پنجمین کنفرانس ملی
کاربرد فناوری های نوین در علوم مهندسی
5th National Conferences
Application of Novel Technologies
in Engineering Sciences
۵ و ۶ اسفندماه ۱۳۹۹
دانشگاه تربت حیدریه
www.engconfs.ir

کاربرد داده‌کاوی در تشخیص اسکناس‌های جعلی

وحیده بابائیان [*۱]، ابراهیم ابراهیم زاده [۲]

۱- عضو هیات علمی گروه مهندسی کامپیوتر، دانشگاه صنعتی بیرجند، بیرجند، ایران

۲- دانش آموخته مقطع کارشناسی، گروه مهندسی کامپیوتر، دانشگاه صنعتی بیرجند، بیرجند، ایران

خلاصه

امروزه به دلیل اهمیت بررسی داده‌ها در سازمان‌های دولتی مخصوصاً بانک‌ها اهمیت داده‌کاوی روز به روز افزایش یافته است. در سال‌های اخیر تلاش‌های زیادی در زمینه ساخت دستگاه‌هایی که قادر به خواندن ارزش اسکناس باشند انجام شده است. در موارد پیشرفته‌تر با به کارگیری الگوریتم‌های کارا ساخت دستگاه‌هایی که توانایی تشخیص جعلی بودن اسکناس را داشته باشند نیز فراهم شده است. در این مقاله سعی شده است با تکنیک‌های داده‌کاوی و از طریق کلاس‌بندی به تشخیص هویت پول پرداخته شود. جامعه ی آماری استفاده شده در پژوهش، پایگاه داده UCI می‌باشد که از از کل مجموعه داده های موجود در آن، ۱۵۰ داده به طور تصادفی انتخاب و مورد استفاده قرار گرفته است. ویژگی‌های مورد استفاده در این مجموعه داده‌ها شامل واریانس تبدیل ویولت (موجک) تصویر، چولگی (skewness) تبدیل ویولت (موجک) تصویر، برجستگی تبدیل ویولت (موجک) تصویر و آنتروپی تصویر می‌باشد. هدف از انجام این مطالعه تشخیص تقلبی یا اصلی بودن اسکناس‌ها می‌باشد. در این مقاله، از کلاس‌بندهای KNN ، SVM، Decision tree و Bayes استفاده شده است و مقایسه بر اساس پارامتر کیفی دقت صورت گرفته است. نتایج بدست آمده حاکی از دقت بالای دو الگوریتم KNN و SVM می باشد.

کلمات کلیدی: اعتبار سنجی اسکناس، داده کاوی، K نزدیکترین همسایه ، ماشین بردار پشتیبان، درخت تصمیم، بیز

۱. مقدمه

امروزه به دلیل اهمیت بررسی داده‌ها در سازمان‌های دولتی مخصوصاً بانک‌ها اهمیت داده‌کاوی روز به روز افزایش یافته است. بررسی داده‌ها و استخراج اطلاعات از داده‌ها از اهمیت بالایی در دنیای کنونی برخوردار می‌باشد. در سیستم بانکی تشخیص پول‌های حقیقی و تقلبی بسیار مهم می‌باشد زیرا به دلیل جعل اسناد و وارد شدن پول‌های تقلبی در چرخه مالی کشور صدمات جبران ناپذیری به سیستم مالی کشور وارد می‌شود[۱].

در سال‌های اخیر تلاش‌های زیادی در زمینه ساخت دستگاه‌هایی که قادر به خواندن ارزش اسکناس باشند انجام شده است. در موارد پیشرفته‌تر با به کارگیری الگوریتم‌های کارا ساخت دستگاه‌هایی که توانایی تشخیص جعلی بودن اسکناس را

* نویسنده مسئول: babaiyan@birjandut.ac.ir

داشته باشند نیز فراهم شده است[۲]. در این مقاله سعی شده است با تکنیک‌های داده‌کاوی و از طریق کلاس‌بندی به تشخیص هویت پول پرداخته شود. کاربردهای داده‌کاوی در زمینه‌های مختلف علوم بصورت گسترده قابل مشاهده است. کاربردهای داده‌کاوی به طور گسترده ای در تجزیه و تحلیل داده‌های مالی[۳]، صنعت خرده فروشی[۴]، صنعت مخابرات[۵]، تجزیه و تحلیل داده‌های زیستی[۶] و ... استفاده می‌شود. داده‌های مالی به عنوان اطلاعات بسیار قابل اعتماد شناخته می‌شوند. به همین دلیل کاربردهای داده‌کاوی در این نوع داده‌ها بسیار گسترده است. نمونه هایی از کاربرد داده‌کاوی از تجزیه و تحلیل داده‌های مالی شامل پیش بینی پرداخت وام[۷]، تجزیه و تحلیل سیاست‌های اعتباری مشتری[۸]، طبقه‌بندی و خوشه‌بندی مشتریان برای بازاریابی هدفمند[۹]، تشخیص پول شویی و جعل اسکناس[۱۰] و ... می‌باشد.

داده‌کاوی از چند مرحله اساسی به شرح ذیل تشکیل می‌شود[۱۱]:

مرحله اول (تشکیل انبارداده): این مرحله برای انجام مراحل بعدی و داده‌کاوی در آن الزامی‌است. انبارداده محیطی است در حال تغییر که برای پژوهش آماده می‌شود.

مرحله دوم (انتخاب داده‌ها): برای آنکه هزینه‌های عملیات داده‌کاوی را کاهش دهیم نیاز است داده‌هایی را که از پایگاه داده انتخاب کنیم تا حجم داده ها کوچکتر شود.

مرحله سوم (تبدیل داده‌ها): برای انجام عملیات داده‌کاوی می‌بایست تبدیلات خاصی روی داده‌ها انجام گیرد، مانند تبدیل byte به integer و یا تبدیل رشته به عدد و ...

مرحله چهارم (کاوش در داده‌ها): کاوش داده از طریق تجهیزات مخصوصی که عملیات کاوش را بر اساس مدل‌های تجزیه و تحلیل انجام می‌دهد، می‌باشد. بررسی داده‌ها با انگیزه کشف نکات ارزشمند و دریافت اطلاعات مفید در حجم قابل توجهی داده که در طول زمان در کار و تجارت به دست آمده است را کاوش داده می‌گوییم.

مرحله پنجم (تفسیر نتیجه): در این مرحله نتایج و الگوهای ارائه شده توسط ابزار داده‌کاوی مورد بررسی قرار گرفته و نتایج مفید معین می‌شود.

۲. تحقیقات پیشین

در طی سالیان اخیر تحقیقات مختلفی در زمینه اعتبار سنجی اسکناس صورت گرفته است که در ادامه به معرفی این تحقیقات می‌پردازیم.

در مطالعه [۱۲] تحت عنوان تأیید اعتبار اسکناس با استفاده از طبقه بندی جنگل تصادفی بیان شده است که مؤسسات مالی، سیستم‌های مختلف بانکی خودکار را با استفاده از تشخیص ارز، به عنوان فعالیت اصلی خود اتخاذ نموده، که باعث می‌شود تشخیص خودکار ارز از اهمیت قابل توجهی برخوردار شود. تشخیص اسکناس‌های واقعی و جعلی برای انسان دشوار است، به خصوص به دلیل اینکه ویژگی‌های مشابه زیادی دارند. نوشته‌های جعلی با دقت ایجاد می‌شوند، از این رو نیاز به الگوریتم کارآمدای وجود دارد که بطور دقیق پیش بینی کند که آیا یک اسکناس اصیل است یا خیر. در این مقاله تکنیک‌های داده‌کاوی و یادگیری ماشین برای ارزیابی اعتبار اسکناس پیشنهاد شده است که از یک الگوریتم طبقه بندی با نظارت (طبقه بندی جنگل تصادفی) برای تمایز اسکناس‌های اصلی از جعلی استفاده می‌شود. نتایج ارزیابی در این مقاله پارامترهای، صحت (۹۴,۲٪)، حساسیت(۹۴,۳٪) و دقت (۹۲,۷٪) را نشان می دهد.

در [۱۳] مقاله ای با عنوان تجزیه و تحلیل سیستم تعیین هویت اسکناس با استفاده از تکنیک‌های یادگیری ماشین، بیان شده است که اسکناس یکی از مهمترین دارایی‌های یک کشور است. برخی از افراد متخلف، یادداشتهای جعلی را ارائه می‌دهند که شباهت زیادی به نوشته ی اصلی دارد که سبب ایجاد تناقض در پول می‌شود. در این مقاله تکنیک‌های یادگیری ماشین برای ارزیابی اعتبار اسکناس پیشنهاد شده است. الگوریتم‌های یادگیری نظارت شده مانند شبکه عصبی پس انتشار (BPN) و ماشین بردار پشتیبانی (SVM) برای تمایز اسکناس‌های اصلی از جعلی مورد استفاده قرار می‌گیرند. نرخ تشخیص در شبکه عصبی پس انتشار ۱۰۰٪ است و در SVM، ۹۸،۶۸% است.

در [۱۴] مقاله ای با عنوان شناسایی اسکناس‌های جعلی به طور کامل توسط یک استراتژی طبقه‌بندی، بیان شده است. اسکناس‌های جعلی برای همه ملت‌ها بلای جان هستند. فرآیندهای خودکار برای شناسایی سریع یادداشت‌های تقلبی با دقت بالا یک نیاز اساسی است. برای دستیابی به طبقه‌بندی کامل، از شبکه عصبی مصنوعی، ماشین بردار پشتیبانی و ماشین طبقه‌بندی نزدیکترین همسایه و روش‌های کاهش ابعاد استفاده شده است.

در [۱۵] مقاله ای با عنوان تجزیه و تحلیل و مقایسه ابزارها و روش‌های داده‌کاوی برای طبقه بندی تأیید اعتبار اسکناس بیان شده است که احراز هویت اسکناس برای تأمین اطلاعات از هر شخص غیرمجاز برای هر بانک و بخش مالی بسیار مهم و چالش برانگیز است. احراز هویت اسکناس از کلاهبرداری یا سوء استفاده از سیستم مالی جلوگیری می‌کند و از داده‌های مالی در برابر کاربران غیر مجاز محافظت می‌کند. کلاس‌بندی نقش بسیار مهمی‌در طبقه‌بندی نوشته‌های معتبر و تقلبی دارد. در این مقاله به دقت طبقه‌بندی‌های مختلف پرداخته شده است و دقت طبقه‌بندها را با استفاده از سه ابزار داده‌کاوی مقایسه کرده است. از نرم افزار Rapid miner، Tanagra و WEKA استفاده شده است. در حالت اول، ابزارهای داده‌کاوی Rapid miner و WEKA مقایسه شده است که کلاس‌بند جنگل تصادفی در ابزار داده‌کاوی WEKA، دقت بهتری برابر ۹۹/۳۰٪ دارد. در مورد دوم ، دقت طبقه‌بندی با استفاده از ابزار داده‌کاوی Rapid Miner و Tanagra مقایسه شده است که در ابزار داده‌کاوی Tanagra، KNN دقت ۹۹،۵۶% بدست آمده است. سرانجام با استفاده از ابزار داده‌کاوی WEKA و Tanagra، دقت طبقه‌بندی را مقایسه کرده است که در آن MLP با دقت ۹۹،۱٪ در ابزار داده‌کاوی WEKA بدست آمده است.

در [۱۶] مقاله ای به عنوان تجزیه و تحلیل عملکرد روش‌های مختلف داده‌کاوی در تأیید اعتبار اسکناس، با استفاده از ویژگی‌های عملکردی، از الگوریتم‌های مختلف داده‌کاوی مانند KMeans، Naive Bayes، Multilayer Perceptron ، درخت تصمیم‌گیری (J48) و (EM)Expectation-Maximization برای طبقه‌بندی مجموعه داده‌های تأیید اسکناس استفاده شده است. این آزمایشات در WEKA انجام شده است. هدف از این پروژه دستیابی به میزان احراز هویت بالاتر در طبقه بندی اسکناس است.

در [۱۷] مقاله‌ای با عنوان طبقه‌بندی اسکناس با استفاده از رویکرد شبکه‌عصبی مصنوعی جهت کلاس‌بندی اسکناس‌های واقعی و تقلبی از شبکه عصبی مصنوعی (ANN) استفاده شده است. جهت استخراج ویژگی از تبدیل ویولت استفاده شده است. رگرسیون آموزش ۰،۹۹۹۱۴ و رگرسیون تست ۰،۹۹۷۸۶ و رگرسیون اعتبار ۰،۹۹۵۳ بدست آمده است. نتایج نشان می‌دهد که استفاده از شبکه عصبی می‌تواند به موفقیت چشمگیری برسد.

در [۱۸] مقاله ای با عنوان مدل درخت تصمیم برای طبقه بندی اسکناس‌های جعلی و واقعی با استفاده از SPSS بیان شده است که جعل یک مشکل جامع است که بطور گسترده، عملاً و همچنین در واقعیت، در هر بخش در سراسر جهان رخ می‌دهد. به منظور شناسایی و طبقه‌بندی اسکناس‌های جعلی و واقعی، روش‌ها و مدل‌های مختلفی ارائه و ساخته شده است. در این مقاله یک مدل پیش بینی مؤثر بر اساس تکنیک یادگیری ماشین برای احراز هویت اسکناس ارائه شده است، که می‌تواند با دقت خوب پیش بینی کند که آیا اسکناس مورد نظر جعلی است یا اصیل. مدل درخت تصمیم‌گیری با

استفاده از ابزار IBM SPSS ساخته شده است. اندازه گیری عملکرد مدل با استفاده از نمودارهای دستیابی و نمودارهای شاخص انجام می‌شود و مشخص می‌شود که مدل درخت‌تصمیم پیشنهادی برای پیش بینی طبقه بندی اسکناس به صورت جعلی یا واقعی به اندازه کافی مناسب است.

در [۱۹] مقاله‌ای با عنوان تشخیص ارز جعلی با استفاده از پردازش تصویر بیان شده است پیشرفت تکنولوژی چاپ رنگ باعث افزایش نرخ چاپ نوشته‌های جعلی و کپی برداری از نوشته‌ها در مقیاس بسیار بزرگ شده است. این امر منجر به طراحی سیستمی می‌شود که در زمان کمتری و به شیوه ای کارآمدتر، اسکناس ارزی تقلبی را تشخیص دهد. سیستم پیشنهادی روشی برای تأیید اسکناس‌های ارزی ارائه می‌دهد. تأیید ارز توسط مفاهیم پردازش تصویر انجام می‌شود. در این مقاله به استخراج ویژگی‌های مختلف اسکناس‌های ارز پرداخته شده است. از نرم افزار متلب برای استخراج ویژگی‌ها استفاده می‌شود. سیستم پیشنهادی مزایایی مانند سادگی و سرعت کارایی بالا دارد.

در [۲۰] مقاله ای با عنوان طبقه بندی اسکناس‌های جعلی با استفاده از موجک و کلاس بند Anfis بیان شده است، امروزه، حجم معاملات در حال افزایش است و برخی از افراد در جوامع مبادله پول خود از اسکناس کلاهبرداری استفاده می‌کنند که این یک جرم است و غیرقانونی است و از نظر اقتصادی آسیب زیادی وارد می‌کند. تشخیص این اسکناس جعلی یک چالش جدی است. این مقاله رویکرد پیشنهادی را برای پیش بینی کلاس اسکناس‌های جعلی نسبت به کلاس واقعی ارائه داده است. شاخص طبقه بندی صحیح، سیستم عصبی فازی (ANFIS) مبتنی بر ویژگی‌های موجک است. این روش در مقایسه با روش‌های دیگر ۱۰۰٪ نتایج بهتر و دقیق‌تری را نشان می‌دهد.

در [۲۱] مقاله ای با عنوان تشخیص ارز با استفاده از یادگیری عمیق بیان شده است که ارز بخشی ضروری از زندگی روزمره ما است. با این حال، چگونگی شناسایی ارزهای واقعی و جعلی در حال حاضر به مهمترین مسئله تبدیل شده است. اگر از رایانه برای شناسایی ارز استفاده کنیم، دقت تشخیص را بسیار بهبود می‌بخشد و بار کاری افراد را به طور مؤثری کاهش می‌دهد. در سال‌های اخیر، یادگیری عمیق به محبوب‌ترین روش جهت تحقیق تبدیل شده است. این برنامه به طور عمده از طریق شبکه‌های عصبی عمیق یک مجموعه داده را آموزش می‌دهد. در طول این مدل‌ها می‌توان دقت در تشخیص ارز را بهبود بخشید. میانگین دقت تشخیص ارز در این مطالعه حداکثر ۹۶٫۶٪ بدست آمده است.

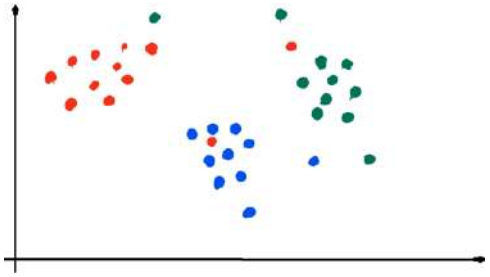
۳. پیاده سازی

نرم افزار مورد استفاده در پژوهش حاضر، نرم افزار متلب(MATLAB) ۲۰۱۹ بوده است. متلب یک محیط نرم‌افزاری برای انجام محاسبات عددی و یک زبان برنامه‌نویسی نسل چهارم است. واژه‌ی متلب به معنی محیط محاسبات رقمی و هم به معنی خود زبان برنامه‌نویسی مورد نظر است که ازترکیب دو واژه‌ی MATrix (ماتریس) و LABoratory (آزمایشگاه) ایجاد شده‌است. این نام حاکی از رویکرد ماتریس محور برنامه است، که در آن حتی اعداد منفرد هم به عنوان ماتریس در نظر گرفته می‌شوند[۲۲]. جامعه ی آماری استفاده شده در پژوهش، پایگاه داده UCI بوده است که در [۲۳] قابل دسترس می‌باشد.

از دیتابیس فوق ۱۵۰ داده به طور تصادفی انتخاب شده و مورد استفاده قرار گرفته است. دیتابیس استفاده شده دارای ۴ ویژگی است که شامل واریانس تبدیل ویولت (موجک) تصویر، چولگی (skewness) تبدیل ویولت (موجک) تصویر، برجستگی تبدیل ویولت (موجک) تصویر و آنتروپی تصویر می‌باشد. در این مجموعه داده‌ها، لیبل با دو مقدار ۰ یعنی

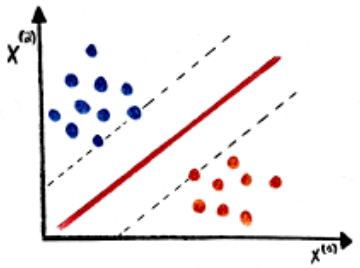
"جعلی نیست" و ۱ یعنی "جعلی است"، مشخص می‌شود. الگوریتم‌های به کار رفته در این مقاله، کلاس‌بندهای KNN، SVM، Tree_decision و Bayes است که در ادامه به معرفی آن ها می‌پردازیم.

KNN: روش K نزدیکترین همسایه (KNN) یک روش یادگیری است و از جمله ساده‌ترین الگوریتم‌های یادگیری ماشین و داده‌کاوی است. در این الگوریتم یک نمونه به دسته ای تعلق می‌گیرد که در همسایگی اش باشد. با تعیین تعداد همسایه‌ها می‌توان تعیین کرد که چند همسایه در اطراف نمونه ی مورد نظر باید در نظر گرفته شود. اگر $k=1$ باشد نمونه به سادگی در کلاس همسایگان نزدیکش تعیین می‌گردد. روش k همسایه نزدیک، برای بسیاری از روش‌ها کاربرد دارد، زیرا اثربخش، غیرپارامتریک و دارای پیاده‌سازی راحت می‌باشد. با این حال زمان دسته‌بندی‌اش طولانی است و یافتن مقدار k بهینه، مشکل است. بهترین انتخاب از k، وابسته به داده‌ها می‌باشد به طور کلی مقدار بزرگ از k، اثر نویز روی دسته‌بندی را کاهش می‌دهد، اما مرز مابین کلاس‌ها کمتر متمایز می‌شود[۲۴]. در شکل۱ نمایی از حاصل اجرای کلاس‌بند KNN با تعداد همسایه‌های متفاوت را مشاهده می کنید.



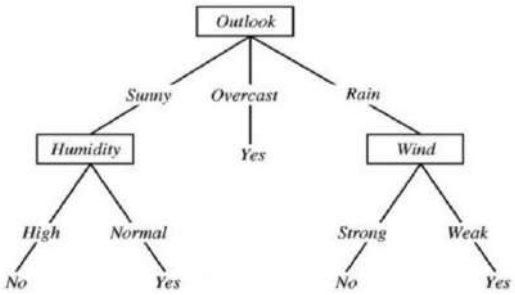
شکل۱- کلاس بند KNN با تعداد همسایه‌های متفاوت

SVM: ماشین بردار پشتیبانی (Support vector machines - SVMs) یکی از روش‌های یادگیری با نظارت است که از آن برای طبقه‌بندی و رگرسیون استفاده می‌کنند [۲۵]. این روش از جمله‌ی روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی نشان داده است. مبنای کاری دسته‌بندی کننده‌ی SVM دسته‌بندی خطی داده‌ها است و در تقسیم خطی داده‌ها سعی می‌کنیم خطی را انتخاب کنیم که حاشیه اطمینان بیشتری داشته باشد. با فرض اینکه دسته‌ها بصورت خطی جداپذیر باشند، ابرصفحه‌هائی با حداکثر حاشیه (maximum margin) را بدست می‌آورد که دسته‌ها را جدا کنند. این فرایند در شکل ۲ به تصویر کشیده شده است.



شکل ۲- کلاس بند SVM

درخت تصمیم (Decision-Tree): الگوریتم درخت تصمیم یک ابزار برای پشتیبانی از تصمیم است که از توصیف محاسبات احتمال شرطی نیز جهت مدل کردن بهره می‌برد. درخت تصمیم به‌طور معمول در تحقیق‌ها و عملیات مختلف استفاده می‌شود. به‌طور خاص در آنالیز تصمیم، برای مشخص کردن استراتژی که با بیشترین احتمال به هدف برسد، بکار می‌رود. یکی از مزایای بکارگیری درخت تصمیم می‌تواند به جلب توجهات بر روی مسئله و رابطه بین رویدادها بطور خیلی فشرده در قالب یک دیاگرام باشد. درخت تصمیم ساختاری شبیه درخت را جهت اخذ تصمیم و تعیین کلاس و دسته‌ی یک داده خاص، ترسیم می‌نماید. همان‌طور که از نام این الگوریتم پیداست، این درخت از تعدادی گره و شاخه تشکیل شده است به‌گونه‌ای که برگ‌ها دسته‌بندی‌ها یا طبقات را نشان می‌دهند و گره‌های میانی هم برای تصمیم‌گیری با توجه به یک یا چند صفت خاصه به کار می‌روند. همچنین، این الگوریتم قادر است علاوه بر متغیرهای کمی، متغیرهای کیفی را نیز پیش‌بینی کند؛ که این موضوع را می‌توان یکی دیگر از نقاط قوت این الگوریتم دانست. درخت تصمیم یک مدل خود توصیفی است؛ یعنی به تنهایی و بدون حضور یک فرد متخصص در آن حوزه، نحوه دسته‌بندی را به صورت گرافیکی نشان می‌دهد و به دلیل همین سادگی و قابل‌فهم بودن، روش محبوبی در داده‌کاوی محسوب می‌شود[۲۶]. در شکل ۳ نمونه ای از درخت تصمیم به تصویر کشیده شده است.



شکل ۳- کلاس‌بند درخت تصمیم

بیز: در یادگیری ماشین به گروهی از دسته‌بندی‌کننده‌های ساده بر پایه احتمالات گفته می‌شود که با فرض استقلال متغیرهای تصادفی و براساس قضیه بیز ساخته می‌شوند. به طور ساده روش بیز روشی برای دسته‌بندی پدیده‌ها، بر پایه احتمال وقوع یا عدم وقوع یک پدیده است. این روش از ساده‌ترین الگوریتم‌های پیش‌بینی است که دقت قابل قبولی هم دارد. شیوه یادگیری در روش بیز ساده از نوع یادگیری با نظارت است. این روش در دهه ۱۹۶۰ در میان دانشمندان بازیابی اطلاعات توسعه یافت و هنوز هم از روش‌های محبوب در دسته‌بندی اسناد به شمار می‌آید.

۴. مقایسه و ارزیابی

در مقاله حاضر، از چهار کلاس‌بند KNN ، SVM، Tree_decision و Bayes استفاده شده است. جهت تعیین میزان دقت در این روش‌ها، از پارامتر ارزیابی Accuracy استفاده می‌شود. این پارامتر از تقسیم تعداد طبقه‌بندی‌های درست به

تعداد کل داده‌های تست بدست می‌آید. مقادیر پارامتر ارزیابی برای الگوریتم های مذکور در جدول ۱ نمایش داده شده است.

جدول ۱- نتایج ارزیابی

الگوریتم	دقت
KNN Classifier	٪ ۹۸,۹۱
SVM Classifier	٪ ۹۹,۹۲
Decision_Tree Classifier	٪ ۸۴,۴۴
Bayes Classifier	٪ ۸۲,۲۲

با توجه به نتایج بدست آمده، با کمک ویژگی‌های مذکور الگوریتم KNN و SVM به خوبی و با دقت نزدیک ۱۰۰٪ می‌توانند پول‌های تقلبی و اصلی را از یکدیگر متمایز کنند. در مقاله ای مشابه [۱۳] که از کلاس بند SVM و شبکه عصبی پس‌انتشار خطا استفاده شده است، دقت SVM، ۹۸,۶۸ ٪ محاسبه شده است، در حالی که پژوهش حاضر به دلیل انجام عملیات داده‌کاوی بر روی دیتابیسی با ویژگی‌های بهتر، SVM، بیشترین دقت نزدیک به ٪۱۰۰ را دارد.

۵. نتیجه گیری

در این مقاله ، ما به بررسی دقت کلاس بندهای مختلف در تشخیص اسکناس تقلبی پرداخته ایم. شناسایی اسکناس جعلی با روش های غیر مدرن و بدون کمک از تکنیک های داده کاوی ممکن است منجر به تشخیص اشتباه شود. هدف از این مقاله شناسایی اسکناس‌های تقلبی با بهترین راه حل‌های ممکن با دقت بالا می باشد. در این مقاله چهار کلاس‌بند KNN، SVM، Tree_decision و Bayes بکار گرفته شد و از جنبه میزان دقت نتیجه حاصل مورد مقایسه قرار گرفت. جهت تعیین میزان دقت در این روش‌ها، از پارامتر ارزیابی Accuracy استفاده شد. نتیجه پیاده سازی حاکی از دقت بالای دو روش KNN و SVM می باشد. پیشنهاد می‌گردد سایر روش های طبقه بندی و همچنین روش های ترکیبی به منظور بررسی جامع تر و همچنین در مجموعه داده های مختلف صورت پذیرد تا ارزیابی جامع تری را داشته باشیم.

۶. مراجع

1. Shefraw, A.A.S.A.,(2020). "*Deep Learning Approach for Ethiopian Banknote Denomination Classification and Fake Detection System*". International Journal of Computer Science and Control Engineering, **7**(4): p. 30.
2. Chang, Y.-F., *Counterfeit paper currency detector*. (2020), Google Patents.
3. Mohammadi, M., et al., (2020). "*Financial reporting fraud detection: an analysis of data mining algorithms*". International Journal of Finance & Managerial Accounting, **4**(16): p. 1-12.

4. Castelo-Branco, F., et al., (2020), "*Business Intelligence and Data Mining to Support Sales in Retail*, in *Marketing and Smart Technologies"*. Springer. p. 406-419.
5. Mosa, M., et al., (2020). "*A Literature Review of Data Mining Techniques for Enhancing Digital Customer Engagement"*. International Journal of Enterprise Information Systems (IJEIS), **16**(4): p. 80-100.
6. Mahmud, M., et al., (2020). "*Deep learning in mining biological data"*. Cognitive Computation, p. 1-33.
7. Kisaka, G.,(2020). "*Assessment of factors leading to loan repayment defaulting in microfinance institutions"*. Mzumbe University.
8. Zou, X. and B. Tian, *Retailer's optimal ordering and payment strategy under two-level and flexible two-part trade credit policy.* Computers & Industrial Engineering, 2020. **142**: p. 106317.
9. Babaiyan, V. and S.A. Sarfarazi, (2019). "*Analyzing customers of South Khorasan telecommunication company with expansion of RFM to LRFM model"*. Journal of AI and Data Mining, **7**(2): p. 331-340.
10. Salazar, J.d.J.R., M.J. Segovia-Vargas, and M. del Mar Camacho-Miñano, (2020). "*Money laundering and terrorism financing detection using neural networks and an abnormality indicator"*. Expert Systems with Applications, p. 114470.
11. Thearling, K., (1999). "*An introduction to data mining"*. Direct Marketing Magazine, p. 28-31.
12. Jaiswal, R. and S Jaiswal, *Banknote Authentication using Random Forest Classifier.*
13. Shahani, S., A. Jagiasi, and R. Priya, (2018). "*Analysis of Banknote Authentication System using Machine Learning Techniques"*. International Journal of Computer Applications, **179**(20): p. ۲.۲۶-۲
14. Ala'raj, M., M. Majdalawieh, and M.F. Abbod, (2020). "*Improving binary classification using filtering based on k-NN proximity graphs"*. Journal of Big Data, **7**(1): p. 1-18.
15. Shrivas, A. and P. Gupta, (2017). "*Analysis and Comparison of Data Mining Tools and Techniques for Classification of Banknote Authentication"*. International Journal of Advanced Research in Computer Science, **8**(5).
16. Nife, N.I., (2016). "*Performance analysis of various data mining techniques on banknote authentication"*. International Journal of Engineering Science Invention, **5**(2): p. 62-71.
17. Kaya, E., A. Yasar, and I. Saritas, (2015). "*Banknote classification using artificial neural network approach"*. training, **10**: p. 1.
18. Upadhyaya, A., V. Shokeen, and G. Srivastava, (2018). "*Decision tree model for classification of fake and genuine banknotes using SPSS"*. World Review of Entrepreneurship, Management and Sustainable Development, **14**(6): p. 683-693.

19. Agasti, T., et al. (2017). "*Fake currency detection using image processing"*. in *IOP Conference Series: Materials Science and Engineering*.
20. Iraji, M.S. and S.A. Moghadam, (2016), "*Fake Banknotes Classification Using Wavelet and Anfis Topology Based on Vision"*. International Journal of Computer Science and Information Security, **14**(8): p. 126.
21. Zhang, Q., (2018). "*Currency recognition using deep learning"*. Auckland University of Technology.
22. Knight, A., *Basics of MATLAB and Beyond*. (2019): CRC press.
23. Lohweg, V., *banknote authentication Data Set*. (2012).
24. Zhang, Z., (2016). "*Introduction to machine learning: k-nearest neighbors"*. Annals of translational medicine, **4**(11).
25. Jakkula, V., (2006). "*Tutorial on support vector machine (svm)"*. School of EECS, Washington State University, **37**.
26. Myles, A.J., et al., (2004). "*An introduction to decision tree modeling.* Journal of Chemometrics: A Journal of the Chemometrics Society, **18**(6): p. 275-285.