز نقاط حداکثری (یعنی نقاط ابتدایی و انتهایی بازه مربوط به تابع درجه دو) و نیز یک نقطه فرضی با لحاظ شرایط مطرح شده در فوق استفاده گردیده است. می‌توان با مراجعه به شکل‌های نشان داده شده در بخش‌های آتی به درک بهتری از موضوع رسید. در ادامه مدل‌های احتمالی خطا براساس پارامترهای موردنظر و نیز تاثیر شرایط جوی بر این احتمال مورد بررسی قرار خواهد گرفت.

## ۲-۱-۱ مدل احتمالی خطا بر اساس توان‌های اکتیو و راکتیو

هنگامی که توان اکتیو عبوری از خط انتقال در محدوده‌ی مجاز و عادی خود باشد تاثیر آن بر احتمال وقوع خطا کم خواهد بود به نحوی که می‌توان این احتمال را برابر با یک مقدار مثبت کوچک قرار داد. اما با افزایش توان عبوری و رسیدن آن به یک مقدار مشخص احتمال وقوع خطا نیز به صورت تصاعدی بالا خواهد رفت. اگر این افزایش ادامه پیدا کند و از محدوده‌ی مجاز خط تجاوز نماید به طور قطع خط دچار نقص خواهد شد و احتمال وقوع خطا برای چنین وضعیتی برابر با ۱ یا ۱۰۰٪ خواهد بود. اگر توان اکتیو نامی خط برابر با  $P_n$  و توان مجاز آن برابر  $P_m$  در نظر گرفته شود می‌توان گفت که احتمال وقوع خطا به سبب عبور توان اکتیو برای توان‌های کمتر از  $P_n$  بسیار کم، برای توان‌های بین  $P_m$  و  $P_n$  افزایشی و برای توان‌های بالاتر از  $P_m$  برابر ۱ خواهد بود. این موضوع در این مقاله به صورت رابطه (۱) مدل‌سازی شده است:

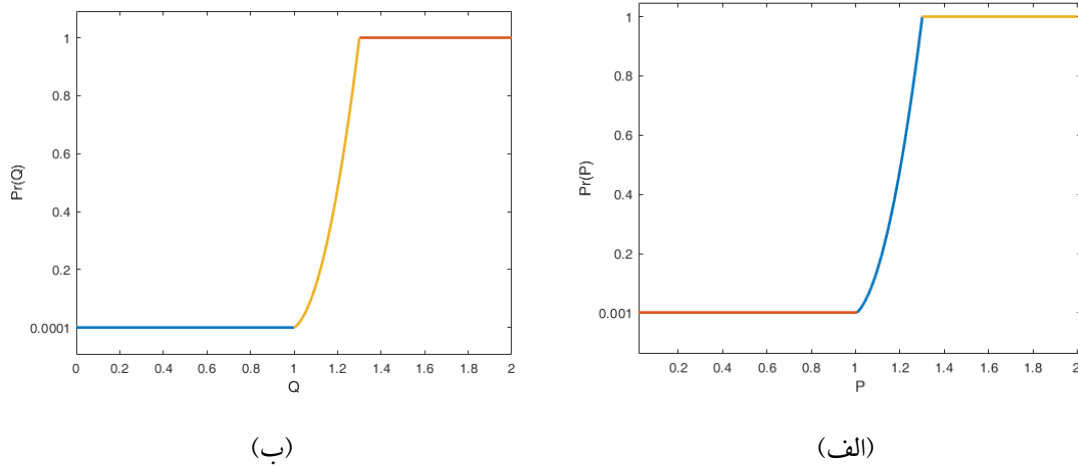
$$\Pr(P) = \begin{cases} 1 \times 10^{-3} & , P < 1 \\ 8/95 \cdot P^2 - 17/26P + 8/306 & , 1 \leq P \leq 1/3 \\ 1 & , P > 1/3 \end{cases} \quad (1)$$

که در این رابطه  $\Pr(P)$  نشان‌دهنده احتمال وقوع خطا در اثر عبور توان اکتیو  $P$  می‌باشد. حداکثر توان اکتیو عبوری از خط در اینجا برابر  $1/3$  در نظر گرفته شده است که برای مقادیر بالاتر از آن خط قطعاً دچار آسیب خواهد شد. برای توان‌های بالاتر از ۱ تا  $1/3$  پیرونیست از یک تابع درجه دوم برای مدل‌سازی خطا استفاده شده است. این تابع باید به نحوی انتخاب شود که حدود احتمالی وارد مقادیر نامعین نشده و رفتار احتمالی خط به نحو مناسب مدل گردد.

به صورت مشابه برای توان راکتیو نیز می‌توان گفت که احتمال بروز خطای ناشی از عبور این نوع توان در محدوده‌ی نرمال ناچیز بوده و سپس با افزایش آن از مقدار نامی، احتمال مربوطه نیز بیشتر خواهد شد. با فرض  $Q_n$  به عنوان توان راکتیو نامی و  $Q_m$  به عنوان حداکثر توان راکتیو مجاز (که بعد از آن خط قطعاً دچار نقص خواهد شد) می‌توان گفت، طبق رابطه (۲) که احتمال وقوع خطا ناشی از عبور توان راکتیو برای توان‌های کمتر از  $Q_n$  ناچیز، برای توان‌های بین  $Q_m$  و  $Q_n$  افزایشی و برای توان‌های بالاتر از  $Q_m$  برابر ۱ خواهد بود. در صورتی که حداکثر توان راکتیو مجاز خط برابر با  $1/3$  پیرونیست در نظر گرفته شود، می‌توان نوشت:

$$\Pr(Q) = \begin{cases} 1 \times 10^{-4} & , Q < 1 \\ 9/528Q^2 - 18/582Q + 9/0541 & , 1 \leq Q \leq 1/3 \\ 1 & , Q > 1/3 \end{cases} \quad (2)$$

که در این رابطه  $\Pr(Q)$  نشان‌دهنده احتمال وقوع خطا در اثر عبور توان راکتیو  $Q$  می‌باشد. برای درک بهتر، چگونگی وابستگی بین توان‌های اکتیو و راکتیو عبوری با خطای خط در شکل ۲ نشان داده شده است.

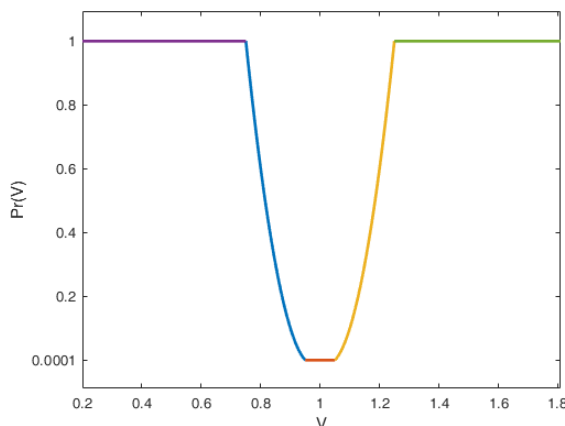


شکل ۲. مدل‌های احتمالی خطا برای (الف) توان اکتیو عبوری از خط و (ب) توان راکتیو عبوری از خط

### ۲-۱-۲- مدل احتمالی خطا بر اساس ولتاژ شین

برخلاف توان‌های اکتیو و راکتیو، افزایش و یا کاهش ولتاژ شین‌های دو طرف خط می‌تواند باعث بروز خطا شود. برای ولتاژ شین در محدوده عادی، احتمال وقوع خطا را می‌توان برابر با یک عدد ثابت کوچک در نظر گرفت اما با افزایش یا کاهش ولتاژ احتمال بروز خطا نیز افزایش خواهد یافت به نحوی که بعد از یک مقدار مشخص قطعاً خط دچار نقص خواهد شد که در این حالت احتمال وقوع خطا برابر ۱ در نظر گرفته می‌شود. اگر محدوده کاری نرمال برای ولتاژ خط بین  $V_{Lmin}$  و  $V_{Lmax}$  باشد، طبق توضیحات ارائه شده رابطه زیر احتمال بروز خطای ناشی از تغییرات ولتاژ شین را به خوبی مدل می‌نماید. محدوده نرمال ولتاژ در این رابطه بین ۰/۹۵ تا ۱/۰۵ پرونیت لحاظ شده است. نواحی افزایشی احتمال بروز خطا در این رابطه نیز به صورت یک تابع درجه دو در نظر گرفته شده‌اند. حدود نهایی برای کران‌های پایین و بالای ولتاژ که پس از آن‌ها وقوع خطا حتمی خواهد بود به ترتیب برابر با ۰/۷۵ و ۱/۲۵ منظور شده‌اند.

$$\Pr(V) = \begin{cases} 1 & , & V \leq 0.75 \\ 20/0.1V^2 - 39/0.2V + 19/0.1 & , & 0.75 < V < 0.95 \\ 1 \times 10^{-4} & , & 0.95 \leq V \leq 1.05 \\ 20/0.1V^2 - 41/0.2V + 21/0.1 & , & 1.05 < V < 1.25 \\ 1 & , & V \geq 1.25 \end{cases} \quad (3)$$



شکل ۳. مدل احتمالی خطا بر اساس ولتاژ شین

شکل ۳ مدل احتمالی بروز خطا در خط بر اساس ولتاژ شین و طبق رابطه (۳) نشان می‌دهد.

### ۲-۱-۳- مدل احتمالی خطا بر اساس شرایط جوی

یکی دیگر از عوامل مهم موثر بر وقوع خطا در خط انتقال شرایط آب و هوایی محیط می‌باشد. وضعیت‌های جوی مختلف به گونه‌های متفاوتی بر احتمال وقوع خطا در خط تاثیر می‌گذارند. به عنوان مثال شرایطی از قبیل رعد و برق، بادهای شدید، یخبندان، برف، سیلاب‌ها، زمین لرزه و ... می‌توانند باعث افزایش شدید خطاها در خط شوند. برای لحاظ کردن اثرات ناشی از حالات مختلف آب و هوایی می‌توان از دسته‌بندی این حالات به سه بخش نرمال، بد و وخیم استفاده کرد. در شرایط نرمال، احتمال وقوع خطا در خط ناشی از اوضاع جوی بسیار اندک خواهد بود، اما با بدتر شدن شرایط احتمال وقوع خطا نیز افزایش خواهد یافت. برای لحاظ کردن این موضوع می‌توان تاثیر شرایط جوی را با اعمال ضرایبی در دو مدل قبلی وارد کرد؛ به این معنی که احتمال وقوع خطا ناشی از توان‌های اکتیو و راکتیو عبوری و نیز ولتاژ شین، در شرایط جوی مختلف مشابه نخواهد بود. نکته مهم دیگر تفاوت در نرخ وقوع حالات مختلف (نرمال، بد و وخیم) می‌باشد که باید در ایجاد یا بررسی داده‌ها به آن توجه کرد. در این مقاله برای هر یک از این سه وضعیت آب و هوایی ضرایبی برای احتمال‌های بحث شده در نظر گرفته شده است. برای تطابق بیشتر این ضرایب با واقعیت از نتایج حاصل از مرجع [۳۲] استفاده گردیده و مقادیر ضرایب مربوطه برای حالات نرمال، بد و وخیم به ترتیب برابر با ۱، ۲ و ۱۵ در نظر گرفته شده است. نحوه اعمال این ضرایب و روابط حاصل در ادامه آورده شده است.

برای حالت نرمال با لحاظ کردن ضریب ۱، همان روابط (۱) تا (۳) بدون تغییر قابل استفاده خواهند بود. در شرایط جوی بد این روابط در ضریب ۲ ضرب شده و علاوه بر این حدود متغیرها نیز باید اصلاح شوند. در شرایط وخیم نیز روابط در ضریب مربوطه ضرب شده و حدود باید اصلاح شوند. روابط (۴) تا (۶) مربوط به شرایط جوی بد و روابط (۷) تا (۹) مربوط به شرایط جوی وخیم می‌باشند.

$$\Pr(P) = \begin{cases} 2 \times 10^{-3} & , P < 1 \\ 17/9P^2 - 34/52P + 16/612 & , 1 \leq P \leq 1/20.4227 \\ 1 & , P > 1/20.4227 \end{cases} \quad (4)$$

$$\Pr(Q) = \begin{cases} 2 \times 10^{-4} & , & Q < 1 \\ 19/0.56Q^2 - 37/164Q + 18/10.82 & , & 1 \leq Q \leq 1/20.553 \\ 1 & , & Q > 1/20.553 \end{cases} \quad (5)$$

$$\Pr(V) = \begin{cases} 1 & , & V \leq 0/81497 \\ 40/0.2V^2 - 78/0.4V + 38/0.2 & , & 0/81497 < V < 0/9499 \\ 2 \times 10^{-4} & , & 0/9499 \leq V \leq 1/0.5 \\ 40/0.2V^2 - 82/0.4V + 42/0.2 & , & 1/0.5 < V < 1/185.0 \\ 1 & , & V \geq 1/185.0 \end{cases} \quad (6)$$

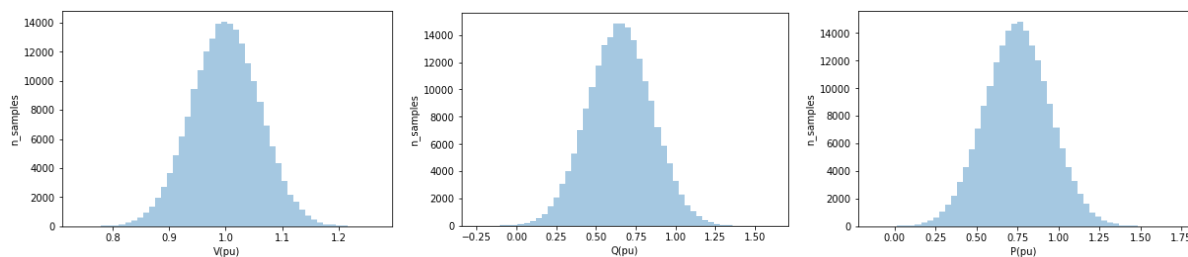
$$\Pr(P) = \begin{cases} 15 \times 10^{-3} & , & P < 0/991255 \\ 134/25P^2 - 258/75P + 124/59 & , & 0/991255 \leq P \leq 1/0.5367 \\ 1 & , & P > 1/0.5367 \end{cases} \quad (7)$$

$$\Pr(Q) = \begin{cases} 15 \times 10^{-4} & , & Q < 1 \\ 142/92Q^2 - 278/73Q + 135/8115 & , & 1 \leq Q \leq 1/0.6223 \\ 1 & , & Q > 1/0.6223 \end{cases} \quad (8)$$

$$\Pr(V) = \begin{cases} 1 & , & V \leq 0/91215 \\ 300/15V^2 - 585/3V + 285/15 & , & 0/91215 < V < 0/949925 \\ 15 \times 10^{-4} & , & 0/949925 \leq V \leq 1/0.5007 \\ 300/15V^2 - 615/3V + 315/15 & , & 1/0.5007 < V < 1/0.8788 \\ 1 & , & V \geq 1/0.8788 \end{cases} \quad (9)$$

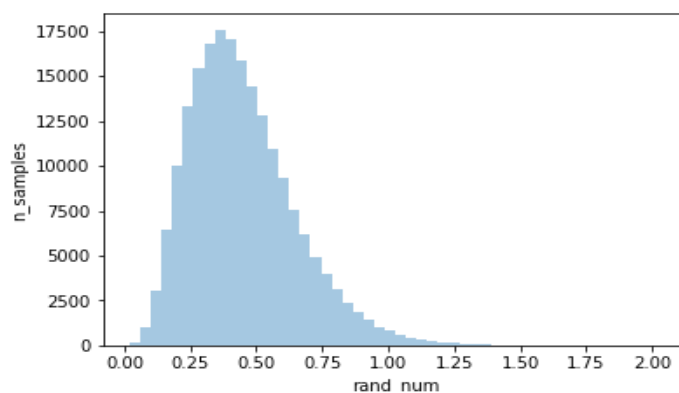
### ۳- متعادل سازی داده‌ها و مطالعات عددی

برای درک بهتر موضوع متعادل سازی داده‌ها و تاثیر آن بر نتایج بدست آمده، بحث نظری و توضیحات تفصیلی مربوطه در کنار مطالعات عددی گنجانده شده است. همان‌طور که در بخش‌های پیشین بحث شد متاسفانه داده‌های مربوط به خطا برای وضعیت‌های بهره‌برداری مختلف و نیز شرایط جوی متنوع در دسترس نبوده یا اطلاعات بسیار ناقصی را می‌توان در این حوزه به دست آورد. البته با توجه به گسترش استفاده از تجهیزات اندازه‌گیری و ارتباطی پیشرفته، هوشمندتر شدن هر چه بیشتر شبکه و افزایش مشاهده‌پذیری و سیلان اطلاعات از سمت شبکه به مراکز کنترلی انتظار می‌رود که در آینده برای مطالعات در حوزه یادگیری ماشین مشکلی از لحاظ تامین داده‌های مورد نیاز وجود نداشته باشد. به هر ترتیب در این بخش از مقاله برای تولید داده‌های مربوط به توان‌های اکتیو و راکتیو عبوری و نیز داده‌های مربوط به ولتاژ از توزیع تصادفی نرمال استفاده شده است. داده‌های تولید شده در شکل ۴ به نمایش درآمده‌اند.



شکل ۴. داده‌های تصادفی برای توان اکتیو، توان راکتیو و ولتاژ

محور افقی نمودارهای فوق مربوط به پارامترهای موردنظر بوده و محور عمودی نشان‌دهنده تعداد نمونه‌ها در هر زیر بازه می‌باشد. تعداد کل داده‌های تصادفی برای هر یک از پارامترها (توان‌های اکتیو، راکتیو و ولتاژ)، برابر با ۲۰۰۰۰۰ نمونه در نظر گرفته شده است. نسبت شرایط آب و هوایی مختلف در طول سال برای مناطق مختلف متفاوت می‌باشد. همچنین نوع وخامت آب‌وهوا نیز می‌تواند برای هر منطقه با مناطق دیگر تمایز داشته باشد؛ به‌عنوان مثال در مناطق مرطوب و کوهستانی طوفان‌های همراه با باران و رعد و برق و در مناطق بیابانی و خشک طوفان‌های دارای گردوغبار شدید را می‌توان جزء شرایط وخیم در نظر گرفت. برای یک منطقه با اقلیم معمولی روزهای با آب‌وهوای وخیم غالباً کمتر از ۱۰٪ ایام سال را شامل می‌شوند. برای برخی مناطق خاص این عدد می‌تواند حتی به ۵۰٪ نیز برسد [۳۳، ۳۴]. به‌هرحال در این مقاله منطقه مورد مطالعه از نوع معمولی در نظر گرفته شده است که می‌توان آن را به‌راحتی برای شرایط مختلف تعمیم داد. برای لحاظ کردن شرایط مختلف جوی از توزیع گاما با پارامتر شکل برابر ۵ و پارامتر مقیاس برابر ۰/۰۹ همانند شکل ۵ استفاده شده است.



شکل ۵. توزیع گاما برای انتخاب تصادفی حالات مختلف شرایط جوی

با توجه به کاهش تعداد داده‌های تولیدی در بازه‌های انتهایی و توضیحات ارائه شده، این توزیع می‌تواند مدل مناسبی از حالات مختلف را به دست دهد. در مدل حاصل از بین داده‌های تصادفی، مقادیر کمتر از ۰/۶ برای حالت عادی، داده‌های بین ۰/۶ تا ۰/۸ برای شرایط جوی بد و مقادیر بالاتر از ۰/۸ برای وضعیت وخیم در نظر گرفته شده‌اند. در این مقاله از روش‌های  $KNN^1$ ،  $SVM^2$  و درخت تصمیم بهبود یافته (که همگی جزء روش‌های معتبر یادگیری ماشینی می‌باشند) جهت پیش‌بینی و هشداردهی خطا در خط انتقال استفاده شده و نتایج حاصل مورد مقایسه و تحلیل قرار گرفته است. نکته مهم در حل مسئله حاضر عدم تعادل بین کمیت

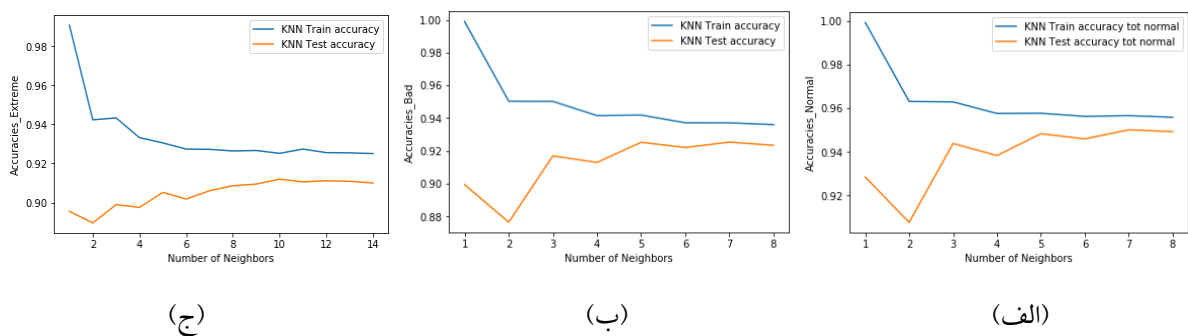
<sup>1</sup> K-Nearest Neighbor

<sup>2</sup> Support-Vector Machine

یا تعداد داده‌های در دسترس از هر برچسب یا دسته به‌خصوص برای شرایط جوی نرمال و بد می‌باشد. روش‌های معمول یادگیری ماشین زمانی پاسخ با دقت مناسب ارائه می‌دهند که تعادل بین داده‌ها در برچسب‌های مختلف برقرار باشد. در صورت عدم برقراری تعادل، ماتریس درهم‌ریختگی<sup>۱</sup> نتایج مطلوبی را ارائه نخواهد داد که نشان‌دهنده ضعف مدل تحت بررسی می‌باشد. برای حل این مشکل در این مقاله از ترکیب روش‌های SMOTE<sup>۲</sup> و Tomek links جهت تنظیم نحوه توزیع داده‌ها در دسته‌ها استفاده گردیده است. همچنین برای بررسی عملکرد مدل‌ها در حالت غیرمتعالی، نتایج حاصل قبل از اعمال روش‌های تعادل دسته‌ها نیز آورده شده و تحت بررسی قرار گرفته‌اند. تولید داده‌ها، شبیه‌سازی‌ها و نمودارهای ارائه شده در این بخش از مقاله با استفاده از زبان برنامه‌نویسی Python پیاده‌سازی شده‌اند.

### ۳-۱ روش KNN

KNN یکی از روش‌های کاربردی در حوزه یادگیری ماشین است که می‌توان از آن برای حل مسائل مربوط به طبقه‌بندی و رگرسیون استفاده کرد. در این روش برای یادگیری مدل از اطلاعات مربوط به K داده اطراف استفاده می‌شود. باید توجه داشت که یادگیری و دقت نتایج حاصل در روش KNN تا حد زیادی به نحوه انتخاب مقدار K بستگی دارد. در روش‌های یادگیری ماشین برای بررسی میزان صحت مدل، داده‌های موجود به دو قسمت داده‌های آموزشی و داده‌های آزمایشی تقسیم‌بندی می‌شوند. معمولاً ۷۰٪ از داده‌ها جهت آموزش مدل و ۳۰٪ باقیمانده به‌منظور آزمایش مدل مورد استفاده قرار می‌گیرند که در اینجا نیز همین نسبت‌ها لحاظ شده‌اند. پس از وارد کردن این اطلاعات به مدل غالباً در ابتدا دو نوع دقت تحت عناوین دقت یادگیری<sup>۳</sup> و دقت آزمایش<sup>۴</sup> بررسی می‌شوند که بالاتر بودن مقدار آن‌ها در اکثر موارد نشان‌دهنده بالا بودن کیفیت مدل به‌دست آمده می‌باشد. در اینجا برای بررسی میزان دقت مدل نتایج مربوط به تعداد همسایگی‌ها از ۱ تا ۹ برای شرایط آب و هوایی نرمال و بد و از ۱ تا ۱۴ برای شرایط وخیم با در نظر گرفتن روابط مربوط به توان‌های اکتیو، راکتیو و ولتاژ در هر یک از این حالات آورده شده است.



شکل ۶. بررسی میزان دقت‌ها در روش KNN با تغییر مقدار K: (الف) شرایط آب و هوایی نرمال (ب) بد (ج) وخیم

در روش KNN باید تعداد نقاط همسایگی به نحوی انتخاب شود که دقت آموزش و آزمایش در آن در یک حد معقول و نزدیک به هم باشند. با توجه به شکل ۶، برای شرایط جوی نرمال و بد می‌توان K را برابر ۷، و برای شرایط وخیم برابر با ۱۰ در نظر گرفت. میانگین دقت پیش‌بینی روش KNN برای داده‌های آزمون در هر یک از حالات فوق به ترتیب برابر ۹۵٪، ۹۲٪/۵ و ۹۱٪/۲.

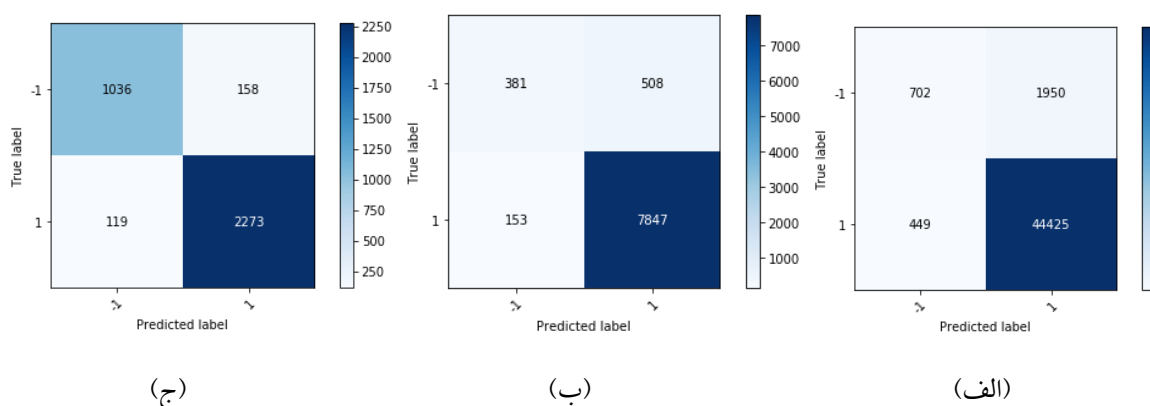
<sup>1</sup> Confusion matrix

<sup>2</sup> Synthetic Minority Over-Sampling Technique

<sup>3</sup> Train accuracy

<sup>4</sup> Test accuracy

به دست آمده است. یکی دیگر از معیارهای مهم در بررسی نتایج پیش‌بینی‌ها در یک مسئله طبقه‌بندی ماتریس درهم‌ریختگی می‌باشد. این ماتریس علاوه بر نشان دادن اشتباهات موجود در طبقه‌بندی، اطلاعات بیشتری نیز در مورد نوع اشتباهات ارائه می‌دهد. در این ماتریس همانند شکل ۷، عناصر روی قطر اصلی نشان‌دهنده تعداد پیش‌بینی‌های صحیح و عناصر روی قطر فرعی نشان‌دهنده تعداد داده‌هایی می‌باشد که به اشتباه برچسب‌گذاری شده‌اند. در حالت کلی ماتریس درهم‌ریختگی چهار مقدار را برای ما مشخص می‌کند: مقدار صحیح مثبت (TP)<sup>۱</sup>، مقدار صحیح منفی (TN)<sup>۲</sup>، مقدار اشتباه مثبت (FP)<sup>۳</sup> و مقدار اشتباه منفی (FN)<sup>۴</sup>. اگر مثبت و منفی به‌عنوان برچسب‌ها در نظر گرفته شوند، مقدار TP و TN در واقع تعداد پیش‌بینی‌هایی را نشان می‌دهند که برچسب‌های مثبت و منفی در آن‌ها به درستی پیش‌بینی شده‌اند. در مقابل FP نشان‌دهنده تعداد پیش‌بینی‌های اشتباهاً مثبت و FN نشان‌دهنده تعداد پیش‌بینی‌های اشتباهاً منفی می‌باشند.



شکل ۷. نمودارهای مربوط به ماتریس‌های درهم‌ریختگی برای روش KNN در هر یک از سه حالت (الف) نرمال (ب) بد (ج) وخیم با توجه به نتایج حاصل از ماتریس درهم‌ریختگی می‌توان کیفیت دسته‌بندی را مورد بررسی قرار داد. چند مورد از مهم‌ترین معیارهای موجود برای این بررسی عبارت‌اند از دقت<sup>۵</sup>، صحت<sup>۶</sup> و معیار "اف-۱"<sup>۷</sup> [۳۵]. مقدار این معیارها برای پیش‌بینی‌های انجام شده در هر یک از شرایط جوی تحت بررسی و برای داده‌های اصلی بدون اعمال متعادل‌سازی برچسب‌ها در جدول ۱ آورده شده است. باید توجه داشت که دقت (precision) تعریف شده با استفاده از ماتریس درهم‌ریختگی با دقت پیش‌بینی (accuracy) که پیش از این بررسی شد متفاوت است. اگر حروف اول هر کدام (یعنی A و P) را به‌عنوان نماینده هر کدام در نظر بگیریم، A نشان‌دهنده متوسط میزان نزدیکی نتایج پیش‌بینی‌شده به نتایج واقعی و P بیان‌کننده میزان قابلیت تکرارپذیری پیش‌بینی‌ها می‌باشد. معیار صحت نسبت به پیش‌بینی‌های صحیح برای برچسب موردنظر را به کل تعداد واقعی موجود از برچسب نشان می‌دهد. معیار اف-۱ تاثیر هر دو معیار دقت (P) و صحت را در خود دارد. این معیار در واقع میزان تعادل بین P و صحت را بیان کرده و در مواقعی بیشتر مورد استفاده قرار

<sup>1</sup> True positive  
<sup>2</sup> True negative  
<sup>3</sup> False positive  
<sup>4</sup> False negative  
<sup>5</sup> Precision  
<sup>6</sup> Recall  
<sup>7</sup> F1-score



می‌گیرد که ایجاد یک موازنه بین دو معیار مذکور مدنظر باشد؛ مقدار این معیار برابر میانگین همساز یا هارمونیک<sup>۱</sup>  $P^1$  و صحت می‌باشد.

جدول ۱- معیارهای استخراج شده از ماتریس درهم‌ریختگی برای روش KNN

برچسب ۱-	دقت (P)	صحت	اف-۱	برچسب ۱	دقت (P)	صحت	اف-۱
نرمال	۰/۶۰	۰/۲۷	۰/۳۷	نرمال	۰/۹۶	۰/۹۹	۰/۹۷
بد	۰/۷۳	۰/۴۷	۰/۵۷	بد	۰/۹۴	۰/۹۸	۰/۹۶
وخیم	۰/۹۰	۰/۸۷	۰/۸۸	وخیم	۰/۹۴	۰/۹۵	۰/۹۴

همان‌طور که از نتایج جدول ۱ می‌توان دریافت، هرچند مدل پیاده شده برای برچسب ۱ نتایج خوبی ارائه کرده اما نتایج به‌دست آمده برای برچسب ۱- به‌خصوص در حالات نرمال و بد مناسب نیستند. دلیل این موضوع کم بودن نسبت وقوع خطاها نسبت به حالات عملکردی عادی در دو وضعیت نرمال و بد می‌باشد که باعث تضعیف کیفیت یادگیری در آن حالات می‌شود. برای حل این مشکل در اینجا از ترکیب روش‌های SMOTE و Tomek Links استفاده شده که با اضافه کردن تعدادی داده به کلاس کوچک‌تر و نیز حذف بخشی از داده‌های کلاس بزرگ‌تر باعث افزایش اثر کلاس کوچک در یادگیری و نیز جداسازی بهتر دو دسته از هم‌دیگر می‌شود. توصیف بیشتر در مورد نحوه عملکرد این روش‌ها در مراجع [۳۶،۳۷] آورده شده است. برای مطالعه حاضر مقادیر معیارهای مربوط به برچسب ۱- (یعنی معیارهای مربوط به پیش‌بینی وقوع خطا) از اهمیت بالایی برخوردار هستند. با اعمال روش مذکور مقادیر دقت (A) برای هر دو حالت نرمال و بد با کاهش جزئی نسبت به حالت قبل برابر ۹۲٪ به‌دست آمده‌اند اما با توجه به نتایج ارائه شده در جدول ۲ مقادیر مربوط به دقت (P)، صحت و معیار اف-۱ برای برچسب ۱- نسبت به حالت قبل بهبود قابل توجهی پیدا کرده‌اند که در اینجا نیز هدف بهبود معیارها برای پیش‌بینی حالات مربوط به وقوع خطا می‌باشد.

جدول ۲- معیارهای استخراج شده از ماتریس درهم‌ریختگی برای روش KNN با اعمال روش متعادل‌سازی برچسب‌ها

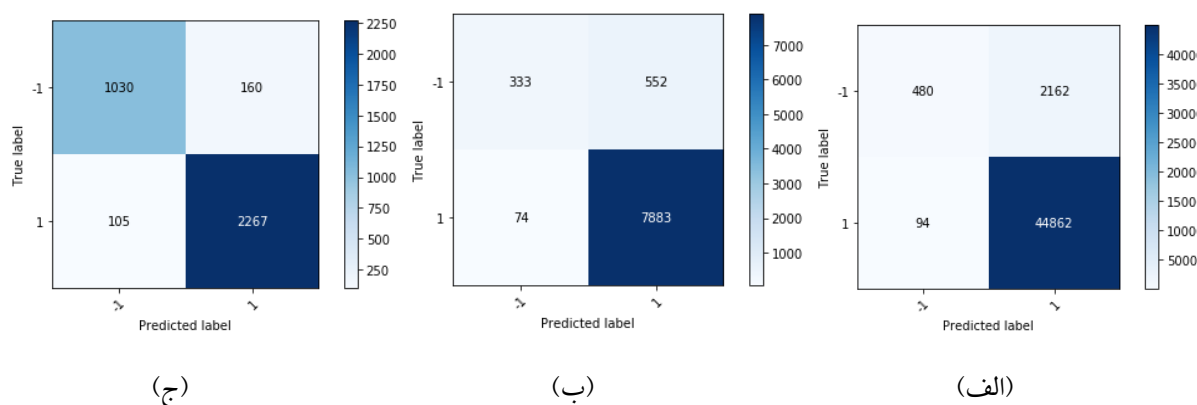
برچسب ۱-	دقت (P)	صحت	اف-۱	برچسب ۱	دقت (P)	صحت	اف-۱
نرمال	۰/۸۹	۰/۹۴	۰/۹۲	نرمال	۰/۹۴	۰/۸۹	۰/۹۱
بد	۰/۹	۰/۹۴	۰/۹۲	بد	۰/۹۴	۰/۹۰	۰/۹۲

### ۲-۳ - روش SVM

روش ماشین بردار پشتیبان روش نسبتاً جدیدی است که توانایی پیاده‌سازی دسته‌بندی‌های خطی و غیرخطی را به نحو مناسبی داراست. هدف این الگوریتم پیدا کردن یک ابرصفحه در فضای  $N$  بعدی ( $N$  ویژگی یا برچسب) است که به‌صورت مجزا داده‌های

<sup>1</sup> Harmonic mean

موجود را دسته‌بندی نماید. این روش از هسته‌های مختلفی برای دسته‌بندی داده‌ها بهره می‌برد که از آن جمله می‌توان به هسته‌های خطی، چندجمله‌ای و  $rbf^1$  اشاره کرد. اگر از هسته چندجمله‌ای یا  $rbf$  استفاده شود باید یک ضریب تحت عنوان گاما نیز در محاسبات وارد شود، همچنین در حالت چندجمله‌ای باید درجه مربوط به آن معین گردد. با توجه به نتایج شبیه‌سازی‌ها و عدم برتری محسوس سایر هسته‌ها برای این مطالعه، در مقاله حاضر از هسته خطی استفاده شده است که نسبت به بقیه، بار محاسباتی کمتری به سیستم تحمیل کرده و زمان کمتری نیاز دارد. در اینجا نیز داده‌ها با همان نسبت ذکر شده در بخش قبل برای آموزش و آزمایش تقسیم‌بندی شده و نتایج بدست آمده مورد بررسی قرار گرفته‌اند. میانگین دقت پیش‌بینی روش SVM نیز برای داده‌های آزمون با داده‌های اولیه غیرمتعادل در هر یک از شرایط جوی نرمال، بد و وخیم به ترتیب برابر با  $0.95/26$ ،  $0.92/92$  و  $0.92/56$  به دست آمده است. نتایج به دست آمده برای ماتریس‌های درهم‌ریختگی در هر یک از حالات در شکل ۸ نشان داده شده است. همچنین معیارهای استخراج شده از ماتریس درهم‌ریختگی برای روش SVM در جدول ۳ آورده شده است.



شکل ۸. نمودارهای مربوط به ماتریس‌های درهم‌ریختگی برای روش SVM در هر یک از سه حالت (الف) نرمال (ب) بد (ج) وخیم

جدول ۳. معیارهای استخراج شده از ماتریس درهم‌ریختگی برای روش SVM

برچسب ۱-	دقت (P)	صحت	اف-۱	برچسب ۱	دقت (P)	صحت	اف-۱
نرمال	۰/۸۲	۰/۱۸	۰/۳۰	نرمال	۰/۹۶	۱/۰۰	۰/۹۸
بد	۰/۸۲	۰/۴۲	۰/۵۵	بد	۰/۹۴	۰/۹۹	۰/۹۶
وخیم	۰/۹۲	۰/۸۴	۰/۸۸	وخیم	۰/۹۳	۰/۹۶	۰/۹۴

مشابه توضیحات ارائه شده برای روش KNN، نتایج به دست آمده از ماتریس درهم‌ریختگی با اعمال روش متعادل‌سازی دسته‌ها در جدول ۴ آورده شده‌اند. هرچند با اعمال این روش مقادیر معیارهای تحت بررسی برای برچسب ۱ به صورت جزئی دچار کاهش گردیده‌اند اما نتایج مربوط به دسته کوچک‌تر (یعنی برچسب ۱-) به مقدار قابل توجهی بهبود یافته است. مقادیر دقت (A) برای وضعیت‌های نرمال و بد در این حالت برابر با  $0.88/3$  و  $0.90/11$  به دست آمده است. همان‌طور که قبلاً نیز توضیح داده شد این

<sup>1</sup> Radial basis function

کاهش‌های جزئی ناشی از حذف برخی از نمونه‌ها از کلاس بزرگ‌تر (برچسب ۱) است. با این وجود توجه اصلی در اینجا بر بهبود معیارهای مربوط به برچسب ۱- می‌باشد.

جدول ۴- معیارهای استخراج‌شده از ماتریس درهم‌ریختگی برای روش SVM با اعمال روش متعادل‌سازی برچسب‌ها

برچسب ۱-	دقت (P)	صحت	اف-۱	برچسب ۱	دقت (P)	صحت	اف-۱
نرمال	۰/۹۰	۰/۸۷	۰/۸۸	نرمال	۰/۸۷	۰/۹۰	۰/۸۸
بد	۰/۸۹	۰/۹۱	۰/۹۰	بد	۰/۹۲	۰/۸۹	۰/۹۰

### ۳-۳- روش درخت تصمیم

درخت تصمیم یکی از مشهورترین الگوریتم‌های موجود برای دسته‌بندی و پیش‌بینی داده‌ها در حوزه یادگیری ماشین است. درخت تصمیم در واقع یک فلوچارت با ساختاری مشابه درخت می‌باشد که در آن هر گره داخلی نشان‌دهنده یک آزمون، هر شاخه نشان‌دهنده خروجی آزمون و هر برگ دارنده یک برچسب (نتیجه حاصل برای داده‌ها) می‌باشد. مسیر حرکت درخت تصمیم از ریشه (اولین گره) شروع شده و در نهایت به برگ‌ها ختم می‌گردد. این روش از معیارهای مختلفی برای جداسازی داده‌ها بهره می‌برد که از آن جمله می‌توان به معیارهای آنترپی<sup>۱</sup>، ضریب جینی<sup>۲</sup> و کای دو<sup>۳</sup> اشاره کرد.

در روش معمول برای بررسی مدل‌های یادگیری ماشین، داده‌ها به دو بخش آموزشی و آزمایشی تقسیم می‌شوند. در این حالت صرفاً یک دسته داده برای آموزش و یک دسته داده برای آزمایش وجود خواهد داشت. به منظور استفاده حداکثری از داده‌های موجود می‌توان آن‌ها را به چند دسته (K دسته) تقسیم‌بندی کرده و هر بار از یک دسته برای آزمایش و از بقیه برای آموزش استفاده نمود که این روش، روش اعتبارسنجی متقابل K بخشی<sup>۴</sup> نامیده می‌شود. در اینجا دقت دسته‌بندی‌ها با استفاده از درخت تصمیم در حالت معمول برای سه حالت نرمال، بد و وخیم به ترتیب برابر ۹۲/۷٪، ۹۰/۶۹٪ و ۹۰/۱۹٪ به دست آمده است درحالی‌که با اعمال روش اعتبارسنجی متقابل با ۴ دسته مقدار این دقت‌ها افزایش یافته و به ترتیب به صورت ۹۴/۴۷٪، ۹۲/۴۷٪ و ۹۲/۳۱٪ به دست می‌آیند. انتخاب مقدار پارامترها برای روش‌های مورد استفاده در یادگیری ماشین یکی از موارد تاثیر گذار و مهم در دقت نهایی مدل می‌باشد. برای به دست آوردن مقادیر مناسب برای این پارامترها روش‌های مختلفی وجود دارد. یکی از این روش‌ها انتخاب و آزمایش تمام مقادیر مختلف (تا حد معین شده) و بررسی نتایج حاصل است. این روش با وجود ارائه نتیجه مناسب می‌تواند به خصوص برای مسائل بزرگ بسیار زمان‌بر باشد. روش دیگر، استفاده از جستجوی تصادفی<sup>۵</sup> برای انتخاب مقادیر پارامترها و بررسی نتایج آن‌هاست. با اعمال این روش به منظور تنظیم پارامترهای درخت تصمیم مقادیر دقت‌ها به ترتیب برابر ۹۴/۶۷٪، ۹۲/۹۲٪ و ۹۲/۴۶٪ به دست آمده است. معیار جداساز انتخاب‌شده با استفاده از الگوریتم جستجوی تصادفی در این حالت، معیار جینی می‌باشد.

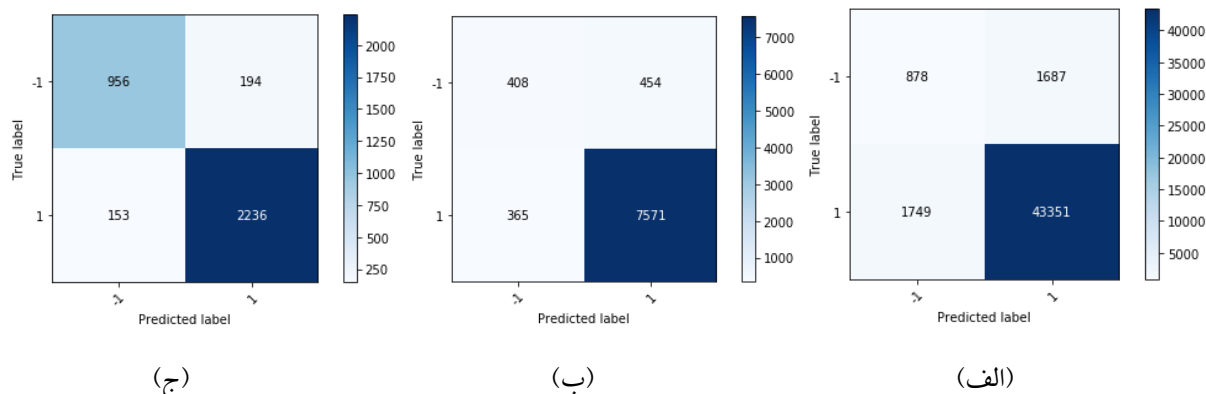
<sup>1</sup> Entropy

<sup>2</sup> Gini index

<sup>3</sup> X<sup>2</sup>

<sup>4</sup> K-fold cross validation

<sup>5</sup> Randomized search



شکل ۹. نمودارهای مربوط به ماتریس‌های درهم‌ریختگی برای روش درخت تصمیم در حالات (الف) نرمال (ب) بد (ج) وخیم

جدول ۵. معیارهای استخراج شده از ماتریس درهم‌ریختگی برای روش درخت تصمیم

برچسب ۱-	دقت (P)	صحت	اف-۱	برچسب ۱	دقت (P)	صحت	اف-۱
نرمال	۰/۳۳	۰/۳۴	۰/۳۴	نرمال	۰/۹۶	۰/۹۵	۰/۹۶
بد	۰/۵۳	۰/۴۷	۰/۵۰	بد	۰/۹۴	۰/۹۵	۰/۹۵
وخیم	۰/۸۶	۰/۸۳	۰/۸۵	وخیم	۰/۹۲	۰/۹۴	۰/۹۳

با توجه به کم بودن تعداد خطاها و داده‌های مربوط به آن‌ها در حالت نرمال درخت تصمیم نسبت به دو الگوریتم بررسی شده دیگر نتیجه ضعیف تری کسب کرده است. این موضوع با توجه به مقادیر P در جدول ۵ نیز مشهود است.

اعمال روش متعادل‌سازی داده‌ها در روش درخت تصمیم نیز باعث بهبود عملکرد مدل شده است؛ نتایج مربوطه در جدول ۶ آورده شده است. مقادیر دقت‌های (A) به دست آمده با اعمال این روش برای دو حالت نرمال و بد به ترتیب برابر ۰/۹۲/۰۸ و ۰/۹۲/۰۷٪ می‌باشند که در حالت بد نسبت به روش درخت تصمیم معمولی، بهبود یافته اما در مقایسه با حالتی که در آن‌ها از اعتبار سنجی متقابل یا جستجوی تصادفی استفاده شده اندکی کمتر می‌باشند. به هر ترتیب مباحث مربوط به بهبود نتایج معیارهای حاصل از ماتریس درهم‌ریختگی و نیز توضیحات ارائه شده برای دو روش KNN و SVM در این خصوص در اینجا نیز معتبرند.

جدول ۶. معیارهای استخراج شده از ماتریس درهم‌ریختگی برای روش درخت تصمیم با اعمال روش متعادل‌سازی برچسب‌ها

برچسب ۱-	دقت (P)	صحت	اف-۱	برچسب ۱	دقت (P)	صحت	اف-۱
نرمال	۰/۹۲	۰/۹۲	۰/۹۲	نرمال	۰/۹۲	۰/۹۲	۰/۹۲
بد	۰/۹۲	۰/۹۳	۰/۹۳	بد	۰/۹۳	۰/۹۲	۰/۹۳

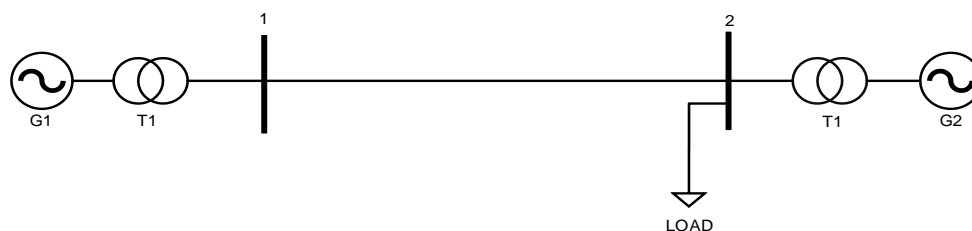
### ۳-۴- تحلیل و مقایسه روش‌ها و معیارهای حاصل برای مجموعه داده‌های اعمالی

با توجه به نتایج حاصل از سه روش مطرح شده یعنی SVM، KNN و درخت تصمیم می‌توان گفت که از لحاظ نزدیکی پیش‌بینی‌ها به مقادیر واقعی خود (یعنی دقت پیش‌بینی‌ها A) در هر سه وضعیت نرمال، بد و وخیم بدون استفاده از روش متعادل‌سازی داده‌ها، SVM نتایج بهتری را ارائه نموده است. از طرف دیگر، از این منظر درخت تصمیم بدون استفاده از اعتبار سنجی متقابل نتایج ضعیف‌تری را نسبت به دو روش دیگر به دست داده است. با اعمال روش اعتبار سنجی متقابل با چهار دسته، نتایج حاصل از درخت تصمیم بهبود یافته‌اند اما غیر از وضعیت جوی وخیم که نسبت به روش KNN دقت بهتری به دست آورده در بقیه موارد هنوز هم در رده سوم قرار می‌گیرند.

در مقایسه‌ی روش‌ها با استفاده از نتایج بدست آمده از ماتریس درهم‌ریختگی، با توجه به مقدار P (دقت) که طبق توضیحات قبلی نشان‌دهنده تکرارپذیری نتایج است، SVM برای برچسب ۱- (وقوع خطا) در هر سه وضعیت جوی به مراتب بهتری کسب کرده است. در اینجا نیز KNN در رتبه دوم و باز درخت تصمیم در جایگاه سوم قرار می‌گیرند. برای برچسب ۱ (عدم وقوع خطا) تقریباً دقت هر سه روش در یک محدوده است هر چند KNN به صورت جزئی عملکرد مطلوب‌تری داشته است. نکته‌ای که باید به آن توجه کرد این است که هر یک از معیارهای مطرح شده برای ماتریس درهم‌ریختگی برای مطالعه‌ای خاص و اهداف آن نسبت به بقیه اهمیت بیشتری خواهد داشت. در اینجا اهمیت بهبود نتایج مربوط به حالت وقوع خطا در دو حالت نرمال و بد لازم داشت تا با اعمال روش‌هایی مقادیر معیارهای تحت بررسی بهبود داده شوند؛ لذا با اعمال روش متعادل‌سازی داده‌ها در کلاس‌های مختلف این مشکل برطرف شده و نتایج بهبود داده شدند. در حالت کلی برای این بخش از مسئله با متعادل‌سازی داده‌ها روش‌های KNN و درخت تصمیم عملکرد بهتری نسبت به روش SVM ارائه داده‌اند.

### ۳-۵- بررسی مدل‌ها برای یک سیستم شبیه‌سازی شده

برای بررسی بهتر عملکرد مدل، از شبیه‌سازی یک سیستم دو ماشینه و داده‌های حاصل از آن استفاده شده است. سیستم به کار گرفته شده به صورت کلی مشابه شبکه‌های مورد استفاده در مراجع [۴۱-۳۸] می‌باشد. داده‌های تولیدی شامل نمونه‌های مربوط به انواع خطاهای خط به زمین (تک‌فاز، دوفاز به هم و سه‌فاز به هم و به زمین) با مشخصات مختلف می‌باشند. از ۶۰۰۰۰ داده‌ی مختلف برای آموزش و آزمایش مدل‌ها استفاده شده است. برای تولید این داده‌ها مکان، مقاومت و زوایای اعمال خطاها بر روی خط انتقال ۱۰۰ کیلومتری تغییر داده شده و پارامترهای مورد نیاز شامل ولتاژ، توان اکتیو و راکتیو در هر حالت ثبت گردیده است. برای شبیه‌سازی این سیستم از سیمولینک MATLAB استفاده شده و سپس داده‌ها وارد مدل ایجاد شده در Python شده‌اند؛ اثرات ناشی از شرایط جوی مختلف نیز مطابق توضیحات بخش‌های قبلی در بخش کدنویسی در Python اعمال گردیده‌اند. نمودار سیستم مورد نظر در شکل ۱۰ و مشخصات عمومی آن در جدول ۷ آورده شده‌اند.



شکل ۱۰. مدل خط انتقال در یک شبکه دو ماشینه

جدول ۷. مشخصات کلی سیستم نشان داده شده در شکل ۱۰

مشخصات کلی سیستم	فرکانس (Hz)	ولتاژ خط (KV)	طول خط (km)	ولتاژ ژنراتورها (KV)	نسبت تبدیل ترانسفورماتورها	بار (متغیر)	مقاومت خط (ohm/km)	اندوکتانس خط (H/km)	کاپاسیتانس خط (F/km)
	۶۰	۲۳۰	۱۰۰	۱۳/۸	۱۳/۸:۲۳۰	MW1-100 MVar1K-1	۰/۰۱۷۵	۳e-۰/۸۷۳۷	۹e-۱۳/۳۳
							ترتیب مثبت	ترتیب مثبت	ترتیب مثبت
							۰/۲۷۸۵	۳e-۳/۲۲۰	۹e-۸/۲۹۷
							ترتیب صفر	ترتیب صفر	ترتیب صفر

نتایج حاصل برای روش‌های ارائه شده در ادامه آورده شده است. این نتایج برای حالات نرمال و بد با اعمال روش‌های متعادل‌سازی داده‌ها ارائه شده‌اند. همانطور که اشاره شد داده‌های اولیه شامل ۶۰۰۰۰ نمونه می‌باشد که از میان آن‌ها صرفاً حدود ۱۱۰۰ داده مربوط به خطاهای مختلف گنجانده شده و با این وجود نتایج حاصل دقت مناسبی در تشخیص خطاها به دست آورده‌اند. با توجه به محدوده تغییرات کمتر پارامترها، با استفاده از داده‌ها با توزیع نرمال مقادیر دقت و سایر معیارها برای حالت شبیه‌سازی شده بهتر از حالت قبلی خواهد بود که این موضوع در نتایج نیز مشهود است. برای حالت وخیم پیش‌بینی‌ها بدون استفاده از روش متعادل‌سازی داده‌ها نزدیک به ۱۰۰ درصد می‌باشد که برای جلوگیری از طولانی شدن بحث در جدول ۸ آورده نشده است.

جدول ۸. معیارهای استخراج شده از ماتریس درهم‌ریختگی برای روش‌های مختلف

	برچسب ۱-	دقت (P)	صحت	اف-۱	برچسب ۱-	دقت (P)	صحت	اف-۱
KNN	نرمال	۰/۹۸۹	۰/۹۶	۰/۹۸	نرمال	۰/۹۶	۰/۹۸	۰/۹۷۹
	بد	۰/۹۹	۰/۹۹	۰/۹۹	بد	۰/۹۹	۰/۹۹	۰/۹۹
	درخت	نرمال	۰/۹۸	۰/۹۶	۰/۹۷۸	نرمال	۰/۹۶	۰/۹۸
تصمیم	بد	۰/۹۹	۰/۹۹	۰/۹۸	بد	۰/۹۶	۰/۹۹	۰/۹۹
	SVM	نرمال	۰/۹۸	۰/۹۶	۰/۹۸	نرمال	۰/۹۶	۰/۹۸
بد		۰/۹۹	۰/۹۹	۰/۹۹	بد	۰/۹۹	۰/۹۹	۰/۹۹

## ۴- نتیجه گیری

در این مقاله چارچوبی برای پیش‌بینی وقوع خطا در خط انتقال با استفاده از مدل‌های یادگیری ماشین ارائه گردید. علاوه بر بررسی تاثیر تغییرات متغیرهای الکتریکی شامل توان‌های اکتیو، راکتیو و ولتاژ شین‌ها در بروز خطا در خط انتقال عوامل ناشی از تغییرات شرایط آب و هوایی نیز در سه وضعیت نرمال، بد و وخیم در مدل‌ها لحاظ شده‌اند. در این مقاله از روش‌های یادگیری ماشینی SVM، KNN و درخت تصمیم که جزء روش‌های یادگیری نظارت شده می‌باشند، استفاده گردیده و از فنون مختلف برای بهبود معیارهای بیان‌کننده کیفیت آن‌ها بهره گرفته شده است. به منظور افزایش کارایی عملی روش‌ها، سعی شده است تا در ایجاد داده‌های مختلف برای بررسی مدل‌ها تا حد امکان اطلاعات به حالات واقعی نزدیک باشند. نتایج تحلیل‌ها در دو حالت داده‌های

نامتعادل و نیز داده‌های متعادل سازی شده ارائه شده و مورد مقایسه قرار گرفته‌اند. برآیند شبیه‌سازی‌ها و نتایج ارائه شده نشان‌دهنده کارایی مدل‌های ارائه شده و نیز دقت مناسب در پیش‌بینی خطاها با اعمال تاثیرات کلیه عوامل مطرح شده می‌باشد. تمامی مدل‌سازی‌ها و روش‌های یادگیری ماشینی مورد بحث در این مقاله با استفاده از نرم‌افزار Python پیاده‌سازی شده‌اند. همچنین برای تولید بخشی از داده‌های مورد استفاده برای بررسی عملکرد مدل‌ها از نرم‌افزار MATLAB استفاده گردیده است. در کارهای آتی، مدل‌ها جهت بهبود برچسب گذاری و تشخیص انواع گسترده‌تر خطا و همچنین تعیین نوع و مکان خطاها تعمیم داده خواهند شد.

## مراجع

- [1] Arun G. Phadke and J. S. Thorp, *Computer Relaying for Power Systems*. 2009.
- [2] A. T. Johns and S. K. Salman, *Digital Protection for Power System*. Stevenage, United Kingdom: Institution of Engineering and Technology, 1997.
- [3] P. M. Anderson, *Power System Protection*. New York: McGraw-Hill : IEEE Press, 1999.
- [4] A. G. Phadke and J. S. Thorp, "Expose hidden failures to prevent cascading outages [in power systems]," *IEEE Comput. Appl. Power*, vol. 9, no. 3, pp. 20–23, Jul. 1996.
- [5] M. S. Sachdev and M. A. Baribeau, "A New Algorithm for Digital Impedance Relays," *IEEE Trans. Power Appar. Syst.*, vol. PAS-98, no. 6, pp. 2232–2240, 1979.
- [6] A. A. Girgis, "A New Kalman Filtering Based Digital Distance Relay," *IEEE Trans. Power Appar. Syst.*, vol. PAS-101, no. 9, pp. 3471–3480, 1982.
- [7] A. Gomez-Exposito, A. J. Conejo, and C. Canizares, *Electric Energy Systems: Analysis and Operation*, 2nd ed. CRC Press, 2018.
- [8] Volker Quaschnig, *Understanding Renewable Energy Systems*. Routledge, 2016.
- [9] K. Amasyali and N. El-Gohary, "A review of data-driven building energy consumption prediction studies," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1192–1205, Jan. 2018.
- [10] Y. Peng, A. Rysanek, Z. Nagy, and A. Schlüter, "Using machine learning techniques for occupancy-prediction-based cooling control in office buildings," *Appl. Energy*, vol. 211, pp. 1343–1358, Feb. 2018.
- [11] R. Sathya and A. Abraham, "Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification," *Int. J. Adv. Res. Artif. Intell.*, vol. 2, no. 2, 2013.
- [12] M. Jamil, S. K. Sharma, and R. Singh, "Fault detection and classification in electrical power transmission system using artificial neural network," *Springerplus*, vol. 4, no. 1, pp. 1–13, Dec. 2015.
- [13] S. S. Nagam, E. Koley, and S. Ghosh, "Artificial Neural Network Based Fault Locator For Three Phase Transmission Line with STATCOM," in *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICIC)*, 2017, pp. 1–4.
- [14] T. Bouthiba, "Fault location in EHV transmission lines using artificial neural networks," *Int. J. Appl. Math. Comput. Sci.*, vol. 14, pp. 69–78, Jan. 2004.
- [15] N. Zhang and M. Kezunovic, "A real time fault analysis tool for monitoring operation of transmission line protective relay," *Electr. Power Syst. Res.*, vol. 77, no. 3–4, pp. 361–370, Mar. 2007.
- [16] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, and D. Xu, "Detection of Power Line Insulator Defects Using Aerial Images Analyzed With Convolutional Neural Networks," *IEEE Trans. Syst. Man, Cybern. Syst.*, pp. 1–13, 2018.
- [17] A. N. Hasan, P. S. P. Eboule, and B. Twala, "The use of machine learning techniques to classify power transmission line fault types and locations," in *2017 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM) & 2017 Intl Aegean Conference on Electrical Machines and Power Electronics (ACEMP)*, 2017, pp. 221–226.
- [18] Joe-Air Jiang, Jun-Zhe Yang, Ying-Hong Lin, Chih-Wen Liu, and Jih-Chen Ma, "An adaptive PMU based fault detection/location technique for transmission lines. I. Theory and algorithms," *IEEE Trans. Power Deliv.*, vol. 15, no. 2, pp. 486–493, Apr. 2000.
- [19] Y. H. Lin, C. W. Liu, and C. S. Chen, "A new PMU-based fault detection/location technique for transmission lines with consideration of arcing fault discrimination - Part II: Performance



- evaluation,” *IEEE Trans. Power Deliv.*, vol. 19, no. 4, pp. 1594–1601, Oct. 2004.
- [20] M. El Chamie, K. G. Lore, D. M. Shila, and A. Surana, “Physics-Based Features for Anomaly Detection in Power Grids with Micro-PMUs,” in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–7.
- [21] S. Ekici and S. Yildirim, *Fault Location Estimation on Transmission Lines Using Wavelet Transform and Artificial Neural Network.*, vol. 1. 2006.
- [22] S. K. Shukla, E. Koley, and S. Ghosh, “Protection Scheme for Shunt Faults in Six-Phase Transmission System Based on Wavelet Transform and Support Vector Machine,” in *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICIC)*, 2017, pp. 1–5.
- [23] S. A. Shaaban and M. A. Abdel-Moamen, “Wavelet signal energy with RBFNN and GRNN for fault classification in transmission line with series compensator,” in *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, 2017, pp. 1–6.
- [24] H. A. Shiddieqy, F. I. Hariadi, and T. Adiono, “Effect of Sampling Variation in Accuracy for Fault Transmission Line Classification Application Based On Convolutional Neural Network,” in *2018 International Symposium on Electronics and Smart Devices (ISESD)*, 2018, pp. 1–3.
- [25] S. R. Samantaray and P. K. Dash, “Transmission line distance relaying using machine intelligence technique,” *IET Gener. Transm. Distrib.*, vol. 2, no. 1, pp. 53–61, 2008.
- [26] A. Oonsivilai and S. Saichoomdee, “Appliance of recurrent neural network toward distance transmission lines protection,” in *TENCON 2009 - 2009 IEEE Region 10 Conference*, 2009, pp. 1–4.
- [27] P. P. K. Chan, J. Zhu, and Z.-W. Qiu, “Three-phase fault location based on Multiple Classifier System in double-circuit transmission lines,” in *2012 International Conference on Wavelet Analysis and Pattern Recognition*, 2012, pp. 250–254.
- [28] Y. Zhang, J. Chen, Q. Fang, and Z. Ye, “Fault Analysis and Prediction of Transmission Line Based on Fuzzy K-Nearest Neighbor Algorithm,” pp. 894–899, 2016.
- [29] A. A. Rahim, I. Z. Abidin, F. Tarlochan, and M. F. Hashim, “Thermal rating monitoring of the TNB overhead transmission line using line ground clearance measurement and weather monitoring techniques,” in *2010 4th International Power Engineering and Optimization Conference (PEOCO)*, 2010, pp. 274–280.
- [30] W. P. Davis, “THE ANALYSIS OF FAULTS IN OVERHEAD TRANSMISSION LINES,” California State University, Sacramento, 2012.
- [31] T. I. A. H. Mustafa, L. H. Meyer, S. H. L. Cabral, H. D. Almaguer, and L. B. Puchale, “Study of the correlation between weather conditions and protection trips in a 230 kV transmission line in southern Brazil,” in *PES T&D 2012*, 2012, pp. 1–5.
- [32] S. YANG *et al.*, “Failure probability estimation of overhead transmission lines considering the spatial and temporal variation in severe weather,” *J. Mod. Power Syst. Clean Energy*, vol. 7, no. 1, pp. 131–138, Jan. 2019.
- [33] [فرج زاده اصل، م.، عزیززاده، خ.، “تحلیل زمانی و مکانی طوفانهای گرد و خاک در ایران،” *برنامه‌ریزی و آمایش فضا*، ۱۳۹۰، دوره ۱۵، شماره ۱، صفحه: ۶۵–۸۴
- [34] “Past weather events - Met Office.” [Online]. Available: <https://www.metoffice.gov.uk/weather/learn-about/past-uk-weather-events#y2015>. [Accessed: 06-Dec-2019].
- [35] Koo Ping Shung, “Accuracy, Precision, Recall or F1? - Towards Data Science.” [Online]. Available: <https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9>. [Accessed: 06-May-2015].
- [36] G. Batista, R. Prati, and M.-C. Monard, “A Study of the Behavior of Several Methods for Balancing machine Learning Training Data,” *SIGKDD Explor.*, vol. 6, pp. 20–29, Jun. 2004.
- [37] G. Batista, A. Bazzan, and M.-C. Monard, *Balancing Training Data for Automated Annotation of Keywords: a Case Study*. 2003.
- [38] Y. Q. Chen, O. Fink, and G. Sansavini, “Combined Fault Location and Classification for Power Transmission Lines Fault Diagnosis With Integrated Feature Extraction,” *IEEE Trans. Ind. Electron.*, vol. 65, no. 1, pp. 561–569, 2017.

- [39] T. Grids and V. W. Synchronized, "Traveling-Wave-Based Fault-Location Technique for Transmission Grids Via Wide-Area Synchronized Voltage Measurements," vol. 27, no. 2, pp. 1003–1011, 2012.
- [40] P. Ray, B. K. Panigrahi, and N. Senroy, "Extreme learning machine based fault classification in a series compensated transmission line," *PEDES 2012 - IEEE Int. Conf. Power Electron. Drives Energy Syst.*, pp. 1–6, 2012.
- [41] P. Ray and D. P. Mishra, "Support vector machine based fault classification and location of a long transmission line," *Eng. Sci. Technol. an Int. J.*, vol. 19, no. 3, pp. 1368–1380, 2016.