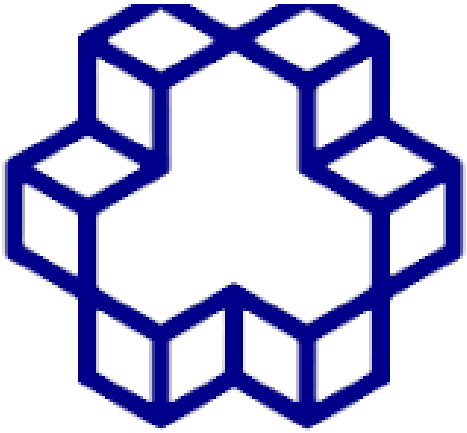

بسم الله الرحمن الرحيم



دانشگاه صنعتی خواجه نصیرالدین طوسی

سمانه اعلائی ۴۰۱۰۲۰۹۴

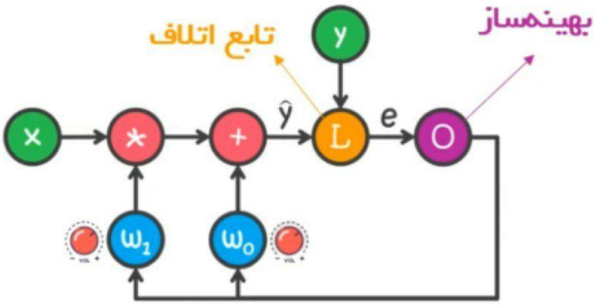
درس یادگیری ماشین

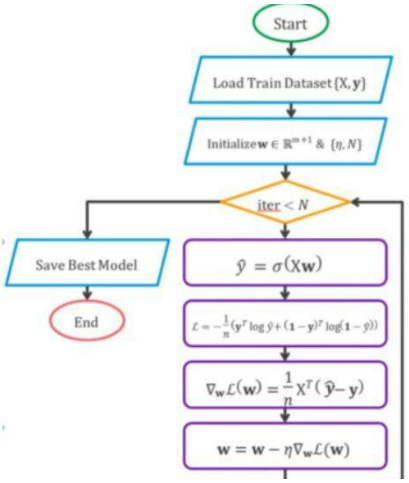
استاد درس : جناب دکتر علیاری

مینی پروژه اول

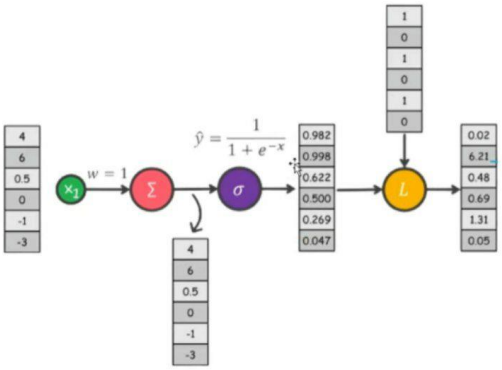
سوال اول

۱_۱)





مدل ما ابتدا یک دیتا به عنوان ورودی و تست دریافت می کنید سپس طبق رابطه خطی $y = w_1x + w_0$ مقدار y_hat را به دست می آورد و با توجه به مقدار اتلاف ما پارامتر های w1 و w0 را تنظیم میکند ما یک مقدار تکرار برای ادامه الگوریتم داریم که هرچه بیشتر باشد دقت ما بیشتر می شود.در حالت چند کلاسه در بلوک تصمیم گیری ما تغییرات داریم و به جای استفاده از سیگموید از توابع دیگری باید استفاده کنیم. زیرا سیگموید برای کلاس بندی دو حالته هست و برای کلاس بندی های بیش تر باید جایگزین شود. بلوک تصمیم در طبقه‌بندی دو کلاسه، از تابع فعال‌سازی سیگموید برای ترسیم خروجی‌های مدل به احتمالات بین ۰ و ۱ استفاده می‌شود که مربوط به احتمال تعلق به یکی از دو کلاس است. در طبقه بندی چند کلاسه، که در آن بیش از دو کلاس وجود دارد، از توابع فعال سازی جایگزین مانند softmax استفاده می شود. تابع softmax احتمالات هر کلاس را محاسبه می‌کند و اطمینان می‌دهد که احتمالات پیش‌بینی‌شده تا یک در همه کلاس‌ها جمع می‌شوند. بنابراین، تغییر تابع فعال‌سازی در بلوک تصمیم‌گیری از سیگموید به softmax هنگام توسعه مدل برای رسیدگی به کلاس‌های متعدد ضروری است و از تخمین‌های احتمال مناسب برای هر کلاس اطمینان حاصل می‌کند.



۱-۲) بله، زیرا کلاس ها به صورت دوتا دوتا باهم ترکسیب شدند که کار ما را چالش برانگیز کرده است. برای چالش برانگیز کردن میتوانیم پارامتر class_sep بین صفر تا یک و یا یک عدد کوچک قرار بدهیم که ۴ کلاس باهم ترکیب شوند و تفکیک آن ها ازهم کار بسیار دشواری است.

۲-۳)

فرآیند انتخاب متا پارامترها، مانند تعداد جلسات آموزشی و میزان یادگیری، بسته به الگوریتم خاص و فراپارامترهای مرتبط با آن متفاوت است.

SGDClassifier

پارامتر max_iter حداکثر تعداد تکرارها (دوران) را که الگوریتم شیب نزولی تصادفی (SGD) در طول آموزش انجام می دهد را تعیین می کند. این پارامتر کنترل می کند که الگوریتم چند بار پارامترهای مدل را با استفاده از داده های آموزشی به روز کند. پارامتر tol (تلرانس) معیار توقف بهینه سازی را مشخص می کند. آموزش زمانی متوقف می شود که فرآیند بهینه سازی به سطح تحمل مشخصی برسد که نشان دهنده همگرایی است. این پارامترها بر اساس ملاحظاتی مانند پیچیدگی مجموعه داده، سطح مطلوب همگرایی مدل و منابع محاسباتی موجود انتخاب می شوند.

LogisticRegressionCV

پارامتر max_iter حداکثر تعداد تکرارها را برای الگوریتم رگرسیون لجستیک مشخص می کند. مشابهSGDClassifier ، تعداد تکرارهای انجام شده در طول بهینه سازی را کنترل می کند. پارامتر cv تعداد تکرار را برای اعتبارسنجی متقاطع در طول آموزش تعیین می کند. اعتبارسنجی متقاطع برای ارزیابی عملکرد مدل و تنظیم فراپارامترهایی مانند قدرت منظم‌سازی استفاده می‌شود. این پارامترها بر اساس ملاحظاتی مانند اندازه و پیچیدگی مجموعه داده و همچنین سطح مطلوب منظم سازی و تعمیم مدل انتخاب می شوند.

RidgeClassifierCV

پارامتر "alphas" محدوده مقادیر آلفا را برای آزمایش در طول رگرسیون پشته تایید متقابل مشخص می کند. رگرسیون ریج نوعی رگرسیون خطی است که شامل تنظیم L2 است که توسط پارامتر آلفا کنترل می شود. پارامتر cv تعداد تکرار را برای اعتبارسنجی متقابل در طول تمرین، مشابهLogisticRegressionCV ، تعیین می کند. انتخاب مقادیر آلفا و تعداد تاها در اعتبارسنجی متقابل معمولاً بر اساس آزمایش تجربی و عملکرد اعتبارسنجی است.

One-vs-Rest (OvR) و One-vs-One (OvO) دو روش استراتژیک برای توسعه الگوریتم‌های دسته‌بندی دودویی به مسائل دسته‌بندی چند کلاسه هستند. One-Vs-Rest (OvR) در استراتژی One-vs-Rest، هر کلاس به عنوان کلاس مثبت در نظر گرفته شده و دسته بندی دودویی برای هر کلاس آموزش داده می شود، در حالی که سایر کلاس ها به عنوان کلاس منفی در نظر

گرفته می شوند. برای یک موضوع با n کلاس، nدسته‌بند دودویی آموزش داده می‌شود. در هنگام پیش‌بینی، هر دسته‌بندی یک پیشبینی انجام می‌دهد و کلاس با بیشترین امتیاز اعتماد به عنوان پیشبینی نهایی انتخاب می‌شود. این طرح مسائلی مناسب است که کلاس‌ها به راحتی جدا شوند و ممکن است اشتراک‌هایی داشته باشند.یک در مقابل یک(OvO) در تدوینOne-vs-One ، برای هر زوج از کلاس‌ها یک دسته بندی دودویی آموزش داده می‌شود. برای یک موضوع با n کلاس، $n * (n-1)/2$ دسته‌بند دودویی آموزش داده می‌شود. هر دسته‌بندی فقط بر روی داده‌های دو کلاس آموزش داده می‌شود که یکی به عنوان کلاس مثبت و دیگری به عنوان کلاس منفی در نظر گرفته می‌شود.در هنگام پیش‌بینی، هر دسته‌بندی یک پیش‌بینی انجام می‌دهد و کلاسی که بیشترین تعداد امتیاز را کسب می‌کند (به عبارتی بیشترین تعداد در مقایسه‌های دو به دو) به عنوان پیش‌بینی نهایی انتخاب می‌شود. این طرح مسائلی با تعداد کمی از کلاس‌ها و زمانی که دسته‌بندی‌های دودویی به لحاظ محاسباتی هستند، مناسب است. برای نتایج بهتز از روش OvO استفاده کردیم و دقت بهتری نسبت OvR به داده شد.

سوال دوم

۱-۲

تشخیص خطای بلبرینگ مبتنی بر یادگیری ماشین:

اهمیت بخش‌های مهمی از ماشین‌آلات چرخشی حمل‌ونقل از طریق بلبرینگ‌های غلتکی است. پیدا کردن نقص‌های بلبرینگ به موقع می‌تواند از تأثیر بر عملکرد کلی تجهیزات جلوگیری کند. فناوری تشخیص عیب مبتنی بر داده‌ها از تازه‌ترین مباحث تحقیقاتی شده و نقطه شروع تحقیقات اغلب دریافت سیگنال‌های ارتعاشی است. دیتاست‌های عمومی بسیاری برای بلبرینگ‌های غلتکی وجود دارند. از میان آن‌ها، محبوب‌ترین دیتاست عمومی مرکز بلبرینگ دانشگاه Case Western Reserve (CWRU) است. از دیتاست CWRU شروع می‌کنیم، برخی از روش‌های پایه‌ای تشخیص عیب بلبرینگ مبتنی بر یادگیری ماشین را مقایسه و تجزیه و تحلیل می‌کند و ویژگی‌های CWRU را خلاصه می‌کند. ابتدا، یک معرفی جامع از CWRU ارائه می‌دهیم و نتایج به‌دست‌آمده را خلاصه می‌کنیم. پس از آن، روش‌ها و اصول پایه‌ای تشخیص عیب بلبرینگ مبتنی بر یادگیری ماشین را خلاصه می‌کنیم.

بخش‌های اساسی‌ترین ماشین‌آلات چرخشی بلبرینگ‌های غلتکی هستند، و با توسعه و تقاضای صنعت، بار کاری بیشتری به اکثر ماشین‌آلات چرخشی تحمیل می‌شود. در شرایط بار بالا، ضربه قوی، بار کاری بالا و محیط پیچیده، بلبرینگ‌های غلتکی اغلب نقص در ناحیه داخلی، ناحیه خارجی و توپ‌های خود را تولید می‌کنند. اگر نقص به موقع شناسایی نشود، معمولاً تجهیزات خاموش می‌شوند که منجر به خسارات اقتصادی عظیم و حتی حوادث ایمنی می‌شود. تشخیص و پیش‌بینی عیب هسته‌ی مدیریت پیش‌بینی و سلامت (PHM) است. هدف اصلی PHM در ماشین‌آلات چرخشی کاهش هزینه‌ها و پشتیبانی، بهبود ایمنی و سلامتی ماشین‌آلات چرخشی است، تا به تعمیرات مبتنی بر شرایط با سرمایه‌گذاری کمتر برسد. روش‌های تشخیص عیب می‌توانند به سه نوع تقسیم شوند، از جمله روش‌های مبتنی بر مدل، روش‌های آماری مبتنی بر قابلیت اطمینان و روش‌های مبتنی بر داده. شرط اولیه تشخیص

عیب بر اساس مدل، شناخت مدل ریاضی سیستم مورد نظر است. این نوع روش‌های تشخیص می‌توانند به عملکرد اساسی سیستم مورد نظر نفوذ کنند و اجرای پیش‌بینی عیب را ممکن سازند. اما برای سیستم‌های پویای پیچیده، ایجاد یک مدل ریاضی با اطمینان بالا دشوار است و کارهای مرتبط با تشخیص عیب به شدت محدود می‌شود. اطلاعات مورد نیاز برای تکنیک‌های تشخیص عیب مبتنی بر قابلیت اطمینان آماری می‌توانند در انواع مختلف توزیع چگالی احتمال (PDFs) یافت شوند. هدف تشخیص عیب می‌تواند با پردازش داده‌های جمع‌آوری شده توسط سنسور و ترکیب روش‌هایی مانند مهندسی ویژگی، یادگیری عمیق یا یادگیری ماشین دست‌یافته شود. تشخیص عیب بلبرینگ مبتنی بر داده، محور تحقیقات فعلی است، مسئله اصلی به دست آوردن داده است. وضعیت بلبرینگ غلتکی و سیگنال ارتعاشات مکمل هستند. وقتی بلبرینگ غلتکی نقص دارد، اغلب همراه با یک سیگنال تصادفی هست. سنسورها در موقعیت‌های مختلف بلبرینگ غلتکی نصب شده‌اند تا سیگنال‌های ارتعاشی را جمع‌آوری کنند. وضعیت بلبرینگ را می‌توان با مشاهده مستقیم این سیگنال‌ها ارزیابی کرد.

یادگیری ماشین به استخراج دانش از داده‌ها، استفاده از یک نمونه داده برای یادگیری، و تعیین و تشخیص خودکار و پیش‌بینی نمونه‌های ورودی پسین و دادن نتایج پیش‌بینی برای نمونه‌های ورودی است. یادگیری ماشین در نهایت می‌تواند خروجی‌های مشابهی را بر اساس ورودی‌های مشابه به‌دست آورد. از شکل ۱ می‌توان مشاهده کرد که فرآیند اساسی یادگیری ماشین معمولاً می‌تواند به ۳ بخش تقسیم شود، شامل استخراج ویژگی‌ها، انتخاب ویژگی‌ها و طبقه‌بندی. در مرحله اول، ویژگی‌های چندگانه از حوزه‌های مختلف با پردازش سیگنال ارتعاشی استخراج می‌شود. در مرحله دوم، بعد فضایی مجموعه ویژگی بر اساس معیارهای مختلف کاهش داده می‌شود. بهترین زیرمجموعه ویژگی با قابلیت تمایز بیشتر و حداقل تعداد، استخراج می‌شود بعد، دقت طبقه‌بندی را بهبود بخشد و زمان طبقه‌بندی را کاهش دهد. و مرحله آخر مرحله طبقه‌بندی است که با وارد کردن یک زیرمجموعه از ویژگی‌ها به طبقه‌بند ورودی، نمونه را طبقه‌بندی می‌کند. این مقاله از دیتاست CWRU شروع می‌کند، مقایسه و تجزیه و تحلیل برخی از روش‌های اصلی تشخیص عیب بلبرینگ غلتکی مبتنی بر یادگیری ماشین را ارائه می‌دهد و ویژگی‌های دیتاست CWRU را خلاصه می‌کند.

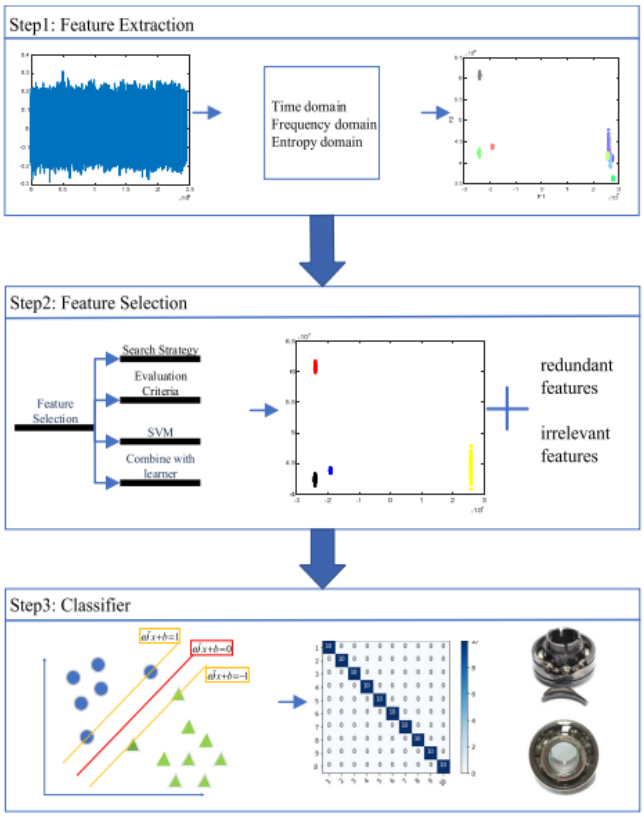


FIGURE 1. The flowchart of machine learning.

دیتاست

با ویژگی‌های آشکار و تشخیص نسبتاً آسان است. این می‌تواند به عنوان یک مجموعه داده اصلی برای آزمون الگوریتم‌ها استفاده شود. تست بنچ دیتاست CWRU در شکل ۲ نشان داده شده است. تست بنچ از یک موتور الکتریکی با توان ۲ اسب بخار، یک سنسور گشتاور و یک دینامومتر برق تشکیل شده است. شتاب‌سنج‌ها به ترتیب بر روی قسمت‌های محفظه انتهای درایو و انتهای فن نصب شده‌اند تا سیگنال‌های ارتعاش را جمع‌آوری کنند. تست بنچ به طور اصلی داده‌های مبنایی عادی، داده‌های عیب انتهای درایو و داده‌های عیب انتهای فن را ثبت می‌کند. شتاب‌سنج‌ها به ترتیب بر روی قسمت‌های محفظه انتهای درایو و انتهای فن نصب شده‌اند تا سیگنال‌های ارتعاش را جمع‌آوری کنند. تست بنچ به طور اصلی داده‌های مبنایی عادی، داده‌های عیب انتهای درایو و داده‌های عیب انتهای فن را ثبت می‌کند. فرکانس نمونه‌برداری داده‌های عیب انتهای درایو ۱۲۰۰۰ sps و ۴۸۰۰۰ sps است و فرکانس نمونه‌برداری داده‌های مبنایی عادی و داده‌های عیب انتهای فن هر دو ۱۲۰۰۰ sps می‌باشد. بنابراین، این دیتاست عمومی شامل چهار دسته داده است. هر نوع داده به طور اصلی شامل عیب‌های ناحیه داخلی، عیب‌های توپ و عیب‌های ناحیه خارجی با بارهای مختلف تحت اقطار عیب مختلف است. عیوب این دیتاست اصلیاً خسارت از پیش‌تراش برق است. این نوع آسیب، نوعی آسیب مصنوعی است. اندازه آسیب شامل ۰،۰۰۷ اینچ، ۰،۰۱۴ اینچ، ۰،۰۲۱ اینچ و ۰،۰۲۸ اینچ می‌شود. در میان آن‌ها، قطر عیب‌های

۰٫۰۰۷، ۰٫۰۱۴ و ۰٫۰۲۱ از بلبرینگ‌های SKF استفاده می‌کنند و قطر عیب ۰٫۰۲۸ از بلبرینگ‌های NTN استفاده می‌کند. و فقط عیب ناحیه داخلی و عیب توپ انتهای درایو ثبت می‌شود وقتی فرکانس نمونه‌برداری ۱۲۰۰۰ sps است و قطر عیب ۰٫۰۲۸ اینچ است. با توجه به توضیحات ارائه شده، دیتاست عمومی بلبرینگ‌های غلتکی شامل داده‌های ارتعاشی برای بررسی و تشخیص عیوب مختلف است. این دیتاست شامل چهار دسته داده است که هرکدام ویژگی‌ها و خصوصیات خاص خود را دارند: بار بلبرینگ دیتاست شامل بارهای مختلف بر روی بلبرینگ است که به ترتیب ۰، ۱، ۲ و ۳ اسب بخار هستند و با سرعت‌های مختلف متناظر است. عیوب ناحیه خارجی عیوب بر اساس موقعیت نقطه عیب به ۶ وقت، ۳ وقت و ۱۲ وقت تقسیم می‌شوند. این عیوب معمولاً در نواحی خارجی بلبرینگ رخ می‌دهند و از جنس‌ها و اندازه‌های مختلفی هستند. هر نوع داده عیب در یک فایل mat ذخیره می‌شود. هر فایل شامل داده‌های ارتعاش انتهای درایو و انتهای فن، همچنین سرعت است. DE به معنای داده‌های انتهای درایو، FE به معنای داده‌های انتهای فن، و RPM به معنای سرعت می‌باشد.

TABLE 1. The rolling bear datasets.

Literature	Dataset	Fault method	Sampling rate	Fault type
[4-5]	FEMTO-ST	Accelerated lifetime tests	25.6kHz	• 3 operating conditions • 6 learning datasets 11 test datasets
[6-8]	IMS	Accelerated lifetime tests	20kHz	• Inner race fault, ball fault, outer race fault • 3 run-to-failure experiments (including 12 bearings)
[9]	XJTU-SY	Accelerated lifetime tests	25.6kHz	• Inner race fault, ball fault, outer race fault, cage fault • 3 load conditions: 11kn, 10kn, 12kn
[10-11]	CWRU	Artificially damaged	12kHz/48kHz	• Inner race fault, ball fault, outer race fault • Fault diameter: 0.007, 0.014, 0.021 • Motor load : 0, 1, 2, 3 (HP)
[12]	MFPT	Artificially damaged/ Accelerated lifetime tests	97656Hz/48828Hz	• Inner race fault, outer race fault • Different load • 3 run-to-failure experiments
[13]	Paderborn	Artificially damaged	64kHz	• 6 Undamaged bearings • 12 Artificially damaged bearings: inner race fault, outer race fault • 2 damage level • 14 run-to-failure experiments

اسمیت و رندال یک معیار مبتنی بر دیتاست CWRU ارائه دادند که بر اساس سه روش تشخیص عیب مرسوم بلبرینگ غلتکی استوار است و از طریق آن می‌توان الگوریتم‌های جدید تشخیص عیب بلبرینگ را آزمایش کرد. همچنین، یونگبو و همکاران دیتاست CWRU را با استفاده از انواع مختلف آنتروپی و طبقه‌بندی‌کننده‌ها مورد ارزیابی قرار دادند و یک روش ارزیابی برای روش‌های جدید طبقه‌بندی پساروش ارائه دادند.

به علاوه، برای تشخیص عیب بلبرینگ، هدف اصلی پیش‌پردازش داده‌ها حل مشکل عدم توازن داده و مشکلات نمونه‌های کوچک است. اینکه نسبت داده‌های عیب به داده‌های سالم ناتوازن است، و داده‌های انواع مختلف عیب نیز ناتوازن هستند، از جمله چالش‌های مطرح است. به علاوه، اندازه نمونه داده هر نوع عیب کوچک است. جیانان و همکاران یک روش بیش‌نمونه‌گیری به نام SCOTE ارائه دادند که مشکل توازن داده‌های چند کلاسی را به مشکلات نامتوازن داده‌های دو کلاسی چندگانه تبدیل می‌کند. این روش با استفاده از LS-SVM چند کلاسی ترکیب می‌شود تا یک مدل جدید برای حل مشکل عدم توازن داده‌های عیب بلبرینگ

غلتکی ایجاد شود. اشرافی و همکاران همچنین مدل مخلوط احتمالاتی (PMM) و مارکوف مونت کارلو (MCMC) را ترکیب کردند تا گسترش مجموعه داده را دست یابند، و از شبکه تراشه پیمانه‌ای نیمه نظارتی (SSLN) برای حل مشکل نمونه‌های برچسب‌گذاری شده کمتر استفاده کردند.

TABLE 2. Time features.

Type	Formula	Type	Formula
Mean	$T_1=\frac{1}{N_S}\sum_{i=1}^{N_S}x_i$	Std	$T_6=\sqrt{\frac{1}{N_S}\sum_{i=1}^{N_S}(x_i-T_1)}$
RMS	$T_2=\sqrt{\frac{1}{N_S}\sum_{i=1}^{N_S}(x_i)^2}$	Shape factor	$T_7=\frac{T_2}{\frac{1}{N_S}\sum_{i=1}^{N_S}\lvert x_i\rvert}$
Kurtosis	$T_3=(\sum_{i=1}^{N_S}(x_i-\max(x_i))^4)/N_S$	Peaking factor	$T_8=\frac{\max(x_i)}{T_2}$
Peak-to-peak	$T_4=\max(x_i)-\min(x_i)$	Pulse factor	$T_9=\frac{\max(x_i)}{\frac{1}{N_S}\sum_{i=1}^{N_S}\lvert x_i\rvert}$
Var	$T_5=\frac{1}{N_S}\sum_{i=1}^{N_S}(x_i-T_1)$	Margin factor	$T_{10}=\frac{\sum_{i=1}^{N_S}(x_i-\max(x_i))^3}{T_2^3N_S}$

TABLE 3. Frequency features.

Type	Formula
CF	$F_1=\left(\sum_{j=0}^{N_f}f_j\times S(f_j)\right)\Big/\sum_{j=0}^{N_f}S(f_i)$
MSF	$F_2=\left(\sum_{j=0}^{N_f}f_j^2\times S(f_j)\right)\Big/\sum_{j=0}^{N_f}S(f_i)$
RMSF	$F_3=\sqrt{\left(\sum_{j=0}^{N_f}f_j^2\times S(f_j)\right)\Big/\sum_{j=0}^{N_f}S(f_i)}$
VF	$F_4=\left(\sum_{j=0}^{N_f}(f_j-P_2)^2)\times S(f_j)\right)\Big/\sum_{j=0}^{N_f}S(f_i)$
RVF	$F_5=\sqrt{\left(\sum_{j=0}^{N_f}(f_j-P_2)^2)\times S(f_j)\right)\Big/\sum_{j=0}^{N_f}S(f_i)}$

استخراج ویژگی‌ها

ویژگی‌ها برای یادگیری ماشین بسیار حیاتی هستند. رویکرد اصلی فعلی این است که ویژگی‌های چند دامنه‌ای را از سیگنال ارتعاش بلبرینگ‌ها استخراج کرده و یک مجموعه ویژگی چند دامنه‌ای شکل دهند. از بعدهای مختلفی مانند دامنه فرکانس، دامنه زمان و دامنه آنتروپی، تعدادی بیشینه از ویژگی‌ها را شکل داده‌اند. متغیر در دامنه زمان (t) است، و معمولاً (t) برای مشاهده تغییرات در سیگنال ارتعاش در دامنه زمان استفاده می‌شود. ویژگی‌های معمولاً استفاده شده در دامنه زمان شامل میانگین، میانه مربعاتی (RMS)، کرتوزیس، اوج به اوج، واریانس (Var)، انحراف معیار (Std)، فاکتور شکل، فاکتور نوسان، فاکتور پالس، و فاکتور حاشیه است. فرمول محاسبه خاص در جدول ۲ نشان داده شده است.

متغیر در دامنه فرکانس فرکانس f است. از طریق تغییر دامنه، تغییرات دامنه فرکانس سیگنال ارتعاش با فرکانس را با فرکانس مشاهده کنید. نسبت به ویژگی‌ها در دامنه زمان، مزیت ویژگی‌ها در دامنه فرکانس آشکار است. ویژگی‌های دامنه فرکانس معمولاً شامل میانگین مربعاتی فرکانس (RMSF)، فرکانس مرکزی (CF)، میانگین مربعاتی فرکانس (MSF)، واریانس فرکانس (VF)، و واریانس مربعاتی فرکانس (RVF) است. فرمول محاسبه خاص در جدول ۳ نشان داده شده است.

آنتروپی برای توصیف ابهام سیستم یا اطلاعات استفاده می‌شود. طیف قدرت توزیع قدرت سیگنال ارتعاش را در دامنه فرکانس توصیف می‌کند. آنتروپی طیفی تک مقداری با انجام تجزیه مقدار ویژه بر روی سیگنال ارتعاش محاسبه می‌شود، و ویژگی‌های محلی سیگنال ارتعاش قابل استخراج است. آنتروپی طیفی تک مقداری یک ویژگی عیب در دامنه زمان است. آنتروپی انرژی موجک یک ویژگی

عیب در دامنه زمان-فرکانس است. آنتروپی دوبلان از تجزیه انحراف سیگنال ارتعاش در دامنه فرکانس برای توصیف عیب استفاده می‌شود.

علاوه بر این، دونگفانگ و همکاران مشـکل دقت کافی آنتروپی چند مقیاسـی را با بهبود عامل مقیاس آنتروپی چند مقیاسـی حل کردنـد. ویژگی‌هـای حـاصـل از این روش می‌تواننـد برای طبقه‌بندی‌ها بردار ویژگی دقیقتری فراهم کنند. این باعث افزایش دقت تشخیص می‌شود. کهنگ و همکاران آنتروپی سلسله مراتبی سیگنال ارتعاش را محاسبه کرده و آنتروپی سلسله مراتبی را به عنوان یک بردار ویژگی ورودی به یک طبقه‌بندی‌کننده که بهینه‌سـازی گلوله‌ای (PSO) و SVM را ترکیب می‌کند، اسـتفاده کردند. این روش نسبت به روشی که از آنتروپی چند مقیاسی به عنوان بردار ویژگی استفاده می‌کند، متفوق است.

نایانا و گیتانجالی ۱۲ ویژگی آماری در دامنه زمان و ۶ ویژگی طیفی وابسـته به زمان (TDSFs) استخراج کردند. از الگوریتم انتخاب ویژگی که WBDE و PSO را ترکیب می‌کند برای پردازش مجموعه اولیه ویژگی‌ها اسـتفاده شده است، و زیرمجموعه ویژگی نهایی بیشتر ویژگی‌های TDSFs را شامل می‌شود. زهرا و همکاران یک روش تشخیص عیب برای شناسایی درجه شکست برای عیب‌های عنصـر غلتکی ارائه دادند. در این روش، ابتدا از EMD برای پیش‌پردازش سـیگنال ارتعاش اسـتفاده شـده و سـپس از KLD برای پردازش بیشـتر IMF به منظور تشـکیل یک بردار ویژگی اسـتفاده می‌شـود. سـه طبقه‌بندDAG-SVM ، KNN، و درخت تصمیم (DT)برای مقایسـه و تأیید نتایج از این روش اسـتفاده می‌شـود. روی و همکاران از تجزیه مود امپریکال مجموعه (EEMD) برای پیش‌پردازش سـیگنال ارتعاش و محاسـبه آنتروپی سلسله مراتبی نمونه استفاده کردند. در این روش، از یک CS-SVM بهبود یافته به عنوان طبقه‌بند مدل استفاده می‌شود.

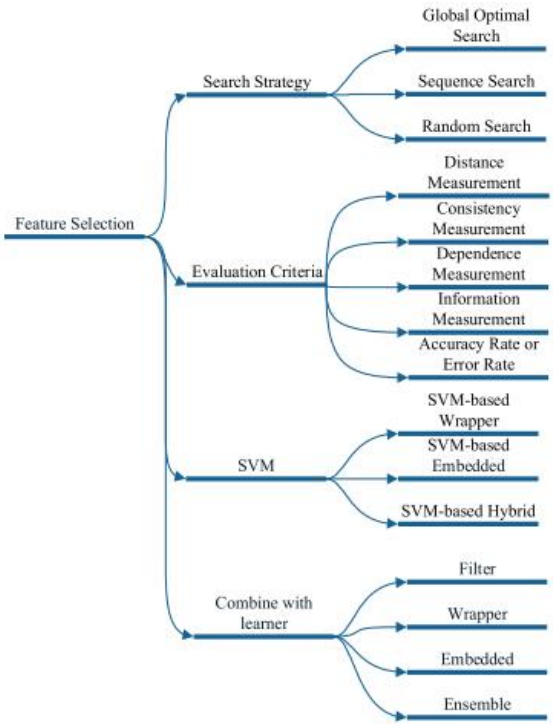


FIGURE 3. Classification of feature selection.

انتخاب ویژگی

ویژگی‌ها می‌توانند بر اساس نیازهای طبقه‌بندی به ویژگی‌های مرتبط، ویژگی‌های تکراری و ویژگی‌های بی‌ربط تقسیم شوند. هدف از انتخاب ویژگی این است که به اندازه‌ی ممکن ویژگی‌های تکراری و ویژگی‌های بی‌ربط را حذف کند و ویژگی‌های مرتبط را حفظ کند، بدین ترتیب بعد بردار ویژگی را کاهش دهد و از وقوع فاجعه بعدی و بیش‌برازش جلوگیری کند. فرایند اصلی انتخاب ویژگی از چهار مرحله تشکیل شده است، به عنوان مثال تولید زیرمجموعه‌های ویژگی، ارزیابی زیرمجموعه‌های ویژگی، شرایط متوقف کردن و نتایج تأییدی. بر اساس استراتژی جستجو، انتخاب ویژگی می‌تواند به سه دسته تقسیم شود، به عنوان مثال، جستجوی بهینه گلوبال، جستجوی توالی و جستجوی تصادفی. بر اساس معیارهای ارزیابی، انتخاب ویژگی می‌تواند به اندازه‌گیری فاصله، اندازه‌گیری پایداری، اندازه‌گیری وابستگی، اندازه‌گیری اطلاعات و نرخ دقت طبقه‌بندی یا اندازه‌گیری خطا در طبقه‌بندی تقسیم شود. بر اساس ترکیب انتخاب ویژگی و یادگیرنده، می‌توان آن را به چهار دسته تقسیم کرد، به عنوان مثال، فیلتر، وراپر، جاسازی و انجمن. برخی مقالات همچنین ماشین‌های بردار پشتیبان را با الگوریتم‌های انتخاب ویژگی ترکیب می‌کنند، که عمدتاً شامل سه دسته است، به عنوان مثال، وراپر مبتنی بر SVM، جاسازی مبتنی بر SVM، و هیبرید مبتنی بر SVM. جزئیات در شکل ۳ نشان داده شده است.

در زمینه تشخیص عیب بلبرینگ، بسیاری از مقالات در مورد انتخاب ویژگی و استخراج ویژگی منتشر شده است و مجموعه داده CWRU برای تحقیقات استفاده شده است. یانگهونگ و همکاران یک مدل انتخاب ویژگی با عنوان GL-mRMR-SVM پیشنهاد دادند که از همبستگی حداکثر و تناقض حداقل به عنوان معیار انتخاب ویژگی استفاده می‌کند، و از ویژگی‌های جهانی در دامنه فرکانس و دامنه زمانی و ویژگی‌های محلی استخراج شده توسط RNN به عنوان مجموعه اولیه ویژگی استفاده می‌کند، طبقه‌بند نهایی SVM استفاده می‌شود. یانگ و همکاران یک روش پردازش پوشش جدید با نام ICIE با بهبود CIE ارائه دادند و ICIE را با تجزیه متوسط محلی (LMD) ترکیب کرده و مدل ICIELMD را پیشنهاد دادند که یک ایده جدید برای استخراج ویژگی بلبرینگ‌های چرخان ارائه می‌دهد. یوه و همکاران PCA و شبکه عصبی BP را ترکیب کردند تا یک مدل جدید تشخیص عیب را پیشنهاد دهند. PCA برای کاهش بعد مجموعه ویژگی چند منبعی که شامل ویژگی‌های دامنه زمان، دامنه فرکانس و آنتروپی است، استفاده می‌شود و زیرمجموعه ویژگی به شبکه عصبی BP برای تشخیص عیب وارد می‌شود.

طبقه‌بندها

طبقه‌بندها نوعی از الگوریتم‌های یادگیری ماشین در طبقه‌بندی هستند. برخی از طبقه‌بندهای کلاسیک نظارت‌شده شامل نزدیک‌ترین همسایه(KNN) ، طبقه‌بند نیو بیز، ماشین بردار پشتیبان(SVM) ، درجه ارتباط خاکستری(GRD) ، و درخت تصمیم می‌شوند. همچنین برخی از طبقه‌بندهای کلاسیک بی‌نظارت شامل روش‌های خوشه‌بندی مانند خوشه‌بندیk-means ، DBSCAN، و خوشه‌بندی تجمعی می‌شوند. پارامترهای طبقه‌بند بر اساس مسائل مختلف بهبود می‌یابند تا توانایی تعمیم‌پذیری طبقه‌بند را افزایش دهند.

ماشین بردار پشتیبان(SVM)

SVM یک طبقه‌بند خطی است که مسائل طبقه‌بندی دودویی را حل می‌کند. این با پیدا کردن هایپرپلان حداکثر فاصله، طبقه‌بندی داده را انجام می‌دهد. SVM به دنبال یافتن یک هایپرپلان مانند $\omega^T x + b = 0$ است. هدف بهینه‌سازی SVM این است که با رعایت

طبقه‌بندی صحیح، فاصله بین بردار پشتیبان و هایپرپلان را به حداکثر برساند، به عبارت دیگر، یافتن صفحه هایپرحداکثر است، مسئله بهینه‌سازی SVM به فرمول ۱ تبدیل می‌شود.

$$\max \frac{1}{\|\omega\|} \quad s.t.\ y_i(\omega^T x_i + b) \geq 1 \qquad (1)$$

بسیاری از مقالات از مجموعه داده CWRU برای مطالعه کاربرد ماشین بردار پشتیبان (SVM) در تشخیص عیب بلبرینگ‌های چرخان استفاده می‌کنند .

جیانان و همکاران یک روش اضافی‌نمونه‌گیری به نام SCOTE پیشنهاد دادند و از SVM به عنوان طبقه‌بند اعتبارسنجی استفاده کردند SCOTE مسئله تعادل داده‌های چند کلاس را به مسائل عدم تعادل داده‌های دو کلاس تبدیل می‌کند و با ماشین بردار پشتیبان LS چند کلاسی ترکیب می‌شود تا مدل جدیدی برای حل مسئله عدم تعادل داده‌های عیب بلبرینگ‌های چرخان شکل گیرد .

یانگ و همکاران از تبدیل بسته‌های موجی (WPT) برای پیش‌پردازش داده‌ها استفاده کردند، تا توزیع انرژی سیگنال را به دست آورند و ویژگی‌ها را استخراج کرده و بردار ویژگی را تشکیل دهند. سپس از الگوریتم بهبود یافته انبوه ذرات (IPSO) پیشنهاد شده در مقاله برای بهینه‌سازی پارامترهای ماشین بردار پشتیبان (SVM) استفاده کردند .

ونتائو و همکاران روش جدید انتخاب ویژگی مبتنی بر فاصله را پیشنهاد دادند، با معرفی یک ماتریس شناسایی گروه برای به دست آوردن ضریب هر ویژگی طبقه‌بندی‌کننده اعتبارسنجی ماشین بردار پشتیبان (SVM) است.

B. K-NEAREST NEIGHBOR (KNN)

KNNیک الگوریتم طبقه‌بندی نظارت شده برای طبقه‌بندی چند کلاسی است و لازم است که قبل از همه برچسب‌های نمونه‌های موجود را به دست آوریم. برای یک نمونه ناشناخته، ما باید فاصله بین نمونه ناشناخته و تمام نمونه‌های موجود را محاسبه کنیم، و kنمونه با کمترین فاصله را انتخاب کنیم، و سپس بر اساس تعداد انواع مختلف نمونه‌ها در k نمونه، کلاس نمونه ناشناخته را تشخیص دهیم. اصل اساسی در شکل ۵ نشان داده شده است.

اولین نکته کلیدی از اصل اساسی KNN، کمی‌سازی ویژگی‌های مجموعه آموزش است. از آنجا که فاصله بین نمونه موجود و نمونه ناشناخته محاسبه می‌شود، لازم است اطمینان حاصل شود که ویژگی‌های موجود در هر نمونه به عنوان اعداد کمی‌سازی شوند. نکته دوم الگوریتم KNN، نرمال‌سازی داده‌ها است. محدوده مقادیر داده‌های ویژگی نمونه تأثیر مستقیمی بر محاسبه فاصله دارد، بنابراین لازم است هر داده ویژگی را به یک محدوده مشخص نرمال‌سازی کنیم. سومین نکته کلیدی الگوریتم KNN، تعیین تابع فاصله است. توابع فاصله موجود شامل فاصله اقلیدسی، فاصله کوسینوسی، فاصله همینگ و فاصله منهتن است. از این توابع، گسترده‌ترین استفاده از فاصله اقلیدسی است، که در فرمول ۲ نشان داده شده است.

$$d(r, R) = \sqrt{\sum_{i=1}^{n} (r_i - R_i)^2} \qquad (2)$$

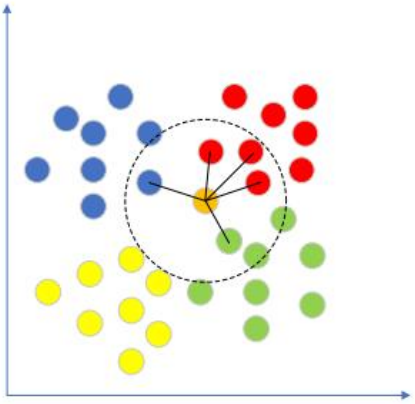


FIGURE 5. KNN.

بسیاری از مقالات از مجموعه داده CWRU برای مطالعه کاربرد KNN در تشخیص عیب بلبرینگ استفاده می‌کنند. Xin و همکاران از ارزش‌های تکیه گاه به عنوان ورودی مدل استفاده کردند و تئوری گراف را با SVD ترکیب کردند و روش جدیدی از مدل‌سازی گراف پیشنهاد دادند. اعتبارسنجی با طبقه‌بند KNN کارآیی این روش را در تشخیص زودهنگام عیب اثبات می‌کند. Qingfeng و همکاران روش تشخیص عیب WKNN با وزن‌دهی را پیشنهاد دادند و از الگوریتم انتخاب ویژگی ReliefF برای پردازش زیرمجموعه ویژگی‌هایی که توسط مجموعه ویژگی‌های چند حوزه‌ای تشکیل شده‌اند به عنوان ورودی WKNN استفاده کردند که توانایی عمومی‌سازی ضعیف طبقه‌بند تحت شرایط کاری متغیر را حل کرد.

۲,۲.ب)

استخراج ویژگی یک مرحله بسیار مهم در فرآیند یادگیری ماشین است که تأثیر قابل توجهی بر عملکرد مدل‌های یادگیری دارد. این فرایند شامل تبدیل داده‌های ورودی به یک مجموعه ویژگی‌های قابل استفاده برای آموزش مدل است. اهمیت استخراج، ویژگی به دلایل زیر است.کاهش ابعاد، با استخراج ویژگی‌های مناسب، ابعاد داده‌ها کاهش می‌یابد. این کاهش ابعاد موجب کاهش پیچیدگی مدل و افزایش سرعت آموزش و پیش‌بینی می‌شود.افزایش دقت، با استفاده از ویژگی‌های مناسب و مفید، دقت مدل‌های یادگیری افزایش می‌یابد. ویژگی‌های خوب و مرتبط می‌توانند اطلاعات مهمی را از داده‌ها استخراج کرده و باعث بهبود دقت و کارایی مدل شوند.از بین بردن اطلاعات، غیرضروری، این فرایند می‌تواند اطلاعات غیرضروری یا تکراری را از داده‌ها حذف کرده و بهبود کارایی مدل را فراهم آورد.افزودن اطلاعات جدید، استخراج ویژگی می‌تواند به ما امکان اضافه کردن اطلاعات جدید و مفید به داده‌ها را بدهد که ممکن است در اصل در داده‌های اولیه موجود نباشد. قابلیت تفسیرپذیری ویژگی‌های مناسب به ما امکان تفسیر و تبیین نحوه عملکرد مدل را می‌دهد، که این امر برای فهم بهتر عملکرد مدل و اعتماد به نتایج بسیار حیاتی است.بنابراین، استخراج ویژگی

از داده‌ها مرحله‌ای بسیار حیاتی در فرآیند یادگیری ماشین است که به بهبود کارایی و دقت مدل‌ها و افزایش قابلیت تفسیرپذیری آن‌ها کمک می‌کند.

۲,۲.ج

فرآیند برزدن (مخلوط کردن) و تقسیم داده دو مرحله اساسی در پردازش داده و ساخت مدل در یادگیری ماشین هستند که اهمیت بسیاری دارند. برزدن داده (مخلوط کردن) در این مرحله، داده‌ها را از منابع مختلف جمع‌آوری می‌کنیم و آن‌ها را با یکدیگر ترکیب می‌کنیم تا یک مجموعه داده کامل و یکپارچه بسازیم. این فرآیند اهمیت زیادی دارد زیرا داده‌ها ممکن است از منابع مختلف با کیفیت‌ها و فرمت‌های مختلفی باشند، بنابراین لازم است آن‌ها را به یک فرمت استاندارد و همگن تبدیل کنیم. همچنین، این فرآیند از اهمیت بسیاری برخوردار است زیرا داده‌های مختلفی را به یک مجموعه ترکیب می‌کند که می‌تواند به دقت و کارایی مدل‌های یادگیری ماشین کمک کند. تقسیم داده پس از برزدن داده، مجموعه داده را به دو بخش آموزشی و آزمون (یا همچنین معمولاً به عنوان مجموعه‌ی ارزیابی یا اعتبارسنجی شناخته می‌شود) تقسیم می‌کنیم. معمولاً یک قسمت از داده‌ها را برای آموزش مدل استفاده می‌کنیم تا مدل بتواند الگوهایی را از این داده‌ها بیاموزد و سپس از داده‌های باقی‌مانده برای ارزیابی عملکرد مدل استفاده می‌شود. این تقسیم داده اهمیت زیادی دارد چرا که ما باید مطمئن شویم که مدلی که آموزش می‌دهیم، به درستی عمل می‌کند و قادر است الگوها را در داده‌های جدید و ناآشنا تعمیم دهد.همچنین، این فرآیند به ما کمک می‌کند که از بروز مشکلاتی مانند بیش‌برازش یا کم‌برازش جلوگیری کنیم و مدل را بهبود بخشیم.

۲,۲.د

نرمال‌سازی یکی از مراحل اساسی و مهم در پیش‌پردازش داده در یادگیری ماشین است. هدف اصلی این فرآیند، تبدیل ویژگی‌های موجود در یک مجموعه داده به یک محدوده مشترک و مشابه است. با انجام نرمال‌سازی، تمام ویژگی‌ها به یک مقیاس مشابه تغییر می‌کنند، که این امر باعث می‌شود تا همه ویژگی‌ها در فرآیند یادگیری مساوی اهمیت داشته باشند و از بروز مشکلاتی مانند تسلط ناخواسته بر ویژگی‌های با مقیاس‌های بزرگ‌تر جلوگیری شود.دو روش متداول برای انجام نرمال‌سازی داده‌ها عبارتند از مقیاس‌گذاری حداقل حداکثر و استانداردسازی امتیاز است در روش مقیاس‌گذاری حداقل حداکثر، داده‌ها به گونه‌ای مقیاس می‌شوند که در محدوده ثابتی، معمولاً بین ۰ و ۱، قرار گیرند. این روش مناسب است زمانی که توزیع داده‌ها گاوسی نیست و دامنه ویژگی‌ها ثابت است.اما در روش استانداردسازی امتیاز ، داده‌ها به میانگین ۰ و انحراف استاندارد ۱ تبدیل می‌شوند. این روش بیشتر مناسب است زمانی که توزیع داده‌ها گاوسی است و مقیاس ویژگی‌ها گسترده و متفاوت است.استفاده از نرمال‌سازی داده‌ها، به عنوان یک مرحله پیش‌پردازش، می‌تواند بهبود همگرایی الگوریتم‌های بهینه‌سازی و جلوگیری از ناپایداری‌های عددی کمک کند. خیر از (x_train) برای تعیین مقیاس و پارامترهای نرمال‌سازی استفاده شده است. سپس همین پارامترها برای نرمال‌سازی داده‌های آموزشی و ارزیابی (x_test)استفاده شده‌اند. اطلاعات بخش ارزیابی به طور مستقیم در فرآیند نرمال‌سازی استفاده نشده است چون مقیاس‌ها و پارامترهای نرمال‌سازی از داده‌های آموزشی تعیین شده و این مقیاس‌ها برای نرمال‌سازی داده‌های ارزیابی به‌کار گرفته شده‌است.

سوال سوم

۳-۱

هیت مپ ماتریس همبستگی نشان می‌دهد که هرچه نقشه حرارتی تیره‌تر باشد، ما همبستگی بیشتری را مشاهده می‌کنیم. اعداد همبستگی بین ۱ و -۱ قرار دارند و می‌توان دید که دما و رطوبت همبستگی عکس دارند، اما میزان همبستگی خوبی با دید ندارند. همچنین، دما همبستگی معکوسی با رطوبت دارد و می‌توانیم این همبستگی را در هیت‌مپ مشاهده کنیم. در هیستوگرام، ما توزیع آماری داده‌ها را داریم و در هر بخش از داده‌ها می‌توانیم توزیع آماری هر ویژگی از آب و هوا را مشاهده کنیم. به عنوان مثال، توزیع آماری دما بیشتر بین فاصله ۰ تا ۲۰ درجه سانتیگراد قرار دارد و بیشترین تعداد نمونه‌ها در این فاصله قرار دارند. برای رطوبت، در نزدیکی یک تقریبا دوهزار نمونه داریم و می‌توانیم توزیعات آماری را در این هیستوگرام مشاهده کنیم.

۳-۲

روش کمترین مربعات (LS) به کمینه کردن مجموع مربع خطاهای بین مقادیر مشاهده شده و پیش‌بینی شده می‌پردازد. با MSE برابر با ۵۴,۷۶، مدل LS به خطای کمتری دست پیدا می‌کند که نشان می‌دهد مدل LS به خوبی با داده‌ها سازگار است و رابطه اصلی بین متغیرهای پیش‌بینی‌کننده و متغیر هدف را به خوبی بازتاب می‌دهد. در مقابل، تکنیک حداقل مربعات منظم‌شده (RLS)، یا رگرسیون ریج، یک عبارت منظم‌سازی را به تابع هدف LS اضافه می‌کند تا برازش بیش از حد را کاهش دهد. MSE بالاتر از ۱۶۲,۷۸ برای مدل RLS نشان می‌دهد که این تکنیک ممکن است منجر به یک مدل با انعطاف‌پذیری کمتر شود و باعث افزایش خطا شود. با این وجود، RLS می‌تواند در شرایطی که برازش بیش از حد یا چندخطی‌بودن یک مشکل است، مفید باشد. در نهایت، انتخاب بین LS و RLS باید با در نظر گرفتن عوامل مختلفی مانند توانایی تفسیر مدل، کارایی محاسباتی و ویژگی‌های داده‌ها انجام شود.

۳-۳

حداقل مربعات وزنی (WLS) یک تکنیک رگرسیونی است که برای تخمین پارامترهای یک مدل رگرسیون خطی در حالی که ناهمسانی یا واریانس نابرابر خطاها در داده‌ها را در نظر می‌گیرد، استفاده می‌شود. درWLS ، به هر نقطه داده بر اساس واریانس یا قابلیت اطمینان آن وزن اختصاص داده می شود. به نقاط داده با واریانس کمتر یا پایایی بالاتر وزن بیشتری داده می شود، در حالی که به نقاطی که واریانس بالاتر یا پایایی کمتر دارند وزن کمتری داده می شود.

مزیت اصلی WLS توانایی آن در محاسبه قابلیت اطمینان متغیر نقاط داده است که منجر به تخمین پارامترهای دقیق تر و تناسب مدل بهتر می شود. با تخصیص وزن‌های بالاتر به نقاط داده قابل‌اعتمادتر و وزن‌های پایین‌تر به نقاط کمتر قابل اعتماد، WLSتأثیر نقاط پرت و داده‌ها را با واریانس بالا کاهش می‌دهد و در نتیجه یک مدل رگرسیونی قوی‌تر ایجاد می‌کند. این امر WLS را به ویژه در شرایطی که فرض همسویی (واریانس ثابت) در رگرسیون حداقل مربعات معمولی نقض می شود مفید است.

علاوه بر این، WLS می‌تواند کارایی تخمین پارامتر را بهبود بخشد و با تنظیم مناسب برای تغییرپذیری خطاها در داده‌ها، به استنتاج دقیق‌تری منجر شود. این می‌تواند در مقایسه با رگرسیون حداقل مربعات معمولی، به فواصل اطمینان باریک‌تر و آزمون‌های فرضیه قابل‌اعتمادتر منجر شود، به‌ویژه زمانی که با داده‌هایی سروکار داریم که ناهمسانی را نشان می‌دهند. به طور کلی، WLS یک رویکرد انعطاف‌پذیر و مؤثر برای مدل‌سازی رگرسیون با گنجاندن اطلاعات مربوط به قابلیت اطمینان نقاط داده‌ای فردی در فرآیند تخمین ارائه می‌کند.