بهبود نرخ تشخیص حسابهای ارسال کننده هرزنامه در شبکه های اجتماعی

مهدی سالخورده حقیقی[۱]، ایوب شفیعی فرد[۲]
[۱]دانشگاه صنعتی سجاد، haghighi@sadjad.ac.ir
[۲] دانشگاه صنعتی سجاد، a.shafieifard234@sadjad.ac.ir

چکیده – شبکه‌های اجتماعی با وجود آنکه کاربرد زیادی در زندگی روزمره افراد دارند اما به بستر مناسبی برای ارسال هرزنامه‌ها به منظور تبلیغات و نیز نشر پیامهای مخرب تبدیل شده‌اند. اثرات هرزنامه‌های اجتماعی بسیار قابل توجه است، یک هرزنامه توسط تمام دنبالگران و دوستان دیده می‌شود و بدتر از آن ممکن است موجب گمراهی و کج فهمی در موضوعات جذاب و بحث‌ها گردد. با توجه به تأثیر بسیار مضر این نوع رفتارها، تشخیص حساب افرادی که هرزنامه در شبکه‌های اجتماعی ارسال می‌کنند، به عنوان یک مؤلفه بسیار مهم در انجام این تحقیق در نظر گرفته شده است. در این مقاله برای تشخیص حساب کاربری افرادی که هرزنامه ارسال می‌کنند یک روش دو مرحله‌ای با استفاده از روش‌های داده کاوی بیز به‌صورت پویا ارائه شده است. مرحله اول، ایجاد پایگاه داده‌ای از کاربران شبکه اجتماعی فیس‌بوک برای شناسایی گره‌های بالقوه ناهنجار در شبکه اجتماعی برای شناسایی و ایجاد الگو می‌باشد. در مرحله دوم، با استفاده از الگوی به دست آمده برای هر پارامتر می‌توان گره‌های جدید را مورد مقایسه و ارزیابی قرار داد و در صورت شناسایی گره مولد هرزنامه، اقدامات لازم را جهت دسته بندی این گره‌ها انجام داد. در پایان روش پیشنهادی با روش‌های دیگر مقایسه و ارزیابی می‌گردد. نتایج حاصل نشان دهنده افزایش دقت روش پیشنهادی در مقایسه با سایر روشهای تشخیص هرزنامه است.

کلید واژه- داده کاوی، شبکه‌های اجتماعی ، هرزنامه، ناهنجاری

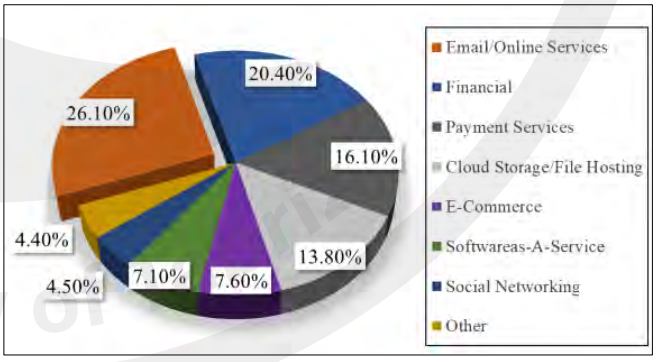
۱-مقدمه

در سالهای اخیر، نظرات جعلی و هرزنامه مشکلی است که به شدت در حال گسترش و افزایش است. امروزه شبکه‌های اجتماعی، هدف مناسبی برای تولیدکنندگان هرزنامه هستند. اغلب این شبکه‌ها قسمتی برای نظرات کاربران دارند و کاربران م توانند دیدگاه‌های خود را انتشار دهند که حاوی اطلاعات ارزشمندی می‌باشد. برای اینکه نظرات و تجربیات واقعی کاربران به درستی منعکس شود، شناسایی نظر جعلی و انتشار دهنده پست‌های جعلی در شبکه‌های اجتماعی بسیار مهم است[۱]. متأسفانه تولید کنندگان هرزنامه نیز از طریق ارسال پست می‌تواند به اهداف خود نظیر دسترسی به اطلاعات شخصی کاربران، انتشار ویروس و غیره دست یابند. پس می‌توان پستها را به دو دسته هرزنامه و غیر هرزنامه دسته بندی کرد. در سالهای اخیر، داده کاوی موفقیت بسیاری در زمینه تشخیص ناهنجاری در شبکه‌های اجتماعی داشته است. به همین دلیل متخصصین زیادی به استفاده از روش‌های داده کاوی برای ساخت سیستم‌های تشخیص ناهنجاری مشغول هستند[۲]. تحلیل شبکه‌های اجتماعی رویکردی برای مطالعه فعل و انفعال شبکه‌های بشری است، و برای بررسی الگوها، ساختار ارتباطی یا سازماندهی شبکه‌های اجتماعی استفاده شود. تحلیل شبکه اجتماعی دو تحلیل ریاضی و بصری، روابط بشری (بین افراد) را نشان می‌دهد[۳]. اثرات هرزنامه‌های اجتماعی بسیار قابل توجه است، یک هرزنامه توسط تمام دنبالگران و دوستان دیده می‌شود و بدتر از آن ممکن است موجب گمراهی و کج فهمی در موضوعات جذاب و بحث‌ها گردد. به عنوان مثال موضوع‌های جذاب همیشه توسط ارسال کنندگان هرزنامه مورد سوء استفاده قرار می گیرد تا کاربران را به وب سایت‌های غیر مرتبط و دلخواه خودشان گمراه نمایند. همچنین هرزنامه‌ها حجم زیادی از پهنای باند را اشغال نموده‌اند، و شناسایی ارسال کنندگان هرزنامه می‌تواند سهم عمده‌ای در کاهش مزاحمت‌های آنها و سربار ناشی از هرزنامه‌ها را داشته باشد. به خاطر حجم زیاد داده برای تجزیه وتحلیل در شبکه‌های اجتماعی بدون استفاده از روشهای داده کاوی این کار تقریبا یک کار غیر ممکن شده است. به عنوان یک فرد خاص و یا گروه‌هایی از افراد، کاربران مولد هرزنامه الگوهای تعاملاتی متفاوتی نسبت به کاربران عادی دارند و تغییرات ناگهانی در رفتار آنها به وجود می‌آید و یا به شیوه‌ای که به طور قابل توجهی از همتایان آن‌ها متفاوت است، برهم اثر می‌گذارند. در نتیجه اثرات این رفتارهای ناهنجار در ساختار شبکه مشاهده می‌گردد. به عنوان مثال، ممکن است افراد کلاه‌بردار در سیستم مزایده آن‌لاین در شبکه‌های اجتماعی با ارسال هرزنامه برای افزایش شهرت خود، استفاده کنند[۴]. یکی از چالش‌های مهم شناسایی افراد مخربی هستند که می‌توانند در شبکه‌های اجتماعی حداکثر استفاده را از

امکانات موجود برده و اقدام به سودجویی نمایند و یا به افراد آسیب‌پذیر ضربات جبران‌ناپذیری تحمیل کنند. بسیاری از کاربران سیستم‌های اجتماعی آنلاین مانند فیس‌بوک، گوگل پلاس و توییتر، به طور منظم در معرض یک سری از هرزنامه‌ها و دیگر موارد تهاجمی هستند[۵]. در این مقاله سعی خواهد شد با ارائه الگوریتم داده‌کاوی پیشنهادی با رویکرد شبکه بیزین حساب‌های کاربری افرادی که به ارسال هرزنامه در شبکه‌های اجتماعی مشغول هستند تشخیص داده شود[۶].

۲-پیشینه تحقیق

امروزه انتشار هرزنامه در شبکه‌های اجتماعی یکی از مشکلات عمده این شبکه‌ها به شمار می‌رود که خسارات زیادی برای کاربران به همراه دارد. یکی از بزرگترین شبکه‌های اجتماعی که امروزه کاربران بسیار زیادی از آن استفاده می‌کنند، توییتر است. توییتر، خود دارای قابلیتهایی برای گزارش این هرزنامه‌ها است و کاربران می‌توانند حساب‌های کاربری که آنها را دچار مشکل کرده‌اند گزارش دهند اما این مسئله از دو بعد دارای اشکال است. اول این که کاربران، خود به‌طور دستی باید اقدام به گزارش نمایند و دوم این که احتمال اینکه گزارش اشتباهی صورت بگیرد نیز وجود دارد لذا به سیستمهایی برای شناسایی این حساب‌ها و مقابله با آنها نیاز است. در نمودار شکل۱، میزان تاثیرگذاری هرزنامه و حملات فیشینگ مبتنی بر آنها در بخش-های مختلف نمایش داده شده است و می‌توان دریافت که هرزنامه‌ها در این موارد بیشتر از جانب وب‌سایتهای مالی برای کاربران ارسال شده و هویت آنها سرقت می‌شود[۱۲و۷]:



شکل۱: هرزنامه و ارتباط آن با سرقتهای آنلاین در اینترنت

تحلیل نمودار نشان می‌دهد که ۲۶,۱۰٪ هرزنامه‌ها بر علیه سرویس-های ایمیل بکار گرفته شده و ۲۴,۴۰٪ نیز برای سرقت مالی از وب-سایتهای مالی و کاربران آن نقش دارند. از این جهت می‌توان دریافت که هرزنامه‌ها تا چه اندازه می‌توانند بر بخش‌های مالی تاثیر مخرب داشته باشند و لذا شناسایی آنها مهم و از اهمیت بالایی برخوردار است. تاکنون برای شناسایی هرزنامه در شبکه اجتماعی مطالعات مختلفی انجام شده که بیشتر آنها بر پایه روش‌های یادگیری ماشین و داده‌کاوی است.

عیسی و همکاران [۸] در سال ۲۰۱۸ روشی برای شناسایی هرزنامه در شبکه اجتماعی توئیتر استفاده کرده‌اند. در این تحقیق به ازای هر ۲۰۰ پیام در شبکه‌های اجتماعی توییتر ۲۱ توییت احتمال می‌رود که ارسال کننده هرزنامه باشند. با توجه به رشد سریع حجم هرزنامه جهانی، این مطالعه رویکرد جدیدی را برای تشخیص هرزنامه در رسانه‌های اجتماعی از غیر هرزنامه ارائه می‌دهد و تشخیص بهتری نسبت به رفتار کاربران مولد هرزنامه در توییتر دارد. در سال ۲۰۱۷ پاروز و همکاران [۹] در مقاله‌ای، از تکنیک ثبت جزئیات تماس برای تشخیص ناهنجاری در شبکه تلفن همراه استفاده کردند. آن‌ها برای این منظور با استفاده از تکنیک خوشه‌بندی در داده‌های حجیم شبکه فعالیت‌های کاربران را مورد بررسی قرار دادند. آن‌ها با ترکیب روش k-means و تکنیک‌های خوشه‌بندی سلسله مراتبی زمان و مکان فعالیت‌های کاربر که به‌طور غیرعادی باعث شده بود تا تقاضای ترافیک شبکه بالا برود را مورد بررسی و تجزیه و تحلیل قراردادند.

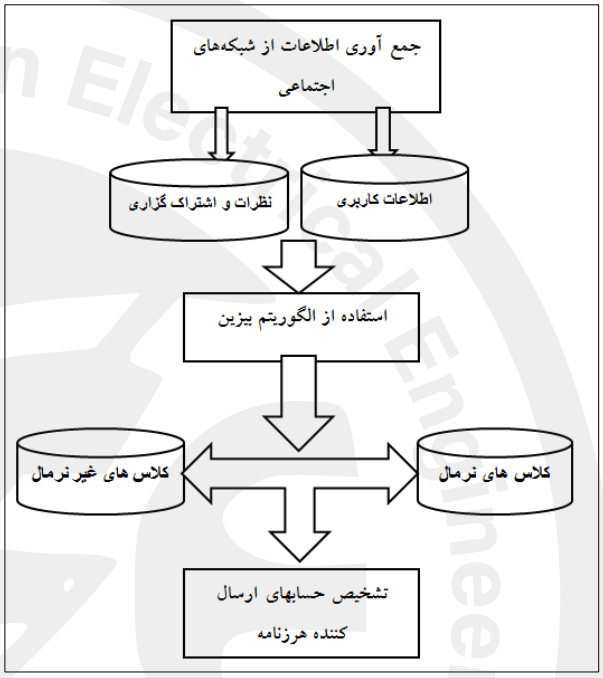
هیمانی چاولا[۱۰] در سال ۲۰۱۴ سیستمی ارائه داد که با استفاده از ماشین بردار پشتیبان قادر به دسته بندی صفحات مجاز از صفحات هرزنامه بود. برای هر صفحه، اطلاعاتی نظیر پستها، توضیحات صفحه، وجود لینک در توضیحات صفحه و یا در پستها و غیره موردبررسی و تجزیه وتحلیل قرار می‌گیرد. در این سیستم، دسته‌بندی و فیلتر کردن صفحات توسط ماشین بردار پشتیبان به خوبی صورت گرفته به طوری-که در یکی از موارد آزمون شده سیستم دارای دقتی برابر با ۸۹,۶٪ بوده است اما این کار هنوز نیازمند بهینه‌سازی است.

کار دیگری که برای محافظت شبکه اجتماعی فیسبوک و کاربرانش از حملات صورت گرفت، سیستمی بود که در سال ۲۰۱۲ توسط رحمان و همکارانش[۱۱] ارائه شد. این روش برای دسته بندی از ماشین بردار پشتیبان استفاده کرده است. نتایج بررسی آنها روی ۴۰ میلیون پست نشان داد که ۴۹٪ از کاربران مورد مطالعه در یک دوره ۴ ماهه در معرض حداقل یک پست هرزنامه قرار دارند. همچنین ۱۳٪ از برنامه موردمطالعه، برنامه‌های مخرب می‌باشند. مشکلی که سیستم ارائه شده توسط رحمان و همکارانش دارد این است که قادر به پاک کردن پستهای هرزنامه به طور خودکار نیست و اگر بخواهد به این مسئله دست یابد با مشکل بالا رفتن مثبت کاذب روبرو خواهد شد.

۳-روش پیشنهادی

در این بخش مراحل پیاده‌سازی تشخیص حساب‌های ارسال کننده هرزنامه در شبکه‌های اجتماعی و اجرای آن به‌صورت مرحله به مرحله و با بیان شرح کامل هر مرحله مانند شکل ۲ توضیح داده می‌شود.

در طراحی و پیاده‌سازی الگوریتم داده‌کاوی برای تشخیص حساب-های ارسال کننده هرزنامه در شبکه‌های اجتماعی در مرحله اول به‌عنوان ورودی، داده‌های واقعی از اطلاعات کاربری و روابط کاربران در شبکه‌های اجتماعی و همچنین اطلاعات مربوط به فعالیت‌های کاربران مانند پست‌های کاربران، ثبت نظرات کاربران و گروه‌هایی که کاربران عضو هستند جمع‌آوری شده است.



شکل۲: چارچوب کلی روش پیشنهادی

در مرحله دوم، الگوریتم بیز ساده با انجام بررسی‌های لازم برای کاربران جدید با توجه به الگو و حد آستانه برای هر پارامتر، حساب کاربران را بین کلاس‌های نرمال و غیر نرمال تقسیم می‌کند و سعی در تشخیص حساب‌های ارسال کننده هرزنامه و ارائه گزارش تعداد کاربران ناهنجار به عنوان خروجی برنامه دارد. در مرحله آخر با استفاده از مدل پیشنهادی حساب‌های ارسال کننده هرزنامه در شبکه‌های اجتماعی شناسایی می‌شود و در اختیار افراد و سازمان‌های مختلف قرار می‌گیرد. جامعه آماری اصلی این تحقیق به کاربران شبکه اجتماعی فیس‌بوک و روابط دوستی کاربران در این شبکه اجتماعی مربوط می‌گردد. در انتخاب جامعه آماری در این تحقیق برحسب موضوع تحقیق لازم است تا افرادی که از تجربه استفاده از شبکه‌های اجتماعی و عضویت در این شبکه برخوردارند به‌عنوان منابع جمع‌آوری داده انتخاب شوند. در مرحله اول پیاده‌سازی روش پیشنهادی در این تحقیق، به‌عنوان ورودی اطلاعات کاربری، اطلاعات مربوط به سلیقه-های کاربران و روابط و نظرات و پست‌های کاربران در شبکه‌های اجتماعی مانند فیس‌بوک توسط برنامه‌های کاربردی موجود مانند برنامه کاربردی Netvizz که در شبکه اجتماعی فیس‌بوک کار می‌کند جمع‌آوری می‌شود. بعد از جمع‌آوری و ذخیره‌سازی این اطلاعات مفید از شبکه‌های اجتماعی مانند فیس‌بوک، این اطلاعات برای استفاده توسط نرم‌افزار تحلیل شبکه‌های اجتماعی به نام GEPHI اجرا می‌شود و درنهایت می‌توان اطلاعات موردنیاز خود در زمینه خاص را با فرمت CSV از این برنامه استخراج نمود.

در ادامه با گردآوری داده‌ها و نمونه‌ها پیش پردازش آنها انجام شده و فاز نرمال‌سازی انجام می‌شود. نرمال‌سازی داده‌ها از جمله مهم‌ترین مراحل پیش‌پردازش در علم داده‌کاوی است. در این روش ساده، هر مجموعه‌ای از داده‌ها به بازه‌ای دلخواه، که کمترین و بیشترین مقدار آن از قبل مشخص است نگاشت می‌شود. در این روش می‌توان هر بازه دلخواه را تنها با یک تبدیل ساده، به بازه‌ای جدید نگاشت کرد. فرض کنید قرار است ویژگی A از مجموعه داده که در بازه بین min_A تا max_A قرار دارد، به بازه جدید new_Min تا new_Max نگاشت شود. برای این منظور، هر مقدار اولیه مانند x در بازه اولیه، طبق رابطه زیر به مقدار جدید x’ در بازه جدید تبدیل خواهد شد که در رابطه(۱)، نمایش داده شده است:

$$x' = \frac{newMax - newMin}{MaxA - MinA}(x - MinA) + (newMin) \quad (۱)$$

تعداد ارتباطات افراد در شبکه‌های اجتماعی در طول زمان به صورت فرآیندهای شمارشی زمان گسسته با استفاده از الگوریتم بیز، مورد بررسی قرار گرفته‌اند. در این الگوریتم برای هر دوره زمانی، تعداد ارتباطات بین افراد، وزن فعلی تجمع آن‌ها را در شبکه اجتماعی نشان خواهد داد. در پیاده سازی روش پیشنهادی ابتدا برخی از روش‌های مختلف شمارش ارتباطات مورد ارزیابی قرار گرفته است. سپس، مدل‌های احتمالی بیز ساده، برای پیاده‌سازی و اجرای این چنین فرآیندهای شمارشی ارائه می‌شوند.

برای هر کاربر i وقتی که فرآیند جمع آوری داده‌ها از شبکه‌های اجتماعی آغاز می‌گردد، از زمان صفر شروع می‌شود. فرض کنید که $N_i(t)$ تعداد ارتباطات ایجاد شده از i با سایر کاربران در زمان گسسته t باشد؛ به همین ترتیب برای مشاهده باینری ساده‌تری از شبکه، فرض کنید $N_j(t)$، تعداد دوره‌های زمانی باشد که در آن j با سایر کاربران در زمان t ارتباط داشته است. با توجه به اینکه الگوریتم بیز بر پایه احتمالات کار می‌کند، فرض کنید p_i یک مدل احتمالی برای توسعه $dN_i(t) = N_i(t) - N_i(t-1)$ در شرایط عادی باشد. در ساده‌ترین حالت، می‌توان $dN_i(1)، dN_i(2)، \cdots$ را مستقل از p_i در نظر گرفت؛ p_i در این مسئله جستجوی توزیع شده‌ای است که مربوط به حالت رفتار ارتباطی عادی و نرمال برای هر فرد در شبکه‌های اجتماعی می‌باشد. از سوی دیگر، رفتار غیرعادی (ناهنجار) در زمان

t، می‌تواند به عنوان یک مقدار $dN_i(t)$ دریافت شده از بررسی و مقایسه دیگری از p_i باشد که باید مورد توجه قرار گیرد. در نتیجه هدف در این مسئله برای شناسایی ارسال کننده هرزنامه، تشخیص مقادیری از $dN_i(t)$ است که از p_i مقداری بیشتر یا نامتعارف‌تر داشته باشد. همچنین در روش پیشنهای برای یک مقدار دریافتی از $dN_i(t)$، یک مقدار P به عنوان آستانه در الگوریتم بیز در نظر گرفته می‌شود که از این مقدار به دست آمده به عنوان معیاری برای شمارش احتمال پسین برای یافتن حساب کاربری ارسال کننده هرزنامه استفاده می‌شود. برای پیاده سازی روش پیشنهادی چارچوب فوق را می‌توان به عنوان یک جزء مستقل از تجزیه و تحلیل اعضای شبکه اجتماعی در نظر گرفت. این چنین فرآیندهایی، بر سطح فعالیت سراسر شبکه نظارت دارند و در صورتی که در شبکه اجتماعی ارسال کننده هرزنامه اتفاق بیفتد با توجه به روابط شمارشی بالا قابل شناسایی و ردیابی می باشد.

الگوریتم بیزین به دلیل سرعت بالا و سادگی پیاده سازی در بسیاری از کاربردها مورد استفاده گسترده قرار گرفته است. از جمله این کاربردها می‌توان به دسته بندی متون اشاره کرد. قضیه بیز یک معادله معروف است که به ما اجازه می‌دهد بر اساس داده، عمل پیش‌بینی را انجام دهیم. در رابطه(۲)، نسخه اولیه قضیه بیز نشان داده شده است.

$$p(class|data) = \frac{p(data|class) * p(class)}{p(data)} \quad (۲)$$

که در آن:

class : یک طبقه‌بندی یا برچسب خاص است.

data : یک داده مشاهده‌شده است.

p(class|data) : احتمال پسین نامیده می‌شود.

p(data|class) : درستی نامیده می‌شود.

p(class): احتمال پیشین نامیده می‌شود.

در طبقه‌بند بیز، ابتدا احتمال پسین را محاسبه می‌کنیم. سپس مشاهدات بر اساس کلاسی با بزرگ‌ترین مقدار پسین طبقه‌بندی می‌شوند. در روش پیشنهادی دو کلاس(نرمال و هرزنامه) و یک مشاهده برای پیش‌بینی داریم. در نتیجه دو احتمال پسین محاسبه خواهیم کرد: یکی برای کلاس نرمال و یکی برای کلاس هرزنامه که به ترتیب در رابطه(۳) و (۴) نمایش و فرموله شده است:

$$(normal|data) = \frac{p(data\,|\,normal) * p(normal)}{p(data)} \quad (۳)$$

$$(spamer\,|\,data) = \frac{p(data\,|\,spamer) * p(spamer)}{p(data)} \quad (۴)$$

در ادامه پیاده سازی و اجرای الگوریتم بیز با دو کلاس (نرمال و اسپم) و پنج ویژگی (تعداد دوستان، تعداد لایک، تعداد گروه، تعداد پست و تعداد کامنت) بر مبنای رابطه (۵) ارائه است.

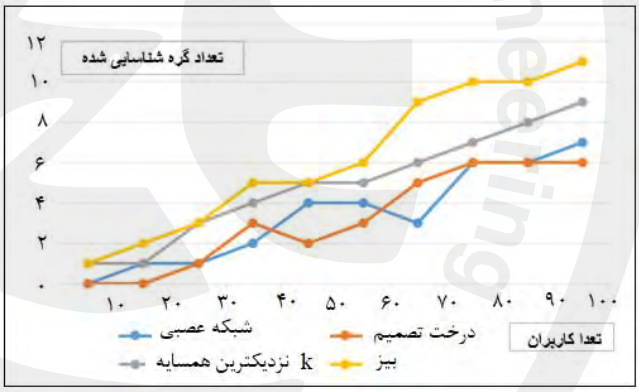
$$p(friends|spamer) * p(like|spamer) * p(group|spamer) * p(post|spamer) * p(comment|spamer) \quad (۵)$$

رابطه (۵) معادله درستی است. این نکته قابل توجه است که در اجرای این الگوریتم به داده مربوط به کاربران دسترسی داریم، بنابراین همه ویژگی‌های موجود در مجموعه داده را در اختیار خواهیم داشت. در هر یک از عبارت‌های رابطه (۵)، هر ویژگی مستقل از ویژگی‌های دیگر است. در نتیجه می‌توان از را بطه بیز ساده که بر مبنای مستقل بودن ویژگی‌ها محاسبه را انجام می‌دهد استفاده نمود. به عنوان مثال: تعداد دوستان با تعداد لایک و کامنت وابستگی ندارد. فرض می‌کنیم که مقدار ویژگی‌ها (مثل: تعداد دوستان با تعداد لایک و کامنت) به طور نرمال توزیع شده‌اند. به این معنی که $p = p(friends|spamer)$ با قرار دادن پارامترهای مورد نیاز در تابع چگالی احتمال محاسبه شده است. الگوریتم بیز با در نظر گرفتن آستانه برای هر پارامتر فعالیت‌های کاربر، رفتار او را با توجه به ارتباطاتی که با دیگران در شبکه‌های اجتماعی دارد شناسایی می‌کند و با تشخیص مناسب هرزنامه جواب بهینه را در اختیار برنامه قرار می‌دهد. در این الگوریتم با توجه به اطلاعات جمع آوری شده از شبکه‌های اجتماعی و جامعه آماری که در این تحقیق در نظر گرفته شده، و تخمین و شناسایی که با استفاده از این منابع اطلاعاتی مفید و نظرات و سلایق دوستان کاربران انجام می‌شود، ارسال کنندگان هرزنامه تشخیص داده شده و در اختیار کاربران قرار می‌گیرد.

۴-آزمایش‌ها و تحلیل نتایج

در این مقاله برای شبیه‌سازی و انجام محاسبات با الگوریتم بیز پیشنهادی، در مرحله اول تعداد ۲۲۷۰ نفر از کاربران عضو شبکه اجتماعی فیس‌بوک به‌عنوان نمونه انتخاب شده‌اند و روابط دوستی، فعالیت‌ها و مشخصات آن‌ها با استفاده از روش‌هایی که در فصل قبل ذکر شد از شبکه‌های اجتماعی جمع آوری شده است. همچنین نظرات دوستان این کاربران با توجه به نظراتی که در باره موضوعات خاص در گروه‌های مختلف مانند گروه‌های فروش در شبکه‌های اجتماعی داده‌اند جمع آوری شده است و در ادامه این بخش مورد ارزیابی و بررسی قرار می‌گیرد. همچنین در ادامه دقت اجرای الگوریتم بیز را برای انجام روش پیشنهادی دو مرحله‌ای با الگوریتم‌های پایه‌ای دیگر مانند الگوریتم شبکه‌های عصبی، درخت‌های تصمیم و k - نزدیک‌ترین همسایه برحسب افزایش تعداد کاربران با یکدیگر مقایسه و ارزیابی می‌شوند. در این تحقیق با توجه به تعداد مقایسه‌ای که در کارهای پیشین در این زمینه انجام شده است به‌عنوان نمونه الگوریتم

پیشنهادی با الگوریتم شبکه‌های عصبی، درخت‌های تصمیم و k - نزدیک‌ترین همسایه مقایسه و ارزیابی می‌شود. در ابتدا و به عنوان مرحله اول در مقایسه این الگوریتم‌ها، تعداد کاربران برای هر یک از الگوریتم‌ها ثابت نگه داشته شده و بر حسب افزایش یکسان تعداد کاربران برای الگوریتم‌ها، تعداد گره‌های مولد هرزنامه تشخیص داده شده توسط برنامه محاسبه شده و اعداد مشخصی بر حسب تعداد این کاربران به دست می‌آید. با توجه به اینکه شرایط و داده‌های مسئله برای هر یک از الگوریتم ها یکسان در نظر گرفته شده است، ارزیابی و مقایسه و آماری که از اجرای الگوریتم‌ها به‌دست می‌آید در این مرحله بر حسب تعداد کاربران هر الگوریتم است. همچنین در ارزیابی این الگوریتم‌ها در این تحقیق تعداد پارامتر با توجه به نوع و عملکرد هر کدام از الگوریتم‌ها با یکدیگر متفاوت در نظر گرفته شده است. در نمودار شکل۳، همانطور که مشاهده می‌کنید با ارزیابی و مقایسه مقادیر به دست آمده از روش‌های مختلف داده کاوی با مقادیر به دست آمده در روش پیشنهادی ، مشخص می‌شود که با افزایش تعداد کاربران، الگوریتم پیشنهادی تعداد بیشتری گره را شناسایی کرده است.



نمودار ۳: تعداد گره‌های مولد هرزنامه تشخیص داده‌شده بر اساس تعداد کاربران

بنابراین با این مقایسه می‌توانیم نتیجه بگیریم که استفاده از روش پیشنهادی می‌تواند باعث افزایش دقت در تشخیص ارسال کنندگان هرزنامه شود. برای ارزیابی بهتر روش پیشنهادی می‌توان از شاخص-های طبقه‌بندی استفاده نمود و به طور کلی معیارهای دقت، صحت و حساسیت از مهمترین معیارهایی هستند که در ارزیابی الگوریتم‌های طبقه بندی مورد استفاده قرار می‌گیرند. بر این اساس در مدل پیشنهادی از این سه معیار به منظور مقایسه طبقه‌بندی پیشنهادی با دو طبقه‌بندی جنگل تصادفی و شبکه عصبی استفاده شده است. در رابطه(۶)، (۷) و (۸)، به ترتیب سه شاخص صحت، دقت و حساسیت در طبقه‌بندی روش پیشنهادی فرموله شده است. برای ارزیابی این شاخص‌ها نیاز است که نمونه‌های صحیح مثبت(TP)، نمونه‌های غلط مثبت(FP)، نمونه‌های صحیح منفی(TN) و نمونه‌های غلط منفی(FN) محاسبه و شمارش شوند:

$$Precision = \frac{TP}{TP + FP} \quad (۶)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (۷)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (۸)$$

پس از انجام آزمایش‌ها برای طبقه‌بندهای مختلف نتایج به‌دست‌آمده‌اندکه بر اساس هر یک از معیارها مشخص‌شده است. نتایج به دست آمده از برنامه در جداول(۱) تا (۳) نشان داده شده است.

جدول(۱): نتایج ارزیابی طبقه بندها براساس معیار دقت

جنگل تصادفی	شبکه عصبی	طبقه بندی پیشنهادی	کاربران
۰,۵۱۲	۰,۸۱۹	۰,۷۷۶	۱۰۰
۰,۵۱۹	۰,۵۱۳	۰,۵۴۸	۱۰۰۰
۰,۵۲۶	۰,۵۱	۰,۵۶	۲۲۷۰

جدول(۲): نتایج ارزیابی طبقه بندها براساس معیار صحت

جنگل تصادفی	شبکه عصبی	طبقه بندی پیشنهادی	کاربران
۰,۵۱۲	۰,۸۱۹	۰,۷۶۳	۱۰۰
۰,۵۲۲	۰,۵۱۲	۰,۵۴۸	۱۰۰۰
۰,۵۰۴	۰,۵۰۶	۰,۵۸	۲۲۷۰

جدول(۳): نتایج ارزیابی طبقه بندها براساس معیار حساسیت

جنگل تصادفی	شبکه عصبی	طبقه بندی پیشنهادی	کاربران
۰,۵۱۷	۰,۸۲	۰,۷۷	۱۰۰
۰,۵۲۴	۰,۵۱۳	۰,۵۴۸	۱۰۰۰
۰,۵۱۴	۰,۵۱۱	۰,۵۵	۲۲۷۰

با مقایسه نتایج در هر سه معیار حساسیت، دقت و صحت مشخص شد با افزایش تعداد داده‌های آزمایشی، طبقه بند بیزین می‌تواند عملکرد بهتری در مقایسه با دو طبقه بند دیگر داشته باشد. دلیل اصلی موفقیت عملکرد الگوریتم پیشنهادی نسبت به دو طبقه بند شبکه عصبی و ماشین بردار پشتیبان در تصمیم‌گیری دقیق‌تر بر روی داده‌های نزدیک به مرز است. با افزایش داده‌های آزمایشی تعداد

بیشتری از داده‌ها در نزدیکی مرز تصمیم‌گیری قرار می‌گیرند که باعث افزایش میزان خطا در طبقه بندها خواهد شد. نتایج ارزیابی اثبات کرده است که طبقه بندی بیزین بهترین عملکرد نسبت به سایر روش-ها را دارد.

نتیجه گیری

در این پژوهش طبقه‌بند پیشنهادی بر مبنای ویژگی‌های استخراج شده از اطلاعات شبکه‌های اجتماعی با روش‌های مشابه دیگر مقایسه می‌شود و به منظور ارزیابی عملکرد، ابتدا سه معیار ارزیابی دقت، صحت و حساسیت تعریف شدند و سپس طبقه‌بند پیشنهادی با دو طبقه‌بند شبکه عصبی و جنگل تصادفی مقایسه شد. نتایج ارزیابی در هر سه معیار ارزیابی نشان داد که طبقه بند شبکه عصبی در مقایسه با دو طبقه بندی دیگر بهترین عملکرد را برای نمونه کاربران کمتر از ۱۰۰ دارد اما زمانی که حجم داده‌ها افزایش یافته است کارایی این طبقه بندی در هر سه معیار ارزیابی افت کرده است. همین روند برای طبقه بند جنگل تصادفی روی داده است. اما خلاف دو طبقه بند شبکه عصبی و جنگل تصادفی با افزایش تعداد کاربران عملکرد طبقه بند پیشنهادی بهتر شده است بطوری که برای ۲۲۷۰ کاربر دارای بهترین عملکرد برای هر سه معیار ارزیابی در هر دو سناریو می‌باشد. زمانی که تعداد داده‌ها افزایش پیدا می‌کند درصد بیشتری از نمونه‌ها در درون مرز یا نزدیک مرز قرار می‌گیرند.

با توجه به نتایج به‌دست‌آمده اثبات می‌شود که طبقه بند پیشنهادی در مقایسه با دو طبقه بندی دیگر زمانی که ابهام و تعداد کاربران افزایش می‌یابد یا به‌عبارت‌دیگر زمانی که تعداد نمونه‌ها در درون مرز یا نزدیک به مرز بیشتر می‌شود عملکرد بهتری داشته است. همچنین نتایج به‌دست آمده نشان می‌دهد که هر سه طبقه بند زمانی که کاربران مجاز در کلاس تائید و کاربران غیرمجاز در کلاس هشدار قرار گرفته‌اند عملکرد بهتری داشته‌اند. از جمله زمینه‌های تحقیقاتی که نیاز به کار بیشتر دارد می‌توان به منطق فازی و پیاده سازی الگوریتمهای فازی در داده کاوی به جهت تشخیص نفوذ اشاره نمود. در این پژوهش سعی بر آن شد تا از سیستم شبکه بیزین در الگوریتمهای داده کاوی بهره برده شود ولی در خصوص الگوریتمهای تشخیص ارسال کنندگان هرزنامه در زمینه داده کاوی ابعاد تحقیقی فراوانی وجود دارد و در پژوهش آتی از روش‌های فازی شده برای تشخیص هرزنامه استفاده می‌گردد.

مراجع

[1] Rojas, E., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2016). Process mining in healthcare: A literature review. *Journal of biomedical informatics*, *61*, 224-236.

[2] Contractor, D., Chawda, B., Mehta, S., Subramaniam, L. V., & Faruquie, T. A. (2015, June). Tracking political elections on social media: applications and experience. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.

[3] Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017, May). Online human-bot interactions: Detection, estimation, and characterization. In *Eleventh international AAAI conference on web and social media*.

[4] Davis, C. A., Varol, O., Ferrara, E., Flammini, A., & Menczer, F. (2016, April). Botornot: A system to evaluate social bots. In *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 273-274). International World Wide Web Conferences Steering Committee.

[5] Wang, B., Zubiaga, A., Liakata, M., & Procter, R. (2015). Making the most of tweet-inherent features for social spam detection on twitter. *arXiv preprint arXiv:1503.07405*.

[6] Inuwa-Dutse, I. (2018, April). Modelling formation of online temporal communities. In *Companion Proceedings of the The Web Conference 2018* (pp. 867-871). International World Wide Web Conferences Steering Committee.

[7] Fang, Y., Zhang, C., Huang, C., Liu, L., & Yang, Y. (2019). Phishing Email Detection Using Improved RCNN Model With Multilevel Vectors and Attention Mechanism. IEEE Access, 7, 56329-56340.

[8] Detection of spam-posting accounts on Twitter, Preprint submitted to Neurocomputing

[9] Parwez, M. S., Rawat, D. B., & Garuba, M. (2017). Big data analytics for user-activity analysis and user-anomaly detection in mobile wireless network. *IEEE Transactions on Industrial Informatics*, *13*(4), 2058-2065.

[10] Chawla, H. (2014). Facebook Page Spam detection using Support Vector Machines based on n-gram model. *International Journal of Computer Science Issues (IJCSI)*, *11*(5), 161.

[11] Rahman, M. S., Huang, T. K., Madhyastha, H. V., & Faloutsos, M. (2012). Efficient and scalable socware detection in online social networks. In *Presented as part of the 21st {USENIX} Security Symposium ({USENIX} Security 12)* (pp. 663-678).

[21] روتیا پاشری تشخیص هرزنامه با استفاده از ویژگی‌های شبکه عصبی مصنوعی و الگوریتم بهینه‌سازی کلونی مورچگان در شبکه‌های اجتماعی، پنجمین کنفرانس بین‌المللی فناوری اطلاعات و دانش در مهندسی برق، کامپیوتر و مکاترونیک، پاییز 2189.