Journal of Soft Computing and Information Technology (JSCIT)
Babol Noshirvani University of Technology, Babol, Iran
Journal Homepage: jscit.nit.ac.ir
Volume 8, Number 2, Summer 2019, pp. 1-16
Received: 08/26/2018, Revised: 12/14/2018, Accepted: 12/31/2018

Feature ranking for Persian Spam Review detection

Neshat Safarian[1], Mohammad Ehsan Basiri[2*] and Hadi Khosravi[3]
1- Department of Computer, Safahan Institute of Higher Educations, Janbazan St., Isfahan, Iran.
2*- Computer Department, Shahrekord University, Shahrekord, Iran.
3- Computer Department, Shahrekord University, Shahrekord, Iran.
[1]neshat.safarian@safahan.ac.ir, [2*]basiri@sku.ac.ir, and [3]khosravi@sku.ac.ir

Corresponding author address: Mohammad Ehsan Basiri, Faculty of Engineering, Shahrekord University, Shahrekord, Iran, Post Code: 115.

Abstract- Using online reviews is one of the main factors in customers' decision making for buying a product or using a service. These reviews are valuable sources of information which can be used for detecting public opinion about products or services. Although online reviews are useful, trusting them blindly is dangerous for both costumers and sellers as they may be manipulated by spammers to earn profit; such reviews are called spam reviews. The current study addresses Persian reviews about cell-phone extracted from Digikala.com and investigates spam type 1 and type 2 which are fake reviews and reviews describing brands' names only, respectively. Features used in this study, due to their efficiency, are review-based and metadata features. These features and their combinations in detecting Persian spam reviews, also their effect on the accuracy of classifier are assessed. Spam classification is performed using decision tree, support vector machines, and naïve Bayes classifiers and their accuracy are compared using different features' combinations. The highest accuracy is obtained using the decision tree classifier which achieves 0.778 in terms of F-measure. In ranking features, again the decision tree outperforms the other two classifiers by achieving 0.824 F-measure by combining the positive feedback, overall score, and review polarity features.

Keywords- Persian Spam Review، Fake Review، Opinion Mining، Sentiment Analysis، Text Categorization، Feature Ranking.

مجله علمی پژوهشی رایانش نرم و فن‌آوری اطلاعات
دانشگاه صنعتی نوشیروانی بابل
صفحه مجله: jscit.nit.ac.ir
جلد ۸، شماره ۲، تابستان ۱۳۹۸، صفحه ۱۶-۱
دریافت: ۱۳۹۷/۰۶/۰۴، بازنگری: ۱۳۹۷/۰۹/۲۳، پذیرش: ۱۳۹۷/۱۰/۱۰

رتبه‌بندی ویژگی‌ها در تشخیص نظرات اسپم فارسی

نشاط صفریان[۱]، محمد احسان بصیری*[۲]، هادی خسروی[۳]
۱- دانشکده مهندسی برق و کامپیوتر، مؤسسه آموزش عالی صفاهان، اصفهان، ایران.
۲- دانشکده فنی و مهندسی، دانشگاه شهرکرد، شهرکرد، ایران.
۳- دانشکده فنی و مهندسی، دانشگاه شهرکرد، شهرکرد، ایران.
[1]neshat.safarian@safahan.ac.ir, [2*]basiri@sku.ac.ir, and [3]khosravi@sku.ac.ir

* نشانی نویسنده مسئول: محمد احسان بصیری، شهرکرد، بلوار رهبر، دانشگاه شهرکرد، دانشکده فنی و مهندسی، صندوق پستی: ۱۱۵

چکیده- یکی از عوامل اصلی در تصمیم‌گیری مشتریان برای خرید یک محصول یا استفاده از خدماتی خاص، استفاده از نظرات برخط است. این‌گونه از نظرات منابعی ارزشمند از اطلاعات هستند که برای تشخیص افکار عمومی در مورد محصول یا خدمات استفاده می شوند. اگرچه نظرات برخط می‌توانند مفید باشند اما اعتماد کورکورانه به آن‌ها، هم برای فروشنده و هم برای خریدار خطرناک است زیرا ممکن است برای به‌دست آوردن سود، دستکاری شده باشند که اصطلاحاً به این گونه نظرات « نظرات اسپم» گفته می‌شود. پژوهش حاضر روی نظرات فارسی ثبت شده در مورد تلفن همراه در وب سایت دیجی‌کالا انجام شده و از میان انواع اسپم، اسپم نوع یک و دو بررسی شده‌اند که نوع اول، نظرات جعلی و نوع دوم نظراتی هستند که تنها در رابطه با مدل کالا نوشته شده‌اند. ویژگی‌های مورد استفاده در این پژوهش به علت کارا بودن آن‌ها در دسته‌بندی، شامل ویژگی‌های مبتنی بر نظر و ویژگی‌های فراداده است. این ویژگی‌ها و نیز ترکیب‌های متفاوت از آن‌ها در تشخیص نظرات اسپم فارسی و تأثیر آن‌ها روی دقت دسته‌بند بررسی شده است. دسته‌بندی توسط درخت تصمیم، دسته‌بند ماشین بردار پشتیبان و دسته‌بند نایو بیز انجام شده و در نهایت دقت آن‌ها روی ترکیب‌های مختلف این ویژگی‌ها با هم مقایسه گردیده است. بالاترین میزان دقت به‌دست آمده از سه دسته‌بند توسط درخت تصمیم حاصل می‌شود که برابر با ۰/۷۷۸ برا ساس معیار اِف ا ست. در رتبه‌دهی به ویژگی‌ها باز هم درخت تصمیم با دقت ۰/۸۲۴ و با ترکیب سه ویژگی بازخوردهای مثبت، امتیاز کلی کالا و قطبیت نظر رتبه‌ی برتر را به خود اختصاص می‌دهد.

واژه‌های کلیدی: نظرات اسپم فارسی، نظرات جعلی، نظرکاوی، تحلیل احساسات، دسته‌بندی متن، رتبه‌بندی ویژگی‌ها.

۱- مقدمه

یکی از در دسترس‌ترین روش‌ها برای ابراز عقیده در مورد یک شی یا محصول و حتی خدماتی خاص، ثبت نظر در وب‌سایت‌های متناظر است. به این‌گونه نوشته‌ها، محتوای تولید شده توسط کاربر می‌گویند؛ به این معنا که این محتوا با آن‌چه که توسط صاحبان و مدیران وب‌سایت‌ها نوشته شده است متفاوت است [۱]. با سرعت گرفتن رشد اینترنت و ایجاد تنوع در انواع وب‌سایت‌های مختلف، مردم نیز از تجارت الکترونیک و وب‌سایت‌های اشتراک عقاید استفاده می‌کنند. اغلب فروشندگان و خدمات برخط برای توسعه‌ی کسب و کار خود از مشتریان می‌خواهند تا عقیده و نظر خود را در مورد تجربه‌شان از خرید کالا یا دریافت خدمات بنویسند [۲].

هیجان، احساس، رفتارهای مختلف و عواطف نه تنها در مورد مسائل سیاسی و اقتصادی بیان می‌شود، بلکه می‌توان آن‌ها را در مورد کالاها، محصولات و خدمات ارائه‌شده در سطح اینترنت نیز مشاهده نمود [۳]. این امر موجب شده است تا حجم پیام‌های ارسالی کاربران به چنین وب‌سایت‌هایی به‌صورت چشمگیری در حال افزایش باشد. این نوع نظرات که از تجربه‌ی کاربران در مورد خرید کالایی خاص سرچشمه می‌گیرد، می‌تواند به‌طور مستقیم بر روی تصمیم خرید

سایر مشتریان در آینده تأثیرگذار باشد [۴].

با توجه به این واقعیت که هیچ کنترل کیفیتی روی این گونه نظرات در اینترنت انجام نمی‌پذیرد [۵]، هر فردی قادر خواهد بود که هر آن‌چه را که می‌خواهد در محیط وب نوشته و توسعه دهد. با وجود این‌که این نوع نظرات در تصمیم‌گیری، کمک شایانی به خریدار می‌کنند، اما اعتماد کورکورانه به آن‌ها، هم برای خریدار و هم برای فروشنده بسیار خطرناک است [۶]. از آن‌جایی که ممکن است این نظرات برای به‌دست آوردن سود، دستکاری شده باشند، هر گونه تصمیم‌گیری بر پایه‌ی نظرات برخط بایستی با دقت و با احتیاط کامل انجام پذیرد. علاوه بر این صاحبان مشاغل معمولاً به سمت و سوی افرادی کشیده می‌شوند که در مورد کالای تجاریشان نظرات خوب ثبت می‌کنند. این در حالی است که حتی به برخی افراد در قبال ثبت نظرات منفی در مورد کالاهای شرکت‌های رقیبشان پول پرداخت می‌کنند. این‌گونه از نظرات تقلبی را به عنوان «نظرات اسپم» شناخته و به افرادی که این نظرات را ثبت می‌کنند «اسپمر» می‌گویند [۷].

از آن‌جایی که نظرات اسپم به مشکلی فراگیر و نابودکننده تبدیل شده است، توسعه‌ی روش‌هایی که بتوانند میان نظرات قابل اعتماد و نظرات اسپم تمایز ایجاد کنند، به مسأله‌ای مهم و رقابتی تبدیل شده است. نظرات اسپم به سه دسته تقسیم می‌شوند [۸]: نظرات غیر قابل اعتماد، نظرات نوشته‌شده در مورد نام شرکت سازنده‌ی کالا و غیرنظرها. دسته اول از نظرات با ایجاد نظرات تقلبی مثبت، خوانندگان یا سیستم‌های نظرکاوی را گمراه می‌کنند تا محصول مورد نظر را ارتقا دهند. این نوع اسپم را آبراسپم[۱] می‌گویند [۹]. از جانب دیگر اسپمرها با ایجاد نظرات منفی دروغین برای برخی کالاهای دیگر می‌خواهند شهرت آن محصول را لکه‌دار کنند. این نوع اسپم را اسپم بدنام‌کننده[۲] گویند [۹]. دسته دوم برای تبلیغ نام شرکت و یا کارخانه‌ی سازنده‌ی یک کالای خاص ثبت شده‌اند. اگرچه این نوع نظرات می‌توانند مفید واقع شوند، آن‌ها جزو نظرات اسپم دسته‌بندی می‌شوند؛ زیرا هدف کالای خاصی نیست و همواره به سمت و سوی نام کارخانه‌ها تمایل دارند. دسته سوم نوعی از نظرات هستند که به هیچ وجه در مورد کالایی خاص نوشته نشده‌اند و نسبت به موضوع مورد بحث بی‌ربط هستند. نویسنده‌ی این نوع نظرات اغلب متنی را به‌منظور گسترش تبلیغات ثبت می‌کند.

از میان این سه نوع اسپم، نظرات اسپم نوع اول بیشترین نگرانی را در مورد نظرات تقلبی ایجاد کرده‌اند؛ زیرا کلیت سیستم نظرات برخط را زیر سوال برده و آن را کاملاً تحت تأثیر قرار می‌دهد. شناسایی این نوع نظرات بسیار سخت است و به‌راحتی قابل تمایز از نظرات معمولی نیستند [۱۰].

برای این‌که انواع نظرات غیر قابل اعتماد بهتر مشخص شوند، رابطه‌ی بین این نظرات و جهت احساسی آن‌ها در جدول ۱ نشان داده شده است. نظرات اسپمی که در نواحی ۳ و ۵ وجود دارند، توسط سازندگان کالا و افرادی نوشته می‌شوند که سهم اقتصادی مستقیم و منافع مادی زیادی در محصول داشته و هدفشان ارتقای کالا در دید مشتریان است. نظرات اسپم موجود در قسمت‌های ۲ و ۴ معمولاً توسط رقیبان نوشته شده‌اند. نظرات اسپم موجود در بخش‌های ۲، ۳، ۴ و ۵ بسیار مضر بوده و روش‌های تشخیص اسپم باید روی این‌گونه نظرات متمرکز شوند.

جدول ۱: ارتباط بین نظرات اسپم و کیفیت کالا.

نوع کالا	نظر مثبت	نظر منفی
کالا با کیفیت بالا	۱	۲
کالا با کیفیت متوسط	۳	۴
کالا با کیفیت پایین	۵	۶

روش‌های مختلفی در تشخیص نظرات اسپم وجود دارد. یادگیری نظارت‌شده متداول‌ترین روش یادگیری ماشین در تشخیص این‌گونه از نظرها است. مشکل بودن تشخیص نظرات اسپم و نیز برچسب‌گذاری آن‌ها موجب شده است تا در بسیاری از پژوهش‌ها از مجموعه داده‌های کوچک و یا مصنوعی استفاده شود. در پژوهش پیشِ رو برای غلبه بر این مشکل در برچسب‌گذاری نظرات از رویکرد پیشنهادشده در [۱۱] کمک گرفته شده و مسأله‌ی برچسب‌گذاری سلیقه‌ای نظرات اسپم حل شده است.

علیرغم تحقیقات زیادی که به وسیله‌ی داده‌های مصنوعی انجام شده است، دیده می‌شود که مجموعه داده‌های ساختگی یا مصنوعی شاخص بسیار ضعیفی برای نمایش کارایی در داده‌های دنیای واقعی هستند [۱۲]. این داده‌ها معمولاً تنها برای استفاده در پژوهش‌ها و آزمایش‌ها کاربرد دارند و در دنیای واقعی بسیار ضعیف عمل می‌کنند. به این منظور پژوهش حاضر بر روی نظرات واقعی نوشته‌شده در وب‌سایت دیجی‌کالا[۳] انجام شده است.

یکی دیگر از مشکلات مربوط به داده‌های تشخیص اسپم این است که داده‌های دنیای واقعی ممکن است عدم توازن زیادی در دسته‌های خود داشته باشند [۱]. زیرا در حال حاضر تعداد نظرات واقعی و صادقانه بسیار بیشتر از نظرات تقلبی برخط است که با برداشتن تعداد نمونه‌های یکسان از هر دو دسته این مشکل نیز قابل حل خواهد بود. گفتنی است که در پژوهش حاضر آزمایش روی هر دو نوع نمونه‌ی متوازن و نامتوازن انجام شده است.

آخرین مشکل مربوط به داده‌ها، کیفیت آن‌ها و وجود نویز در

آن‌ها است. نویز می‌تواند در داده‌های پردازش‌نشده به‌وجود آید. این داده‌ها باعث وجود دقت پایین در دسته‌بند می‌شوند. بدترین نوع نویز مربوط به دسته‌ای است که به‌دلیل نمونه‌های به‌اشتباه برچسب‌گذاری شده به‌وجود می‌آید. برای حل این مشکل در این پژوهش از راهکار موجود در [۱۱] استفاده شده است تا بتوان نمونه‌ها را به‌درستی دسته‌بندی کرد.

تحقیقاتی که تاکنون صورت پذیرفته‌است روی تشخیص اسپم تمرکز کرده‌اند و در مورد این‌که چه ترکیبی از ویژگی‌ها نامفهوم است و چه ترکیبی از ویژگی‌ها در چه مجموعه داده‌ای بهترین خروجی را خواهد داشت، صحبت نشده است. پژوهش‌های انجام شده‌ی [۹] و [۱۳] نشان می‌دهد که افزودن ویژگی‌های مبتنی بر نظردهنده نسبت به ویژگی‌های *n-gram* منجر به کارایی بالاتر دسته‌بند خواهد شد و سایر آزمایش‌ها و بررسی‌ها این نتیجه‌گیری را تأیید و حمایت می‌کند [۱۲]، [۱۴] و [۱۵]. بالاترین کارایی مشاهده‌شده مربوط به ترکیب ویژگی‌های متنی و غیرمتنی است. ویژگی‌های مبتنی بر نظر نیز در تشخیص دقیق نظرات اسپم مهم هستند. با وجود پژوهش‌های زیادی که با ویژگی‌های مختلف انجام شده است، اما پژوهشی در زمینه‌ی ترکیب مختلف ویژگی‌ها و کارایی آن‌ها روی دسته‌بند گزارش نشده است. از طرف دیگر کارهای انجام‌شده‌ی قبلی روی مجموعه داده‌های انگلیسی بوده و پژوهشی در زمینه‌ی نظرات اسپم فارسی گزارش نشده است.

در پژوهش حاضر پس از انجام پیش‌پردازش و پالایش داده‌ها، از ویژگی‌های استخراج‌شده از متن نظر (مبتنی بر نظر) و اطلاعات نوشته‌شده در کنار نظر (فراداده[۴]) کمک گرفته شده است. برای دسته‌بندی از درخت تصمیم، دسته‌بند نایو بیز و ماشین بردار پشتیبان[۵] (SVM) استفاده شده است؛ زیرا این دسته‌بندها کارایی بالایی در تشخیص نظرات اسپم انگلیسی و عربی داشته‌اند. سپس تأثیر ترکیب‌های مختلف ویژگی‌ها روی دقت این دسته‌بندها بررسی و تحلیل شده‌اند.

در ادامه، در بخش دوم مروری بر پژوهش‌های انجام‌شده در زمینه‌ی تشخیص اسپم انجام شده و مسائل به‌وجود آمده در خلال این آزمایش‌ها بررسی می‌شود. در بخش سوم، توصیف دقیق مراحل پژوهش حاضر آورده شده است. در بخش چهارم نتایج به‌دست آمده از پژوهش، بررسی و تحلیل می‌شود. در بخش پایانی نیز جمع‌بندی و پیشنهادهایی جهت انجام بهتر کار آمده است.

۲- مروری بر پژوهش‌های پیشین

پژوهش‌های انجام شده در زمینه‌ی تشخیص اسپم را می‌توان براساس نوع الگوریتم یادگیری استفاده شده به سه دسته‌ی نظارت شده[۶]، بدون ناظر[۷] و نیمه‌نظارتی[۸] تقسیم کرد [۱۶]. روش نظارت شده معمول‌ترین روش یادگیری است که دسته‌بندی را از نمونه‌های برچسب‌گذاری‌شده یاد می‌گیرد اما روش‌های دوم وسوم نیز مزایایی دارند که باعث شده برخی پژوهشگران آن‌ها را ترجیح دهند [۵]. در ادامه منتخبی از پژوهش‌های انجام شده در تشخیص نظرات اسپم در هر دسته ارائه می‌شود.

۲-۱- روش‌های نظارت شده

یکی از اولین پژوهش‌های گزارش شده در زمینه اسپم پژوهش جیندال و لیو [۹] است که در آن از روش‌های دسته‌بندی نظارت‌شده استفاده شده است. پژوهش آن‌ها روی استخراج عقاید از متن با استفاده از پردازش زبان طبیعی[۹] (NLP) متمرکز شده است. آن‌ها برای یافتن عقاید اسپم در میان نظرات، از ۵/۸ میلیون نظر نوشته‌شده توسط ۲/۱۴ میلیون نفر در مورد کالاهای ثبت‌شده در وب‌سایت آمازون استفاده کرده و نظرات دسته اسپم را به سه دسته‌ی نظرات غیر قابل اعتماد و دروغین، نظرات نوشته شده در مورد شرکت سازنده‌ی کالاها یا غیرنظرها تقسیم کردند.

روشی دیگر برای یافتن نظرات نیمه مشابه با استفاده از مدل‌های زبانی-معنایی در [۱۷] معرفی شد. آن‌ها ویژگی‌های مبتنی بر نظر، نظردهنده و محصولات را استخراج کردند و سپس یک مدل رگرسیون منطقی با استفاده از این ویژگی‌ها به‌وجود آوردند. این مدل را با استفاده از اعتبارسنجی مورب ده‌گانه[۱۰] آزمودند و با به‌کارگیری همه‌ی ویژگی‌ها، سطح زیر منحنی برابر با ۰/۷۸ شد؛ درحالی‌که با استفاده از ویژگی‌های متنی به‌تنهایی این مقدار برابر با ۰/۶۳ گردید.

چارچوب دیگری در دسته‌بندی در [۱۸] پیشنهاد شد. نظریه‌ی آن‌ها این بود که مجموعه داده‌ی آموزشی استفاده‌شده در پژوهش‌ها به دامنه‌ی خاصی از واژگان وابسته است. آن‌ها یک دامنه‌ی ترکیبی از مجموعه داده‌ها ایجاد کردند که در آن نظرات از سه دامنه‌ی مختلف (هتل، رستوران و پزشکی) برداشته شده بود. آن‌ها ویژگی‌های متفاوتی را روی مدل ساخته‌شده آزمودند و متوجه شدند که ویژگی اجزاء سخن (POS)[۱۱] خیلی قوی‌تر از ویژگی‌های یونی‌گِرَم[۱۲] به‌تنهایی عمل می‌کند. درحالی‌که ویژگی یونی‌گِرَم به‌تنهایی دقت ۵۲ درصد را حاصل کرد، استفاده از ویژگی اجزاء سخن منجر به دقت ۶۳/۴ درصد شد.

در پژوهش دیگری [۱۹] تنها روی ویژگی‌های زبانی[۱۳] و تأثیر آن‌ها روی کارایی تشخیص نظرات اسپم در زبان انگلیسی تمرکز شده است. آن‌ها نشان دادند که با استفاده از این ویژگی‌ها به‌تنهایی

به دقتی بیش از ۹۳ درصد دست یافته‌اند و بیان کردند که اگر از تعداد بسیار زیادی از ویژگی‌های زبانی استفاده کنند این امر منجر به افزایش کارایی تشخیص نظرات اسپم خواهد شد. در این روش دسته‌بند ماشین بردار پشتیبان بالاترین دقت را ایجاد کرد.

در پژوهش دیگری که اخیراً گزارش شده، [۲۰] از روش‌های یادگیری برخط برای دسته‌بندی نظرات اسپم استفاده شده است. در این پژوهش با مقایسه دسته‌بندهای مختلفی مانند بیزساده، نزدیک‌ترین همسایه و ماشین بردار پشتیبان نشان داده شده که درصورتی که یادگیری برخط و براساس استفاده تدریجی از نمونه‌های آموزشی باشد، دقت کاهش می‌یابد. همچنین نشان دادند که ترتیب زمانی ارائه نمونه‌ها و نوع کالا یا خدمتی که برای آن نظر نوشته شده باعث تغییر در دقت سامانه می‌شود.

۲-۲- روش‌های بدون نظارت

از آن‌جایی که ایجاد مجموعه داده‌های برچسب‌گذاری شده‌ی دقیق که نظرات اسپم را توصیف می‌کنند سخت است، یادگیری بدون نظارت روشی برای حل این مشکل است؛ چون در این روش نیازی به داده‌های برچسب‌گذاری‌شده نیست. در این روش از شباهت میان نمونه‌ها برای دسته‌بندی و ایجاد خوشه‌ها استفاده می‌شود.

یک مدل بدون نظارت برای تشخیص نظرات اسپم توسط لاو و همکاران [۱۷] معرفی و با روش‌های یادگیری نظارت‌شده مقایسه شد. مدل آن‌ها بر اساس نظرات تکراری[۱۴] کار می‌کرد که به کمک میزان همپوشانی معنایی متن‌ها با استفاده از مدل‌های زبانی-معنایی پیدا می‌شدند.

پژوهش دیگری در [۲۱] و [۲۲] انجام گزارش شده که در آن، روش بدون ناظری برای امتیازدهی نظرات با استفاده از آستانه‌گذاری روی میانگین امتیازات ویژگی‌های مختلف پیشنهاد شد. ویژگی‌های استفاده شده در این پژوهش شامل ویژگی‌های مبتنی بر محتوا مانند شباهت نظر با نظرات نویسنده یا سایر نظرات و ویژگی‌های مبتنی بر نظردهنده مانند تناوب نظردهی نویسنده و تناوب نظرات ارسال شده روی یک کالا است.

در پژوهشی که اخیراً انجام شده [۲۳] از روش‌های خوشه‌بندی برای تشخیص نظرات نامتعارف استفاده شده است. در این پژوهش با استفاده از مدل تبدیل سند به بردار (doc2vec) و روش‌های تحلیل احساسات به خوشه‌بندی نظرات پرداخته شده است.

۲-۳- یادگیری نیمه‌نظارتی

در سایر دامنه‌های استخراج داده، پژوهشگران متوجه شدند که استفاده از نمونه‌های بدون برچسب در ترکیب با مقدار کوچکی از داده‌های برچسب‌گذاری‌شده می‌تواند مفید واقع شود. به عنوان مثال در پژوهشی که توسط لی و همکاران [۱۴] انجام شد، یک روش دو منظوره‌ی نیمه نظارتی در تشخیص نظرات اسپم با استفاده از الگوریتم هم‌آموزش[۱۵] ایجاد گردید تا از تعداد بسیار زیادی نظرات برچسب‌گذاری‌نشده که در دسترس بودند استفاده کند. در این پژوهش ۱۳۹۴ نمونه به‌عنوان اسپم برچسب‌گذاری شد. سپس ویژگی‌های مبتنی بر نظر مثل ویژگی‌های احساسی، محتوایی و فراداده ایجاد شدند. علاوه بر این ویژگی‌ها آن‌ها از ویژگی‌های مبتنی بر نظردهنده مثل ویژگی‌های پروفایلی و رفتاری استفاده کردند.

در مقاله‌ی ارائه‌شده در [۲۴] از روش یادگیری پی‌یو[۱۶] استفاده شد. این نوع یادگیری یک روش تکراری و متناوب است که سعی دارد مجموعه نمونه‌های منفی قابل اعتماد را از میان داده‌های بدون برچسب شناسایی کند. مدل با استفاده از همه‌ی داده‌های بدون برچسب به‌عنوان دسته منفی آموزش داده شده و چنان‌چه نمونه‌ای در دسته مثبت دسته‌بندی شود، دور ریخته می‌شود و تا جایی که شرط توقف برقرار شود، فرایند تکرار می‌گردد. یادگیری پی‌یو توسط نایو بیز به دقتی برابر با ۸۳/۷ درصد دست یافت که در آن مجموعه داده، نمونه‌های مثبت وجود داشت.

یکی دیگر از پژوهش‌های مرتبط که در دسته‌ی نیمه نظارتی قرار می‌گیرد، پژوهش شحنه‌پور و همکاران در سال ۲۰۱۷ [۲۵] است که در آن مسأله‌ی تشخیص اسپم توسط یک شبکه مدل شده و برای دسته‌بندی گراف مربوطه از روش نیمه‌نظارتی و بدون نظارت استفاده شده است. در نهایت ویژگی‌های مبتنی بر رفتار نظر و نویسنده و مبتنی بر ویژگی زبانی نظرات و نویسندگان برای دسته‌بندی نظرات استفاده شده است.

در پژوهش جدیدی که توسط روت و همکاران ارائه شده [۲۶] روش‌های یادگیری نیمه نظارتی با استفاده از ویژگی‌های زبانی، نحوی و احساسی آزموده شده‌اند. برای رفع معایب روش یادگیری نظارت شده از مجموعه داده برچسب خورده‌ی کوچکی برای برچسب زدن کل مجموعه داده استفاده کردند. در جدول ۲ خلاصه‌ای از روش‌های اشاره شده در این بخش آورده شده است.

با وجود پژوهش‌های زیادی که با ویژگی‌های مختلف انجام شده‌است، اما تاکنون پژوهشی در زمینه‌ی ترکیب مختلف ویژگی‌ها و کارایی آنها روی دسته‌بند گزارش نشده است. بالاترین کارایی مشاهده‌شده مربوط به ترکیب ویژگی‌های متنی و غیرمتنی است. ویژگی‌های مبتنی بر نظر نیز در تشخیص دقیق نظرات اسپم مهم هستند. از طرف دیگر همه‌ی کارهای انجام‌شده‌ی قبلی روی

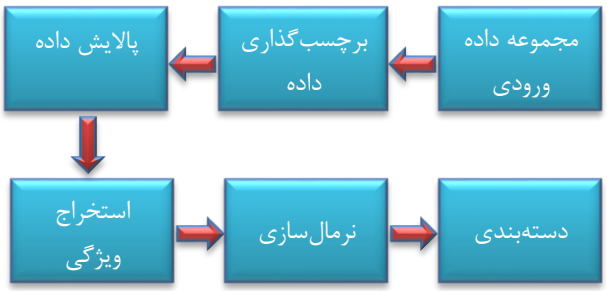
مجموعه داده‌های انگلیسی بوده و مطالعه‌ای در زمینه‌ی نظرات اسپم فارسی گزارش نشده است.

جدول ۲: مقایسه‌ی پژوهش‌های پیشین.

شماره مرجع	سال	موضوع	ویژگی
۹	۲۰۰۸	نظارت‌شده	استفاده از پردازش زبان طبیعی
۱۷	۲۰۱۱	نظارت‌شده	استفاده از مدل‌های زبانی-معنایی
۱۸	۲۰۱۴	نظارت‌شده	به کارگیری دامنه‌ی ترکیبی از مجموعه داده‌ها
۱۹	۲۰۱۵	نظارت‌شده	استفاده از ویژگی‌های زبانی
۲۰	۲۰۱۸	نظارت‌شده	استفاده از روش‌های یادگیری برخط
۱۷	۲۰۱۱	بدون نظارت	استفاده از مدل‌های زبانی-معنایی
۲۱ و ۲۲	۲۰۱۴	بدون نظارت	استفاده از آستانه‌گذاری روی میانگین امتیازات ویژگی‌های مختلف
۲۳	۲۰۱۸	بدون نظارت	استفاده از روش‌های خوشه‌بندی
۱۴	۲۰۱۱	بدون نظارت	استفاده از الگوریتم هم‌آموزش
۲۴	۲۰۱۳	بدون نظارت	استفاده از روش یادگیری پی‌یو
۲۵	۲۰۱۷	بدون نظارت	مدل‌سازی با استفاده از شبکه
۲۶	۲۰۱۷	بدون نظارت	استفاده از ویژگی‌های زبانی، نحوی و احساسی

۳- مدل پیشنهادی

نمای کلی مدل پیشنهادی برای دسته‌بندی نظرات اسپم در شکل ۱ نشان داده شده است.



شکل ۱: نمای کلی مدل پیشنهادی.

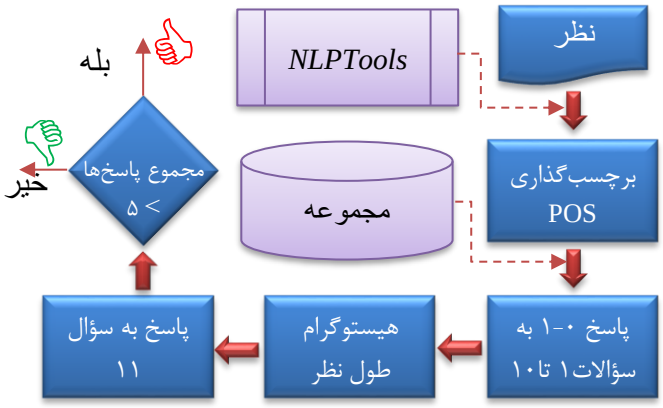
در این پژوهش وب‌سایت دیجی‌کالا [۲۷] به‌عنوان منبع استخراج نظرات در نظر گرفته شده است. نظرات برخط مورد نیاز برای انجام پژوهش، در بازه‌ی زمانی فروردین سال ۱۳۹۵ تا اردیبهشت سال ۱۳۹۶ به مدت ۱۴ ماه استخراج شده است که این بازه‌ی زمانی ۳۰۰۰ نمونه از نظرات نوشته‌شده در مورد گوشی همراه را شامل می‌شود. از میان شرکت‌های تولید گوشی تلفن همراه، پنج شرکت برتر که محبوبیت خاصی در میان طرفداران ایرانی دارند انتخاب شده‌اند و از بین گوشی‌های تولید این شرکت‌ها چندین نمونه‌ی خاص و معمولی برگزیده شده است. نمونه‌ای از نظرات موجود در این مجموعه داده در جدول ۳ نشان داده است.

جدول ۳: نمونه‌ای از داده‌های خام استخراج‌شده از وب‌سایت دیجی‌کالا.

ویژگی	مقدار
شناسه (ID)	کاربر فرضی ۱
عنوان نظر	فوق العادس
متن نظر	گوشی بسیار عالییه. من تا حالا هنگ یا لگ ازش ندیدم. فوق العادس. واسه من این گوشی هیچوقت قدیمی نمیشه.
تعداد بازخورد مثبت	۱۳
تعداد بازخورد منفی	۷
امتیاز کالا	۳/۹
امتیاز کاربر	۵
شرکت سازنده کالا	Apple iPhone
مدل	5S
تاریخ	۱۳۹۵/۰۱/۲۶

۳-۱- برچسب‌گذاری داده

به این دلیل که استفاده از یادگیری نظارت‌شده در تشخیص اسپم دقت بالاتری دارد [۷]، [۲۸] و [۲۹]، در پژوهش حاضر این نوع یادگیری مورد استفاده قرار گرفته است. برای استفاده از یادگیری نظارت‌شده، به داده‌های دارای برچسب نیاز است. همان‌طور که در مقدمه بیان شد، یکی از چالش‌های پیش روی پژوهشگران در تشخیص نظرات اسپم، برچسب‌گذاری نمونه‌های اسپم و غیراسپم است؛ زیرا تشخیص این نوع نظرات بسیار سخت بوده و به اطلاعات تخصصی زیادی در این زمینه نیاز است. برای این منظور از روشی که فلوچارت فرآیند برچسب‌زنی نظرها براساس آن در شکل ۲ آورده شده است، استفاده می‌گردد.



شکل ۲: فلوچارت برچسب‌زنی نظرها به دو دسته اسپم و غیراسپم.

برای برچسب‌زنی نظرات یازده سؤال که در پژوهش [۱۱] پیشنهاد شده است، استفاده شدند. این سؤالات در جدول ۴ آورده شده‌اند.

جدول ۴: پرسش‌های مطرح‌شده برای برچسب‌گذاری نمونه‌ها.

شماره سؤال	متن سؤال
۱	آیا در نظر از ویژگی‌های عمومی کالا که در توصیف کالا نیز می‌آید استفاده شده است یا ویژگی‌های خاص کالا که هنگام استفاده مشخص می‌شوند؟
۲	آیا نظر، غیر مرتبط با کالا است؟
۳	آیا نظر پر از صفات بی‌ربط و کلمات نامربوط است؟
۴	آیا فهم و درک نظر سخت است؟
۵	آیا در این نظر نام کالا بارها و بارها تکرار شده‌است؟
۶	آیا در این نظر از آدرس رایانامه، شماره تلفن، آدرس وب‌سایت و کد خاصی استفاده شده‌است؟
۷	آیا در این نظر تعداد ضمایر شخصی و افعال بیش از اسم، صفت و حروف اضافه است؟
۸	آیا در این نظر از کارخانه‌ی سازنده‌ی کالا نام برده شده‌است؟
۹	آیا این نظر نوعی بازاریابی است؟
۱۰	آیا در این نظر دعوای رقابتی وجود دارد؟
۱۱	آیا این نظر کوتاه است؟

در این پژوهش هر یک از این سوال‌ها از تک تک نظرات پرسیده شده و با صفر یا یک مقداردهی می‌شوند. برای درک بهتر این پاسخ‌ها در زیر چگونگی برچسب‌زدن نمونه‌ها آورده شده است:

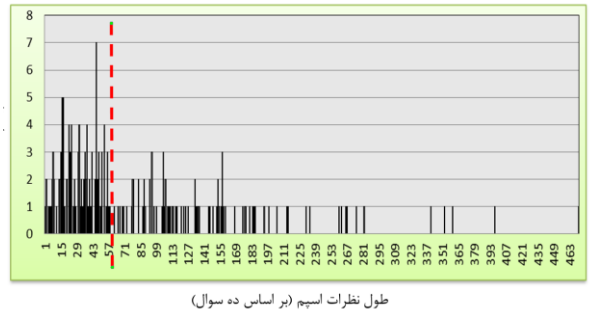
- یکی از قابلیت‌های خوب دیجی‌کالا این است که اگر فردی کالایی را از این وب‌سایت خریداری کرده باشد و نظری را درج کند، در کنار نام این شخص عبارت «خریدار این محصول» درج می‌شود. به این ترتیب می‌توان کسانی را که واقعاً خریدار کالا بوده و در مورد آن به نوشتن نظر مبادرت می‌ورزند شناسایی کرد. برای آنکه دقیقاً مشخص شود که فرد از کالا استفاده کرده‌است، پاسخ اولین پرسش برای نظراتی که نویسنده‌ی آن‌ها خریدار کالا بوده است، با عدد صفر مشخص گردیده است. علت این امر آن است که فردی که کالا را خریداری کرده، قادر خواهد بود اطلاعات تخصصی‌تری در مورد کالای خریداری‌شده داشته باشد.
- در تمام حالت‌هایی که نظر صددرصد به کالای مورد نظر ربط ندارد و همچنین حالتی که بخش اعظم نظر بی‌ربط به محصول باشد، آن نمونه به عنوان بی‌ربط به کالا در نظر گرفته شده و به پرسش شماره‌ی ۲، عدد یک اختصاص داده شده‌است. در این نوع از نظرات، محوریت نظر حول موضوع دیگری است.
- با توجه به اینکه نظرات نوشته‌شده در مورد گوشی تلفن همراه شرکت‌های اِل‌جی، هوآوی، اِیسوس، سامسونگ و اَپِل هستند، در صورت وجود نام این شرکت‌ها به زبان‌های فارسی و انگلیسی و نحوه نگارش گوناگون آن‌ها در پاسخ به این پرسش عدد یک درج می‌شود. در برخی از نظرات از کلمه‌ی آی‌فون که نوعی از گوشی‌های اَپِل می‌باشد، استفاده شده است و واضح است که هدف از بیان آی‌فون طرفداری و جانبداری از شرکت اَپِل بوده و گوشی‌های آی‌فون را در شرکت اَپِل مورد مدح و ستایش قرار داده‌اند. به بیان دیگر به جای استفاده از کلمه‌ی اَپِل از کلمه‌ی آی‌فون استفاده نموده‌اند. در چنین مواردی پاسخ پرسش شماره‌ی ۸، عدد یک در نظر گرفته شده‌است.
- در نظراتی که عنوان شده گوشی عالی، خوب یا فوق‌العاده است و حتماً بخرید یا به شما پیشنهاد می‌دهم خریداری کنید، سوال ۹ با عدد یک مقداردهی شده است. برعکس آن هم صادق است؛ مثلاً تلفن همراه مورد نظر با گوشی‌هایی از شرکت یا شرکت‌های دیگر مقایسه می‌شود و در نظر بیان می‌شود که اصلاً گوشی خوبی نیست، یا اصلاً نخرید و یا به جای این گوشی، گوشی مدل «الف» از شرکت «ب» یا گوشی‌های شرکت «ب» را خریداری کنید. این نوع از نظرات به نوعی برای شرکت‌های دیگر بازاریابی کرده و شرکت مورد نظر و آن محصول خاص را در نظر مردم خراب می‌کنند، بدون این‌که دلیلی قانع‌کننده داشته باشند.

۲-۳- برچسب‌گذاری اجزاء سخن

برای استخراج اجزاء سخن نظرها، باید در ابتدا نوع هر کلمه‌ی نظر از لحاظ دستور زبان مشخص شود. همان‌طور که شکل ۲ اشاره شد،

برای حل این مساله، از ابزار *NLPTools* استفاده شده است [۳۰]. از آن‌جایی که برای استخراج اجزاء سخن توسط این ابزار باید نظرات صاف و یک‌دست شده باشند تا تشخیص نوع کلمات به‌آسانی انجام گیرد، ابتدا پیش‌پردازش نمونه‌ها انجام شد پس از پیش‌پردازش نمونه‌ها، نظرات وارد ابزار شده و نوع هر کلمه‌ی درون نظر استخراج شده است. خروجی این مرحله وارد قطعه کدی می‌شود که در آن تعداد صفات، اسم‌ها، حروف اضافه، فعل‌ها و ضمیرهای شخصی در هر نظر شمارش می‌شود.

برای برچسب‌گذاری نظرات ابتدا پرسش‌های شماره ۱ تا ۱۰ موجود در جدول ۴ پاسخ داده شده و در مرحله‌ی بعد مجموع پاسخ‌های داده‌شده محاسبه گردیدند. برای تشخیص کوتاه بودن یا بلند بودن نظر ابتدا نیاز است تعداد کلمات هر نظر شمارش شود. شمارش تعداد کلمات با استفاده از تشخیص فاصله میان آن‌ها با استفاده از *NLPTools* انجام شده و تعداد کلمات هر نظر به‌دست آمد. بر اساس تعداد کلمات به‌دست آمده و برچسب اولیه‌ی نظرات اسپم، نمودار فراوانی نظرات براساس طول (شکل ۳) آن‌ها ترسیم شده و با توجه به فراوانی داده‌ها، آستانه ۵۹ کلمه برای تشخیص کوتاه بودن نظر از این نمودار انتخاب شد.

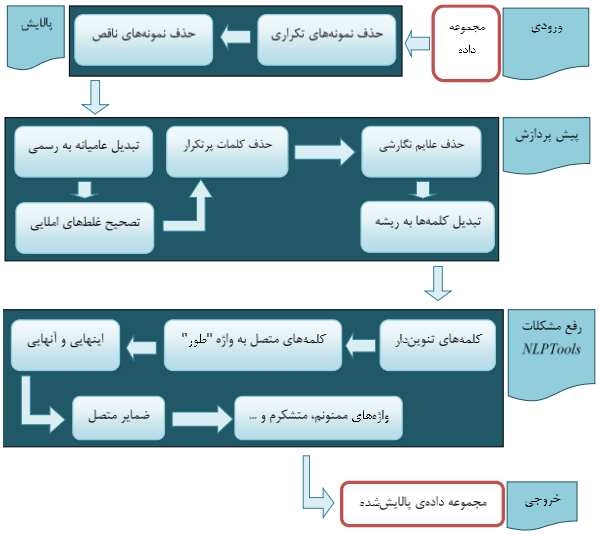


شکل ۳: نمودار تشخیص حد تعداد کلمات در شناسایی نظرات اسپم. سمت چپ خط چین قرمز نظرات کوتاه درنظر گرفته می‌شود.

پس از تعیین آستانه برای تعداد کلمات نظر، در صورت کوتاه بودن عدد یک و در غیر این صورت مقدار اختصاص داده شده به این پرسش صفر خواهد بود. با مشخص شدن پاسخ پرسش شماره ۱۱، می‌توان بر اساس مقاله‌ی [۱۱] برچسب نهایی نظر را مشخص کرد. اگر مجموع پاسخ‌های داده‌شده به پرسش‌ها، بزرگ‌تر یا مساوی با عدد ۵ باشد، آن نظر در نهایت برچسب اسپم می‌خورد و درصورتی‌که این مقدار از ۵ کمتر باشد، برچسب نهایی نظر غیراسپم خواهد بود. در رابطه با اسپم نوع سه که غیرنظرها هستند باید خاطرنشان کرد که وب‌سایت دیجی‌کالا این نوع از نظرات را به‌صورت خودکار حذف می‌کند و امکان تشخیص آن‌ها در پژوهش حاضر وجود ندارد.

۳-۳- پالایش داده

قبل از ورود داده‌ها به دسته‌بند، برای گرفتن نتایج بهتر باید روی نمونه‌ها پیش‌پردازش انجام داد [۳۱]. فرآیند انجام پالایش و پیش‌پردازش نمونه‌ها در شکل ۴ آورده شده است.



شکل ۴: فرآیند پالایش مجموعه داده.

مراحل تبدیل کلمات عامیانه به رسمی، حذف کلمات پرتکرار و تبدیل کلمات به ریشه‌ی آن‌ها توسط ابزار *NLPTools* انجام شد و تصحیح غلط‌های املایی با استفاده از نرم‌افزار مایکروسافت وُرد. برای حذف علایم نگارشی نیز با توجه به عدم توانایی ابزار مذکور کد ساده‌ای به زبان پایتون نوشته شد.

لازم به ذکر است که اگرچه از روش یادگیری ماشین با ناظر در این پژوهش استفاده شده است، تبدیل کلمات عامیانه به رسمی ضروری است. دلیل این امر این است که در قسمت‌های بعدی مانند «حذف کلمات پرتکرار» یا «تبدیل کلمات به ریشه‌ی آن‌ها» از کلمات رسمی استفاده شده است.

۴-۳- برطرف کردن کاستی‌های ابزار *NLPTools*

در خلال کار با ابزار *NLPTools* برخی کاستی‌ها در انجام مرحله‌ی پیش‌پردازش مشخص شد که در ادامه به آن‌ها اشاره خواهد شد.

- در تشخیص اجزاء سخن، برخی واژه‌های در نظر گرفته‌شده در واقع صفت هستند اما ابزار آن‌ها را به‌عنوان اسم تشخیص می‌دهد.
- در مورد ضمایر چسبان، چون این ضمیر به‌صورت چسبیده به کلمه استفاده شده است، ابزار آن را تشخیص نداده و کل کلمه را به‌صورت اسم در نظر می‌گیرد.
- برخی از کلمات که قید هستند به صورت اسم نمایش

داده می‌شوند.

- کلماتی که دارای تنوین در انتها بوده و نقش قیدی دارند اگر چنان‌چه در متن نظر بدون تنوین آمده باشند، اسم نمایش داده می‌شوند.
- کلمات چسبیده به کلمه‌ی «طور» به‌صورت اسم نام‌گذاری شدند.
- نقش دو کلمه‌ی «این‌هایی» و «آن‌هایی» اسم نمایش داده می‌شوند، درصورتی‌که با حذف «یی» از انتهای آن‌ها می‌توان برچسب درست ضمیر را به آن‌ها اختصاص داد.
- ضمایر متصل «م، ت، ش، مان، تان، شان» در انتهای اسم‌ها به عنوان ضمیر در نظر گرفته است.
- واژه‌هایی مانند «ممنونم، متشکرم، متاسفم و امیدوارم» به‌صورت اسم نمایش داده می‌شوند که با تفکیک آن‌ها به شکل ممنون هستم، متشکر هستم، متاسف هستم و امیدوار هستم آن‌ها به صفت+فعل تبدیل می‌شوند.

برای برطرف کردن موارد اشاره شده روش مبتنی بر قانون[۱۷] انتخاب شده و مجموعه قوانینی برای تبدیل موارد اشتباه به صحیح تهیه شد. در ادامه در زبان پایتون این قوانین پیاده‌سازی شدند. دو نمونه از کدهای مربوط به اصلاح مشکلات بالا به زبان پایتون در شکل ۵ و ۶ آورده شده است.

```
list_of_adverbs = {'کاملاً':'کاملا' ,'دقیقاً':'دقیقا'
{'قبلاً':'قبلا' ,'اتفاقاً':'اتفاقا','حتماً':'حتما' ,
for adv in list_of_adverbs:
    if adv in review:
        review = review.replace(adv,list_of_adverbs[adv])
```

شکل ۵: کد اضافه کردن تنوین به انتهای قیود دارای تنوین در انتهای کلمه. فرض شده لیستی از قیدهای دارای تنوین در دیکشنری list_of_adverbs قرار دارد.

```
for word in review.split():
    if len(word) > 3 and 'طور' in word:
        w = word[:word.index('طور')]
        + ' ' + word[word.index('طور'):]
        review = review.replace(word, w)
```

شکل ۶: کد ایجاد فاصله بین کلمات دارای «طور» برای برطرف کردن مشکل برچسب‌گذاری ابزار NLPTools.

ا

ستخراج ویژگی

در این پژوهش از ویژگی‌های مبتنی بر نظر و فراداده استفاده شده است. ویژگی‌های مبتنی بر نظر که در این پژوهش از آن‌ها استفاده شده است نیز عبارتند از:

- کلمات احساسی و امتیاز آن‌ها،
- کلمات پرتکرار و تعداد آن‌ها در نظر،
- نقش کلمات نظر و تعداد هر یک از آن‌ها.

برای تشخیص کلمات احساسی درون متن نظرات، از واژه‌نامه‌ای که توسط بصیری و کبیری [۳۲], [۳۳] استخراج شده است استفاده و به‌روزرسانی شد. بعد از بررسی، بسیاری از کلمات احساسیِ درون نظرات جمع‌آوری‌شده به واژه‌نامه افزوده و امتیازدهی شدند. سپس تعداد زیادی از کلمات موجود در این واژه‌نامه حذف گردیدند؛ زیرا با دامنه‌ی مورد استفاده در نظرات جمع‌آوری‌شده سازگار نبودند. نمونه‌ای از واژه‌نامه مذکور در جدول ۵ آورده شده است.

برای شناسایی کلمات پرتکرار واژه‌نامه‌ای که ابزار *NLPTools* از آن بهره برده است استفاده شد [۳۰]. بر اساس این واژه‌نامه کلمات پرتکرار درون نظرها شناسایی شده و تعداد آن‌ها محاسبه می‌شود. تشخیص نقش کلمات هر نظر نیز به کمک همین ابزار انجام می‌شود.

ویژگی‌های فراداده استفاده شده در این پژوهش شامل تعداد بازخوردهای مثبت، تعداد بازخوردهای منفی، طول عنوان نظر، طول نظر، امتیاز کاربر به کالا، امتیاز کلی کالا و میزان اختلاف امتیاز کلی کالا و امتیاز کاربر است.

جدول ۵: نمونه‌ای از واژه‌نامه احساسی ارائه شده در [۳۲] و [۳۳]. امتیازات در بازه ۱ تا ۵ از منفی به مثبت نوشته شده‌اند.

کلمه	امتیاز
اشتباه	۱
گمراه	۱
نامربوط	۲
گرفتار	۲
جدی	۳
ظاهری	۳
زیبا	۴
مطلوب	۴
عاشق	۵
عالی	۵

۵-۳- نرمال‌سازی

هر یک از ویژگی‌ها با مقادیری در بازه‌های گوناگون مقداردهی

شده‌اند. پس از استخراج تمامی ویژگی‌ها، آن‌ها با استفاده از نرم‌افزار وکا در بازه‌ی [۰و ۱] نرمال‌سازی شده‌اند تا داده‌های ورودی به دسته‌بند یک‌دست شوند. دلیل این کار این است که مقیاس ویژگی‌ها یکسان شده و الگوریتم‌های یادگیری نظارت شده بر روی ویژگی‌ها به درستی آموزش ببینند.

۶-۳- دسته‌بندی

آخرین مرحله از انجام پژوهش (دسته‌بندی)، از سه دسته‌بند معروف استفاده شده است. دسته‌بندهای مورد استفاده در انجام پژوهش شامل دسته‌بند نایو بیز، دسته‌بند ماشین بردار پشتیبان و درخت تصمیم است. علت انتخاب این است که این دسته‌بندها در شناسایی نظرات اسپم در پژوهش‌های پیشین کارایی خوبی را از خود نشان داده‌اند [۷]، [۳۴] و [۳۵].

در این پژوهش، عمل دسته‌بندی توسط نرم‌افزار وکا انجام شده است و تأثیر ویژگی‌های مبتنی بر نظر و ویژگی‌های فراداده به‌صورت جداگانه و نیز ترکیب آن‌ها با هم روی هر دسته‌بند مورد بررسی قرار گرفته‌است. از روش اعتبارسنجی مورب ده‌گانه برای دسته‌بندی استفاده می‌شود؛ زیرا بیش‌انطباق[۱۸] دسته‌بند روی نمونه‌ها را به میزان قابل توجهی کاهش می‌دهد.

برای دسته‌بند ماشین بردار پشتیبان از پیاده سازی الگوریتم بهینه‌سازی کمینه[۱۹] (SMO) با تابع کرنل چندجمله‌ای استفاده شده است. برای دسته‌بند درخت تصمیم از پیاده‌سازی الگوریتم C4 در تابع J48 ابزار وکا استفاده شده است. همچنین از پیاده‌سازی پیش‌فرض مربوط به الگوریتم نایو بیز در ابزار وکا برای این دسته‌بند استفاده شده است.

پس از آماده‌سازی ورودی‌های دسته‌بند، دسته‌بندی انجام می‌شود. برای انجام این پژوهش نمونه‌ها دو بار دسته‌بندی می‌شوند. ابتدا با تعداد ۳۰۰۰ نمونه، دسته‌بندی انجام می‌شود. از میان ۳۰۰۰ نمونه‌ی ورودی به دسته‌بند، تعداد ۱۹۴ نمونه به‌عنوان اسپم نوع یک برچسب خورده و بقیه نمونه‌ها غیراسپم هستند. از آن‌جایی که وزن نمونه‌های غیراسپم بسیار بالاتر از نمونه‌های اسپم است و توازن میان آن‌ها وجود ندارد، ابتدا نمونه‌ها توسط نرم‌افزار وکا وزن‌دار می‌شوند و سپس به دسته‌بند وارد می‌گردند. این کار با استفاده از تابع classBalancer در ابزار وکا انجام می‌شود. این تابع نمونه‌ها را به گونه‌ای وزندهی می کند که هر دسته در نهایت وزن کلی یکسانی با دسته‌ی دیگر داشته باشد. در مرحله‌ی دوم، دسته‌بندی با تعداد ۳۸۸ نمونه انجام می‌پذیرد؛ به این ترتیب که ۱۹۴ نمونه به‌عنوان اسپم نوع یک و ۱۹۴ نمونه از نوع غیراسپم برداشته می‌شود.

۴- نتایج

نتایج در دو بخش دسته‌بندی اسپم و بررسی ترکیبات مختلف ویژگی‌ها ارائه خواهد شد.

۴-۱- معیار سنجش

در این پژوهش برای سنجش دقت دسته‌بند معیارهای اِف وان (F1-measure)، دقت (Precision) و بازخوانی (Recall) در نظر گرفته شده است؛ این معیارهای کارایی توسط رابطه‌های ۱ تا ۳ مشخص می‌شوند. در این روابط TP تعداد نمونه‌های مثبتی است که توسط دسته‌بند به‌درستی در دسته‌ی مثبت قرار گرفته و FP تعداد نمونه‌های منفی است که به‌صورت اشتباه در دسته مثبت قرار گرفته‌اند. FN نیز تعداد نمونه‌هایی هستند که در واقع مثبت بوده‌اند اما دسته‌بند آن‌ها را به اشتباه در دسته منفی قرار داده است.

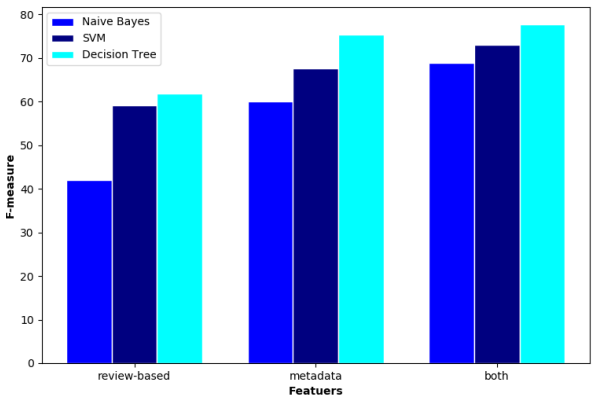
$$F-measure = \frac{2\times\pi\times\rho}{\pi+\rho} \quad (۱)$$

$$\pi = \frac{TP}{TP+FP} \quad (۲)$$

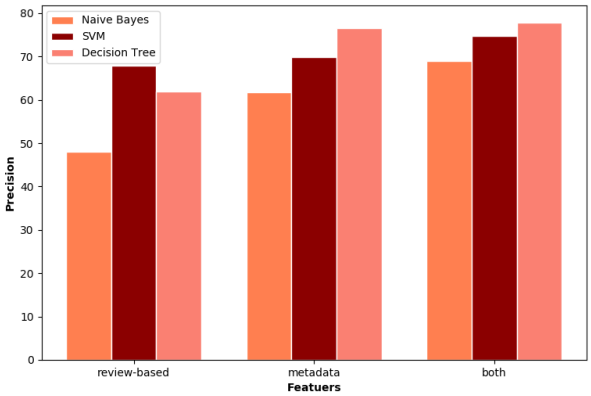
$$\rho = \frac{TP}{TP+FN} \quad (۳)$$

۴-۲- دسته‌بندی اسپم نوع ۱

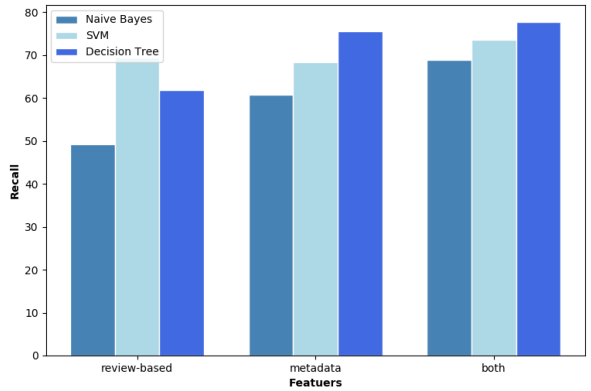
نتایج حاصل از دسته‌بندی با استفاده از سه نوع ویژگی مبتنی بر نظر، ویژگی‌های فراداده و ترکیب هر دو نوع ویژگی در شکل‌های ۷ تا ۱۲ آورده شده است.



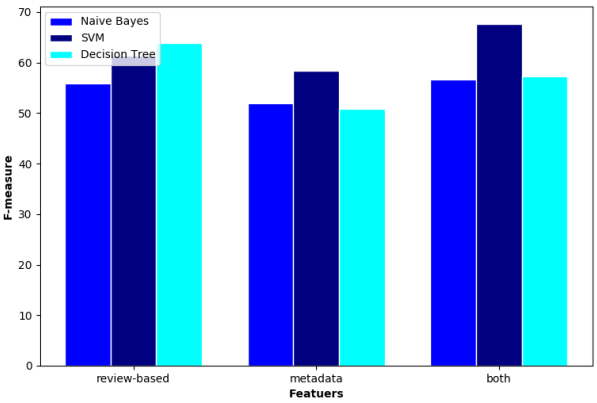
شکل ۷: کارایی الگوریتم‌های دسته‌بندی با ویژگی‌های مبتنی بر نظر (review-based)، فراداده (meta-data) و ترکیب آن‌ها (both) روی داده متوازن براساس معیار F1.



شکل ۸: کارایی الگوریتم‌های دسته‌بندی با ویژگی‌های مبتنی بر نظر (review-based)، فراداده (meta-data) و ترکیب آن‌ها (both) روی داده متوازن براساس معیار دقت (Precision).

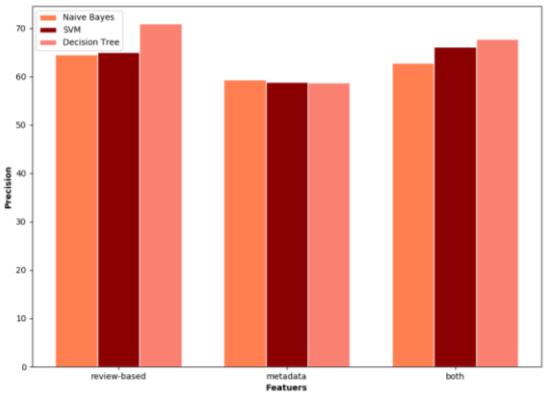


شکل ۹: کارایی الگوریتم‌های دسته‌بندی با ویژگی‌های مبتنی بر نظر (review-based)، فراداده (meta-data) و ترکیب آن‌ها (both) روی داده متوازن براساس معیار بازخوانی (Recall).

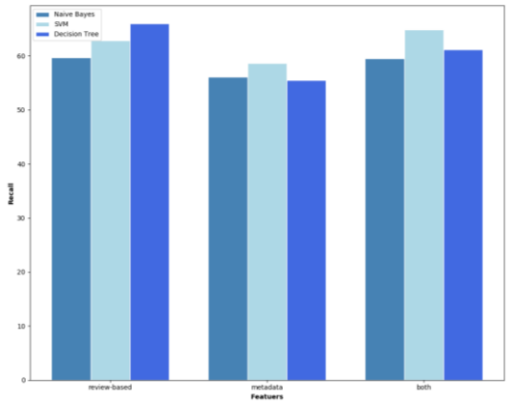


شکل ۱۰: کارایی الگوریتم‌های دسته‌بندی با ویژگی‌های مبتنی بر نظر (review-based)، فراداده (meta-data) و ترکیب آن‌ها (both) روی داده نامتوازن براساس معیار F1.

نتایج نشان می‌دهند که دسته‌بند نایو بیز روی ویژگی‌های فراداده نسبتاً بهتر از ویژگی‌های مبتنی بر نظر عمل می‌کند و ترکیب ویژگی‌های مبتنی بر نظر و فراداده دقت بالاتری را رقم می‌زند.



شکل ۱۱: کارایی الگوریتم‌های دسته‌بندی با ویژگی‌های مبتنی بر نظر (review-based)، فراداده (meta-data) و ترکیب آن‌ها (both) روی داده نامتوازن براساس معیار دقت (Precision).



شکل ۱۲: کارایی الگوریتم‌های دسته‌بندی با ویژگی‌های مبتنی بر نظر (review-based)، فراداده (meta-data) و ترکیب آن‌ها (both) روی داده نامتوازن براساس معیار بازخوانی (Recall).

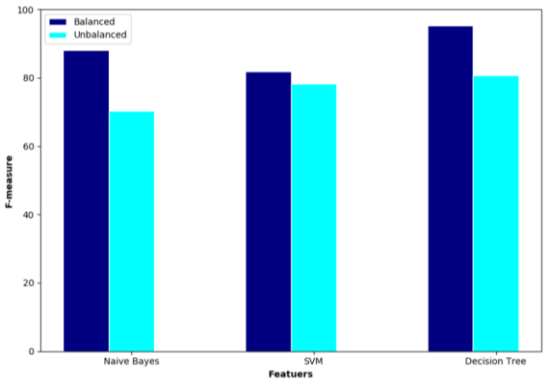
علاوه براین می‌توان دید که همواره ترکیب ویژگی‌ها منجر به افزایش دقت دسته‌بندها می‌شود. مشاهده می‌شود که به‌طور متوسط دسته‌بند ماشین بردار پشتیبان روی نمونه‌های نامتوازن و ویژگی‌های فراداده و نیز ترکیب ویژگی‌ها نسبتاً خوب عمل می‌کند.

۳-۴- دسته‌بندی اسپم نوع ۲

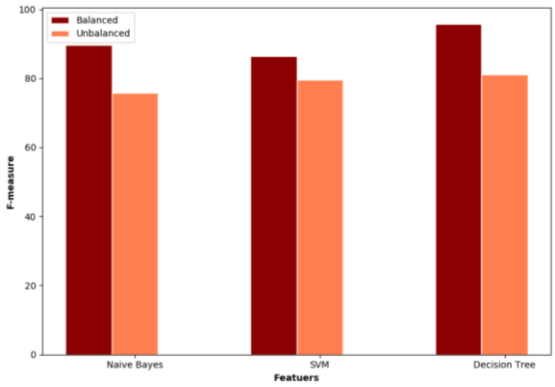
همان‌گونه که در بخش‌های قبلی گفته‌شد، نوع دوم نظرات اسپم آن‌هایی هستند که در مورد کالایی خاص سخنی به میان نیاورده و تنها از کارخانه‌ی سازنده‌ی کالا صحبت می‌کنند. مثلاً نظری مثل «من از اپل بدم می‌آید. اپل تنها پول برند خود را می‌گیرد و هیچ نوآوری در زمینه‌ی گوشی‌هایش ندارد» نمونه‌ای از این نوع اسپم است. مشاهده می‌شود که در این نظر هیچ کالایی هدف قرار نگرفته است و به‌طور کلی در مورد کارخانه‌ی سازنده‌ی کالا صحبت می‌شود.

همانند تشخیص اسپم نوع یک، در شناسایی این نوع اسپم نیز از نمونه‌های متوازن و نامتوازن استفاده شده و دقت آن‌ها روی

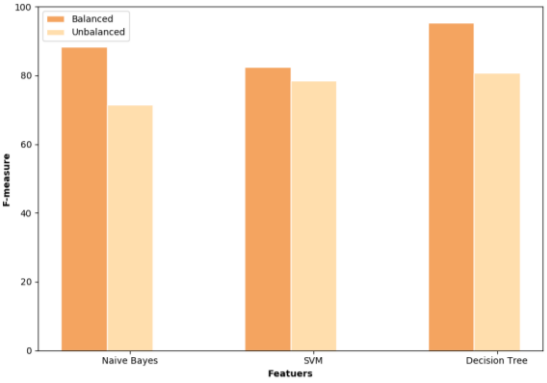
دسته‌بندهای مختلف سنجیده می‌شود. نتایج حاصل از دسته‌بندی در شکل‌های ۱۳ تا ۱۵ آورده شده است.



شکل ۱۳: مقایسه الگوریتم‌ها برای تشخیص اسپم نوع ۲ براساس معیار F1.



شکل ۱۴: مقایسه الگوریتم‌ها برای تشخیص اسپم نوع ۲ براساس معیار دقت (Precision).



شکل ۱۵: مقایسه الگوریتم‌ها برای تشخیص اسپم نوع ۲ براساس معیار بازخوانی (Recall).

همان‌گونه که در نتایج مربوط به دسته‌بندی اسپم نوع دوم مشخص است، درخت تصمیم بهترین کارایی را دارد. همچنین مشخص است که کارایی الگوریتم‌ها روی مجموعه داده متوازن به مراتب بهتر از داده‌ی نامتوازن است.

۴-۴- بررسی ترکیب‌های مختلف ویژگی‌ها

در این بخش ترکیب‌های مختلف زیرمجموعه‌ی ویژگی‌های مبتنی بر نظر و فراداده در افزایش یا کاهش کارایی دسته‌بندها مورد تحلیل و بازبینی قرار می‌گیرند تا بهترین ویژگی که نقشی متمایزکننده در تشخیص نظرات اسپم فارسی دارد و ویژگی که می‌تواند منجر به کاهش کارایی دسته‌بندها گردد مشخص شود. مقایسه کارایی ترکیب‌های مختلف ویژگی‌های مبتنی بر نظر در جداول ۶ و ۷ و ترکیب‌های مختلف ویژگی‌های مبتنی بر فراداده در جداول ۸ و ۹ آورده شده است.

جدول ۶: مقایسه کارایی ترکیب‌های ویژگی‌های مبتنی بر نظر در مجموعه داده متوازن. مقادیر به صورت میانگین وزن‌دار دسته‌ها محاسبه شده‌اند.

نام دسته‌بند	نام ویژگی‌ها	اِف وان	دقت	بازخوانی
NB	اجزاء سخن	۰/۴۰۰	۰/۴۸۶	۰/۴۹۵
SVM		۰/۵۵۶	۰/۵۷۰	۰/۵۶۴
DT		۰/۵۴۹	۰/۵۴۹	۰/۵۴۹
NB	کلمه‌های پرتکرار	۰/۴۱۰	۰/۵۱۵	۰/۵۰۵
SVM		۰/۴۶۵	۰/۵۰۷	۰/۵۰۵
DT		۰/۴۸۰	۰/۴۸۲	۰/۴۸۲
NB	کلمه‌های پرتکرار+ اجزاء سخن	۰/۴۰۸	۰/۵۰۷	۰/۵۰۳
SVM		۰/۵۱۲	۰/۵۵۹	۰/۵۴۴
DT		۰/۵۳۹	۰/۵۳۹	۰/۵۳۹
NB	احساسی	۰/۵۹۷	۰/۶۷۹	۰/۶۲۴
SVM		۰/۵۹۷	۰/۶۷۹	۰/۶۲۴
DT		۰/۵۹۲	۰/۶۷۹	۰/۶۲۴
NB	احساسی + اجزاء سخن	۰/۵۱۴	۰/۵۴۰	۰/۵۳۴
SVM		۰/۵۹۲	۰/۶۷۹	۰/۶۲۴
DT		۰/۶۲۰	۰/۶۲۳	۰/۶۲۱
NB	کلمه‌های پرتکرار+ احساسی	۰/۵۱۴	۰/۴۸۶	۰/۴۹۵
SVM		۰/۵۹۲	۰/۵۷۰	۰/۵۶۴
DT		۰/۵۹۲	۰/۵۴۹	۰/۵۴۹
NB	کلمه‌های پرتکرار + احساسی + اجزاء سخن	۰/۴۲۲	۰/۴۸۵	۰/۴۹۲
SVM		۰/۵۹۲	۰/۶۷۹	۰/۶۲۴
DT		۰/۶۱۸	۰/۶۲۰	۰/۶۱۹

جدول ۷: مقایسه کارایی ترکیب‌های ویژگی‌های مبتنی بر نظر در مجموعه داده نامتوازن. مقادیر به صورت میانگین وزن‌دار دسته‌ها محاسبه شده‌اند.

نام دسته‌بند	نام ویژگی‌ها	اِف وان	دقت	بازخوانی
NB	اجزاء سخن	۰/۵۲۷	۰/۶۲۹	۰/۵۷۶
SVM		۰/۵۴۶	۰/۶۴۷	۰/۵۹۰
DT		۰/۶۰۹	۰/۶۷۰	۰/۶۳۱
NB	کلمه‌های پرتکرار	۰/۵۱۸	۰/۶۳۱	۰/۵۷۲
SVM		۰/۴۸۳	۰/۶۱۸	۰/۵۵۳
DT		۰/۵۷۰	۰/۵۷۹	۰/۵۷۵
NB	کلمه‌های پرتکرار+ اجزاء سخن	۰/۵۲۹	۰/۶۴۰	۰/۵۸۰
SVM		۰/۵۴۶	۰/۶۴۳	۰/۵۸۹
DT		۰/۶۰۲	۰/۶۷۹	۰/۶۳۰
NB	احساسی	۰/۵۳۲	۰/۶۳۷	۰/۵۸۱
SVM		۰/۵۵۳	۰/۵۵۸	۰/۵۵۶
DT		۰/۵۳۲	۰/۶۳۷	۰/۵۸۱
NB	احساسی + اجزاء سخن	۰/۵۵۹	۰/۶۴۹	۰/۵۹۷
SVM		۰/۶۱۲	۰/۶۴۹	۰/۶۲۷
DT		۰/۶۵۸	۰/۷۰۹	۰/۶۷۳
NB	کلمه‌های پرتکرار+ احساسی	۰/۵۹۲	۰/۶۳۷	۰/۶۱۱
SVM		۰/۵۷۴	۰/۶۲۷	۰/۵۹۸
DT		۰/۶۲۴	۰/۶۳۱	۰/۶۲۷
NB	کلمه‌های پرتکرار+ احساسی + اجزاء سخن	۰/۵۵۸	۰/۶۴۸	۰/۵۹۷
SVM		۰/۶۱۳	۰/۶۵۰	۰/۶۲۸
DT		۰/۶۳۹	۰/۷۱۰	۰/۶۶۰

با بررسی جداول ۶ و۷، مشاهده می‌شود که ویژگی‌های قطبیت نظر[۲۰] (احساسی) و اجزاء سخن کلمه‌های درون آن کارایی بالایی دارند. همچنین، ترکیباتی که ویژگی‌های احساسی در آن‌ها حضور دارد نیز کارایی قابل قبولی از خود نشان می‌دهند. درصورتی که نیاز باشد تنها یک نوع ویژگی انتخاب شود، می‌توان از جداول ۶ و ۷ ویژگی احساسی را انتخاب کرد که بازدهی بالاتری نسبت به ویژگی‌های اجزاء سخن و کلمه‌های پرتکرار دارد.

با بررسی جداول ۸ و ۹ مشخص می‌شود که از میان ویژگی‌های مبتنی بر فراداده امتیاز کلی محصول، تعداد بازخوردهای مثبت نظر، تعداد بازخوردهای منفی و طول نظر تأثیر به‌سزایی در تشخیص نظرات اسپم نوع یک دارند. بعد از این بررسی ویژگی‌های فراداده و مبتنی بر نظر بر اساس کاراییشان روی دسته‌بندها رتبه‌بندی می‌شوند و از میان ترکیب‌های برتر، بهترین ویژگی‌ها مورد شناسایی قرار می‌گیرند. با استخراج بهترین ویژگی‌ها در شناسایی اسپم نوع یک، تأثیر تک‌تک زیرمجموعه‌های آن‌ها روی دسته‌بندها بازبینی می‌شود. جداول ۱۰ و ۱۱ ترکیب این ویژگی‌ها و میزان دقتشان در سه دسته‌بند را نشان می‌دهند.

جدول ۸: نمونه‌هایی از ترکیب‌های ویژگی‌های مبتنی فراداده در مجموعه داده متوازن. مقادیر به صورت میانگین وزن‌دار دسته‌ها محاسبه شده‌اند.

نام دسته‌بند	نام ویژگی‌ها	اِف وان	دقت	بازخوانی
NB	امتیاز کاربر	۰/۵۲۷	۰/۵۴۷	۰/۵۴۱
SVM		۰/۵۲۵	۰/۵۵۶	۰/۵۴۶
DT		۰/۴۸۴	۰/۴۸۵	۰/۴۸۵
NB	تعداد بازخوردهای مثبت	۰/۴۶۴	۰/۵۱۱	۰/۵۰۸
SVM		۰/۴۴۹	۰/۵۰۴	۰/۵۰۳
DT		۰/۴۷۳	۰/۶۲۸	۰/۵۵۲
NB	بازخوردهای مثبت+امتیاز کلی کالا	۰/۵۷۶	۰/۶۱۱	۰/۵۹۳
SVM		۰/۵۲۲	۰/۵۹۹	۰/۵۶۴
DT		۰/۷۵۲	۰/۷۵۴	۰/۷۵۳
NB	طول نظر+ طول عنوان نظر	۰/۴۵۲	۰/۵۳۵	۰/۵۱۸
SVM		۰/۴۹۴	۰/۵۳۰	۰/۵۲۳
DT		۰/۵۱۹	۰/۵۳۱	۰/۵۲۸
NB	بازخوردهای منفی+طول نظر+امتیاز کلی کالا	۰/۵۸۱	۰/۴۱۷	۰/۵۹۸
SVM		۰/۵۲۲	۰/۶۳۴	۰/۵۷۵
DT		۰/۷۴۷	۰/۷۵۰	۰/۷۴۷
NB	اختلاف امتیاز کالا و کاربر+بازخوردهای مثبت+بازخوردهای منفی	۰/۴۴۷	۰/۵۰۹	۰/۵۰۵
SVM		۰/۵۱۳	۰/۵۴۴	۰/۵۳۶
DT		۰/۵۸۹	۰/۵۹۱	۰/۵۹۰
NB	امتیاز کلی کالا+امتیاز کاربر+طول نظر+تعداد بازخوردهای مثبت	۰/۶۳۴	۰/۶۴۸	۰/۶۳۹
SVM		۰/۶۵۷	۰/۶۸۲	۰/۶۶۵
DT		۰/۷۴۷	۰/۷۵۳	۰/۷۴۷
NB	امتیاز کلی کالا+امتیاز کاربر+اختلاف امتیاز کالا و کاربر+طول عنوان نظر+طول نظر	۰/۶۵۳	۰/۶۶۵	۰/۶۵۷
SVM		۰/۶۴۴	۰/۶۷۶	۰/۶۵۵
DT		۰/۶۸۵	۰/۶۸۸	۰/۶۸۶
NB	امتیاز کلی کالا+طول نظر+بازخوردهای مثبت+بازخوردهای منفی	۰/۵۵۷	۰/۵۸۹	۰/۵۷۵
SVM		۰/۵۳۷	۰/۶۱۱	۰/۵۷۵
DT		۰/۷۶۰	۰/۷۶۲	۰/۷۶۰

برای نشان دادن کیفیت ویژگی‌های فراداده و مبتنی بر نظر، ده ترکیب برتر ویژگی‌های دسته‌ی اول و ۷ ترکیب برتر ویژگی مبتنی بر نظر برای هر کدام از سه دسته‌بند محاسبه شده و به این شکل امتیازدهی شدند؛ در صورتی که یک ویژگی در بهترین ترکیب حضور داشته باشد امتیاز ۱۰، در دومین بهترین امتیاز ۹ و این فرایند تا حضور در دهمین ترکیب برتر که امتیاز ۱ به ویژگی مربوط تعلق می‌گیرد ادامه می‌یابد. در نهایت براساس جمع امتیازات کسب شده در هر رتبه توسط هر سه دسته‌بند، نقشه حرارتی مربوط به ویژگی‌ها مطابق شکل‌های ۱۶ و ۱۷ به‌دست آمد.

جدول ۹: نمونه‌هایی از ترکیب‌های ویژگی‌های مبتنی فراداده در مجموعه داده نامتوازن. مقادیر به صورت میانگین وزن‌دار دسته‌ها محاسبه شده‌اند.

نام دسته‌بند	نام ویژگی‌ها	اِف وان	دقت	بازخوانی
NB	امتیاز کاربر	۰/۵۳۷	۰/۵۷۰	۰/۵۵۷
SVM		۰/۵۵۰	۰/۵۶۱	۰/۵۵۷
DT		۰/۵۳۶	۰/۵۳۷	۰/۵۳۷
NB	تعداد بازخوردهای مثبت	۰/۵۴۳	۰/۵۴۳	۰/۵۴۳
SVM		۰/۴۱۳	۰/۵۴۹	۰/۵۱۵
DT		۰/۵۳۱	۰/۵۳۷	۰/۵۳۶
NB	بازخوردهای مثبت+امتیاز کلی کالا	۰/۵۲۶	۰/۵۲۷	۰/۵۲۷
SVM		۰/۴۲۹	۰/۵۱۲	۰/۵۰۶
DT		۰/۵۱۱	۰/۵۱۱	۰/۵۱۱
NB	طول نظر+ طول عنوان نظر	۰/۴۹۹	۰/۵۹۴	۰/۵۵۳
SVM		۰/۵۱۰	۰/۶۲۵	۰/۵۶۷
DT		۰/۵۹۷	۰/۵۹۷	۰/۵۹۷
NB	بازخوردهای منفی+طول نظر+امتیاز کلی کالا	۰/۵۰۰	۰/۵۷۳	۰/۵۴۶
SVM		۰/۴۷۸	۰/۵۷۹	۰/۵۴۶
DT		۰/۵۳۰	۰/۵۶۲	۰/۵۵۱
NB	اختلاف امتیاز کالا و کاربر+بازخوردهای مثبت+بازخوردهای منفی	۰/۴۷۰	۰/۵۳۸	۰/۵۲۳
SVM		۰/۵۱۷	۰/۵۱۹	۰/۵۱۹
DT		۰/۵۱۶	۰/۵۴۲	۰/۵۳۵
NB	امتیاز کلی کالا+امتیاز کاربر+طول نظر+تعداد بازخوردهای مثبت	۰/۵۷۵	۰/۶۱۶	۰/۵۹۵
SVM		۰/۵۹۵	۰/۶۰۵	۰/۶۰۰
DT		۰/۴۸۷	۰/۵۴۳	۰/۵۲۹
NB	امتیاز کلی کالا+امتیاز کاربر+اختلاف امتیاز کالا و کاربر+طول عنوان نظر+طول نظر	۰/۵۲۴	۰/۶۰۴	۰/۵۶۷
SVM		۰/۵۸۹	۰/۶۰۱	۰/۵۹۵
DT		۰/۴۹۵	۰/۵۵۷	۰/۵۳۸
NB	امتیاز کلی کالا+طول نظر+بازخوردهای مثبت+بازخوردهای منفی	۰/۵۰۴	۰/۵۵۷	۰/۵۴۹
SVM		۰/۵۳۵	۰/۶۳۱	۰/۵۸۰
DT		۰/۵۳۵	۰/۵۸۲	۰/۵۶۳

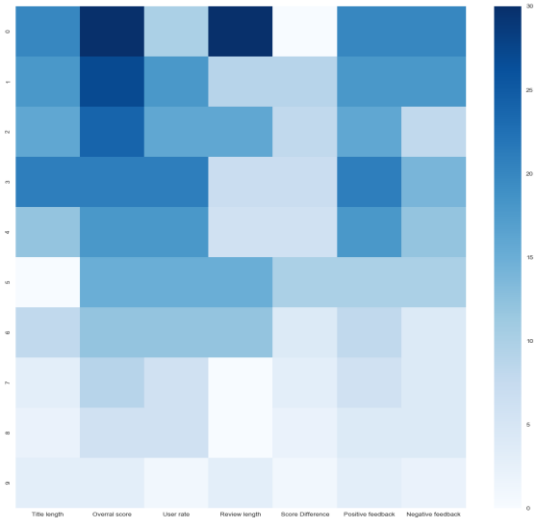
ویژگی امتیاز کلی کالا در تمامی ترکیب‌های برتر تکرار شده است. این ویژگی در دسته‌بند نایو بیز به‌عنوان بهترین ویژگی که قابلیت تمایز بالایی میان نظرات نوع یک اسپم و غیراسپم ایجاد می‌کند شناخته می‌شود. در دسته‌بند ماشین بردار پشتیبان مشاهده می‌شود که علاوه بر ویژگی امتیاز کلی کالا، دو ویژگی امتیاز کاربر و تعداد بازخوردهای مثبت نظر نیز جزو ترکیب‌های برتر ویژگی‌های مبتنی بر نظر هستند. در این دسته‌بند، این سه ویژگی به‌عنوان ویژگی‌های برتر فراداده در تشخیص نظرات اسپم نوع یک شناخته می‌شوند. در درخت تصمیم، دو ویژگی امتیاز کلی کالا و تعداد بازخوردهای مثبت نظر در میان ترکیب برتر ویژگی‌ها مشاهده می‌شوند. در مورد ویژگی‌های مبتنی بر فراداده مشاهده می‌شود که ویژگی احساسی تنها ویژگی است که در ترکیب‌های برتر سه دسته‌بند دیده می‌شود و دو ترکیب ضعیف در کلیه‌ی دسته‌بندها، استفاده از تعداد کلمات پرتکرار در متن نظر به‌تنهایی و ترکیب این ویژگی با اجزاء سخن کلمات نظر در هر دو نوع داده دقت پایینی را رقم می‌زنند.

جدول ۱۰: ترکیب برترین ویژگی‌های مبتنی بر نظر و فراداده در مجموعه داده متوازن. مقادیر به صورت میانگین وزن‌دار دسته‌ها محاسبه شده‌اند.

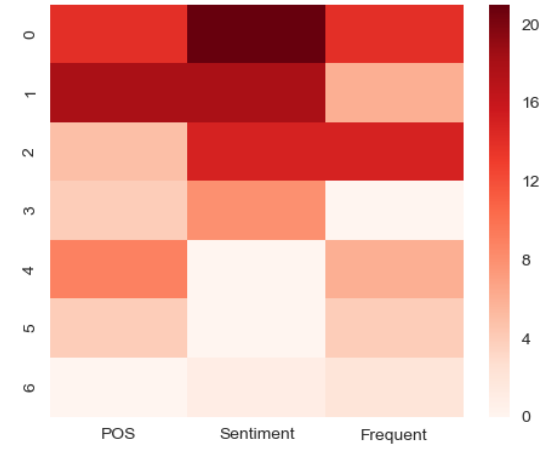
نام دسته‌بند	نام ویژگی‌ها	اِف وان	دقت	بازخوانی
NB	امتیاز کلی کالا+احساسی	۰/۷۳۴	۰/۷۴۰	۰/۷۳۴
SVM		۰/۷۲۵	۰/۷۲۹	۰/۷۲۵
DT		۰/۷۷۹	۰/۷۸۱	۰/۷۷۹
NB	امتیاز کلی کالا+احساسی+تعداد بازخوردهای مثبت	۰/۷۳۶	۰/۷۵۲	/۷۴۰
SVM		۰/۷۳۸	۰/۷۴۴	۰/۷۴۰
DT		۰/۸۲۴	۰/۸۳۳	۰/۸۲۵
NB	POS+تعداد بازخوردهای مثبت+احساسی	۰/۵۵۴	۰/۵۸۱	۰/۵۷۰
SVM		۰/۵۹۳	۰/۶۴۱	۰/۶۱۳
DT		۰/۵۸۳	۰/۵۸۷	۰/۵۸۵
NB	POS+احساسی+امتیاز کلی کالا+تعداد بازخوردهای مثبت	۰/۷۱۹	۰/۶۸۷	۰/۶۸۶
SVM		۰/۷۳۰	۰/۷۴۰	۰/۷۳۵
DT		۰/۸۰۴	۰/۸۰۳	۰/۸۰۲

جدول ۱۱: ترکیب برترین ویژگی‌های مبتنی بر نظر و فراداده در مجموعه داده نامتوازن. مقادیر به صورت میانگین وزن‌دار دسته‌ها محاسبه شده‌اند.

نام دسته‌بند	نام ویژگی‌ها	اِف وان	دقت	بازخوانی
NB	امتیاز کلی کالا+احساسی	۰/۵۳۲	۰/۶۳۷	۰/۵۸۱
SVM		۰/۵۳۲	۰/۶۳۷	۰/۵۸۱
DT		۰/۵۵۹	۰/۶۱۲	۰/۵۸۵
NB	امتیاز کلی کالا+احساسی+تعداد بازخوردهای مثبت	۰/۵۲۹	۰/۶۲۳	۰/۵۷۵
SVM		۰/۵۳۲	۰/۶۳۷	۰/۵۸۱
DT		۰/۵۴۲	۰/۵۴۷	۰/۵۴۶
NB	POS+تعداد بازخوردهای مثبت+احساسی	۰/۵۵۹	۰/۶۴۹	۰/۵۹۷
SVM		۰/۶۲۳	۰/۶۵۹	۰/۶۳۶
DT		۰/۶۵۲	۰/۷۱۴	۰/۶۷۰
NB	POS+احساسی+امتیاز کلی کالا+تعداد بازخوردهای مثبت	۰/۵۵۹	۰/۶۴۹	۰/۵۹۸
SVM		۰/۶۲۳	۰/۶۵۹	۰/۶۳۶
DT		۰/۶۲۸	۰/۷۰۰	۰/۶۵۱

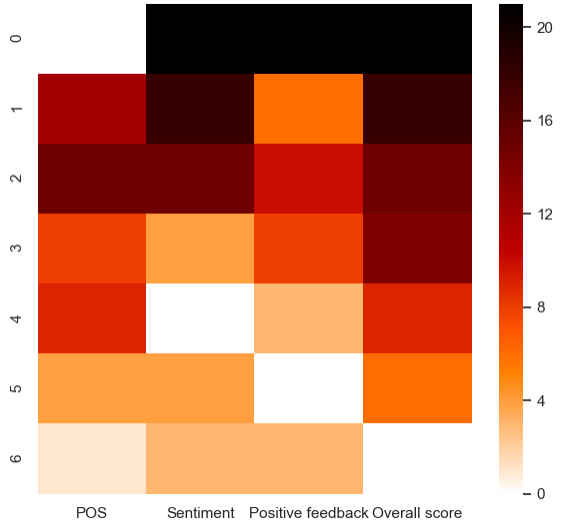


شکل ۱۶: نقشه حرارتی کیفیت ویژگی‌های مبتنی بر فراداده. محور افقی به ترتیب از چپ به راست طول عنوان، امتیاز نظر، امتیاز کاربر، طول نظر، اختلاف امتیازات و بازخورد مثبت. محور عمودی از بالا به پایین رتبه ویژگی در بین ۱۰ ترکیب برتر.



شکل ۱۷: نقشه حرارتی کیفیت ویژگی‌های مبتنی بر نظر. محور افقی به ترتیب از چپ به راست نقش کلمه، قطبیت نظر و کلمات پرتکرار. محور عمودی از بالا به پایین رتبه ویژگی در بین ۷ ترکیب برتر.

برای به‌دست آوردن بهترین ترکیب با استفاده از هر دو نوع ویژگی از این واقعیت استفاده شد که وجود چهار ویژگی متمایزکننده در تشخیص نظرات اسپم نوع یک به چشم می‌خورد. این چهار ویژگی عبارتند از: اجزاء سخن مربوط به کلمات، جهت احساسی، تعداد بازخوردهای مثبت نظر و امتیاز کلی کالا. پس از مطالعه‌ی هر یک از ترکیبات ویژگی‌های فراداده و مبتنی بر نظر به‌تنهایی، به منظور استخراج بهترین ویژگی در میان سه دسته‌بند، ترکیب این چهار ویژگی برتر نیز مورد بررسی قرار گرفت. سپس با روشی که در مورد شکل ۱۶ و ۱۷ توضیح داده شد، نقشه حرارتی مربوط به ترکیب این ویژگی‌ها مطابق شکل ۱۸ ایجاد شد.



شکل ۱۰: نقشه حرارتی کیفیت ویژگی‌های منتخب از نوع مبتنی بر نظر و فراداده. محور افقی به ترتیب از چپ به راست نقش کلمه، قطبیت نظر، بازخورد مثبت و امتیاز کلی نظر. محور عمودی از بالا به پایین رتبه ویژگی در بین ۷ ترکیب برتر.

با توجه به نتایج می‌توان دریافت که درخت تصمیم بالاترین کارایی را در تشخیص اسپم نوع یک دارد. ویژگی متمایزکننده‌ای که در ترکیب‌های برتر از مجموعه داده‌ی متوازن مشاهده می‌شود، ویژگی امتیاز کلی کالا است و بهترین ترکیب ویژگی‌ها در شناخت اسپم نوع یک، ترکیب ویژگی‌های احساسی، امتیاز کلی کالا و تعداد بازخوردهای مثبت است.

برای یافتن ضعیف‌ترین ویژگی فراداده در دسته‌بندی نظرات اسپم و غیراسپم، ده ترکیب ضعیف دسته‌بندها مورد بازبینی قرار گرفتند و از میان آنها ویژگی اختلاف امتیاز کلی کالا و کاربر در اغلب این ترکیب‌ها مشاهده می‌شود. علت این امر آن است که میان این ویژگی با دو ویژگی امتیاز کلی کالا و امتیاز کاربر وابستگی وجود دارد و این وابستگی منجر به کاهش کارایی دسته‌بند خواهد شد.

در میان ویژگی‌های مبتنی بر نظر، دو ترکیب ضعیف در کلیه‌ی دسته‌بندها، استفاده از تعداد کلمات پرتکرار در متن نظر به‌تنهایی و یا ترکیب این ویژگی با نوع کلمات نظر (POS) در هر دو نوع داده دقت بسیار پایینی را رقم می‌زنند. این به آن معنی است که حذف ویژگی ممتازی مثل ویژگی احساسی در شناسایی نظرات اسپم کارایی دسته‌بند را تا حد زیادی پایین می‌آورد؛ به‌طوری‌که دو ویژگی تعداد کلمات پرتکرار به‌تنهایی و ترکیب آن با ویژگی POS به‌عنوان ضعیف‌ترین ترکیب شناخته می‌شوند.

نهایتاً از میان ویژگی‌های ترکیبی (کل ویژگی‌ها)، ضعیف‌ترین ویژگی‌ها ویژگی اختلاف امتیاز کاربر و کالا و تعداد کلمات پرتکرار

و ترکیب آنها با همدیگر است.

۵- نتیجه‌گیری

در این پژوهش تأثیر ویژگی‌ها و تک تک آن‌ها روی دسته‌بندها و ایجاد ترکیب‌های مختلف از ویژگی‌های فراداده و مبتنی بر نظر به تفصیل مورد بحث قرار گرفت. همچنین دید کلی روی نتایج حاصل‌شده از تمامی دسته‌بندها، دقت آن‌ها، ترکیب ویژگی‌های مختلف و تأثیر آن‌ها روی کارایی دسته‌بند ایجاد شد.

برای تشخیص نظرات اسپم، باید سه نوع اسپم شناسایی شوند. از میان سه نوع اسپم، نظرات اسپم نوع یک خطرناک‌ترین نوع نظرات اسپم هستند که اغلب پژوهش‌های انجام‌شده به‌منظور شناسایی این نوع اسپم انجام شده‌اند. در این پژوهش دسته‌بندی توسط نایو بیز، ماشین بردار پشتیبان و درخت تصمیم انجام شده است. علت انتخاب دو دسته‌بند اول، کارایی بالای آن‌ها در شناخت نظرات اسپم سایر زبان‌ها (انگلیسی و عربی) بوده است. با توجه به این‌که توانایی درخت تصمیم در شناسایی نظرات اسپم نوع یک مورد بررسی قرار نگرفته است، این دسته‌بند به‌عنوان دسته‌بند سوم روی نظرات اسپم فارسی استفاده شده است.

با این‌که در پژوهش‌های پیشین ابتدا دسته‌بند ماشین بردار پشتیبان و سپس نایو بیز را به‌عنوان بهترین دسته‌بندها در تشخیص نظرات اسپم می‌دانند، اما در نظرات اسپم فارسی، درخت تصمیم بالاترین کارایی (معیار اِف ۰/۷۷۸) را از خود نشان می‌دهد. از میان سه دسته‌بند استفاده‌شده در این پژوهش، نایو بیز با کسب کم‌ترین دقت به‌عنوان ضعیف‌ترین دسته‌بند و ماشین بردار پشتیبان در هر دو مجموعه داده‌ی متوازن و نامتوازن نظرات اسپم فارسی به‌عنوان دسته‌بندی در حد متوسط معرفی می‌شود.

در میان ویژگی‌های فراداده، ویژگی که موجب تمایز بالا میان نظرات اسپم و غیراسپم می‌شود، امتیاز کلی محصول است. در سطح دوم ویژگی‌های فراداده، تعداد بازخوردهای مثبت نظر جای دارد و ترکیب این دو ویژگی برتر منجر به کارایی بالاتر در دسته‌بندها می‌شود. در هر سه دسته‌بند با حذف ویژگی امتیاز کلی کالا، کارایی به طرز چشمگیری کاهش می‌یابد و علت این امر عدم تشخیص نوع محصول (خوب، متوسط و بد) است. زیرا اغلب نظرات نوشته‌شده در وب‌سایت دیجی‌کالا برای کالاهای خوب ثبت شده‌اند. با بررسی این دو ویژگی روی نمونه‌ها مشخص می‌شود که نظرات اسپم اغلب برای کالاهایی نوشته شده‌اند که میانگین امتیاز کلی آن‌ها بالاتر از ۳/۵ (از میان امتیاز ۵-۱) بوده و در زمره‌ی محصولات خوب به شمار می‌آیند و در واقع این نوع کالاها بیشتر مورد هدف اسپمرها واقع می‌شوند.

دقت دسته‌بندی در ویژگی‌های مبتنی بر نظر در مجموعه داده‌ی نامتوازن که تعداد نمونه‌های بسیار بیشتری نسبت به مجموعه داده‌ی متوازن دارد بالاتر است؛ زیرا برای استخراج ویژگی‌های مبتنی بر نظر به کشف الگوهایی در متن نظر نیاز است و هرچه تعداد نمونه‌ها بالاتر رود، دسته‌بند آن‌ها را بهتر یاد گرفته و در افزایش کارایی آن مفید واقع می‌شود. درخت تصمیم در این نوع ویژگی‌ها نیز بیشترین کارایی را دارد و نایو بیز به‌عنوان ضعیف‌ترین دسته‌بند معرفی می‌شود.

در نوع دیگر دسته‌بندی، از ترکیب بهترین ویژگی‌های مبتنی بر نظر و فراداده استفاده شد. ویژگی امتیاز کلی نظر خیلی خوب عمل می‌کند و دو ویژگی احساسی و اجزاء سخن در ترکیب با ویژگی امتیاز کلی کالا کارایی دسته‌بند را افزایش می‌دهند. پس به‌طور کلی ترکیب سه ویژگی امتیاز کلی کالا، احساسی و اجزاء سخن بیشترین تأثیر را در شناسایی نظرات اسپم و غیراسپم دارند.

مراجع

[1] B. Liu, *Sentiment analysis: Mining opinions, sentiments, and emotions*, First. Cambridge University Press, 2015.

[2] I. Chaturvedi, E. Cambria, R. E. Welsch, and F. Herrera, "Distinguishing between facts and opinions for sentiment analysis: Survey and challenges," *Information Fusion*, vol. 44, pp. 65–77, Nov. 2018.

[3] M. E. Basiri, N. Ghasem-Aghaee, and A. Reza, "Lexicon-based Sentiment Analysis in Persian," *Current and Future Developments in Artificial Intelligence*, p. 154, 2017.

[4] M. E. Basiri, A. R. Naghsh-Nilchi, and N. Ghasem-Aghaee, "Sentiment prediction based on dempster-shafer theory of evidence," *Mathematical Problems in Engineering*, 2014.

[5] A. Heydari, M. ali Tavakoli, N. Salim, and Z. Heydari, "Detection of review spam: A survey," *Expert Systems with Applications*, vol. 42, no. 7, pp. 3634–3642, May 2015.

[6] R. Ghai, S. Kumar, and A. C. Pandey, "Spam Detection Using Rating and Review Processing Method," in *Smart Innovations in Communication and Computational Sciences*, pp. 189–198, 2019.

[7] G. Fei, H. Li, and B. Liu, "Opinion Spam Detection in Social Networks," in *Sentiment Analysis in Social Networks*, pp. 141–156, 2017.

[8] S. Dixit and A. J. Agrawal, "Survey on review spam detection," *Int J Comput Commun Technol ISSN (PRINT)*, vol. 4, pp. 975–7449, 2013.

[9] N. Jindal and B. Liu, "Opinion spam and analysis," in *Proceedings of the international conference on Web search and web data mining - WSDM '08*, p. 219, 2008.

[10] Y. Ren and D. Ji, "Neural networks for deceptive opinion spam detection: An empirical study," *Information Sciences*, vol. 385–386, pp. 213–224, Apr. 2017.

[11] S. Shojaee, A. Azman, M. Murad, N. Sharef, and N. Sulaiman, "A framework for fake review annotation," in *Proceedings of the 2015 17th UKSIM-AMSS International Conference on Modelling and Simulation*, pp. 153–158, 2015.

[12] A. Mukherjee, V. Venkataraman, B. Liu, and N. S. Glance,

"What yelp fake review filter might be doing?," in *ICWSM*, pp. 409–418, 2013.

[13] N. Jindal and B. Liu, "Review spam detection," in *Proceedings of the 16th international conference on World Wide Web - WWW '07*, p. 1189, 2007.

[14] F. Li, M. Huang, Y. Yang, and X. Zhu, "Learning to identify review spam," in *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, vol. 22, no. 3, p. 2488, 2011.

[15] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding Deceptive Opinion Spam by Any Stretch of the Imagination," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, pp. 309–319, 2011.

[16] H. Faris *et al.*, "An Intelligent System for Spam Detection and Identification of the most Relevant Features based on Evolutionary Random Weight Networks," *Information Fusion*, Aug. 2018.

[17] R. Y. K. Lau, S. Y. Liao, R. C.-W. Kwok, K. Xu, Y. Xia, and Y. Li, "Text mining and probabilistic language modeling for online review spam detection," *ACM Transactions on Management Information Systems*, vol. 2, no. 4, pp. 1–30, Dec. 2011.

[18] J. Li, M. Ott, C. Cardie, and E. Hovy, "Towards a general rule for identifying deceptive opinion spam," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, vol. 1, pp. 1566–1576, 2014.

[19] A. Karami and B. Zhou, "Online review spam detection by new linguistic features," *iConference 2015 Proceedings*, 2015.

[20] E. F. Cardoso, R. M. Silva, and T. A. Almeida, "Towards automatic filtering of fake reviews," *Neurocomputing*, vol. 309, pp. 106–116, Oct. 2018.

[21] Y. Lin, T. Zhu, X. Wang, J. Zhang, and A. Zhou, "Towards online review spam detection," in *Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion*, pp. 341–342, 2014.

[22] Y. Lin, T. Zhu, H. Wu, J. Zhang, X. Wang, and A. Zhou, "Towards online anti-opinion spam: Spotting fake reviews from the review sequence," in *Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 261–264, 2014.

[23] W. Chang, Z. Xu, S. Zhou, and W. Cao, "Research on detection methods based on Doc2vec abnormal comments," *Future Generation Computer Systems*, vol. 86, pp. 656–662, Sep. 2018.

[24] D. Hernández, R. Guzmán, M. M. y Gomez, and P. Rosso, "Using PU-learning to detect deceptive opinion spam," in *Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis*, pp. 38–45, 2013.

[25] S. Shehnepoor, M. Salehi, R. Farahbakhsh, and N. Crespi, "NetSpam: A Network-Based Spam Detection Framework for Reviews in Online Social Media," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1585–1595, Jul. 2017.

[26] J. K. Rout, A. Dalmia, K.-K. R. Choo, S. Bakshi, and S. K. Jena, "Revisiting Semi-Supervised Learning for Online Deceptive Review Detection," *IEEE Access*, vol. 5, pp. 1319–1327, 2017.

[27] "Digikala," 2017. [Online]. Available: http://www.digikala.com. [Accessed: 15-Feb-2017].

[28] D. Savage, X. Zhang, X. Yu, P. Chou, and Q. Wang, "Detection of opinion spam based on anomalous rating deviation," *Expert Systems with Applications*, vol. 42, no. 22, pp. 8650–8657, Dec. 2015.

[29] I. Dematis, E. Karapistoli, and A. Vakali, "Fake Review Detection via Exploitation of Spam Indicators and Reviewer Behavior Characteristics," in *nternational Conference on Current Trends in Theory and Practice of Informatics*, pp. 581–595, 2018.

[30] E. Asgarian, R. Saeedi, A. Stiri, B. Bahmadi, and H. Ghaemi, "NLPTools." [Online]. Available: https://wtlab.um.ac.ir. [Accessed: 01-Jul-2016].

[31] M. E. Basiri, A. R. Naghsh-Nilchi, and N. Ghassem-Aghaee, "A framework for sentiment analysis in persian," *Open Transactions on Information Processing*, vol. 1, no. 3, pp. 1–14, 2014.

[32] M. E. Basiri and A. Kabiri, "Sentence-Level Sentiment Analysis in Persian," in *The 3rd International Conference on Pattern Recognition and Image Processing*, no. 1, pp. 84–89, 2017.

[33] M. E. Basiri and A. Kabiri, "Words Are Important: Improving Sentiment Analysis in the Persian Language by Lexicon Refining," *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 17, no. 4, p. 26, 2018.

[34] Y. Liu and B. Pang, "A unified framework for detecting author spamicity by modeling review deviation," *Expert Systems with Applications*, vol. 112, pp. 148–155, Dec. 2018.

[35] L. Li, B. Qin, W. Ren, and T. Liu, "Document representation and feature combination for deceptive spam review detection," *Neurocomputing*, vol. 254, pp. 33–41, Sep. 2017.

زیرنویس‌ها:

[1] Hyper spam
[2] Defaming spam
[3] www.digikala.com
[4] Metadata
[5] Suport vector machine
[6] Supervised
[7] Unsupervised
[8] Semi-supervised
[9] Natural language processing
[10] 10-Fold cross validation
[11] Part of speech
[12] Unigram
[13] Linguistic
[14] Iterative supervision
[15] Co-learning
[16] Positive unlabled
[17] Rule-based
[18] Overfit
[19] Sequential minimal optimization
[20] Polarity