تعیین عوامل مؤثر در پیش‌بینی انواع خرابی روتور بالگرد

هدف: هدف این مقاله بررسی و شناسایی متغیرهای مؤثر در وقوع حوادث بالگردی که ناشی از انواع مختلف خرابی روتور بالگرد هستند، می باشد. این عوامل مؤثر شامل شرایط پرواز، شرایط تعمیر و نگهداری و پیکربندی بالگرد است. با این رویکرد، می‌توان حوادث را به طور مؤثرتری بررسی کرد و ایمنی پرواز را به طور قابل توجهی بهبود بخشید.

روش‌شناسی پژوهش: با تجزیه و تحلیل ۱۳۵ حادثه خرابی روتور از یک مجموعه داده جامع شامل ۵۶۵۲ حادثه مرتبط با هلیکوپتر، هشت کلاس خرابی شناسایی شد. با نظرسنجی از متخصصان و بررسی مقالات در زمینه حوادث بالگردی، نه ویژگی به عنوان عوامل احتمالی موثردر وقوع حوادث بالگردی پیشنهاد شد. اهمیت این عوامل با کمک پنج روش انتخاب ویژگی بررسی شده است. حداکثر وزن برخاستن، تعداد ساعات پرواز بعد از آخرین بازرسی، نوع آخرین بازرسی، توان موتور بالگرد، کل ساعات پروازی، ارتفاع، سرعت باد، جهت باد و فاز پرواز به عنوان ویژگی‌های ورودی در نظر گرفته شده‌اند. پنج روش شناخته‌شده انتخاب ویژگی، شامل ماتریس همبستگی، روش Extreme Gradient Boosting ، اطلاعات متقابل، یادگیری عمیق و روش شبکه عصبی برای یافتن عوامل مهم‌تر استفاده شدند.

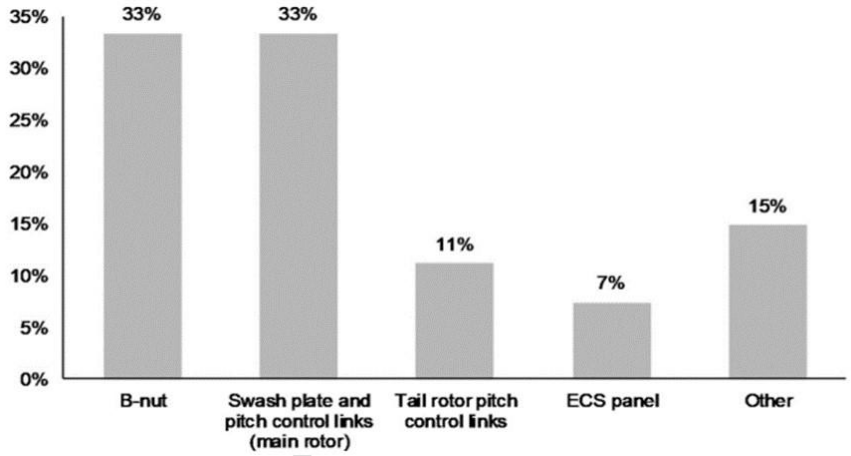
یافته‌ها: در نهایت، «حداکثر وزن »، « توان موتور بالگرد»، «فاز پرواز» و «ساعات پرواز بالگرد » به عنوان متغیرهایی با بالاترین درجه اهمیت در پیش‌بینی کلاس خرابی روتور بالگرد شناسایی شدند که در مکانیک پرواز نیز توجیه قوی و قابل قبولی دارند.

اصالت/ارزش افزوده علمی: تفاوت کار حاضر با مطالعات مشابه این بود که متغیرهای بیشتری، نظیر شرایط پرواز و پیکربندی بالگرد، در نظر گرفته شدند، برخلاف سایر مطالعات که درآن‌ها مجموعه‌ای محدود از متغیرها در نظر گرفته شدند. با اولویت‌بندی این متغیرها، یافته‌ها با هدف افزایش دقت پیش‌بینی، قابلیت اطمینان و ایمنی پرواز، راه را برای اقدامات پیشگیرانه در پیشگیری از خرابی روتور هموار می‌کنند.

کلیدواژه‌ها: انتخاب ویژگی، پیش‌بینی خرابی، روتور هلیکوپتر، شبکه عصبی، یادگیری عمیق

۱- مقدمه

تصادفات هلیکوپتر به اندازه سایر صنایع مانند مواد شیمیایی، نفت، گاز یا هوانوردی مورد مطالعه قرار نگرفته است [۱]. بالگردها برای حمل و نقل هوایی حیاتی هستند، اما سوابق ایمنی نشان می‌دهد که در سال‌های اخیر کمتر به آن پرداخته شده است [۲]. مطالعات عمده در مورد تصادفات هلیکوپتر از جنبه‌های دیگر به این موضوع پرداخته‌اند. به عنوان مثال، برخی تصادفات را در عملیات‌های مختلف [۳،۴] یا مناطق جغرافیایی مختلف [۵،۶] یا خطای انسانی بررسی می‌کنند. هدف از شناسایی عوامل مؤثر بر تصادفات، کاهش جراحات، تلفات و خسارات مالی و افزایش قابلیت اطمینان پرواز است. میزان تصادفات مرگبار هلیکوپترهای غیرنظامی در ایالات متحده تقریباً ۱۷ برابر بیشتر از میزان تصادفات مرگبار خودروهای سواری است، به این معنی که سالانه بیش از ۱۲۰ تصادف هلیکوپتر رخ می‌دهد که صدها جراحت و مرگ و میر و چند صد میلیون دلار خسارت برای حدود ۱۳۰۰۰ نفر به همراه دارد [۲،۷]. به طور متوسط، ۰.۷ تصادف مرگبار در هر ۱۰۰۰۰۰ ساعت پرواز به تصادفات مرگبار هلیکوپترهای غیرنظامی ایالات متحده اختصاص دارد [۷]. بین ۵۰ تا ۷۰ درصد از کل تصادفات بالگرد به خطاهای خلبان و بقیه به عوامل دیگری مانند فرسودگی، خوردگی، سایش قطعات و غیره نسبت داده می‌شود [۸]. شکل ۱ طبقه‌بندی خطاهای احتمالی در سیستم بالگرد را نشان می‌دهد [۲]، که در آن درصد نسبتاً بالایی از تصادفات بالگرد مربوط به قسمت روتور است.



شکل ۱- طبقه‌بندی خطاهای احتمالی در سیستم هلیکوپتر [۲]

برخی مقالات در این زمینه، تعمیر و نگهداری و بازرسی را به عنوان علل مستقیم تصادفات بالگرد مطرح می‌کنند. به جای تمرکز بر خطاهای انسانی در خلبانان و اپراتورها، برخی مطالعات بر خطاهای تعمیر و نگهداری تمرکز کرده‌اند [۲]. بنی و همکاران با استفاده از یک مدل رگرسیون و ویژگی‌هایی مانند معیارهای آب و هوای نامساعد، حداکثر ارتفاع، معیارهای کنترل، تعداد روزهای گذشته از بازرسی و دمای کابین، تصادفات هوایی را پیش‌بینی کردند [۹]. عدم وجود خلبان دوم و گواهی پزشکی درجه دو به جای گواهی درجه یک، از عوامل حیاتی در مطالعات مربوط به تصادفات بالگرد پزشکی هستند [۱۰]. ژو و همکاران از متغیرهای تعداد پره‌های روتور اصلی، تعداد موتورها، قطر روتور و وزن استفاده می‌کنند و با کمک شبکه‌های عصبی عمیق، وقوع یا عدم وقوع تصادف در یک بالگرد را پیش‌بینی می‌کنند. آنها پیکربندی‌های مختلف بالگرد را در وزن‌های مختلف بررسی می‌کنند و تأثیر متغیرهای وزن، قطر روتور و تعداد پره‌ها را بر وقوع تصادف بررسی می‌کنند [۱۱]. چرچول و همکاران به نرخ تصادفات و پیشرفت کم در پیشگیری از تصادفات هلیکوپتر در دهه گذشته پرداخته و سپس رابطه بین تصادفات را با تعداد پره‌های روتور اصلی و نوع موتور تجزیه و تحلیل کردند [۷]. مهندسی ویژگی با هدف تمرکز بر داده‌ها به جای تمرکز بر اعمال الگوریتم‌های پیچیده‌تر برای حل یک مسئله یادگیری ماشینی انجام می‌شود. بسیاری از مقالات در مورد استفاده از روش‌های مختلف انتخاب ویژگی در سیستم‌های مختلف، مانند دستگاه‌های سوئیچ حمل و نقل ریلی [۱۲]، تشخیص نقص خطوط لوله نفت و گاز [۱۳] و روش تشخیص خطای سیستم روتور [۱۴] بحث می‌کنند. او و همکارانش تکنیک‌های مختلف یادگیری ماشین، از جمله انتخاب ویژگی و برخی روش‌های دیگر را برای تشخیص خستگی راننده کامیون با جمع‌آوری ۱۸ ویژگی در سه دسته ویژگی‌های فیزیولوژیکی، خودرویی و

زمانی [۱۵] به کار بردند. وقتی داده‌ها شامل تعداد زیادی ویژگی هستند و حجم بالایی دارند، تکنیک‌های مهندسی ویژگی به طور قابل توجهی به استخراج دانش از انبوه داده‌ها کمک می‌کنند. با مهندسی ویژگی، قبل از اعمال مدل‌های یادگیری ماشین، می‌توان داده‌های خام را به ویژگی‌های مؤثرتری برای پیش‌بینی‌پذیری و تفسیرپذیری تبدیل کرد [۱۶]. انجام صحیح مرحله مهندسی ویژگی می‌تواند منجر به خروجی دقیق‌تری در الگوریتم‌های یادگیری ماشین شود. روش‌های انتخاب ویژگی به سه دسته طبقه‌بندی می‌شوند: روش‌های درون‌سازی‌شده[1] ، فیلتر[2] و روش‌های پوششی[3] [۱۷]. روش‌های درون‌سازی‌شده ، روش‌های انتخاب ویژگی را ادغام می‌کنند و روش‌های فیلتر بر اساس رتبه‌بندی ویژگی‌ها بر اساس مقیاس‌های مختلف، مانند استقلال ویژگی و ویژگی‌های واریانس، کار می‌کنند [۱۸]. روش‌های پوششی، اهمیت ویژگی را بر اساس عملکرد مدل پیش‌بینی‌کننده با استفاده از انتخاب متغیرهای متنوع تشخیص می‌دهند [۱۹]. روش‌های فیلتر از یک تابع امتیازدهی برای رتبه‌بندی ویژگی‌ها استفاده می‌کنند. امتیاز بالا نشان دهنده کارایی یک متغیر است، اطلاعات متقابل و ضریب پیرسون معیارهای شناخته شده‌ای هستند [۲۰]. بنابراین، در کار حاضر علاوه بر سایر روشها، از این روشها نیز استفاده شد. همانطور که گفته شد، موارد زیر را می‌توان از مقالاتی که به بررسی حوادث بالگرد پرداخته‌اند، دریافت: اولاً، هر یک از آنها از یک جنبه خاص مثلا عوامل انسانی یا صرفا پیکربندی بالگرد به حادثه پرداخته است.ثانیاً، مطالعه‌ای که عوامل جامع‌تری مانند شرایط پرواز، شرایط تعمیر، نگهداری و ویژگی‌های فیزیکی بالگرد را بررسی کند، یافت نشد. بنابراین، مطالعه عوامل مختلف در بررسی حوادث بالگرد برای دستیابی به نتایج متمایز جهت بهبود ایمنی مورد نیاز است. از سوی دیگر، عوامل زیادی را می‌توان در حوادث بالگرد در نظر گرفت. با این حال، گنجاندن همه این عوامل در مطالعه می‌تواند دقت مدل‌های پیش‌بینی را کاهش داده و منجر به نتایج نادرست شود. بنابراین، روش‌های انتخاب ویژگی می‌توانند به تعیین دقیق‌تر عواملی که مهم‌تر هستند، کمک کنند. در مدل یادگیری ماشین، ارزیابی این نکته ضروری است که آیا همه متغیرها نقش اساسی در پیش‌بینی مدل دارند یا خیر [۲۱]. هدف این مطالعه بررسی عوامل مهم و تأثیرگذار دروقوع حوادث بالگرد است. این عوامل تأثیرگذار می‌توانند شامل شرایط پرواز، شرایط تعمیر و نگهداری، حالت‌های خرابی و انواع آنها، شرایط فیزیکی بالگرد و بسیاری از موارد دیگر باشند. در کار حاضردر مقایسه با تحقیقات قبلی که رابطه بین حوادث بالگردی و مجموعه‌ای از ویژگی‌ها را در نظر می‌گیرند، میزان متغیرها متنوع تر است. در این مقاله ، با کمک مشورت با متخصصان و بررسی مقالات مرتبط با حوادث بالگردی، برخی از متغیرهای مؤثر بر وقوع حوادث انتخاب می‌شوند تا میزان اثربخشی و اهمیت هر متغیر مشخص شود. سپس، با استفاده از انتخاب ویژگی، ویژگی‌های ضروری شناسایی می‌شوند. ادامه مقاله به شرح زیر سازماندهی شده است. در بخش مواد و روشها، داده‌ها، روش‌ها و متغیرهای ورودی معرفی می‌شوند. در بخش نتایج، خروجی روش‌های انتخاب ویژگی و مهم‌ترین ویژگی‌ها مورد بحث قرار می‌گیرند. در نهایت، در بخش‌های بحث و نتیجه گیری، نتایج پژوهش ، تفسیر و تحلیل شد و پیشنهادات برای تحقیقات آینده بیان شده است.

۲- مواد و روشها

مطالعه حاضر از داده‌های عمومی موجود در مورد حوادث بالگرد از هیئت ملی ایمنی حمل و نقل ایالات متحده آمریکا (NTSB) [4] استفاده می‌کند[۲۲]. به دلیل تفاوت در ویژگی‌های پروازی و نگهداری میان بالگردهای شخصی و بالگردهای تجاری، داده‌های مربوط به این نوع سوانح از بررسی کنار گذاشته شدند. حوادث در این مطالعه شامل آسیب‌های جدی به افراد و آسیب‌های جدی به بالگرد است، بنابراین

[1] Embedded methods

[2] Filter methods

[3] Wrapper methods

[4] National Transportation Safety Board

حوادث بدون آسیب یا با حداقل آسیب در نظر گرفته نمی‌شوند. مجموعه داده‌های مورد استفاده در این مطالعه شامل ۶۳۹۱۳ حادثه هوایی مربوط به انواع مختلف ماشین‌های پرنده مانند هواپیما، بالگرد، پاراگلایدر، بالن بود که در این میان ۵۶۵۲ حادثه آن، مربوط به بالگردها میشد که گزارش مربوط به هر یک، به طور جداگانه و یک به یک بررسی و تجزیه و تحلیل شدند. زمان قابل توجهی صرف مطالعه روایت‌ها و جزئیات همه ۵۶۵۲ حادثه و بررسی دقیق آنها شد تا فقط مواردی که به دلیل خرابی روتور رخ داده است، استخراج شود. همه حوادث ناشی از علل دیگر، مانند خرابی موتور و خطای خلبان، حذف شدند. بنابراین، ۱۳۵ حادثه خرابی روتور به عنوان مطالعات موردی بدست آمد و مورد تجزیه و تحلیل قرار گرفتند. با مطالعه دقیق شرح حادثه روتور، تمام حالت‌های خرابی روتور در هشت کلاس جداگانه طبقه‌بندی شدند:

- کلاس ۱ : شامل ترک و فرسودگی در محور اصلی روتور ، تیغه روتور اصلی یا دم ، و یا توپی[۱]
- کلاس ۲ شامل فرسودگی میلهٔ تغییر گام، لقی در سامانهٔ کنترل[۲]، گیرکردگی در کنترل‌ها[۳] و فرسودگی قطعات مختلف
- کلاس ۳ شامل پدیدهٔ واگرایی روتور و حالت حلقهٔ گردابه‌ای[۴]
- کلاس ۴ شامل عدم تعادل روتور و لرزش بیش از اندازه[۵]
- کلاس ۵ شامل جداشدن وزنهٔ تعادل تیغهٔ روتور[۶]
- کلاس ۶ شامل جدا شدگی پوسته روتور[۷]
- کلاس ۷ شامل فرسودگی در تسمه یا میلهٔ پیچشی– کششی روتور[۸]
- کلاس ۸ شامل میراکنندهٔ پس‌روی یا میراگر پس‌روی[۹] (قطعه‌ای است که حرکت پس‌روی و پیش‌روی تیغه‌های روتور را در صفحهٔ چرخش کنترل می‌کند تا ارتعاش کاهش یابد). میراکنندهٔ پایدارساز[۱۰] (به کاهش نوسانات دم یا پایدارکنندهٔ افقی کمک می‌کند و پایداری پرواز را بهبود می‌بخشد).

---

[1] Hub

[2] Control looseness

[3] Control Seizure

[4] Vortex Ring State

[5] Rotor unbalance and excessive vibration

[6] Rotor blade weight separation

[7] Skin separation

[8] Torsion tension strap or bar fatigue

[9] Lag damper

[10] Stabilizer damper

۱-۲- متغیرهای ورودی

همانطور که قبلاً ذکر شد، بر اساس نظرات کارشناسان و بررسی متون، ویژگی‌های اساسی در وقوع حوادث بالگرد انتخاب و در مطالعه گنجانده شده‌اند. ژو و همکارانش پنج ویژگی را برای پیش‌بینی حوادث بالگرد بررسی کردند: حداکثر وزن برخاستن[1] ؛ قطر پره‌های روتور؛ تعداد پره‌های روتور اصلی [۷،۱۱]؛ نوع موتور [۷]؛ تعداد موتورها [۷] مطالعات نشان می‌دهد که احتمال حادثه در یک بالگرد در ساعات اولیه پس از تعمیرات و نگهداری بیشتر است. با گذشت زمان، احتمال حادثه بالگرد کاهش می‌یابد [۲۳]. حدود ۳۴٪ از حوادث بالگرد در ۱۰ ساعت اول پرواز و ۵۱٪ در ۲۰ ساعت اول پرواز رخ می‌دهد [۲]. بنابراین، زمان سپری شده از آخرین بازرسی می‌تواند به عنوان یک متغیر ورودی در پیش‌بینی وقوع حادثه در نظر گرفته شود. جدول ۱ این ویژگی‌های انتخاب شده را نشان می‌دهد. متغیر خروجی، کلاس خرابی است که در ۸ کلاس در انتهای بخش «مواد و روش‌ها» توضیح داده شده اند.

جدول ۱. متغیرهای ورودی شامل مشخصه های حین پرواز

	نام ویژگی	نوع ویژگی	توصیف ویژگی	نام اختصاری
۱	حداکثر وزن برخاستن	عددی	حداکثر وزن مجازواقعی حین برخاستن برای هلی کوپتردرگیر در حادثه	Cert_max_gr_wt[2]
۲	نوع آخرین بازرسی	کیفی	نوع آخرین بازرسی روی هلی کوپترا مشخص میکند..	Type_last_insp
۳	ساعات کاری بعد از آخرین بازرسی	عددی	تعداد کل ساعت‌های عملیاتی ثبت شده برای بدنهٔ هلی کوپتراززمان آخرین بازرسی تا کنون	Afm_hrs_last_insp
۴	توان موتور	عددی	به توان خروجی موتور اشاره دارد	Power_units
۵	ساعات کاری	عددی	تعداد کل ساعات عملیاتی روی بدنه هواپیما	Afm_hrs
۶	ارتفاع	عددی	فشار بارومتریک همچنین به‌عنوان تنظیم ارتفاع‌سنج (Altimeter Setting) شناخته می‌شود؛ یعنی همان مقدار فشاری که برای تنظیم ارتفاع‌سنج با تغییرات فشار جو مورد استفاده قرار می‌گیرد.	Altitude
۷	سرعت باد	عددی	سرعت باد ثبت‌شده در محل، در زمان وقوع رویداد	Wind_vel_kts
۸	جهت باد (بر حسب درجه	عددی	جهت باد نشان‌داده‌شده در محل، در زمان وقوع رویداد	Wind_dir_deg

[1] Maximum Take-Off Weight

[2] Certified Maximum Gross Weight

	نسبت به شمال مغناطیسی)			
۹	فاز پرواز	کیفی	مرحلهٔ پرواز که رویداد در آن رخ داده است.	Phase of flight
۱۰	کلاس خرابی	کیفی	نقصی که منجر به حادثه شده مربوط به کدام قسمت بالگرد و از چه نوعی است	Fault class

• حداکثر وزن برخاستن: در این مطالعه، این متغیر به عنوان یک متغیر پیوسته و عددی در نظر گرفته شده است و در دسته‌های سبک، متوسط و سنگین قرار نمی‌گیرد، زیرا طبقه‌بندی متغیرهای عددی باعث خطا و از دست رفتن بخشی از اطلاعات می‌شود [۲۴، ۲۵]. به عنوان مثال، سازمان هوانوردی فدرال ایالات متحده آمریکا [۱] بالگردها را بر اساس حداکثر وزن برخاست به دو دسته طبقه‌بندی می‌کند: ۱. حمل و نقل با وزن بیش از ۷۰۰۰ پوند یا ۳۱۷۵ کیلوگرم ۲. کمتر از ۷۰۰۰ پوند یا ۳۱۷۵ کیلوگرم (وزن معمولی)؛ یا به عنوان مثال دیگر، مرکز ملی آتش نشانی بین سازمانی[۲] این طبقه‌بندی را در سه حالت زیر انجام می‌دهد: سبک (کمتر از ۶۰۰۰ پوند یا ۲۷۲۲ کیلوگرم)، متوسط (بین ۶۰۰۰ تا ۱۲۵۰۰ پوند یا ۲۷۲۲ تا ۵۶۷۰ کیلوگرم) و بالگردهای سنگین (بیش از ۱۲۵۰۰ پوند یا ۵۶۷۰ کیلوگرم) [۱].

• نوع آخرین بازرسی: این نوع بازرسی، جدیدترین نوع بازرسی انجام شده روی هلی کوپتر مربوطه را نشان می‌دهد، درمطالعه حاضربه شکل یک متغیر کیفی در ۵ کلاس برنامه سالانه، هر ۱۰۰ ساعت یکبار، **COAW** [۳](بر حسب تقویم، ساعات کاری یا مقدار تجمعی — هرکدام که زودتر اتفاق بیفتد)، برنامهٔ نگهداری و بازرسی اختصاصی برای هر نوع بالگرد **AAI**[۴] ، وکلاس **UNK** [۵] یا نامشخص، طبقه بندی شده است.

• ساعات کاری بعد از آخرین بازرسی: صالح و همکارانش گزارش دادند که خطای تعمیر و نگهداری و بازرسی بین ۱۴ تا ۲۱ درصد به عنوان یکی از علل تصادفات بالگرد در بانک اطلاعاتی **NTSB** تخمین زده می‌شود، که به معنای ۲۰ تا ۳۰ تصادف تعمیر و نگهداری بالگرد در سال است [۲]. بنابراین، نویسندگان تعداد ساعات عملیاتی از آخرین بازرسی را به عنوان یکی از متغیرهای ورودی در نظر گرفتند.

• توان خروجی موتور: یک متغیر عددی که معمولا با یکی از واحدهای اسب بخار (**HP**) یا پوند نیرو (**lbf**) بیان میشود.

• ساعت کارکرد بدنه هواپیما: این متغیر تعداد کل ساعات کارکرد بدنه هواپیما است.

• ارتفاع سنج: فشار بارومتریک در محل مشاهده رویداد در زمان رویداد را گزارش میکند. فشار بر حسب اینچ جیوه (در واحد جیوه) گزارش می‌شود. فشار استاندارد ۲۹.۹۲ اینچ جیوه است، کمتر از ۲۹.۹۲ اینچ جیوه فشار کم و بالاتر از ۲۹.۹۲ اینچ جیوه فشار زیاد است.

• سرعت باد : سرعت باد گزارش شده محلی در طول رویداد. سرعت باد گزارش شده با استفاده از سرعت متوسط طی ۲ دقیقه تعیین می‌شود. اگر در طول دوره ارزیابی ۲ دقیقه‌ای، ۶۰ درجه یا بیشتر تغییر کند، جهت ممکن است متغیر در نظر گرفته شود.

• جهت باد (درجه مغناطیسی): جهت باد محلی مشخص شده در زمان رویداد. این پارامتر جهت باد را در زمان رویداد نشان می‌دهد. جهت باد به صورت میانگین جهت در طول ۲ دقیقه ثبت می‌شود. جهت با افزایش ۱۰ درجه نسبت به شمال واقعی گزارش می‌شود.

---

[1] Federal Aviation Administration

[2] National Interagency Fire Center (NIFC)

[3] Calendar, Operating, or Accumulated Whichever (Occur)

[4] Approved Aircraft Inspection Program

[5] Unknown

- مرحله پرواز: شامل اطلاعاتی در مورد مرحله پروازی است که در آن حادثه رخ داده است. در مطالعه حاضر به شکل یک متغیر کیفی در۷ کلاس دسته بندی شده است شامل پرواز رو به جلو[۱] ، کارکرد زمینی موتور[۲] ، برنامهٔ نهایی فرود یا نزدیک‌شدن نهایی [۳]، شناوری در هوا [۴]، چرخش در پرواز[۵] ، فرود در زمین شیب‌دار[۶] و برخاستن[۷] .

۲-۲- مهندسی ویژگی برای انتخاب متغیرهای تاثیرگزار

ابعاد بالای داده‌ها ممکن است باعث تولید داده‌های آموزشی با کیفیت پایین و تخریب مدل‌های پیش‌بینی شود. بنابراین، مهندسی ویژگی با انتخاب زیرمجموعه‌ای مناسب از ویژگی‌ها، کیفیت مدل‌های پیش‌بینی را بهبود و افزایش می‌دهد [۲۶].

با وجود توانایی بالای الگوریتم‌های یادگیری ماشین در زمینه نگهداری پیش‌بینانه، داده‌های با حجم و ابعاد بالا و وجود ویژگی‌های اضافی منجر به عملکرد ضعیف و قابلیت اطمینان پایین این مدل‌ها می‌شود [۲۷-۳۱]. متأسفانه، داده‌های خام حسگر به دلیل دشواری در انتخاب یک نماینده مناسب از داده‌هایی با ابعاد زیاد و تفاوت در تفسیر اطلاعات، به ندرت توسط مدل‌های داده‌محور مفید واقع می‌شوند [۳۲-۳۷]. معمولاً اولین قدم در مهندسی ویژگی، پاکسازی داده‌ها است که شامل پرداختن به داده‌های گمشده، نرمال بودن و داده‌های پرت است. وجود داده‌های گمشده به دلایل مختلفی مانند خطاهای جمع‌آوری داده‌ها یا بررسی‌های ناقص، انتظار می‌رود. یک رویکرد استاندارد برای جایگزینی داده‌های گمشده با یک مقدار مشخص، میانگین یا میانه داده‌ها یا مقادیر درون‌یابی شده بر اساس داده‌های مجاور است. مدیریت نادرست داده‌های از دست رفته می‌تواند منجر به تحلیل نادرست توسط مدل شود. در این مطالعه، از یک الگوریتم جایگذاری میانگین برای پر کردن داده‌های از دست رفته با مقدار میانگین ستون داده‌ها استفاده شده است.

بین مراحل انتخاب ویژگی (مشخص کردن زیرمجموعه‌ای از ویژگی‌های اصلی)، استخراج ویژگی (ایجاد ویژگی‌های جدید، به عنوان مثال، تشکیل یک ترکیب خطی از ویژگی‌های اصلی) و مهندسی ویژگی تفاوت وجود دارد که گاهی اوقات نمی‌توان بین آنها تمایز قائل شد [۱۶]. هیچ ساختار مشخصی برای انجام مهندسی ویژگی وجود ندارد، بنابراین باید روش‌های مختلفی روی داده‌ها امتحان شود. با توجه به دشواری اجرای دستی مهندسی ویژگی، مهندسی ویژگی را می‌توان با کدنویسی نرم‌افزار انجام داد. خلاصه‌ای بصری از ساختار این کار در شکل ۲ ارائه شده است.

---

[1] Forward Flight

[2] Ground Run

[3] Final Approach

[4] Hover

[5] Turn

[6] Slope Landing

[7] Take-off

شکل ۲. فرآیند تحلیل این مطالعه

۳-۲- تکنیک های مهندسی ویژگی

در کارحاضر، از طریق کدنویسی درنرم‌افزارپایتون، تلاش شد تا کاربردی‌ترین ویژگی‌ها برای وقوع سوانح بالگردی از بین نه متغیر ورودی ذکر شده در جدول ۱ انتخاب شوند.

در مطالعه حاضر، از پنج روش انتخاب ویژگی استفاده شده است:

- **یادگیری عمیق** [۳۸]
- **شبکه عصبی**
- **ماتریس همبستگی**[1]: روش ماتریس همبستگی راهی برای انجام انتخاب ویژگی، است. ماتریس همبستگی جدولی است که ضرایب همبستگی بین ویژگی‌ها و متغیرهای هدف را نشان می‌دهد و از ۱- تا ۱+ متغیر است. مقادیر ۱+، ۱ - و ۰ به ترتیب همبستگی همبستگی خطی کامل مستقیم ، همبستگی خطی کامل معکوس و عدم وابستگی خطی را نشان می‌دهند.

- **روش Extreme Gradient Boosting** [۳۹]: الگوریتمی است که اخیراً در یادگیری ماشین مورد استفاده قرار گرفته است. پس از ساخت درخت‌های تقویت‌شده، این روش برای هر ویژگی امتیازی محاسبه می‌کند تا نشان دهد هر ویژگی چقدر در ساخت درخت‌های تصمیم تقویت‌شده مهم بوده است. این امتیاز امکان رتبه‌بندی ویژگی‌ها را فراهم می‌کند [۴۰، ۴۱]. الگوریتم XGBoost ، تقویت گرادیان درخت تصمیم را پیاده‌سازی می‌کند و برای سرعت و کارایی بالا طراحی شده است. به دلیل توانایی مدیریت طیف گسترده‌ای از انواع داده‌ها و به دلیل مقیاس‌پذیری و عملکرد بهینه، در هر دو مدل رگرسیون و طبقه‌بندی اعمال می‌شود. این روش، یک الگوریتم قوی است که می‌تواند روابط خطی و غیرخطی بین متغیرها و متغیر هدف را در نظر بگیرد. این روش می‌تواند از نظر محاسباتی پرهزینه باشد و ممکن است برای حجم بالای داده‌ها مناسب نباشد.

اطلاعات متقابل[2] [۲۰] : در نظریه احتمال [۴۲] مقیاسی برای نشان دادن میزان وابستگی متقابل دو متغیر تصادفی وجود دارد. روش اطلاعات متقابل یکی از روش‌های انتخاب ویژگی مبتنی بر نظریه اطلاعات است که هدف آن پیدا کردن ویژگی‌هایی است که بیشترین ارتباط آماری را با متغیر هدف دارند. این کمیت اندازه می‌گیرد که چقدر دانستن یک متغیر، باعث کاهش عدم قطعیت در مورد متغیر دیگر

[1] Correlation matrix

[2] Mutual information

می‌شود. این روش تعیین میکند که دو متغیر تا چه حد با هم وابسته اطلاعاتی هستند.از جمله مزایای این روش این است که می‌تواند وابستگی‌های غیرخطی بین ویژگی‌ها و خروجی را کشف کند (برخلاف ضریب همبستگی) ونیازی به فرض توزیع خاصی برای داده‌ها ندارد. اگر دو متغیر کاملاً مستقل باشند، اطلاعات متقابل بینشان صفر است.اگر دانستن یکی از آن‌ها دقیقاً مقدار دیگری را مشخص کند، اطلاعات متقابل مقدار بالایی خواهد بود. این روش می‌تواند روابط خطی و غیرخطی بین متغیرها و متغیر هدف را به دست آورد. برای دو متغیر تصادفی گسسته X و Y و تابع چگالی (بزرگنمایی) احتمال مشترک ، اطلاعات متقابل به صورت رابطه ۱ تعریف می‌شود:

$$I(X;Y) = \iint_{xy} f_{X,Y}(x.y)\, log\left(\frac{f_{X,Y}(x.y)}{f_X(x) f_Y(y)}\right) dxdy \qquad (۱)$$

که در آن $f_{X,Y}(x.y)$ چگالی احتمال مشترک در نقطه $(x.y)$، $I(X;Y)$ مقدار اطلاعات متقابل بین دو متغیر تصادفی، $f_X(x)$ چگالی احتمال حاشیه‌ای متغیرX است. $f_Y(y)$ چگالی احتمال حاشیه‌ای متغیرYاست.

انتگرال دوگانه روی فضای مقادیر انجام می‌شود. وجود و مقدار $I(X;Y)$ مستلزم آن است که چگالی‌ها وجود داشته باشند و انتگرال فوق مطلقاً مقید باشد (قابل تعریف).

۳- نتایج

۳-۱- روش یادگیری عمیق

این رویکرد از یک مدل شبکه عصبی عمیق متناسب با طبقه‌بندی دودویی استفاده می‌کند و اهمیت ویژگی‌ها را که از طریق روش اهمیت جایگشتی انجام می‌شود، بررسی می‌کند. علاوه بر این، نمرات اهمیت ویژگی‌ها به صورت بصری نمایش داده می‌شوند و تصویر روشنی از تأثیرات آنها بر قابلیت‌های پیش‌بینی مدل ارائه می‌دهند.

مدل بکار رفته مدل شبکه عصبی عمیق با معماری ترتیبی است که به ترتیب از لایه‌هایی با ۶۴ و ۳۲ نورون تشکیل شده است و از یک تابع فعال‌سازی ReLU و یک لایه خروجی سیگموئید استفاده می‌کند که با یک تابع زیان آنتروپی متقاطع دودویی و بهینه‌ساز Adam بهینه شده است.

مدل شبکه عصبی عمیق مورد استفاده در این کد، یک شبکه عصبی پیش‌خور با سه لایه است. لایه ورودی دارای ۶۴ گره و لایه پنهان دارای ۳۲ گره است. لایه خروجی دارای یک گره واحد با تابع فعال‌سازی سیگموئید است. تابع سیگموئید هر مقدار ورودی را به مقداری بین ۰ و ۱ نگاشت می‌کند که می‌تواند به عنوان احتمال تعلق به یک کلاس خاص تفسیر شود. تابع زیان مورد استفاده در کد، تفاوت بین احتمال پیش‌بینی‌شده و احتمال واقعی را اندازه‌گیری می‌کند که معمولاً در زمینه‌های طبقه‌بندی دودویی به عنوان زیان آنتروپی متقاطع شناخته می‌شود. هدف بهینه‌ساز در این کد، به حداقل رساندن این معیار زیان برای افزایش عملکرد مدل است. اهمیت جایگشت، یک تکنیک مستقل از مدل است که برای تعیین اهمیت نسبی ویژگی‌های ورودی استفاده می‌شود. این روش با به هم زدن تصادفی مقادیر یک ویژگی واحد در مجموعه داده‌ها و سپس محاسبه تغییر در عملکرد مدل کار می‌کند. تغییر بیشتر در عملکرد مدل به این معنی است که آن ویژگی برای پیش‌بینی مدل حیاتی‌تر است. تابع امتیازدهی سفارشی مورد استفاده در کد برای محاسبه اهمیت هر ویژگی تعریف شده است. این تابع، عملکرد مدل را در یک مجموعه اعتبارسنجی ارزیابی می‌کند و بر اساس تفاوت بین ۱ و مقدار زیان، امتیازی را برمی‌گرداند. امتیاز بالاتر به معنای ویژگی‌های بحرانی‌تر است. به طور خلاصه، این کد یک مدل شبکه عصبی عمیق را برای طبقه‌بندی دودویی با استفاده از Keras پیاده‌سازی می‌کند وهمچنین شامل تجزیه و تحلیل اهمیت ویژگی با استفاده ازاهمیت جایگشت از Scikit - Learn است. این مدل از یک شبکه عصبی پیش‌خور با سه لایه، تابع زیان آنتروپی متقاطع دودویی و بهینه‌ساز Adam برای بهبود عملکرد استفاده می‌کند. تکنیک اهمیت جایگشت، یک رویکرد مستقل از مدل است که اهمیت نسبی ویژگی‌های ورودی را با تغییر مقادیر آنها ومحاسبه تغییرحاصل، درعملکرد مدل تعیین می‌کند. جدول ۲ اهمیت ویژگی را برای هر متغیر ورودی نشان می‌دهد. همانطور که ضرایب اهمیت نشان می‌دهند، چهار متغیرCert_max_gr_wt، Afm_hrs، Power_units و Wind_dir_deg بیشترین اهمیت را در هلیکوپترهای سانحه دیده دارند.

جدول ۲. انتخاب ویژگی بر اساس روش یادگیری عمیق

نام ویژگی	همبستگی با متغیر هدف

Cert_max_gr_wt	0.78861
Afm_hrs	0.13210
Power_units	0.07186
Wind_dir_deg	0.00509
Afm_hrs_last_insp	0.00217
Type_last_insp	0.00009
Wind_vel_kts	0.00005
phase of flight	0.00003
Altimeter	0.00000

۳-۲- روش شبکه عصبی

مدل بکار رفته یک طبقه‌بندی‌کننده پرسپترون چندلایه (**MLP**) با اندازه لایه پنهان ۱۰۰ و یک نوع عمیق‌تر با دو لایه پنهان با اندازه‌های ۱۰۰ و ۵۰ است. برای هر دو مدل **MLP** ۱۰۰۰ تکرار در نظر گرفته شد. اهمیت ویژگی در این مدل‌ها با استفاده از تکنیک اهمیت جایگشت ارزیابی می‌شود و نشان می‌دهد که چگونه هر ویژگی بر قدرت پیش‌بینی مدل تأثیر می‌گذارد.
این روش با بارگذاری مجموعه داده‌ها در یک قاب داده پاندا و جایگذاری مقادیر از دست رفته با میانگین ستون شروع شد. سپس ویژگی‌ها و متغیر هدف جدا شدند و مدل شبکه عصبی با استفاده از طبقه‌بندی‌کننده MLP از Scikit-Learn تعریف شد.

در مرحله بعد، تعداد ویژگی‌های مورد نظر تعریف شدند و یک لیست خالی برای ذخیره ویژگی‌های انتخاب شده مقداردهی اولیه شدند. سپس تعداد ویژگی‌های مورد نظر و ویژگی‌های باقی مانده برای هر تکرار مرورو به‌روزرسانی شدند. بهترین ویژگی به لیست ویژگی‌های انتخاب شده اضافه گردید. جدول ۳ اهمیت ویژگی را برای هر متغیر ورودی نشان می‌دهد. همانطور که ضرایب اهمیت نشان می‌دهند، چهار متغیر phase of flight ، Wind_vel_kts ، Altimeter و Afm_hrs_last_insp ، بیشترین اهمیت را در هلیکوپترهای حادثه دیده دارند.

جدول ۳. انتخاب ویژگی بر اساس روش شبکه عصبی

نام ویژگی	همبستگی با متغیر هدف
phase of flight	0.3777
Wind_vel_kts	0.3555
Altimeter	0.3407
Afm_hrs_last_insp	0.3333
Wind_dir_deg	0.2666

۳-۳- روش ماتریس همبستگی

طبق این روش، در بین متغیرهای ورودی، خطی‌ترین رابطه با Y به ترتیب برای Cert_max_gr_wt، Power_units و Afm_hrs مشاهده می‌شود. با این حال، ممکن است روابط غیرخطی یا تعاملات بین متغیرها را نشان ندهد. همچنین فرض می‌کند که رابطه بین عوامل و متغیر خروجی یکنواخت است. جدول ۴ نتیجه این روش را نشان می‌دهد.

جدول ۴. انتخاب ویژگی بر اساس روش ماتریس همبستگی

نام ویژگی	همبستگی با متغیر هدف
Y	1
Cert_max_gr_wt	0.255944
Power_units	0.243484
Afm_hrs	0.114294
Wind_dir_deg	0.084512
Wind_vel_kts	0.075650
Type_last_insp	0.058101
Phase of Flight	0.041841

۳-۴- روش XGBoost (eXtreme Gradient Boosting)

طبق این روش، در بین متغیرهای ورودی، به ترتیب فاز پرواز، حداکثر وزن، توان موتور و ساعات پرواز تأثیرگذارترین متغیرها بودند. جدول ۵ نتیجه این روش را نشان می‌دهد.

جدول ۵. انتخاب ویژگی بر اساس روشXGBoost

نام ویژگی	همبستگی با متغیر هدف
Phase of Flight	0.207487
Cert_max_gr_wt	0.144031
Power_units	0.110489
Afm_hrs	0.095156
Wind_vel_kts	0.093962
Afm_hrs_last_insp	0.093376
Wind_dir_deg	0.092764
Altitude	0.090474
Type_last_insp	0.072261

۵-۳- روش اطلاعات متقابل

طبق جدول ۶، توان موتور، حداکثر وزن، فاز پرواز و ساعات پرواز به ترتیب پربازده‌ترین عوامل بودند.

جدول ۶. انتخاب ویژگی بر اساس روش اطلاعات متقابل

نام ویژگی	همبستگی با متغیر هدف
Power_units	0.267140
Cert_max_gr_wt	0.174941
Phase of Flight	0.124138
Afm_hrs	0.077117
Wind_vel_kts	0.056578
Type_last_insp	0.035869
Afm_hrs_last_insp	0.018324
Altitude	0.000000
Wind_dir_deg	0.000000

هیچ روشی به عنوان «بهترین»، برای انتخاب ویژگی وجود ندارد که عملکرد ایده‌آلی برای همه انواع مجموعه داده‌ها داشته باشد و انتخاب روش آن به مسئله و مجموعه داده خاص بستگی دارد. در جدول ۷. با بررسی و مقایسه پنج تکنیک، متوجه می‌شویم که متغیرهای Cert_max_gr_wt، Power_units، phase of flight و Afm_hrs دراکثر تحلیل‌ها مهم شناخته شده‌اند. در شکل ۳، نمای گرافیکی متغیرها، روش‌های انتخاب ویژگی و ضرایب اهمیت هر متغیر برای هر تکنیک قابل مشاهده است.

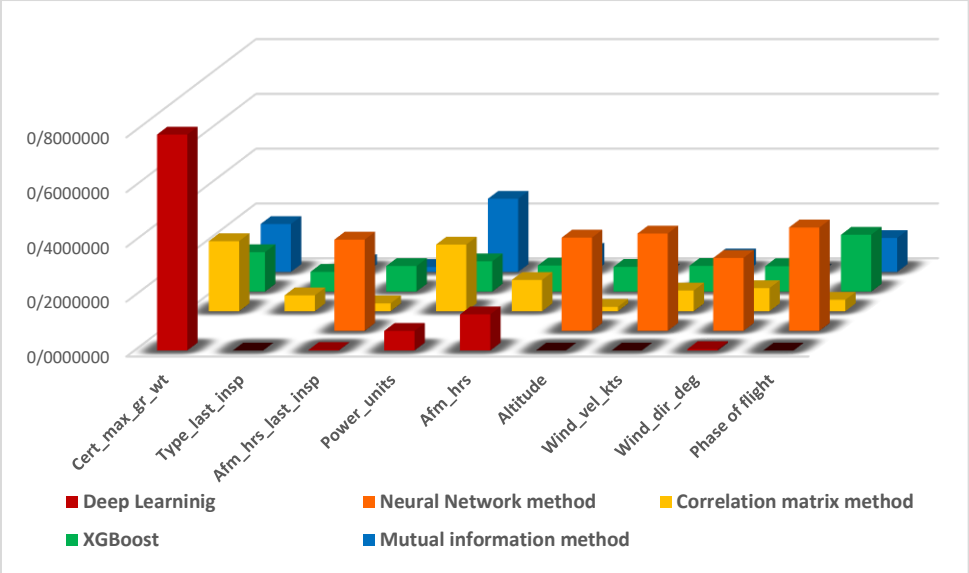
۴- بحث

همانطور که جدول ۷ نشان می‌دهد، برای هر یک از متغیرها، بررسی شد که در چند روش به عنوان متغیر اول، دوم، سوم، چهارم و پنجم شناخته می‌شود. همچنین بررسی شد که هر متغیر در چند روش از پنج روش انتخاب ویژگی، جزو پنج متغیر اول ظاهر شده است. به عنوان مثال، متغیر Cert_max_gr_wt در ۴ روش از ۵ روش، جزو پنج متغیر اول: در ۲ روش به عنوان متغیر اول و در ۲ روش به عنوان متغیر

دوم ظاهر شده است. همچنین، متغیر Power_units در۴ روش از ۵ روش، در بین پنج متغیر اول ظاهر شده است: در ۱ روش به عنوان اولین متغیر مهم، در ۱ روش به عنوان دومین متغیر مهم و در ۲ روش به عنوان سومین متغیر مهم. هایلایت زرد رنگ متغیرهایی را نشان میدهد که در مجموع نسبت به بقیه متغیرها و در اکثر تکنیکها ، تاثیرگزار شناخته شدند.

جدول ۷. اهمیت متغیرها در پنج مدل ویژگی

نام ویژگی	تعداد دفعاتی که ویژگی به عنوان اولین متغیر مهم شناخته شده است	تعداد دفعاتی که ویژگی به عنوان دومین متغیر مهم شناخته شده است	تعداد دفعاتی که ویژگی به عنوان سومین متغیر مهم شناخته شده است	تعداد دفعاتی که ویژگی به عنوان چهارمین متغیر مهم شناخته شده است	تعداد دفعاتی که ویژگی به عنوان پنجمین متغیر مهم شناخته شده است	تعداد مدلهایی که ویژگی در آنها مهم ظاهر شده است
Cert_max_gr_wt	2	2	0	0	0	4
Power_units	1	1	2	0	0	4
Afm_hrs	0	1	1	2	0	4
Wind_vel_kts	0	1	0	0	3	4
Phase of Flight	2	0	1	0	0	3
Wind_dir_deg	0	0	0	2	1	3
Afm_hrs_last_insp	0	0	0	1	1	2
Altitude	0	0	1	0	0	1
Type_last_insp	0	0	0	0	0	0



شکل ۳. ضرایب اهمیت متغیرها با استفاده از پنج روش انتخاب ویژگی

۵- نتیجه گیری و اقدامات پیشگیرانه

در این مقاله، عوامل تاثیرگزاری که منجر به تصادفات بالگرد، بدلیل خرابی روتور، می‌شدند شناسایی شد. تفاوت کار حاضر با مطالعات مشابه این بود که متغیرهای بیشتری، نظیر شرایط پرواز و پیکربندی بالگرد، در نظر گرفته شدند، برخلاف سایر مطالعات در این زمینه که در آن‌ها مجموعه‌ای از متغیرهای محدود در نظر گرفته شدند. این مطالعه تلاش کرد از پنج روش، از جمله یادگیری عمیق و شبکه‌های عصبی مصنوعی، برای شناسایی این متغیرهای بحرانی استفاده کند.

به این ترتیب، مقایسه نتایج همه روش‌ها، نشان داد که متغیرهای"حداکثر وزن بالگرد هنگام برخاستن"، "قدرت موتور بالگرد"، "تعداد ساعات کاری تا لحظه تصادف" و "فاز پرواز" پرتکرارترین متغیرهای مهم ظاهر شده در همه روش‌ها هستند. علاوه بر روشهای ریاضی و نرم افزاری به کار رفته، متغیرهای مذکوربه لحاظ منطق و مکانیک پرواز هم قابل توجیه هستند و بیشترین تأثیر را در پیش‌بینی نوع خرابی روتور دارند. با تمرکز بر متغیرهایی که بیشترین تأثیر را بر وقوع حوادث دارند، میتوان حوادث را با دقت و قابلیت اطمینان بالاتری پیش‌بینی کرد و ایمنی پرواز را بهبود بخشید. پس از تعیین متغیرهای مهم در تحقیقات آینده، می‌توان از طریق مدل‌های مختلف یادگیری ماشین، حوادث را پیش‌بینی کرد، آنها را با یکدیگر مقایسه کرد و زمان حادثه را پیش‌بینی نمود. مطالعات آینده می‌توانند بر طراحی و پیاده‌سازی سیستم‌های هوشمند پایش سلامت روتور در زمان واقعی متمرکز شوند تا بتوان خرابی را قبل از وقوع تشخیص داد.

اعلام منافع رقابتی

نویسندگان اعلام می‌کنند که هیچ منافع مالی رقابتی شناخته شده یا روابط شخصی ای که می‌توانسته بر کار گزارش شده در این مقاله، تأثیر بگذارد، ندارند.

تشکر و قدر دانی

نویسندگان از سرکار خانم شمیمه صنیع ثالث برای مشاوره علمی و نرم‌افزاری صمیمانه سپاسگزارند.

منابع مالی

این تحقیق هیچ کمک مالی خاصی از سازمان‌های تأمین مالی در بخش‌های دولتی، تجاری یا غیرانتفاعی دریافت نکرده است و صرفاً یک تحقیق علمی است.

منابع

[1] Subagia, R., Homer Saleh, J., Churchwell, J.S., & Zhang, K.S.(2020). Statistical learning for turboshaft helicopter accidents using logistic regression. *PLoS One*, 15(1), e0227334. DOI: 10.1371/journal.pone.02273342020

[2] Homer Saleh, J., Tikayat Ray, A., Zhang, K.S., & Churchwell, J.S. (2019). Maintenance and inspection as risk factors in helicopter accidents: Analysis and recommendations. *PLoS One*, 14(2), e0211424. DOI: 10.1371/journal.pone.0211424. eCollection 2019

[3] Herrera, I., Håbrekke, S., Krakenes, T., Hokstad, P.,& Forseth, U. (2010). Helicopter Safety Study 3, Trondheim, Norway: SINTEF Technology and Society, Report No. A15753.

[4] Kråkenes, T., Evjemo, T.E., Håbrekke, S., &Hoem, Å.S. (2017). Helicopter Safety Study 3b, Trondheim, Norway: SINTEF Technology and Society, Report No. 2017:00079

[5] Hokstad, P., Jersin, E., & Sten, T.A. (2001). A risk influence model applied to North Sea helicopter transport. *Reliab Eng Syst Safe*, 74(3), 311–322.

[6] Vinnem, J.E. (2014). Offshore Risk Assessment Vol 2 Principles, Modelling and Applications of QRA Studies. Springer, 3th Edn.

[7] Churchwell, J.S., Zhang, K.S., & Homer Saleh, J. (2018). Epidemiology of helicopter accidents: Trends, rates, and covariates. *Reliab Eng Syst Safe*, 180, 373–384.

[8] Fox, R.G. (2005). The history of helicopter safety. *Proceedings of the 1st International Helicopter Safety Symposium, Montreal*, (pp. 26–29).

[9] Benny, A., Johny, M., & Mathew, L. S. (2020). Prediction of Aviation Accidents using Logistic Regression Model. *IJIRT*, 7(6), 241-245.

[10] Greenhaw, R., &Jamali, M. (2021). Medical Helicopter Accident Review: Causes and Contributing Factors. Civil Aerospace Medical Institute Federal Aviation Administration Oklahoma City, Office of Aerospace Medicine, Washington, DC 20591, Final Report.

[11] Xu, Z.h., Homer Saleh, J., & Subagia, R. (2020). Machine learning for helicopter accident analysis using supervised classification: Inference, prediction, and implications. *Reliab Eng Syst Safe*, 204, 10721.

[12] Lao, Z., He, D., Wei, Z., Shang, H., Jin, Z., Miao, J., & Ren, C. (2023). Intelligent fault diagnosis for rail transit switch machine based on adaptive feature selection and improved LightGBM. *Eng Fail Anal,* Vol.148, 107219.

[13] Wu, L., Liang, W., & Sha, D. (2023). Cross-domain feature selection and diagnosis of oil and gas pipeline defects based on transfer learning. *Eng Fail Anal,* Vol.143, Part A, 106876.

[14] Wang, S., Wang, Q., Xiao, Y., Liu, W., & Shang, M. (2022). Research on rotor system fault diagnosis method based on vibration signal feature vector transfer learning, *Eng Fail Anal*, Vol.139, 106424.

[15] He, C., Xu, P., Pei, X., Wang, Q., Yue, Y., & Han, Ch. (2024). Fatigue at the wheel: A non-visual approach to truck driver fatigue detection by multi-feature fusion. *Accident analysis and prevention*, Vol.199, 107511.

[16] Verdonck, T., Baesens, B., Óskarsdóttir, M., & Broucke, S. V. (2023). Special issue on feature engineering editorial*, Machine Learning*, Vol. 113, 3917–3928.

[17] Jie, C., Jiawei, L., Shulin, W., & Sheng, Y. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing,* Vol.300, 70–79.

[18] Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T.,& Vapnik, V. (2000). Feature selection for SVMs, *Adv. Neural inf. Process syst*, Vol.13, 668–674.

[19] Das, S. (2001, June). Filters, wrappers and a boosting-based hybrid for feature selection. *Proceedings of the Eighteenth International Conference on Machine Learning (ICML'01)*, (pp.74–81).

[20] Aremua, O. O., Cody, R. A., Wood, D. H., & McAreea, P.R. (2020). A relative entropy-based feature selection framework for asset data in predictive maintenance. *Comput Ind Eng*, Vol.145, 106536.

[21] Ali, Y., Hussain, F., & Haque, Md.M. (2024). Advances, challenges, and future research needs in machine learning-based crash prediction models: A systematic review. *Accident Analysis & Prevention*, Vol.194, 107378.

[22] https://data.ntsb.gov/avdata. [accessed 10 October 2025].

[23] Rausand, M. (2009). System Reliability Theory. John Wiley & Sons.

[24] Altman, D.G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, Vol. 332 (1080). https://doi.org/10.1136/bmj.332.7549.1080

[25] Austin, P.C., & Brunner, L.J. (2004). Inflation of the type I error rate when a continuous confounding variable is categorized in logistic regression analyses. *Stat Med*, 23(7), 1159-78. https://doi.org/10.1002/sim.1687.

[26] Khoshgoftaar, T., Gao, K., Chen, Y., & (2015). Napolitano, A. Comparing Feature Selection Techniques for Software Quality Estimation Using Data-Sampling-Based Boosting Algorithms. *International Journal of Reliability, Quality and Safety Engineering*, 22(3), 1550013.

[27] Ellefsen, A. L., Bjørlykhaug, E., Æsøy, V., Ushakov, S., & Zhang, H. (2019). Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture. *Reliab Eng Syst safe*, Vol.183, 240–251.

[28] Ghodsi, A. (2006). Dimensionality reduction a short tutorial, Department of Statistics and Actuarial Science, Univ. of Waterloo, Ontario, Canada, pp.38

[29] Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *J Mach Learn Res*, Vol. 3, 1157–1182.

[30] Russell, S. J., & Norvig, P. (2016). Artificial intelligence: A modern approach, Third ed., Malaysia: Pearson Education Inc

[31] Rotari, M., & Kulahci, M. (2024). Variable selection wrapper in presence of correlated input variables for random forest models. *Quality and Reliability Engineering International*, 40(1), 297-312.

[32] Fang, X., Paynabar, K., & Gebraeel, N. (2017). Multistream sensor fusion-based prognostics model for systems with single failure modes. *Reliab Eng Syst Safe*, Vol. 159, 322–331.

[33] Khan, S., & Yairi, T. (2018). A review on the application of deep learning in system health management. *Mech Syst Signal Process*, Vol.107, 241–265.

[34] Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mech Syst Signal Process*, Vol. 104, 799–834.

[35] Qiu, G., Huang, S., & Chen, Y. (2017). Automatic segmentation and prognostic method of a turbofan engine using manifold learning and spectral clustering algorithms. *Adv Mech Eng*, 9(9). https://doi.org/10.1177/1687814017722712.

[36] Trappey, A. J., Trappey, C. V., Ma, L., & Chang, J. C. (2015). Intelligent engineering asset management system for power transformer maintenance decision supports under various operating conditions. *Comput Ind Eng,* Vol. 84, 3–11.

[37] Yang, F., Habibullah, M. S., Zhang, T., Xu, Z., Lim, P., & Nadarajan, S. (2016). Health index-based prognostics for remaining useful life predictions in electrical machines. *IEEE Transactions on Industrial Electronics*, Vol. 63, 2633–2644.

[38] Wang, Y., & Zhao, Y. (2023). Three-stage feature selection approach for deep learning-based RUL prediction methods. *Quality and Reliability Engineering International*, 39(4), 1223-1247.

[39] Leevy, J., Khoshgoftaar, T,&Hancock, J. (2024). Using Random Undersampling and Ensemble Feature Selection for IoT Attack Prediction. *International Journal of Reliability, Quality and Safety Engineering,* 31, (01), 2350012.

[40] Brownlee, J. (2020). Feature Importance and Feature Selection with XGBoost in Python. machine learning mastery. Available at: https://machinelearningmastery.com/feature-importance-and-feature-selection-with-xgboost-in-python/ (accessed 10 August 2024).

[41] Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed, Springer, New York.

[42] Tang, J., Wu, Z., Jia, M., & Liu, Z. (2015). Mutual Information-Based Modified Randomized Weights Neural Networks. *J.Comput.Commun,* 3(11). DOI: 10.4236/jcc.2015.311030.

Determining the Factors Influencing the Prediction of Helicopter Rotor Failures

Purpose: The purpose of this paper is to investigate and identify the variables that influence the occurrence of helicopter accidents caused by different types of rotor failures. These crucial factors include flight conditions, maintenance conditions, and helicopter configuration. With this approach, accidents can be investigated more effectively and flight safety can be significantly improved.

Methodology: By analyzing 135 rotor faults accident from a comprehensive dataset containing 5652 helicopter-related accidents, eight classes of rotor faults were identified. Based on expert surveys and a review of studies in the field of helicopter accidents, nine features were proposed as crucial factors to such accidents. The significance of these factors was assessed using five feature selection methods. The input features included maximum takeoff weight, flight hours since the last inspection, type of last inspection, engine power, flight hours, altitude, wind speed, wind direction, and flight phase. Five well-known feature selection techniques—Correlation Matrix, Extreme Gradient Boosting (XGBoost), Mutual Information, Deep Learning, and Neural Network—were employed to identify the most essential factors.

Findings: "Maximum weight", "helicopter engine power", "flight phase" and "flight hours" were identified as variables with the highest degree of importance in predicting faults class of helicopter rotor, which also have a strong and acceptable justification in flight mechanics.

Originality/Value: The distinction of the present study from similar works lies in the inclusion of a broader range of variables, such as flight conditions and helicopter configuration, in contrast to previous studies that considered

only a limited set of variables. By prioritizing these variables, the findings pave the way for proactive measures to prevent rotor faults, aiming to enhance prediction accuracy, reliability, and flight safety.

Keywords: Helicopter Rotor; Feature Selection; Deep Learning; Neural Network; Failure Prediction

آماده انتشار