

## ارائه روشی جدید مبتنی بر قطبیت برای بهبود نظر کاوی فارسی

محبوبه شمس<sup>۱</sup>، عبدالرضا رسولی کناری<sup>۲</sup>، مهدیه واحدی پور<sup>۳</sup>

<sup>۱</sup>گروه مهندسی کامپیوتر، دانشکده برق و کامپیوتر، دانشگاه صنعتی قم

<sup>۲</sup>گروه مهندسی کامپیوتر، دانشکده برق و کامپیوتر، دانشگاه صنعتی قم

<sup>۳</sup>گروه مهندسی کامپیوتر، دانشکده برق و کامپیوتر، دانشگاه صنعتی قم

\*نویسنده مسئول مکاتبات: [shamsi@Qut.ac.ir](mailto:shamsi@Qut.ac.ir)

\*نویسنده مسئول مکاتبات: [Rasouli@qut.ac.ir](mailto:Rasouli@qut.ac.ir)

\*نویسنده مسئول مکاتبات: [vahedipoor.M@Qut.ac.ir](mailto:vahedipoor.M@Qut.ac.ir)

### چکیده

در سال‌های اخیر رشد انبوه محتوای تولید شده توسط کاربر مانند شبکه‌های اجتماعی و سایت‌های بازاریابی آنلاین به مردم اجازه می‌دهد تا احساسات و نظرات خود را در بسیاری از نظرهای مربوط به محصولات و خدمات مختلف به اشتراک بگذارند. تجزیه و تحلیل احساسات منبع مهمی برای تصمیم‌گیری بهتر می‌باشد که با استفاده از پردازش زبان طبیعی (NLP)<sup>۱</sup>، تکنیک‌های محاسباتی و تجزیه و تحلیل متن برای استخراج قطبیت اسناد بدون ساختار و روش‌های یادگیری ماشین تحت نظارت و روش‌های یادگیری عمیق بررسی می‌شود. پیچیدگی زبان‌های انسانی و تحلیل احساسات یک زمینه تحقیقاتی چالش برانگیز در علوم رایانه و زبان‌شناسی محاسباتی ساخته‌اند. در این مقاله روشی جدید مبتنی بر قطبیت نظرات فارسی تلفن همراه سایت دیجی کالا برای تعیین گروه‌های قطبی مختلف مانند مثبت و منفی ارائه می‌کنیم و با الگوریتم‌های یادگیری ماشین<sup>۲</sup> از جمله بیز ساده، ماشین بردار پشتیبان، تبادلات گرادیان تصادفی، برگشت منطقی، جنگل تصادفی و یادگیری عمیق<sup>۳</sup> مانند شبکه عصبی کانولوشن و حافظه طولانی کوتاه مدت به منظور دستیابی به عملکرد آن‌ها براساس پارامترهایی مانند صحت، فراخوان، اندازه F و دقت این روش مورد توجه قرار داده‌ایم. روش پیشنهادی با الگوریتم‌های ماشین بردار پشتیبان بین ۵ تا ۲۴ درصد و تبادلات گرادیان تصادفی بین ۷ تا ۳۸ درصد و برگشت منطقی بین ۵ تا ۳۸ درصد در تمام روش‌های n-gram و شبکه عصبی کانولوشن به میزان ۴ درصد دقت نسبت به روش قبلی به دست آورده است.

### واژگان کلیدی:

تحلیل احساسات؛ نظر کاوی؛ پیش‌پردازش متن؛ یادگیری ماشین؛ یادگیری عمیق

## ۱ مقدمه

در حال حاضر شبکه‌ها و رسانه‌های اجتماعی نقش مهمی در ارائه اطلاعات در مورد هر محصول از نظرات مختلف و وبلاگ‌ها نقش دارند. تجزیه و تحلیل احساسات به عنوان نظرکاوی<sup>۴</sup> شناخته شده است و نظرات مردم و همچنین احساسات نسبت به محصولات، سازمان‌ها و ویژگی‌های مربوط به آن‌ها را تحلیل می‌کند و به منظور استخراج اطلاعات معنادار از احساسات مردم با روش‌های مختلف یادگیری ماشین توسط دانشمندان مورد استفاده قرار گرفته است.

تشخیص ساختار متن به سه شکل مختلف سند، جمله و کلمه امکان‌پذیر است. در سطح سند، اسناد به نظرات مثبت<sup>۵</sup> (موافق)، منفی<sup>۶</sup> (مخالف) و خنثی<sup>۷</sup> (بی‌طرف) طبقه‌بندی می‌شود. در سطح جمله تعیین می‌کند که آیا جمله نظر مثبت، منفی و یا خنثی را بیان می‌کند و در سطح ویژگی، برای تعیین نظر مثبت، منفی و یا خنثی به جزئی‌ترین شکل ممکن (کلمه) انجام می‌گیرد و نظرات را در مورد ویژگی‌های خاص موضوع بررسی می‌کند [۱]. در این مطالعه، طبق داده‌های جمع‌آوری شده نظرات تلفن همراه سایت دیجی کالا تحلیل احساسات در سطح جمله مورد توجه قرار گرفته است.

به طور عمده دو نوع تکنیک یادگیری ماشین برای دسته‌بندی متون وجود دارد از جمله تکنیک مبتنی بر یادگیری ماشین تحت نظارت<sup>۸</sup> و تکنیک مبتنی بر یادگیری ماشین بدون نظارت<sup>۹</sup> می‌باشد. در روش یادگیری نظارت شده، مجموعه داده‌ها توسط کارشناسان به صورت دستی یا استفاده از فرهنگ لغت‌نامه برچسب‌گذاری<sup>۱۰</sup> می‌شوند. بنابراین یادگیری مدل‌ها با بخشی از داده‌ها به عنوان داده‌های آموزشی<sup>۱۱</sup> صورت گرفته و با بخش دیگری از داده‌ها به عنوان داده‌های آزمایشی<sup>۱۲</sup> یک خروجی معقول را به دست می‌آورند [۲].

بر خلاف یادگیری نظارت شده، فرآیند یادگیری بدون نظارت هیچ اطلاعاتی در مورد برچسب ندارند و نمی‌توانند به راحتی پردازش شوند زیرا آموزشی صورت نمی‌گیرد. الگوریتم‌های خوشه‌بندی<sup>۱۳</sup> با دسته‌بندی داده‌های مشابه جهت حل مشکل پردازش داده‌های بدون برچسب به کار می‌روند.

الگوریتم‌های یادگیری عمیق زیرمجموعه‌ای از الگوریتم‌های یادگیری ماشین هستند که هدف آنها کشف چندین سطح از نمایش‌های توزیع شده از داده ورودی است. در سال‌های اخیر الگوریتم‌های یادگیری عمیق زیادی برای حل مسائل هوش مصنوعی سنتی ارائه شده‌اند. کاربردهای آن‌ها در زمینه‌های مختلف بینایی همانند دسته‌بندی تصاویر، شناسایی اشیاء، استخراج تصاویر، قطعه‌بندی معنایی و برآورد ژست انسان می‌باشد [۳].

نظرات، متون بدون ساختار هستند. بنابراین، برای ساختاردهی به ورودی‌های متنی (معمولا تجزیه، همراه با افزودن برخی ویژگی‌ها، تفاسیر زبانی و حذف موارد اضافی و درج موارد بعدی در پایگاه داده انجام می‌گیرد)، استخراج الگوهای درون داده‌های ساختاریافته، و در نهایت ارزیابی و تفسیر خروجی‌ها است.

عدم تعادل داده‌ها با اختلاف در تعداد نمونه‌ها در هر کلاس از یک مجموعه داده مشخص می‌شود. این پدیده شناخته شده است که عملکرد طبقه‌بندی‌ها را خراب می‌کند، زیرا آن‌ها کمتر توانایی ویژگی‌های کلاس‌های کمتر را می‌آموزند.

برای اکثر ناسازگاری مجموعه داده‌ها، با استفاده از تکنیک‌های نمونه‌گیری از جمله نمونه‌گیری بیش از حد<sup>۱۴</sup> و نمونه‌گیری کمتر از حد<sup>۱۵</sup> عملکرد طبقه‌بندی را بهبود می‌بخشند. برای مجموعه داده‌های کوچک، نمونه‌گیری بیش از حد انتخاب شده است که مناسب‌ترین استراتژی است به دلیل این که مجموعه اولیه نمونه را افزایش می‌دهد و این کار، بر روی وظایف دودویی یا چند ستونی تمرکز دارد و در این مطالعه از نمونه‌گیری بیش از حد برای سازگاری داده‌ها استفاده شده است [۴].

تجزیه و تحلیل احساسات مبتنی بر فرهنگ‌نامه، برای ارزیابی سطح نظرات تلفن همراه رضایت مشتریان استفاده می‌شود. یک پایگاه داده از کلمات احساسات فارسی برای شناسایی قطبیت کلمات به عنوان یک منبع واژگانی ساخته شده است. فرهنگ‌نامه واژگانی ارائه شده شامل کلمات نظرات در حوزه تلفن همراه برای دستیابی به نتیجه بهتر است. هر کلمه‌ای در پایگاه داده ارزش دارد. ارزش احساسات از ۵- تا ۵+ متغیر است. این روش به طور خودکار نظرات مشتریان را به دسته‌های کلمات خیلی ضعیف منفی، کلمات ضعیف منفی، خصوصیات منفی، کلمات متوسط منفی، کلمات نادرست منفی، کلمات نامفهوم، کلمات نادرست مثبت، کلمات متوسط مثبت، خصوصیات مثبت، کلمات قوی مثبت و کلمات خیلی قوی مثبت با استفاده از فرهنگ‌نامه ایجاد شده و تجزیه و تحلیل می‌کند [۵].

در این پژوهش، ابتدا کلمات متوقف<sup>۱۶</sup> و سایر اطلاعات ناخواسته از نظرات موردنظر حذف می‌گردد و سپس طبق فرهنگ‌نامه موردنظر ارزش نظرات، محاسبه شده و قطبیت جدیدی به نظرات داده می‌شود و به دلیل ناهمگن بودن داده‌های مثبت و منفی با روش نمونه‌گیری بیش از حد، داده‌ها متعادل می‌شوند و از طریق فرآیند بردارسازی<sup>۱۷</sup> داده‌های متن ساختاریافته استخراجی از مرحله پیش‌پردازش به ماتریس اعداد تبدیل می‌شوند. سپس این ماتریس‌های عددی به الگوریتم‌های مختلف یادگیری ماشین تحت نظارت و یادگیری عمیق برای طبقه‌بندی این نظرات داده شده است

سپس پارامترهای مختلف برای ارزیابی بهتر عملکرد الگوریتم‌های یادگیری ماشین تحت نظارت و یادگیری عمیق مانند صحت، فراخوان، اندازه F و دقت روش ارائه شده جهت بهبودی نظر کاوی نظرات فارسی مورد توجه قرار داده‌ایم [۶].

ساختار مقاله به شرح زیر تعریف می‌شود: بخش ۲، بررسی ادبیات را ارائه می‌دهد. بخش ۳، مجموعه داده مورد استفاده در پژوهش و جزئیات آن را نشان می‌دهد. در بخش ۴، پیاده‌سازی رویکرد پیشنهادی توضیح داده شده است. بخش ۵، ارزیابی عملکرد روش پیشنهادی انجام شده است. بخش ۶، مقاله را به پایان می‌رساند و دامنه کارهای آینده را ارائه می‌دهد.

## ۲ کارهای گذشته

تجزیه و تحلیل احساسات در حوزه وسیعی مانند بررسی فیلم، بررسی تدریس، بررسی محصول، یادگیری الکترونیکی، بررسی هتل و بسیاری دیگر مورد مطالعه قرار گرفته است. بیشتر دانشمندان متمرکز به تجزیه و تحلیل داده‌های کمی هستند. با این حال، برخی از مطالعات انجام شده بر روی داده‌های کیفی با استفاده از تجزیه و تحلیل احساسات انجام شده است و همچنین مطالعه خوبی توسط نویسندگان مختلف براساس طبقه‌بندی احساسات در سطح سند انجام شده است که در زیر به برخی از آن‌ها اشاره شده است.

Lee و Pang طبقه‌بندی احساسات در سطح سند با احساسات مثبت و منفی را تحقیق و بررسی کرده‌اند. آن‌ها با سه الگوریتم مختلف یادگیری ماشین نظیر  $^{18}NB$ ،  $^{19}SVM$  و  $^{20}ME$ ، آزمایش کرده‌اند و فرآیند طبقه‌بندی را با استفاده از تکنیک‌های n-gram مانند تک‌واژه‌ای<sup>۲۱</sup>، دوواژه‌ای<sup>۲۲</sup> و ترکیب تک‌واژه‌ای و دوواژه‌ای انجام می‌دهند. آن‌ها چارچوب ویژگی‌های کیفی را برای اجرای الگوریتم‌های یادگیری ماشین استفاده کرده‌اند. در نتیجه تجزیه و تحلیل آن‌ها، الگوریتم NB نتایج ضعیف و الگوریتم SVM نتایج بهتری را نشان می‌دهد [۷].

Salveti و همکارانش، در مفهوم کلی قطبیت عقیده‌ای ( $^{23}OvOp$ ) با استفاده از الگوریتم‌های یادگیری ماشین نظیر مدل بیز ساده و مدل مارکوف<sup>۲۴</sup> برای طبقه‌بندی متون مورد بحث قرار داده‌اند و آن‌ها توسط وردنت ( $^{25}WN$ ) و بخشی از گفتار ( $^{26}POS$ ) متون را برچسب‌گذاری کرده‌اند و به عنوان فیلتر لغوی<sup>۲۷</sup> برای طبقه‌بندی عمل می‌کنند. آزمایش آن‌ها نشان می‌دهد که نتیجه به دست آمده از فیلتر وردنت نسبت به فیلتر POS کمتر است. آن‌ها براساس عملکرد احتمال نظر کاوی رتبه‌بندی می‌کنند. رویکرد آن‌ها، نتیجه بهتری روی داده‌های وب نشان می‌دهد [۸].

Beineke و همکارانش از مدل بیز ساده برای طبقه‌بندی احساسات استفاده کرده‌اند. آن‌ها ویژگی‌های مشتق‌شده<sup>۲۸</sup> را برای پیش‌بینی احساسات استخراج کرده‌اند. آن‌ها ویژگی‌های مشتق‌شده دیگری را به منظور بهبود نتیجه دقت<sup>۲۹</sup> به مدل اضافه کرده‌اند و داده‌های موردنظر را برای تعیین تاثیر نسبی استفاده می‌کنند. این ایده به سیستم اجازه می‌دهد که به عنوان یک مدل احتمالی در مقیاس منطقی خطی عمل کند. در این روش هر واژه الحاقی با ضریب k همراه است که این ضرایب باید شناخته شده باشند. با این حال، زمانی که اسناد برچسب‌دار در دسترس هستند، ممکن است تخمین آن‌ها مفید باشد [۹].

Mullen و Collier با استفاده از الگوریتم ماشین بردار پشتیبان ارزش‌ها را به کلمات انتخاب شده اختصاص می‌دهند تا یک مدل برای طبقه‌بندی ایجاد کنند. علاوه بر این، کلاس‌های مختلفی از ویژگی‌های نزدیک به موضوع اختصاص داده شده با مقادیر بهتر به کار برده‌اند که به طبقه‌بندی کمک می‌کند. نویسندگان مقایسه‌ای از رویکرد پیشنهادی خود را با داده‌ها، حاشیه‌نویسی موضوع<sup>۳۰</sup> و حاشیه‌نویسی دستی<sup>۳۱</sup> ارائه کرده‌اند. رویکرد پیشنهادی در مقایسه با حاشیه‌نویسی موضوع نشان داده شده است که در مقایسه با داده‌های حاشیه‌نویسی دستی نتایج بهبودی بیشتری دارد. یک مشکل این روش با محدود کردن دامنه با اضافه کردن محدودیت‌های کلمه مرتبط به موضوع این است که تعداد شمارش نتیجه به شدت کاهش می‌یابد و هرگونه افزایش بالقوه را از بین می‌برد. همچنین به نظر می‌رسد که جنبه‌های موضوع روابط تحقیق حاضر فقط سطحی بیان می‌شود [۱۰].

Dave و همکارانش تکنیک‌های بازیابی اطلاعات مورد استفاده برای بازیابی ویژگی و نتیجه روش‌های مختلف مورد آزمایش قرار داده‌اند. آن‌ها از یک ابزار برای سنتز نظرات استفاده کرده‌اند، سپس آن‌ها را تغییر داده و با استفاده از سایت‌های تجمعی طبقه‌بندی می‌کنند. آن‌ها با استفاده از این نظرات ویژگی‌ها را شناسایی کرده و در نهایت روش‌هایی برای تعیین مثبت یا منفی بودن نظرات به کار می‌برند و برای طبقه‌بندی جملات حاصل از جستجوی وب با استفاده از نام محصول به عنوان شرایط جستجو استفاده می‌کنند. در این روش مواردی چون عدم تطابق درجه‌بندی، هم‌بستگی، مقایسه و داده‌های انعطاف‌پذیری را دشوار می‌سازد [۱۱].

Matsumoto و همکارانش از ارتباطات نحوی میان کلمات هم‌چون مبنای تحلیل احساسات استفاده کرده‌اند. در پژوهش خود، توالی کلمه مکرر و وابستگی جملات با استفاده از الگوریتم SVM استخراج می‌کنند. آن‌ها روش‌های تک‌واژه‌ای، دوواژه‌ای، توالی کلمه و وابستگی زیردرخت از هر جمله در مجموعه داده‌ها استخراج می‌کنند. این روش در سطح سند امکان‌پذیر نیست [۱۲].

Su و Xu و Zhang روشی برای به دست آوردن ویژگی‌های معنایی با استفاده از word2vec برای گرفتن ویژگی‌های مشابه و سپس طبقه‌بندی نظرات از طریق الگوریتم SVM ارائه داده‌اند. رویکرد آن‌ها براساس دو بخش است. در بخش اول، آن‌ها از ابزار word2vec برای خوشه‌بندی ویژگی‌های

مشابه استفاده کرده‌اند تا ویژگی‌های معناداری در دامنه انتخاب شده را ذخیره کنند. سپس در بخش دوم، با انتخاب ویژگی مبتنی بر واژگان برای دریافت داده‌های آموزشی مورد استفاده قرار می‌دهند و این روش روی زبان چینی انجام شده است [۱۳].

Chen و Liu چندین طبقه‌بندی مختلف را در تحلیل احساسات پیشنهاد کرده‌اند. آن‌ها از یازده روش طبقه‌بندی چندسطحی در دو مجموعه داده‌های میکرو بلاگ<sup>۳۲</sup> و هشت ماتریس ارزیابی مختلف برای تجزیه و تحلیل استفاده کرده‌اند. آن‌ها همچنین از سه فرهنگ لغت مختلف احساسات برای طبقه‌بندی چندسطحی استفاده می‌کنند. این روش از فرهنگ‌نامه از پیش تعریف شده استفاده کرده است در حالی که می‌توان در دامنه‌های مختلف فرهنگ‌نامه موردنظر را ایجاد کرد و به دلیل برچسب‌های چندگانه مقایسه آزمایشات مختلف گران است [۱۴].

Lu و همکارانش از روشی برای تبدیل متن به فضای عاطفی کمتر ( $ESM^3$ )، با استفاده از تکنیک‌های یادگیری ماشین جهت طبقه‌بندی<sup>۳۴</sup> متون پیشنهاد کرده‌اند. آن‌ها متن را به کلماتی که معنی قطعی و روشن دارند تبدیل می‌کنند و برای طبقه‌بندی کلمات آن را به شش دسته اصلی مانند خشم، ترس، انزجار، غم و اندوه، شادی و تعجب تقسیم می‌کنند. آن‌ها دو روش متفاوت برای تخصیص وزن به کلمات را با برچسب‌های عاطفی در نظر گرفته‌اند و تمام وزن کلمات عاطفی را محاسبه کرده‌اند و براساس این مقادیر پیام‌ها به گروه‌های مختلف طبقه‌بندی می‌شوند. این روش می‌تواند نتایج منطقی را در هر مجموعه داده یا دامنه استفاده کند ولی تبدیل کلمات به معنای قطعی و تقسیم آن‌ها به شش دسته اصلی زمان‌بر است [۱۵].

Niu و همکارانش مجموعه‌ای از داده‌های ( $MVSA^35$ ) که حاوی مجموعه‌ای از جفت متن-تصویر با حاشیه‌نویسی دستی از توییتر است را جمع‌آوری کرده‌اند. رویکرد آن‌ها از تجزیه و تحلیل احساسات دارای دو بخش یادگیری مبتنی بر واژگان و آماری می‌باشد که در مورد تجزیه و تحلیل مبتنی بر واژگان، مجموعه‌ای از کلمات یا عبارات در نظر گرفته شده است و نمره احساسات از پیش تعریف شده دارند در حالی که در یادگیری آماری، تکنیک‌های مختلف یادگیری ماشین با ویژگی‌های متن مورد استفاده قرار می‌گیرد [۱۶].

Junming و Caiqiang مدل ارزیابی معلم آموزشی آنلاین را براساس نظر کاوی انجام می‌دهند. آن‌ها با استفاده از خزنده وب، نظرات دانش‌آموزان را که در زبان چینی منتشر شده در سیستم مدیریت یادگیری ( $LMS^36$ ) نوشته شده است، جمع‌آوری می‌کنند. آنها اصطلاح قطبی را در کار خود نشان ندادند. اگر کلمات احساسی متن ذهنی در لغت‌نامه قطب‌واژه نیستند، آن‌ها از روش متقابل اطلاعات متناوب ( $PMI^37$ ) برای قضاوت قطبیت استفاده کردند. این مدل یک ارزیابی کلی از هر معلم را می‌کند [۱۷].

Ortigosa, Martin و Carro روش دیگر تحلیل احساسات برای محیط یادگیری الکترونیکی را با استفاده از روش ترکیبی از تکنیک‌های یادگیری زبان اسپانیایی پیشنهاد دادند. در آزمایش آن‌ها ترکیبی از تکنیک‌های مبتنی بر واژگان و SVM بالاترین دقت را به دست می‌آورند. در زمینه یادگیری الکترونیکی، ممکن است اطلاعاتی در مورد احساسات دانش‌آموز از پیام‌هایی که در فیس‌بوک می‌نویسند استخراج شود. احساسات دانش‌آموزان نسبت به یک دوره می‌تواند به عنوان بازخورد برای معلمان، به ویژه در مورد یادگیری آنلاین، مفید باشد. با این حال، این کار هم‌چنان دارای محدودیتی برای تحلیل احساسات است، تمام کلمات به عنوان یک قطب، نظرات را برچسب‌گذاری می‌کنند، تمام کلمات مثبت نمره ۱، تمام کلمات منفی نمره -۱ و تمام کلمات خنثی نمره صفر را می‌گیرند آن‌ها وزن‌های مختلف را به کلمات مختلف اختصاص نمی‌دهند و برای بدتر و بد امتیاز مشابه به دست می‌آورند [۱۸].

Pong-inwong و Rungworawut پیشنهاد ساخت تدریس تئوری ارزیابی احساسات را برای قطبیت احساسات خودکار ارائه دادند. کار آن‌ها متشکل از دو بخش است، بخش اول تهیه داده‌ها و بخش دوم مدل‌سازی داده‌ها و ارزیابی است. نویسندگان انتخاب ID3، بیزین و ماشین بردار پشتیبان را برای طبقه‌بندی ارزیابی احساسات آموزشی انتخاب کردند. نتایج تجربی آن‌ها نشان می‌دهد که SVM دارای بالاترین دقت است. در این کار، نمره وزن توسط یک متخصص تعریف شده است که دارای تجربه در ارزیابی تدریس است. وزن احساسات از -۱،۰۰ به +۱،۰۰ متغیر بود. روش پیشنهادی آنها نمی‌تواند واژه تقویت‌کننده را در نظر بگیرد و در زبان تایلندی ساخته شده است [۱۹].

Hee Yong Youn و Yili Wang روش وزن‌دهی تطبیقی جدید به نام میان‌رده قدرت تشخیصی با مدل Chi-square (CDS-C) پیشنهاد کرده‌اند که به مقیاس صحیح میان‌رده و اهمیت داخل دسته‌ای را مشخص می‌کند. یک مدل ریاضی جدید به نام ( $CDS^38$ ) برای اندازه‌گیری قدرت تشخیصی میان‌رده‌ها از ویژگی‌ها ارائه کرده‌اند، در حالی که یک مدل Chi-square اصلاح شده برای اندازه‌گیری وابستگی ویژگی‌ها در داخل دسته‌بندی ارائه شده است. هم‌چنین استراتژی خوشه‌بندی ریزنمونه برای تعیین حاکمیت ویژگی‌های تشخیصی و تعیین ویژگی‌های توزیع مشابه برای وزن‌دهی کارایی پیشنهاد شده است. علاوه بر این، یک استراتژی وزن سازگاری پیشنهاد شده است تا وزن هر ویژگی را به درستی تعیین کند. به طور خاص، توزیع یک کلمه در یک دسته خاص نشان‌دهنده وابستگی آن به آن دسته است. با توجه به وابستگی وزن، دقیق‌تر می‌تواند باشد. از آنجا که طرح پیشنهادی بر اساس وابستگی هر یک از ویژگی به رده مربوطه است، طول مدت جمله به طور مستقیم بر عملکرد روش پیشنهادی تاثیر نمی‌گذارد و همان‌طور که هر کلمه دارای وزن مشخصی است منعکس‌کننده درجه اهمیت آن برای دسته‌های مختلف است، قطبیت یک جمله دارای دو دیدگاه مختلف است که توسط وزن‌های به دست آمده از طبقه‌بندی متن  $MNB^39$  تعیین می‌شود. برای عقاید مختلف یا متضاد ایجاد نشده است. در طرح

پیشنهادی شکلک‌ها با استفاده از تابعی در نرم‌افزار متلب شناسایی و حذف می‌شوند در حالی که احساسات به طور معمول احساسات مردم را دربرمی‌گیرند و این روش مناسب نیست [۲۰].

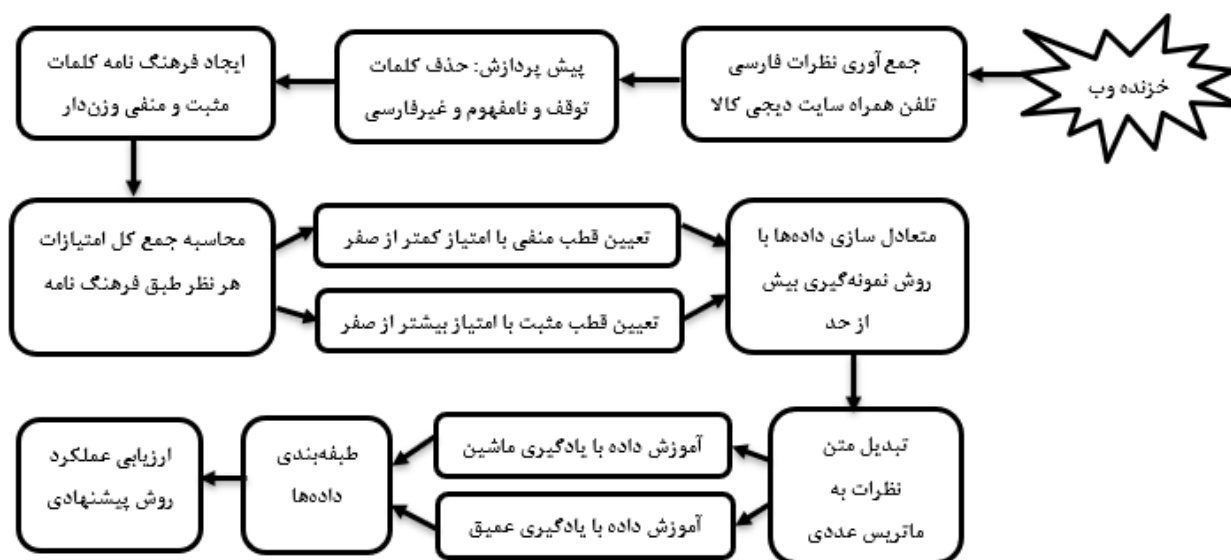
### ۳ مجموعه داده

در این پژوهش، مجموعه داده‌های فارسی نظرات تلفن همراه سایت دیجی کالا برای تحلیل احساسات در نظر گرفته شده است. این مجموعه داده را با استفاده از خزنده وب<sup>۴۰</sup> جمع‌آوری کرده‌ایم که شامل ۳۴۶۲۵ نظرسنجی با برچسب مثبت و ۵۰۸۵ نظرسنجی با برچسب منفی است. در این پژوهش برای بهبودی نظرکاوی یک فرهنگ‌نامه از کلمات مثبت و منفی داده‌های نظرات تلفن همراه ایجاد کرده‌ایم و به هر یک از کلمات با توجه به ارزش آن‌ها وزنی اختصاص داده‌ایم. این فرهنگ‌نامه شامل ۱۴۰۹۰ لغت با وزن‌های مختلف از ۵- تا ۵+ می‌باشد که دارای ۷۲۷۰ کلمه مثبت و ۶۸۲۰ کلمه منفی است. مجموعه داده بعد از به دست آوردن قطبیت جدید نظرات طبق فرهنگ‌نامه ارائه شده دارای ۳۱۸۹۵ نظرسنجی با برچسب مثبت و ۶۶۱۵ نظرسنجی با برچسب منفی می‌باشد. سپس به دلیل ناهمگن بودن داده‌های مثبت و منفی با استفاده از روش نمونه‌گیری بیش از حد، کل داده‌ها را به بالاترین مقدار سوق می‌دهیم و داده‌های با برچسب مثبت و منفی هر یک دارای ۳۱۸۹۵ نظرسنجی می‌شوند. بعد از آماده‌سازی داده‌ها از ۷۰ درصد آن‌ها برای آموزش و ۳۰ درصد برای آزمایش استفاده کرده‌ایم. داده‌های آزمایش از روش ۱۰ دور<sup>۴۱</sup> استفاده می‌کند یعنی داده‌های آزمایش در ۱۰ جای مختلف از کل داده آزمایش می‌شود و میانگین دقت را برمی‌گرداند.

### ۴ رویکرد پیشنهادی

نظرات در سطح جمله به دلیل این که ممکن است در مورد دو یا چند موضوع مختلف بحث شده باشد و چون در زبان فارسی فرهنگ‌نامه مناسبی جهت تعیین قطبیت وجود ندارد در نتیجه نمی‌توان قطبیت مناسبی برای جمله در نظر بگیریم. در این پژوهش نظرات فارسی تلفن همراه از سایت دیجی کالا پردازش شده است و روشی ارائه داده شده است که با حذف کلمات اضافی و نامفهوم، کلمات توقف، علائم و اعداد و حروف غیرفارسی و حروف تکراری، متن هر نظر را تا جای ممکن پاک‌سازی کرده و کاهش داده‌ایم. سپس فرهنگ‌نامه‌ای از واژگان کل داده‌ها با اختصاص یک امتیاز بین ۵- تا ۵+ به هر یک از آن‌ها ایجاد کرده‌ایم و طبق این فرهنگ‌نامه امتیاز کل کلمات هر نظر را محاسبه می‌کنیم تا قطبیت نظرات را به درستی تشخیص دهیم و نظرکاوی در سطح جملات فارسی را بهبود ببخشیم.

بعد از متعادل‌سازی داده‌های مثبت و منفی به کمک روش نمونه‌گیری بیش از حد داده‌های متنی به کمک تکنیک بردار ویژگی که در هر سند با استفاده از شمارش تعداد دفعاتی که هر کلمه از این فرهنگ لغت در آن ظاهر شده است را به ماتریس عددی تبدیل می‌کنیم. علاوه بر این، آموزش مجموعه داده‌ها با استفاده از الگوریتم‌های یادگیری ماشین از جمله بیز ساده، ماشین بردار پشتیبان، تبدلات گرادین تصادفی، برگشت منطقی و جنگل تصادفی و یادگیری عمیق مانند شبکه عصبی کانولوشن و حافظه طولانی کوتاه مدت انجام می‌شود. مراحل رویکرد پیشنهادی در شکل ۱ مورد بحث قرار گرفته است.



شکل ۱- نمایی از رویکرد پیشنهادی

**مرحله ۱:** مجموعه داده‌های فارسی نظرات تلفن همراه سایت دیجی کالا با استفاده از خزنده وب به زبان پایتون جمع‌آوری کرده‌ایم که شامل ۳۴۶۲۵ نظرسنجی با برچسب مثبت و ۵۰۸۵ نظرسنجی با برچسب منفی است و برای بهبود نظرکاوی در سطح جملات فارسی مورد توجه قرار گرفته است.

**مرحله ۲:** نظرات گاهی اوقات از داده‌های پوچ و بی‌معنا هستند که باید قبل از طبقه‌بندی آن‌ها حذف شوند. برای پیش‌پردازش متون نظرات تلفن همراه سایت دیجی کالا عملیات زیر انجام شده است:

- ✓ حذف اعداد و علائم و کلمات غیرحروف فارسی مانند علائم نقطه‌گذاری، تعجب، اصطلاحات انگلیسی و غیره.
- ✓ حذف کلمات پایانی مانند اگر، به، چون و غیره که آن‌ها در تعیین احساسات نقش مهمی ایفا نمی‌کنند.
- ✓ حذف حروف تکراری بیش از یک‌بار در کلمه به طور مثال به جای کلماتی چون عالللاللی کلمه عالی جایگزین می‌شود.
- ✓ حذف کلمات یک حرفی که در جمله بی‌معنا هستند.
- ✓ نرمال‌سازی متن که به منظور حذف فاصله‌ها می‌باشد.

**مرحله ۳:** پس از پیش‌پردازش نظرات، از کلمات باقی‌مانده کل داده‌ها یک فرهنگ‌نامه‌ای ایجاد می‌کنیم و به هر یک از کلمات با توجه به کاربرد آن‌ها امتیازی بین ۵- تا ۵+ می‌دهیم. در واقع کلمات خیلی ضعیف منفی با نمره (۵-)، کلمات ضعیف منفی با نمره (۴-)، خصوصیات منفی تلفن همراه با نمره (۳-)، کلمات متوسط منفی با نمره (۲-)، کلمات تقریباً نادرست منفی با نمره (۱-)، کلمات نامفهوم با نمره (۰)، کلمات تقریباً نادرست مثبت با نمره (۱)، کلمات متوسط مثبت با نمره (۲)، خصوصیات مثبت تلفن همراه با نمره (۳)، کلمات قوی مثبت با نمره (۴) و کلمات خیلی قوی مثبت با نمره (۵) امتیازدهی می‌شوند.

**مرحله ۴:** در این مرحله، هر نظر را به طور جداگانه به صورت کلمه به کلمه تفکیک کرده و امتیاز هر کلمه را از فرهنگ‌نامه ساخته‌شده در مرحله قبل پیدا می‌کنیم و امتیاز همه کلمات یک نظر را با هم جمع می‌کنیم اگر این مجموع کل امتیازات یک نظر بیشتر از صفر باشد آن نظر برچسب مثبت و اگر کمتر از صفر باشد برچسب منفی را به خود اختصاص می‌دهد. بنابراین کل مجموعه داده به این روش مجدد برچسب‌گذاری می‌شوند.

**مرحله ۵:** در ابتدا مجموعه داده نظرات مثبت و منفی استخراجی از سایت دیجی کالا در این تحقیق یکسان نمی‌باشد و برای این کار از روش‌های متعادل‌سازی از جمله روش نمونه‌گیری بیش از حد و روش نمونه‌گیری کمتر از حد برای بهبودی عملکرد طبقه‌بندی استفاده می‌کنیم. در این پژوهش از روش نمونه‌گیری بیش از حد استفاده شده است زیرا این روش برای مجموعه داده‌های کوچک مناسب است و مناسب‌ترین استراتژی است که مجموعه اولیه نمونه را افزایش می‌دهد و این کار بر روی وظایف دودویی یا چند ستونی تمرکز دارد.

**مرحله ۶:** برای استفاده از الگوریتم‌های یادگیری ماشین و یادگیری عمیق برای ارزیابی صحت روش پیشنهادی داده‌های متنی باید به ماتریس بردارهای عددی تبدیل شوند. در این تحقیق با استفاده از روش بردار ویژگی که اهمیت کلمه را به سند نشان می‌دهد برای ایجاد ماتریس عددی تبدیل کرده‌ایم.

**مرحله ۷:** پس از تبدیل متن به ماتریس اعداد، این ماتریس‌ها را با پنج الگوریتم مختلف یادگیری ماشین تحت نظارت از جمله بیز ساده، ماشین بردار پشتیبانی، تبادلات گرادیان تصادفی، برگشت منطقی و جنگل تصادفی و همچنین الگوریتم‌های یادگیری عمیق مانند شبکه عصبی کانولوشن و حافظه طولانی کوتاه مدت برای اهداف طبقه‌بندی آموزش و آزمایش برای رسیدن به نتایج بهتر استفاده می‌شوند.

**مرحله ۸:** مجموعه نظرات فارسی تلفن همراه سایت دیجی کالا با استفاده از الگوریتم‌های یادگیری ماشین و یادگیری عمیق که در مرحله قبل بیان شد طبقه‌بندی شده‌اند. سپس نتایج پارامترهای ارزیابی با استفاده از الگوریتم‌های یادگیری ماشین توسط تکنیک‌های n-gram از جمله تک‌واژه‌ای (جدول ۱)، دوواژه‌ای (جدول ۲)، سه‌واژه‌ای (جدول ۳)، ترکیب تک‌واژه‌ای و دوواژه‌ای (جدول ۴)، ترکیب دوواژه‌ای و سه‌واژه‌ای (جدول ۵) و ترکیب هر سه تکنیک (جدول ۶) و از الگوریتم‌های یادگیری عمیق روش CNN (جدول ۱۵) و روش LSTM (جدول ۱۶) نشان داده شده است.

## ۵ ارزیابی عملکرد رویکرد پیشنهادی

جدول ۱- نتایج پارامترهای ارزیابی عملکرد برای تک‌واژه‌ای مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

الگوریتم	ماتریس درهم ریختگی	پارامترهای ارزیابی	دقت
بیز ساده	مثبت منفی	صحت	۰.۸۴۷
مثبت	۱۶۸۵۲	۰.۳۵۶	۰.۲۸۳
منفی	۶۳۶	۰.۶۰۵	۰.۸۲۱
ماشین بردار پشتیبانی	مثبت منفی	صحت	۰.۹۷۰
		فراخوان	اندازه F



مثبت	۲۵۱۷۹	۲۱۳۸	۰,۳۹۸	۰,۳۵۹	۰,۳۷۸
منفی	۳۷۲	۹۱۴۳	۰,۷۹۴	۰,۹۹۷	۰,۹۳۶
تبادلات گرادیان تصادفی	مثبت	منفی	صحت	فراخوان	اندازه F
مثبت	۲۲۴۹۵	۴۸۲۲	۰,۳۹۸	۰,۳۳۸	۰,۳۶۱
منفی	۳۴۰	۹۱۷۵	۰,۷۷۲	۰,۹۹۱	۰,۹۱۳
برگشت منطقی	مثبت	منفی	صحت	فراخوان	اندازه F
مثبت	۲۳۶۵۲	۳۶۶۵	۰,۳۹۹	۰,۳۴۳	۰,۳۶۹
منفی	۲۹۴	۹۲۲۱	۰,۸۳۲	۰,۹۹۷	۰,۹۱۸
جنگل تصادفی	مثبت	منفی	صحت	فراخوان	اندازه F
مثبت	۲۰۶۴۳	۶۶۷۴	۰,۳۹۶	۰,۹۹۶	۰,۳۰۷
منفی	۳۰۶	۹۲۰۹	۰,۶۷۳	۰,۲۵۲	۰,۸۵۱

ماتریس درهم ریختگی و پارامترهای مختلف ارزیابی مانند صحت، فراخوان، اندازه F و دقت به دست آمده پس از اعمال روش پیشنهادی و طبقه‌بندی با تکنیک تک‌واژه‌ای و استفاده از الگوریتم‌های مختلف یادگیری ماشین در جدول ۱ نشان داده شده است. همان‌طور که در جدول ۱ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده با تکنیک تک‌واژه‌ای در الگوریتم ماشین بردار پشتیبان بین ۲ تا ۱۳ درصد بهتر از مقدار حاصل شده در الگوریتم‌های دیگر است. روش تک‌واژه‌ای به دلیل این که به صورت کلمه به کلمه پردازش می‌شود در نتیجه نسبت به روش‌های دیگر n-gram نتایج بهتری را به دست می‌آورد.

جدول ۲- نتایج پارامترهای ارزیابی عملکرد برای دوواژه‌ای مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

الگوریتم	ماتریس درهم ریختگی	پارامترهای ارزیابی	دقت
بیز ساده	مثبت ۱۹۰۲۴ منفی ۱۱۲۳	صحت ۰,۳۶۷ فراخوان ۰,۲۴۷ اندازه F ۰,۲۹۳	۰,۸۷۰
ماشین بردار پشتیبان	مثبت ۲۰۸۸۴ منفی ۹۰۰	صحت ۰,۳۹۳ فراخوان ۰,۲۶۷ اندازه F ۰,۳۱۸	۰,۹۰۵
تبادلات گرادیان تصادفی	مثبت ۱۷۱۵۲ منفی ۴۶۳	صحت ۰,۳۹۰ فراخوان ۰,۲۵۱ اندازه F ۰,۳۰۶	۰,۸۹۲
برگشت منطقی	مثبت ۱۷۳۵۶ منفی ۳۵۳	صحت ۰,۲۳۲ فراخوان ۰,۲۳۱ اندازه F ۰,۲۹۱	۰,۸۸۰
جنگل تصادفی	مثبت ۱۹۸۵۶ منفی ۸۶۱	صحت ۰,۳۸۹ فراخوان ۰,۲۲۱ اندازه F ۰,۲۸۵	۰,۸۷۴

ماتریس درهم ریختگی و پارامترهای مختلف ارزیابی مانند صحت، فراخوان، اندازه F و دقت به دست آمده پس از اعمال روش پیشنهادی و طبقه‌بندی با تکنیک دوواژه‌ای و استفاده از الگوریتم‌های مختلف یادگیری ماشین در جدول ۲ نشان داده شده است. همان‌طور که در جدول ۲ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده با تکنیک دوواژه‌ای در الگوریتم ماشین بردار پشتیبان بین ۱ تا ۳ درصد بهتر از مقدار حاصل شده در الگوریتم‌های دیگر است. روش دوواژه‌ای به دلیل این که به صورت دو کلمه‌ای پردازش می‌شود در نتیجه نسبت به روش‌های دیگر n-gram به غیر از تک‌واژه‌ای نتایج بهتری را به دست می‌آورد.

جدول ۳- نتایج پارامترهای ارزیابی عملکرد برای سه‌واژه‌ای مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

الگوریتم	ماتریس درهم ریختگی	پارامترهای ارزیابی	دقت
بیز ساده	مثبت	صحت	۰,۸۱۲
	مثبت	فراخوان	۰,۲۱۱
	مثبت	اندازه F	۰,۷۹۹
ماشین بردار پشتیبان	مثبت	صحت	۰,۸۲۵
	مثبت	فراخوان	۰,۲۱۳
	مثبت	اندازه F	۰,۸۰۷
تبادلات گرادیان تصادفی	مثبت	صحت	۰,۷۹۳
	مثبت	فراخوان	۰,۱۵۰
	مثبت	اندازه F	۰,۷۹۴
برگشت منطقی	مثبت	صحت	۰,۷۸۷
	مثبت	فراخوان	۰,۱۵۱
	مثبت	اندازه F	۰,۷۹۲
جنگل تصادفی	مثبت	صحت	۰,۸۲۵
	مثبت	فراخوان	۰,۲۱۱
	مثبت	اندازه F	۰,۸۰۷

ماتریس درهم ریختگی و پارامترهای مختلف ارزیابی مانند صحت، فراخوان، اندازه F و دقت به دست آمده پس از اعمال روش پیشنهادی و طبقه‌بندی با تکنیک سه‌واژه‌ای و استفاده از الگوریتم‌های مختلف یادگیری ماشین در جدول ۳ نشان داده شده است. همان‌طور که در جدول ۳ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده با تکنیک سه‌واژه‌ای در الگوریتم‌های ماشین بردار پشتیبان و جنگل تصادفی بین ۱ تا ۴ درصد بهتر از مقدار حاصل شده در الگوریتم‌های دیگر هستند. روش سه‌واژه‌ای به دلیل این که به صورت سه کلمه‌ای پردازش می‌شود و کلمات چند بار تکرار می‌شوند؛ بنابراین، بر احتمال سند اثر می‌گذارد در نتیجه نسبت به روش‌های دیگر n-gram دقت طبقه‌بندی را کاهش می‌دهد.

جدول ۴- نتایج پارامترهای ارزیابی عملکرد برای تک‌واژه‌ای و دوواژه‌ای مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

الگوریتم	ماتریس درهم ریختگی	پارامترهای ارزیابی	دقت
بیز ساده	مثبت	صحت	۰,۸۹۶
	مثبت	فراخوان	۰,۳۲۰
	مثبت	اندازه F	۰,۸۵۵
ماشین بردار پشتیبان	مثبت	صحت	۰,۹۶۶
	مثبت	فراخوان	۰,۳۷۳
	مثبت	اندازه F	۰,۹۲۹
تبادلات گرادیان تصادفی	مثبت	صحت	۰,۹۶۴
	مثبت	فراخوان	۰,۳۶۸
	مثبت	اندازه F	۰,۹۱۷
برگشت منطقی	مثبت	صحت	۰,۹۵۵
	مثبت	فراخوان	۰,۳۶۴
	مثبت	اندازه F	۰,۳۳۶



منفی	۲۸۷	۹۲۸۸	۰,۸۲۰	۰,۹۹۷	۰,۹۱۲
جنگل تصادفی	مثبت	منفی	صحت	فراخوان	اندازه F
مثبت	۲۰۰۸۶	۷۱۷۱	۰,۳۹۵	۰,۲۲۲	۰,۲۸۶
منفی	۲۱۹	۹۳۵۶	۰,۷۲۰	۰,۹۹۷	۰,۷۳۷

ماتریس درهم ریختگی و پارامترهای مختلف ارزیابی مانند صحت، فراخوان، اندازه F و دقت به دست آمده پس از اعمال روش پیشنهادی و طبقه‌بندی با ترکیب تکنیک‌های تک‌واژه‌ای و دوواژه‌ای و استفاده از الگوریتم‌های مختلف یادگیری ماشین در جدول ۴ نشان داده شده است. همان‌طور که در جدول ۴ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده با تکنیک‌های تک‌واژه‌ای و دوواژه‌ای در الگوریتم‌های ماشین بردار پشتیبان و تبادلات گرادیان تصادفی بین ۱ تا ۹ درصد بهتر از مقدار حاصل شده در الگوریتم‌های دیگر هستند. روش ترکیبی تک‌واژه‌ای و دوواژه‌ای به دلیل وجود روش تک‌واژه‌ای و دوواژه‌ای در نتیجه نسبت به روش‌های دیگر n-gram نیز نتایج بهتری را به دست می‌آورند.

جدول ۵- نتایج پارامترهای ارزیابی عملکرد برای دوواژه‌ای و سه‌واژه‌ای مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

الگوریتم	ماتریس درهم ریختگی	پارامترهای ارزیابی	دقت
بیز ساده	مثبت منفی	صحت فراخوان	اندازه F ۰,۸۷۷
	مثبت منفی	۰,۳۶۷ ۰,۴۵۰	۰,۳۰۱ ۰,۸۳۹
ماشین بردار پشتیبان	مثبت منفی	صحت فراخوان	اندازه F ۰,۹۰۲
	مثبت منفی	۰,۳۹۲ ۰,۴۸۸	۰,۳۱۶ ۰,۸۵۵
تبادلات گرادیان تصادفی	مثبت منفی	صحت فراخوان	اندازه F ۰,۸۹۱
	مثبت منفی	۰,۳۹۱ ۰,۵۶۸	۰,۵۷۸ ۰,۸۴۷
برگشت منطقی	مثبت منفی	صحت فراخوان	اندازه F ۰,۸۷۸
	مثبت منفی	۰,۳۹۲ ۰,۶۳۴	۰,۲۹۱ ۰,۸۳۷
جنگل تصادفی	مثبت منفی	صحت فراخوان	اندازه F ۰,۸۷۰
	مثبت منفی	۰,۳۹۴ ۰,۳۹۹	۰,۲۷۵ ۰,۸۳۰

ماتریس درهم ریختگی و پارامترهای مختلف ارزیابی مانند صحت، فراخوان، اندازه F و دقت به دست آمده پس از اعمال روش پیشنهادی و طبقه‌بندی با ترکیب تکنیک‌های دوواژه‌ای و سه‌واژه‌ای و استفاده از الگوریتم‌های مختلف یادگیری ماشین در جدول ۵ نشان داده شده است. همان‌طور که در جدول ۵ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده با تکنیک‌های دوواژه‌ای و سه‌واژه‌ای در الگوریتم ماشین بردار پشتیبان بین ۱ تا ۳ درصد بهتر از مقدار حاصل شده در الگوریتم‌های دیگر است. روش ترکیبی دوواژه‌ای و سه‌واژه‌ای به دلیل تاثیر روش سه‌واژه‌ای باعث می‌شود که مقدار دقت نسبت به روش‌های دیگر n-gram نیز نتایج ضعیف‌تری را به دست آورد.

جدول ۶- نتایج پارامترهای ارزیابی عملکرد برای هر سه تکنیک مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

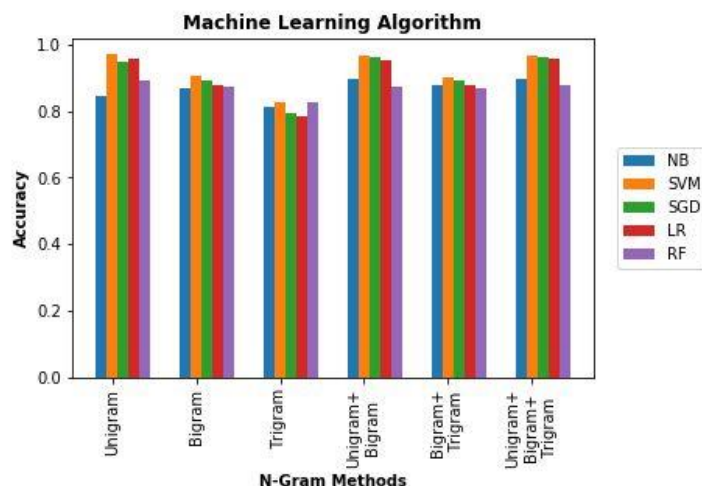
الگوریتم	ماتریس درهم ریختگی	پارامترهای ارزیابی	دقت
بیز ساده	مثبت منفی	صحت فراخوان	اندازه F ۰,۸۹۹
	مثبت منفی	۰,۳۷۷ ۰,۵۸۳	۰,۳۲۳ ۰,۸۵۸

ماشین بردار پشتیبان	مثبت	مثبت	منفی	صحت	فراخوان	اندازه F	۰,۹۶۷
	مثبت	۲۴۸۲۹	۲۴۶۱	۰,۳۹۷	۰,۳۵۳	۰,۳۷۴	
	منفی	۳۹۹	۹۱۴۳	۰,۷۶۴	۰,۹۹۷	۰,۹۳۱	
تبادلات گرادیان تصادفی	مثبت	۲۳۳۰۶	۳۹۸۴	۰,۳۹۷	۰,۳۴۹	۰,۳۶۹	۰,۹۶۲
	منفی	۲۵۷	۹۲۸۵	۰,۷۷۸	۰,۹۹۶	۰,۹۲۲	
برگشت منطقی	مثبت	۲۴۱۶۶	۳۱۲۴	۰,۳۹۷	۰,۳۴۰	۰,۳۶۶	۰,۹۵۷
	منفی	۲۸۹	۹۲۵۳	۰,۸۱۳	۰,۹۹۸	۰,۹۱۵	
جنگل تصادفی	مثبت	۱۹۸۲۶	۷۴۶۴	۰,۳۹۶	۰,۲۲۲	۰,۲۸۷	۰,۸۷۷
	منفی	۲۱۲	۹۳۳۰	۰,۷۱۴	۰,۹۹۷	۰,۸۳۶	

ماتریس درهم ریختگی و پارامترهای مختلف ارزیابی مانند صحت، فراخوان، اندازه F و دقت به دست آمده پس از اعمال روش پیشنهادی و طبقه‌بندی با ترکیب تکنیک‌های تک‌واژه‌ای و دوواژه‌ای و سه‌واژه‌ای و استفاده از الگوریتم‌های مختلف یادگیری ماشین در جدول ۶ نشان داده شده است. همان‌طور که در جدول ۶ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده با سه تکنیک در الگوریتم‌های ماشین بردار پشتیبان و تبادلات گرادیان تصادفی و برگشت منطقی بین ۶ تا ۹ درصد بهتر از مقدار حاصل شده در الگوریتم‌های دیگر هستند. روش ترکیبی تک‌واژه‌ای و دوواژه‌ای و سه‌واژه‌ای به دلیل تاثیر روش سه‌واژه‌ای باعث می‌شود که مقدار دقت نسبت به روش‌های دیگر n-gram نیز نتایج ضعیف‌تر و تاثیر روش تک‌واژه‌ای و دوواژه‌ای نتایج بهتری را به دست آورند. بنابراین نتایج خوبی را به ارمغان می‌آورند.

جدول ۷- نتایج دقت روش‌های n-gram مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

روش n-gram	بیز ساده	ماشین بردار پشتیبان	تبادلات گرادیان تصادفی	برگشت منطقی	جنگل تصادفی
تک‌واژه‌ای	۰,۸۴۷	۰,۹۷۰	۰,۹۵۱	۰,۹۵۸	۰,۸۹۲
دوواژه‌ای	۰,۸۷۰	۰,۹۰۵	۰,۸۹۲	۰,۸۸۰	۰,۸۷۴
سه‌واژه‌ای	۰,۸۱۲	۰,۸۲۵	۰,۷۹۳	۰,۷۸۷	۰,۸۲۵
تک‌واژه‌ای و دوواژه‌ای	۰,۸۹۶	۰,۹۶۶	۰,۹۶۴	۰,۹۵۵	۰,۸۷۵
دوواژه‌ای و سه‌واژه‌ای	۰,۸۷۷	۰,۹۰۲	۰,۸۹۱	۰,۸۷۸	۰,۸۷۰
هر سه تکنیک	۰,۸۹۹	۰,۹۶۷	۰,۹۶۲	۰,۹۵۷	۰,۸۷۷

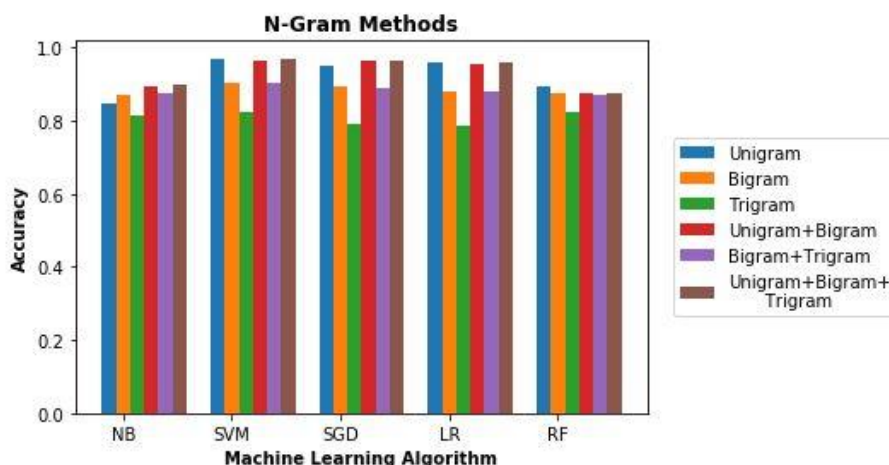


شکل ۲- نتایج دقت روش‌های n-gram مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دقت به دست آمده پس از اعمال روش پیشنهادی و طبقه‌بندی با تکنیک‌های روش n-gram و استفاده از الگوریتم‌های مختلف یادگیری ماشین در جدول ۷ و شکل ۲ نشان داده شده است. همان‌طور که در جدول ۷ و شکل ۲ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده در الگوریتم‌های ماشین بردار پشتیبان در همه روش‌های n-gram به غیر از روش سه واژه‌ای به میزان ۷ درصد و تبدلات گرادیان تصادفی و برگشت منطقی در همه روش‌های n-gram به غیر از دو واژه‌ای و سه واژه‌ای و روش ترکیبی دو واژه‌ای و سه واژه‌ای به میزان ۵ درصد بهتر از مقدار حاصل شده در الگوریتم‌های دیگر در روش‌های n-gram هستند. تاثیر روش سه واژه‌ای باعث می‌شود که مقدار دقت نسبت به روش‌های دیگر n-gram در الگوریتم‌های مختلف نیز نتایج ضعیف‌تری را به دست آورند.

جدول ۸- نتایج دقت الگوریتم‌های یادگیری ماشین مجموعه داده با استفاده از روش‌های n-gram

الگوریتم	تک‌واژه‌ای	دو واژه‌ای	سه واژه‌ای	تک‌واژه‌ای و دو واژه‌ای	دو واژه‌ای و سه واژه‌ای	هر سه تکنیک
بیز ساده	۰.۸۴۷	۰.۸۷۰	۰.۸۱۲	۰.۸۹۶	۰.۸۷۷	۰.۸۹۹
ماشین بردار پشتیبان	۰.۹۷۰	۰.۹۰۵	۰.۸۲۵	۰.۹۶۶	۰.۹۰۲	۰.۹۶۷
تبدلات گرادیان تصادفی	۰.۹۵۱	۰.۸۹۲	۰.۷۹۳	۰.۹۶۴	۰.۸۹۱	۰.۹۶۲
برگشت منطقی	۰.۹۵۸	۰.۸۸۰	۰.۷۸۷	۰.۹۵۵	۰.۸۷۸	۰.۹۵۷
جنگل تصادفی	۰.۸۹۲	۰.۸۷۴	۰.۸۲۵	۰.۸۷۵	۰.۸۷۰	۰.۸۷۷



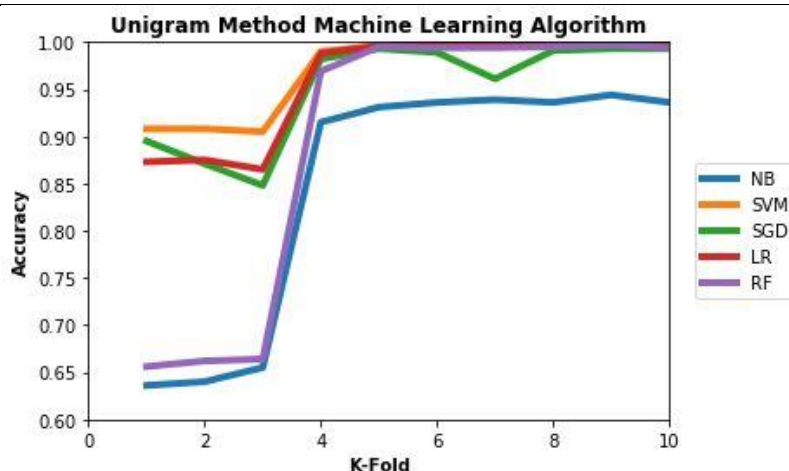
شکل ۳- نتایج دقت الگوریتم‌های یادگیری ماشین مجموعه داده با استفاده از روش‌های n-gram

دقت به دست آمده پس از اعمال روش پیشنهادی و الگوریتم‌های مختلف یادگیری ماشین و استفاده از طبقه‌بندی با تکنیک‌های روش n-gram در جدول ۸ و شکل ۳ نشان داده شده است. همان‌طور که در جدول ۸ و شکل ۳ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده در الگوریتم‌های ماشین بردار پشتیبان در همه روش‌های n-gram به غیر از روش سه واژه‌ای به میزان ۷ درصد و تبدلات گرادیان تصادفی و برگشت منطقی در همه روش‌های n-gram به غیر از دو واژه‌ای و سه واژه‌ای و روش ترکیبی دو واژه‌ای و سه واژه‌ای به میزان ۵ درصد بهتر از مقدار حاصل شده در الگوریتم‌های دیگر در روش‌های n-gram هستند. تاثیر روش سه واژه‌ای باعث می‌شود که مقدار دقت نسبت به روش‌های دیگر n-gram در الگوریتم‌های مختلف نیز نتایج ضعیف‌تری را به دست آورند.

جدول ۹- نتایج دقت روش تک‌واژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دور	بیز ساده	ماشین بردار پشتیبان	تبدلات گرادیان تصادفی	برگشت منطقی	جنگل تصادفی
دور اول	۰.۶۳۶	۰.۹۰۸	۰.۸۹۵	۰.۸۷۳	۰.۶۵۶
دور دوم	۰.۶۴۰	۰.۹۰۸	۰.۸۷۱	۰.۸۷۵	۰.۶۶۲
دور سوم	۰.۶۵۵	۰.۹۰۵	۰.۸۴۸	۰.۸۶۵	۰.۶۶۴
دور چهارم	۰.۹۱۵	۰.۹۹۰	۰.۹۸۲	۰.۹۸۸	۰.۹۶۹

دور پنجم	۰,۹۳۱	۰,۹۹۶	۰,۹۹۳	۰,۹۹۶
دور ششم	۰,۹۳۶	۰,۹۹۷	۰,۹۸۹	۰,۹۹۴
دور هفتم	۰,۹۳۹	۰,۹۹۸	۰,۹۶۱	۰,۹۹۴
دور هشتم	۰,۹۳۶	۰,۹۹۸	۰,۹۹۱	۰,۹۹۵
دور نهم	۰,۹۴۴	۰,۹۹۸	۰,۹۹۳	۰,۹۹۷
دور دهم	۰,۹۳۶	۰,۹۹۷	۰,۹۹۳	۰,۹۹۴

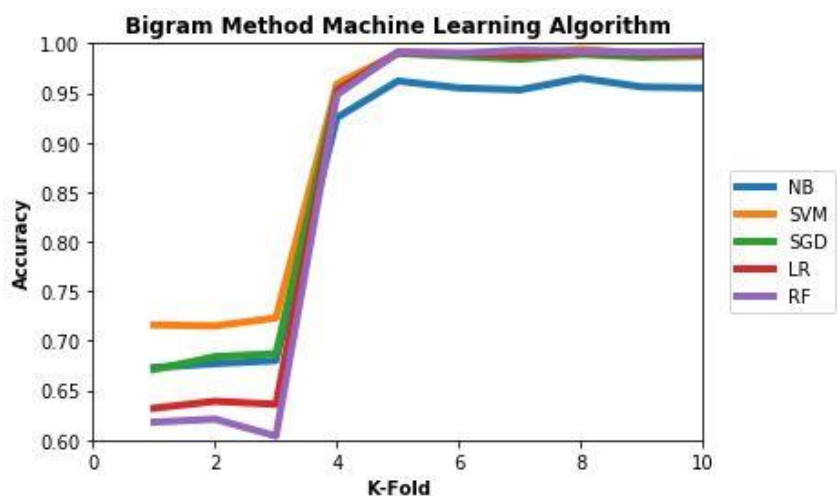


شکل ۴- نتایج دقت روش تک‌واژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دقت به دست آمده در روش تک‌واژه‌ای پس از اعمال روش پیشنهادی و با استفاده از الگوریتم‌های مختلف یادگیری ماشین با آزمایش در ۱۰ دور متفاوت در جدول ۹ و شکل ۴ نشان داده شده است. همان‌طور که در جدول ۹ و شکل ۴ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده در الگوریتم‌های ماشین بردار پشتیبان به میزان ۲۵ درصد و تبادلات گرادیان تصادفی به میزان ۲۴ درصد و برگشت منطقی به میزان ۲۲ درصد از دور اول و الگوریتم‌های بیز ساده به میزان ۲۶ درصد و جنگل تصادفی به میزان ۳۰ درصد از دور چهارم به سمت بالا بهبود یافته‌اند.

جدول ۱۰- نتایج دقت روش دوواژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دور	بیز ساده	ماشین بردار پشتیبان	تبادلات گرادیان تصادفی	برگشت منطقی	جنگل تصادفی
دور اول	۰,۶۷۳	۰,۷۱۶	۰,۶۷۱	۰,۶۳۲	۰,۶۱۸
دور دوم	۰,۶۷۷	۰,۷۱۵	۰,۶۸۴	۰,۶۳۹	۰,۶۲۱
دور سوم	۰,۶۸۰	۰,۷۲۳	۰,۶۸۷	۰,۶۳۶	۰,۶۰۴
دور چهارم	۰,۹۲۵	۰,۹۵۹	۰,۹۵۳	۰,۹۵۲	۰,۹۴۸
دور پنجم	۰,۹۶۲	۰,۹۹۰	۰,۹۹۰	۰,۹۹۱	۰,۹۹۱
دور ششم	۰,۹۵۵	۰,۹۸۸	۰,۹۸۷	۰,۹۹۰	۰,۹۹۰
دور هفتم	۰,۹۵۳	۰,۹۸۸	۰,۹۸۴	۰,۹۸۸	۰,۹۹۳
دور هشتم	۰,۹۶۵	۰,۹۹۴	۰,۹۸۹	۰,۹۹۱	۰,۹۹۲
دور نهم	۰,۹۵۶	۰,۹۹۰	۰,۹۸۶	۰,۹۹۰	۰,۹۹۱
دور دهم	۰,۹۵۵	۰,۹۹۰	۰,۹۸۷	۰,۹۸۹	۰,۹۹۲

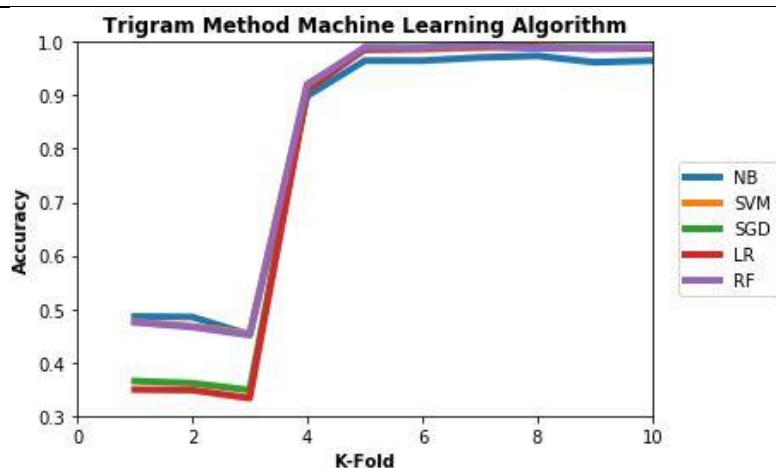


شکل ۵- نتایج دقت روش دوواژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دقت به دست آمده در روش دوواژه‌ای پس از اعمال روش پیشنهادی و با استفاده از الگوریتم‌های مختلف یادگیری ماشین با آزمایش در ۱۰ دور متفاوت در جدول ۱۰ و شکل ۵ نشان داده شده است. همان‌طور که در جدول ۱۰ و شکل ۵ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده در همه الگوریتم‌های یادگیری ماشین از دور چهارم بین ۲۳ تا ۳۴ درصد به سمت بالا بهبود یافته‌اند.

جدول ۱۱- نتایج دقت روش سه‌واژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دور	بیز ساده	ماشین بردار پشتیبان	تبادلات گرادیان تصادفی	برگشت منطقی	جنگل تصادفی
دور اول	۰,۴۸۶	۰,۴۷۶	۰,۳۵۶	۰,۳۴۹	۰,۴۷۵
دور دوم	۰,۴۸۵	۰,۴۶۸	۰,۳۶۱	۰,۳۴۸	۰,۴۶۶
دور سوم	۰,۴۵۲	۰,۴۵۴	۰,۳۴۸	۰,۳۳۳	۰,۴۵۱
دور چهارم	۰,۸۹۹	۰,۹۲۱	۰,۹۰۸	۰,۹۱۱	۰,۹۲۰
دور پنجم	۰,۹۶۵	۰,۹۸۶	۰,۹۹۰	۰,۹۸۶	۰,۹۹۰
دور ششم	۰,۹۶۵	۰,۹۸۶	۰,۹۹۰	۰,۹۸۸	۰,۹۸۹
دور هفتم	۰,۹۷۱	۰,۹۸۹	۰,۹۹۴	۰,۹۹۲	۰,۹۹۳
دور هشتم	۰,۹۷۴	۰,۹۹۰	۰,۹۹۰	۰,۹۹۰	۰,۹۸۸
دور نهم	۰,۹۶۲	۰,۹۸۹	۰,۹۹۰	۰,۹۸۷	۰,۹۸۸
دور دهم	۰,۹۶۵	۰,۹۸۸	۰,۹۸۹	۰,۹۸۸	۰,۹۹۰

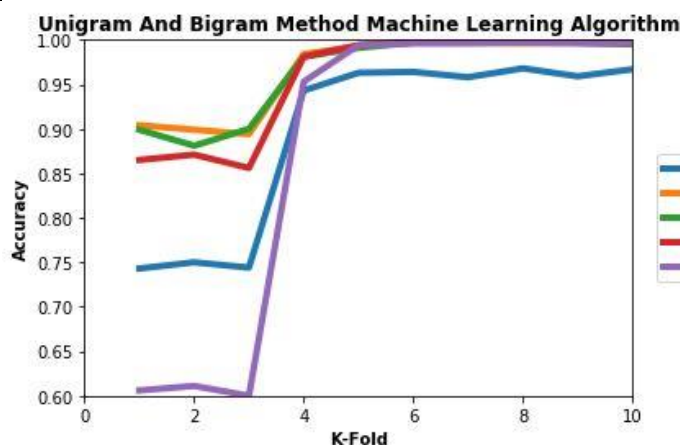


شکل ۶- نتایج دقت روش سه‌واژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دقت به دست آمده در روش سهواژه‌ای پس از اعمال روش پیشنهادی و با استفاده از الگوریتم‌های مختلف یادگیری ماشین با آزمایش در ۱۰ دور متفاوت در جدول ۱۱ و شکل ۶ نشان داده شده است. همان‌طور که در جدول ۱۱ و شکل ۶ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده در همه الگوریتم‌های یادگیری ماشین از دور چهارم بین ۴۷ تا ۵۸ درصد و الگوریتم بیز ساده از دور پنجم به میزان ۷ درصد به سمت بالا بهبود یافته‌اند.

جدول ۱۲- نتایج دقت روش تکواژه‌ای و دوواژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دور	بیز ساده	ماشین بردار پشتیبان	تبادلات گرادیان تصادفی	برگشت منطقی	جنگل تصادفی
دور اول	۰,۷۴۳	۰,۹۰۴	۰,۸۹۹	۰,۸۶۵	۰,۶۰۶
دور دوم	۰,۷۵۰	۰,۸۹۹	۰,۸۸۱	۰,۸۷۱	۰,۶۱۱
دور سوم	۰,۷۴۴	۰,۸۹۴	۰,۹۰۰	۰,۸۵۶	۰,۶۰۰
دور چهارم	۰,۹۴۳	۰,۹۸۴	۰,۹۸۱	۰,۹۸۱	۰,۹۵۳
دور پنجم	۰,۹۶۳	۰,۹۹۳	۰,۹۹۱	۰,۹۹۴	۰,۹۹۴
دور ششم	۰,۹۶۴	۰,۹۹۷	۰,۹۹۷	۰,۹۹۸	۰,۹۹۶
دور هفتم	۰,۹۵۸	۰,۹۹۷	۰,۹۹۸	۰,۹۹۸	۰,۹۹۶
دور هشتم	۰,۹۶۸	۰,۹۹۶	۰,۹۹۷	۰,۹۹۷	۰,۹۹۷
دور نهم	۰,۹۵۹	۰,۹۹۷	۰,۹۹۷	۰,۹۹۷	۰,۹۹۶
دور دهم	۰,۹۶۷	۰,۹۹۵	۰,۹۹۵	۰,۹۹۷	۰,۹۹۶



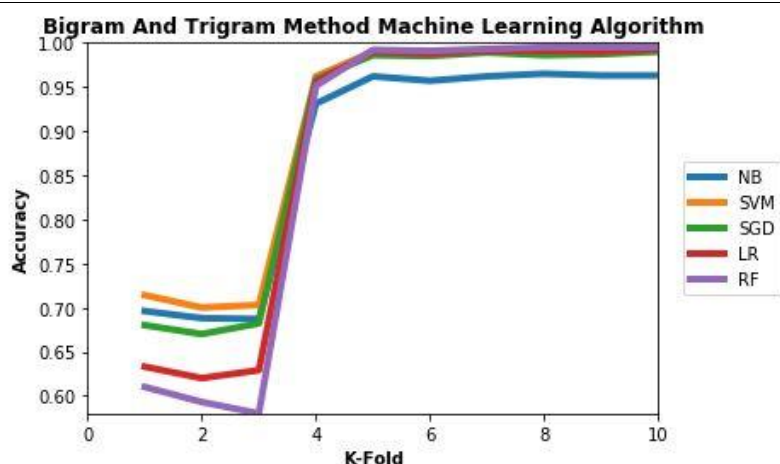
شکل ۷- نتایج دقت روش تکواژه‌ای و دوواژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دقت به دست آمده در روش ترکیبی تکواژه‌ای و دوواژه‌ای پس از اعمال روش پیشنهادی و با استفاده از الگوریتم‌های مختلف یادگیری ماشین با آزمایش در ۱۰ دور متفاوت در جدول ۱۲ و شکل ۷ نشان داده شده است. همان‌طور که در جدول ۱۲ و شکل ۷ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده در الگوریتم‌های ماشین بردار پشتیبان به میزان ۳۰ درصد و تبادلات گرادیان تصادفی به میزان ۲۹ درصد و برگشت منطقی به میزان ۲۶ درصد از دور اول و الگوریتم‌های بیز ساده به میزان ۲۰ درصد و جنگل تصادفی به میزان ۳۵ درصد از دور چهارم به سمت بالا بهبود یافته‌اند.

جدول ۱۳- نتایج دقت روش دوواژه‌ای و سهواژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دور	بیز ساده	ماشین بردار پشتیبان	تبادلات گرادیان تصادفی	برگشت منطقی	جنگل تصادفی
دور اول	۰,۶۹۶	۰,۷۱۴	۰,۶۸۰	۰,۶۳۳	۰,۶۱۰
دور دوم	۰,۶۸۸	۰,۷۰۰	۰,۶۷۰	۰,۶۲۰	۰,۵۹۳
دور سوم	۰,۶۸۷	۰,۷۰۳	۰,۶۸۲	۰,۶۲۹	۰,۵۸۰

دور چهارم	۰,۹۳۱	۰,۹۶۲	۰,۹۵۹	۰,۹۵۶	۰,۹۵۱
دور پنجم	۰,۹۶۲	۰,۹۹۰	۰,۹۸۶	۰,۹۹۰	۰,۹۹۲
دور ششم	۰,۹۵۷	۰,۹۸۹	۰,۹۸۵	۰,۹۸۸	۰,۹۹۱
دور هفتم	۰,۹۶۲	۰,۹۹۱	۰,۹۸۹	۰,۹۹۱	۰,۹۹۳
دور هشتم	۰,۹۶۵	۰,۹۹۲	۰,۹۸۶	۰,۹۹۱	۰,۹۹۵
دور نهم	۰,۹۶۳	۰,۹۹۱	۰,۹۸۷	۰,۹۹۰	۰,۹۹۵
دور دهم	۰,۹۶۳	۰,۹۹۱	۰,۹۹۰	۰,۹۹۳	۰,۹۹۵



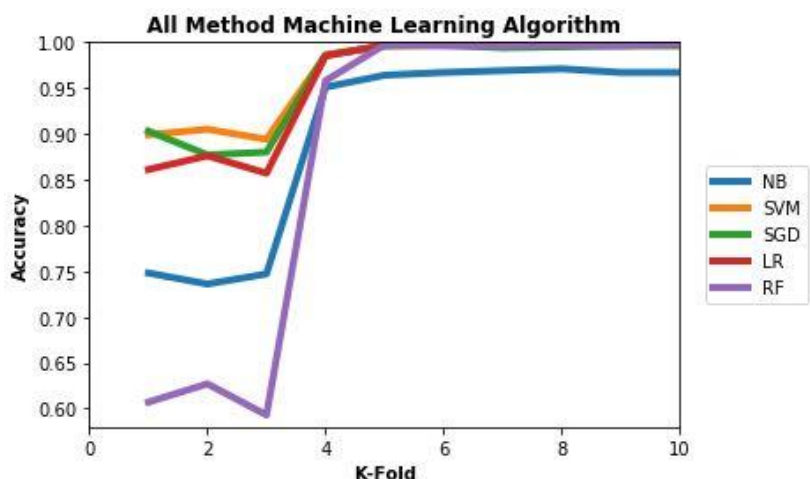
شکل ۸- نتایج دقت روش دوواژه‌ای و سه‌واژه‌ای روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دقت به دست آمده در روش ترکیبی دوواژه‌ای و سه‌واژه‌ای پس از اعمال روش پیشنهادی و با استفاده از الگوریتم‌های مختلف یادگیری ماشین با آزمایش در ۱۰ دور متفاوت در جدول ۱۳ و شکل ۸ نشان داده شده است. همان‌طور که در جدول ۱۳ و شکل ۸ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده در همه الگوریتم‌های یادگیری ماشین از دور چهارم بین ۲۵ تا ۳۲ درصد به سمت بالا بهبود یافته‌اند.

جدول ۱۴- نتایج دقت روش هر سه تکنیک، روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

دور	بیز ساده	ماشین بردار پشتیبان	تبادلات گرادیان تصادفی	برگشت منطقی	جنگل تصادفی
دور اول	۰,۷۴۸	۰,۸۹۹	۰,۹۰۳	۰,۸۶۱	۰,۶۰۷
دور دوم	۰,۷۳۶	۰,۹۰۵	۰,۸۷۷	۰,۸۷۶	۰,۶۲۷
دور سوم	۰,۷۴۷	۰,۸۹۴	۰,۸۸۰	۰,۸۵۷	۰,۵۹۳
دور چهارم	۰,۹۵۱	۰,۹۸۶	۰,۹۸۶	۰,۹۸۵	۰,۹۵۸
دور پنجم	۰,۹۶۴	۰,۹۹۶	۰,۹۹۶	۰,۹۹۷	۰,۹۹۷
دور ششم	۰,۹۶۷	۰,۹۹۷	۰,۹۹۷	۰,۹۹۹	۰,۹۹۶
دور هفتم	۰,۹۶۹	۰,۹۹۶	۰,۹۹۴	۰,۹۹۷	۰,۹۹۶
دور هشتم	۰,۹۷۱	۰,۹۹۶	۰,۹۹۵	۰,۹۹۷	۰,۹۹۷
دور نهم	۰,۹۶۷	۰,۹۹۹	۰,۹۹۶	۰,۹۹۸	۰,۹۹۶
دور دهم	۰,۹۶۷	۰,۹۹۷	۰,۹۹۶	۰,۹۹۸	۰,۹۹۸



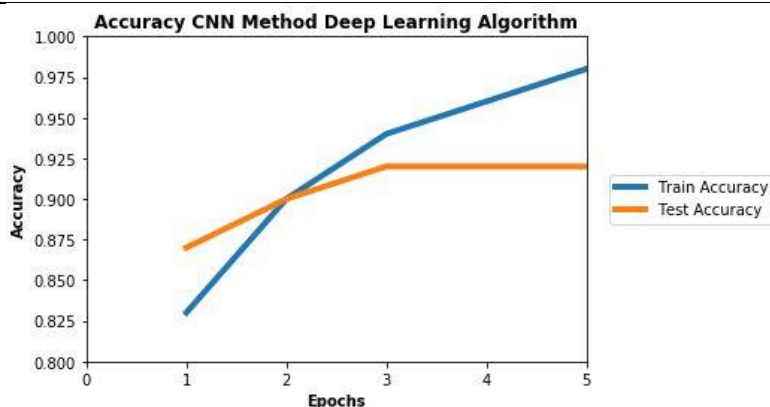


شکل ۹- نتایج دقت روش هر سه تکنیک، روند 10-Fold مجموعه داده با استفاده از الگوریتم‌های یادگیری ماشین

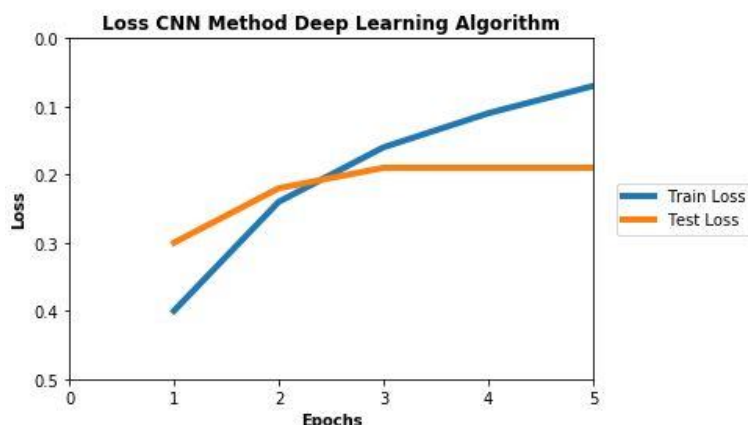
دقت به دست آمده در روش ترکیبی سه تکنیک پس از اعمال روش پیشنهادی و با استفاده از الگوریتم‌های مختلف یادگیری ماشین با آزمایش در ۱۰ دور متفاوت در جدول ۱۴ و شکل ۹ نشان داده شده است. همان‌طور که در جدول ۱۴ و شکل ۹ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده در الگوریتم‌های ماشین بردار پشتیبان و تبادلات گرادینان تصادفی و برگشت منطقی از دور اول به میزان ۳۰ درصد و الگوریتم‌های بی‌ساده به میزان ۲۱ درصد و جنگل تصادفی به میزان ۳۵ درصد از دور چهارم به سمت بالا بهبود یافته‌اند.

جدول ۱۵- نتایج دقت و خطای آموزش و آزمایش روش شبکه عصبی کانولوشن یادگیری عمیق

دوره	دقت آموزش	دقت آزمایش	خطای آموزش	خطای آزمایش
دوره اول	۰.۸۳	۰.۸۷	۰.۴۰	۰.۳۰
دوره دوم	۰.۹۰	۰.۹۰	۰.۲۴	۰.۲۲
دوره سوم	۰.۹۴	۰.۹۲	۰.۱۶	۰.۱۹
دوره چهارم	۰.۹۶	۰.۹۲	۰.۱۱	۰.۱۹
دوره پنجم	۰.۹۸	۰.۹۲	۰.۰۷	۰.۱۹



شکل ۱۰- نتایج دقت روش شبکه عصبی کانولوشن یادگیری عمیق

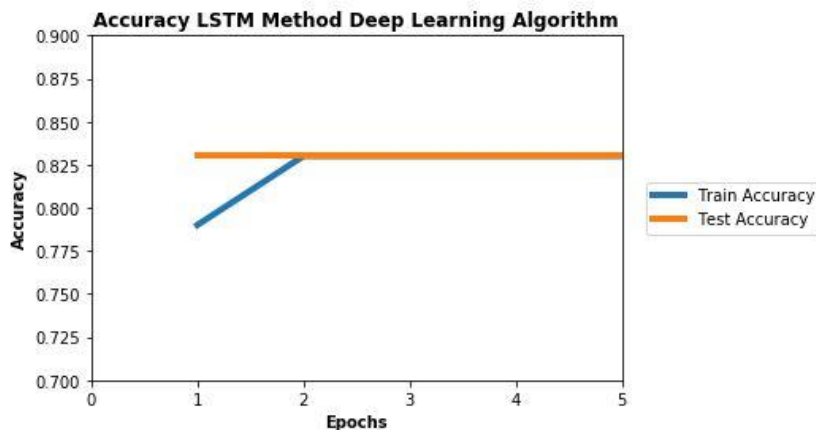


شکل ۱۱- نتایج درصد خطای روش شبکه عصبی کانولوشن یادگیری عمیق

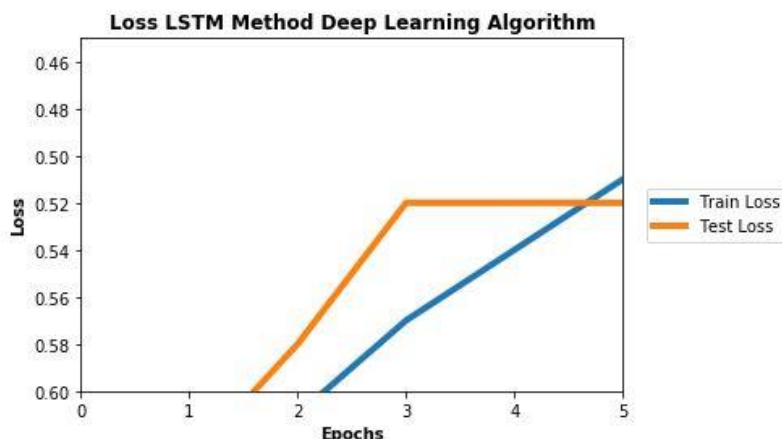
دقت به دست آمده در روش شبکه عصبی کانولوشن از روش‌های یادگیری عمیق پس از اعمال روش پیشنهادی با آزمایش در ۵ دور متفاوت در جدول ۱۵ و شکل ۱۰ نشان داده شده است. همان‌طور که در جدول ۱۵ و شکل ۱۰ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده بین آموزش و آزمایش تا حدودی از بین رفته است و هر چند در دور سوم به میزان ۲ درصد افزایش دقت در داده‌های آموزش داریم ولی در داده‌های آزمایش ثابت شده است و حالت سرریز<sup>۴۲</sup> پیدا کرده است و از آموزش کمتر شده است. همچنین درصد خطای روش شبکه عصبی کانولوشن از روش‌های یادگیری عمیق پس از اعمال روش پیشنهادی با آزمایش در ۵ دور متفاوت در جدول ۱۵ و شکل ۱۱ نشان داده شده است. همان‌طور که در جدول ۱۵ و شکل ۱۱ نشان داده شده است، می‌توان آن را تحلیل کرد که درصد خطای به دست آمده در داده‌های آموزش در هر دوره بین ۴ تا ۲۵ درصد کمتر شده است ولی برای داده‌های آزمایش در دوره سوم با ۱۹ درصد خطا ثابت شده است و بیشتر از داده‌های آموزش می‌باشد. به طور کلی روش CNN بر روی داده‌های موردنظر و بعد از اعمال روش پیشنهادی نسبت به الگوریتم‌های جنگل تصادفی و بیز ساده با ۹۲ درصد دقت بهتر عمل کرده است.

جدول ۱۶- نتایج دقت و خطای آموزش و آزمایش روش حافظه طولانی کوتاه مدت یادگیری عمیق

دوره	دقت آموزش	دقت آزمایش	خطای آموزش	خطای آزمایش
دوره اول	۰٫۷۹	۰٫۸۳	۰٫۶۷	۰٫۶۳
دوره دوم	۰٫۸۳	۰٫۸۳	۰٫۶۱	۰٫۵۸
دوره سوم	۰٫۸۳	۰٫۸۳	۰٫۵۷	۰٫۵۲
دوره چهارم	۰٫۸۳	۰٫۸۳	۰٫۵۴	۰٫۵۲
دوره پنجم	۰٫۸۳	۰٫۸۳	۰٫۵۱	۰٫۵۰



شکل ۱۲- نتایج دقت روش حافظه طولانی کوتاه مدت یادگیری عمیق



شکل ۱۳- نتایج خطای آموزش و آزمایش روش حافظه طولانی کوتاه مدت یادگیری عمیق

دقت به دست آمده در روش حافظه طولانی کوتاه مدت از روش‌های یادگیری عمیق پس از اعمال روش پیشنهادی با آزمایش در ۵ دور متفاوت در جدول ۱۶ و شکل ۱۲ نشان داده شده است. همان‌طور که در جدول ۱۶ و شکل ۱۲ نشان داده شده است، می‌توان آن را تحلیل کرد که مقدار دقت به دست آمده بین آموزش و آزمایش در همان دور اول هم در آموزش و هم در آزمایش با مقدار ۸۳ درصد ثابت شده‌اند و حالت سرریز پیدا کرده‌اند. همچنین درصد خطای روش حافظه طولانی کوتاه مدت از روش‌های یادگیری عمیق پس از اعمال روش پیشنهادی با آزمایش در ۵ دور متفاوت در جدول ۱۶ و شکل ۱۳ نشان داده شده است. همان‌طور که در جدول ۱۶ و شکل ۱۳ نشان داده شده است، می‌توان آن را تحلیل کرد که درصد خطای به دست آمده در داده‌های آموزش و آزمایش در هر دوره بین ۲ تا ۶ درصد کمتر شده است و در دوره‌های آخر در داده‌های آزمایش درصد خطا پایین‌تر نیز آمده است. به طور کلی روش LSTM بر روی داده‌های موردنظر و بعد از اعمال روش پیشنهادی نسبت به الگوریتم‌های یادگیری ماشین و CNN با ۸۳ درصد دقت ضعیف‌تر عمل کرده است.

## ۶ نتیجه‌گیری و کارهای آینده

این پژوهش تلاشی برای طبقه‌بندی نظرات فارسی تلفن همراه سایت دیجی کالا با استفاده از الگوریتم‌های مختلف یادگیری ماشین تحت نظارت مانند بیز ساده (NB)، تبادلات گرادیان تصادفی (SGD)، ماشین بردار پشتیبان (SVM)، برگشت منطقی (LR) و جنگل تصادفی (RF) و از الگوریتم‌های یادگیری عمیق مانند شبکه عصبی کانولوشن (CNN) و حافظه طولانی کوتاه مدت (LSTM) می‌باشند. این الگوریتم‌ها با استفاده از روش‌های n-gram در مجموعه داده‌های تلفن همراه به کار می‌روند. مشاهده می‌شود که با افزایش مقدار n در n-gram، دقت طبقه‌بندی کاهش می‌یابد، یعنی برای تک‌واژه‌ای و دوواژه‌ای، نتیجه استفاده از الگوریتم‌ها بهتر از سه‌واژه‌ای می‌باشد و همچنین الگوریتم‌های ماشین بردار پشتیبان بین ۴ تا ۲۴ درصد و تبادلات گرادیان تصادفی بین ۷ تا ۳۸ درصد و برگشت منطقی بین ۵ تا ۳۸ درصد و روش شبکه عصبی کانولوشن به میزان ۴ درصد عملکرد بهتری نسبت به الگوریتم‌های دیگر داشتند. از روش‌های نمونه‌گیری بیش از حد برای همگن‌سازی داده‌ها و روش امتیازدهی پیشنهادی برای بهبودی تعیین قطبیت و از تکنیک‌های بردارسازی برای تبدیل متن به ماتریس اعداد استفاده شده است و نیز به منظور به دست آوردن مقدار صحت، فراخوان، اندازه F و دقت در روش بهبودیافته، از تکنیک‌های یادگیری ماشین و یادگیری عمیق استفاده می‌شود. در زبان فارسی مشکلاتی از قبیل عامیانه بودن کلمات و رعایت نکردن فاصله بین کلمات و رعایت نکردن قاعده ساختار جمله و استفاده از شکلک‌های تصویری دارا می‌باشد که می‌توان این مشکلات را در کارهای آینده برای بهبودی طبقه‌بندی تحلیل احساسات رفع کرد.

## منابع

1. Feldman, R. "Techniques and applications for sentiment analysis", 2013.
2. Gautam, G., Yadav, D. "Sentiment analysis of twitter data using machine learning approaches and semantic analysis", 2014.
3. Alex, K., Sutskever, I., Geoffrey, E.H. "ImageNet Classification with Deep Convolutional Neural Networks.", 2012.
4. Amorim, P., Abreu, H., Duarte, H., Santos, J. "Evaluation of oversampling data balancing techniques in the context of ordinal classification", 2018.
5. Zezawar, K., Nyein, N. "Sentiment Analysis of Students' Comment Using Lexicon Based Approach", 2017.

6. Hastie, T., Tibshirani, R., Friedman, J. "Unsupervised learning", 2009.
7. Pang, B., Lee, L., Vaithyanathan, S. "Thumbs up?: sentiment classification using machine learning techniques", 2002.
8. Salveti, F., Lewis, S., Reichenbach, C, "Automatic opinion polarity 19classify- cation of movie", 2004.
9. Beineke, P., Hastie, T., Vaithyanathan, S. "The sentimental factor: improving review classification via human-provided information", 2004.
10. Mullen, T., Collier, N. "Sentiment analysis using support vector machines with diverse information sources", 2004.
11. Dave, K., Lawrence, S., Pennock, D.M. "Mining the peanut gallery: opinion extraction and semantic classification of product reviews", 2003.
12. Matsumoto, S., Takamura, H., Okumura, M. "Sentiment classification using word sub-sequences and dependency sub-trees", 2005.
13. Zhang, D., Xu, H., Su, Z., Xu, Y. "Chinese comments sentiment classifica- tion based on word2vec and svm perF", 2015.
14. Liu, S.M. Chen, J-H." A multi-label classification based approach for sen- timent classification", 2015.
15. Luo, B., Zeng, J., Duan, J." Emotion space model for classifying opinions in stock message board", 2016.
16. Niu, T., Zhu, S., Pang, L., El Saddik, A. "Sentiment analysis on multi-view social data", 2016.
17. Caiqiang, L., Junming, M. "Research on online education teacher evaluation model based on opinion mining", 2012.
18. Ortigosa, A., Martín, J.M., Carro, R.M. "Sentiment analysis in Facebook and its application to e-learning", 2014.
19. Pong-inwong, C., Songpan W. (Rungworawut), "TeachingSenti-Lexicon for Automated Sentiment Polarity Definition in Teaching Evaluation.", 2014.
20. Wang , Y., Yong Youn, H., "Feature Weighting Based on Inter-Category and Intra-Category Strength for Twitter Sentiment Analysis", 2019.

Natural Language Processing <sup>۱</sup>

Machine Learnin <sup>۲</sup>

Deep Learning<sup>۳</sup>

Opinion Minin <sup>۴</sup>

Positive <sup>۵</sup>

Negative <sup>۶</sup>

Neutral <sup>۷</sup>

Supervised Machine Learning <sup>۸</sup>

Unsupervised Machine Learning <sup>۹</sup>

Labeled <sup>۱۰</sup>

Train Datas <sup>۱۱</sup>

Test Datas <sup>۱۲</sup>

Clustring Algorithms <sup>۱۳</sup>

Random Over Sampler <sup>۱۴</sup>

Random Under Sampler <sup>۱۵</sup>

Stop Words <sup>۱۶</sup>

Vectorizer Process	۱۷
Naïve Bays	۱۸
Support Vector Machine	۱۹
Maximum Entropy	۲۰
Unigram	۲۱
Bigram	۲۲
Overall Opinion Polarity	۲۳
Markov Model	۲۴
WordNet	۲۵
Part-of-speech	۲۶
Lexical Filter	۲۷
Derived Features	۲۸
Accuracy	۲۹
Topic Annotation	۳۰
Hand Annotation	۳۱
Microblog	۳۲
Emotional Space	۳۳
Classification	۳۴
Multi-View Sentiment Analysis	۳۵
Learning Management System	۳۶
Pointwise Mutual Information	۳۷
Category Difficulty Strength	۳۸
Multinomial Naïve-Bayes	۳۹
Crawler	۴۰
10-fold	۴۱
overfitting	۴۲