پژوهش‌های نوین در تصمیم‌گیری
دوره ۸، شماره ۱، بهار ۱۴۰۲، صص ۱۱۲-۹۱
نوع مقاله: پژوهشی

مدل‌سازی و پیش‌بینی قصد خرید تلفن همراه کاربران توییتر بر مبنای تحلیل احساسات

مینا نوروزی[۱]، آمنه خدیور *[۲]، فاطمه عباسی[۳]

۱. کارشناسی‌ارشد مدیریت کسب‌وکار، دانشکده علوم اجتماعی و اقتصادی، دانشگاه الزهرا (س)، تهران، ایران
۲. دانشیار، گروه مدیریت، دانشکده علوم اجتماعی و اقتصادی، دانشگاه الزهرا (س)، تهران، ایران
۳. استادیار، دانشکده فناوری اطلاعات، دانشگاه مهر البرز، تهران، ایران

تاریخ دریافت: ۱۴۰۱/۰۸/۲۸ تاریخ پذیرش: ۱۴۰۲/۰۱/۱۵

چکیده

با ظهور رسانه‌های اجتماعی مردم ایده‌ها، تجربیات، نظرات و مقاصد خود را در توییتر، فیس‌بوک و دیگر پلتفرم‌های رسانه اجتماعی به اشتراک می‌گذارند. تحلیل و بررسی این نظرات در رسانه‌های اجتماعی اطلاعات سودمندی را در رابطه با بازار، مشتریان و رقبا به سازمان‌ها می‌دهد. یکی از این اطلاعات، قصد خرید کاربران است که برای استفاده‌های تجاری مانند بازاریابی هدفمند و شخصی‌سازی تبلیغات موردتوجه قرار می‌گیرد. در این پژوهش سیستمی مبنی بر هوش مصنوعی جهت پیش‌بینی قصد خرید کاربران در رسانه توییتر ارائه شده‌است. تحقیق حاضر از لحاظ هدف کاربردی است و جامعه مورد بررسی شامل نظرات کاربران در سایت توییتر می‌باشد و نمونه مورد بررسی شامل توییت های کاربران در ارتباط با آیفون ۱۳ می باشد. جهت انجام پژوهش ۲۳۰۰۰ توییت انگلیسی کاربران جمع‌آوری شد و احساسات آن‌ها با استفاده از روش مبتنی بر واژگان تحلیل شد و نتایج قابل قبولی بادقت ۰،۸۱ ارائه شد. نتایج حاصل از تحلیل احساسات، نشان‌دهنده احساسات عمومی مثبت در مورد آیفون ۱۳ است. قصد خرید کاربران توییتر با استفاده از شبکه عصبی عمیق پیچشی حافظه طولانی کوتاه‌مدت بادقت ۰،۸۱ استخراج شد. سپس برای سنجش اثرگذاری احساسات کاربران و ویژگیهای استاتیکی از مدل شبکه عصبی عمیق چندلایه پرسپترون استفاده شد که نتایج حاصل دقت ۰،۸۰ را نمایش داد. استفاده از این مدل و پیش‌بینی قصد خرید به مدیران بازاریابی کمک می‌کند تا برنامه‌ریزی مناسبی جهت شناسایی مشتریان خود انجام دهند و هزینه‌های بازاریابی را کاهش دهند.

کلیدواژه‌ها: یادگیری عمیق، شبکه عصبی پیچشی حافظه طولانی کوتاه‌مدت، توییتر، تحلیل احساسات، قصد خرید

---

* نویسنده مسئول مقاله E-mail: a.khadivar@alzahra.ac.ir

۱- مقدمه و بیان مسئله

در عصری زندگی می‌کنیم که اینترنت بخش حیاتی زندگی روزمره ماست. عصر اینترنت نحوه بیان نظرات و عقاید مردم را تغییر داده است و امروزه بیان نظرات از طریق پست‌های وبلاگ، انجمن‌های آنلاین، سایت‌های بررسی محصول، رسانه‌های اجتماعی و غیره انجام می‌شود[۱] و یکی از بهترین منابع اطلاعاتی، محتوای تولید شده کاربران در مورد شرکت، محصول و رقبا است که در رسانه‌های اجتماعی به اشتراک گذاشته می‌شود و شرکت‌ها می‌توانند با استفاده از تکنیک‌های تحلیلی در کلان داده‌ها از الگوی دانش پنهان در این اطلاعات بهره‌مند شوند[۲]. پیش‌بینی نتیجه‌های معرفی محصول جدید به بازار، باتوجه به هزینه‌های بالای توسعه محصول و تأثیر زیاد شکست احتمالی آن روی سازمان، موضوعی است که دهه‌های گذشته بسیاری از مدیران بازاریابی به آن توجه کرده‌اند[۳].

در توییتر[۱] افراد تمایل دارند که علاوه بر فعالیت‌های روزانه، افکار و احساسات مقاصد خود را در مورد هر کاری که انجام می‌دهند به اشتراک بگذارند. در نتیجه، توییت‌ها می‌توانند منعکس‌کننده بسیاری از افکار عمومی و قصد کاربران باشند؛ بنابراین، تجزیه‌وتحلیل آن‌ها یک رویکرد مؤثر برای شرکت‌ها و کسب‌وکارهاست تا بفهمند مشتریان بالقوه آن‌ها چه چیزهایی می‌خواهند و به چه چیزهایی اهمیت می‌دهند و به آن‌ها کمک می‌کند تا برنامه بازاریابی آنلاین بهتری داشته باشند و سریع‌تر و کارآمدتر به بازار نفوذ کنند[٤]. پیش‌بینی و تجزیه‌وتحلیل قصد خرید معمولاً توسط مدیران بازاریابی به‌عنوان ورودی برای تصمیم‌گیری در مورد خدمات و محصولات فعلی و آتی در نظر گرفته می‌شود و مورداستفاده قرار می‌گیرد. بسیاری از شرکت‌های فناوری از فرم بازخورد نظرسنجی کاربران استفاده می‌کنند و محاسبات لازم را برای پیش‌بینی قصد خرید انجام می‌دهند[٥]. این فرایندهای دستی، بسیار زمان‌بر و پرزحمت است و دارای چالش‌های زیادی مانند سوگیری تفسیرها و عدم مدیریت حجم عظیمی از داده‌های متنی است[٦]. جایی که اندازه مجموعه متن گسترده است و حجم عظیم اطلاعات وجود دارد و تجزیه‌وتحلیل احساسات محتوا به‌صورت دستی امکان‌پذیر نیست، متن‌کاوی و پردازش زبان طبیعی و تحلیل خودکار متون ضروری می‌شود[۷].

شناسایی قصد خرید محصولات می تواند به شرکت‌ها در بهبود کیفیت خدمات و حفظ موقعیت رقابتی کمک نماید و توسعه سیستم‌های هوش تجاری مبتنی بر قصد کاوی می‌تواند عاملی

شتاب دهنده باشد [۷]. باتوجه به بررسی‌های صورت گرفته، کمتر پژوهشی به پیش‌بینی قصد خرید[۲] یک محصول توسط کابران با استفاده از متن‌کاوی پرداخته‌است. ازاین‌رو، در این پژوهش سیستمی جهت پیش‌بینی قصد خرید کاربران بر مبنای احساسات آن‌ها ارائه می‌شود تا شرکت‌ها احساسات کاربران را نسبت به محصولات خود تحلیل کرده، نقاط ضعف و قوت آنها را شناسایی کرده و مشتریان بالقوه خود را پیش‌بینی کنند و برنامه بازاریابی هدفمند با هزینه کمتری ارائه دهند. پرسش‌های اصلی پژوهش حاضر عبارتند از: چه تکنیکی جهت تحلیل احساسات[۳] کاربران توییتر مناسب است؟ کدام مدل قصد خرید کاربران را با دقت مناسبی پیش‌بینی می‌کند؟ آیا با استفاده از تحلیل احساسات و ویژگی‌های استاتیکی قصد خرید کاربران قابل پیش‌بینی است؟ در بخش های بعدی به ترتیب مبانی نظری، روش شناسی، یافته‌های پژوهش و در نهایت نتیجه گیری و پیشنهادات ارائه می‌گردد.

۲- مبانی نظری

یک پست قصد خرید به‌عنوان پستی تعریف می‌شود که دارای نشانه‌های آشکار خرید فوری یا آینده برخی محصولات است. قصد ممکن است به‌صورت صریح یا ضمنی در پست‌های اجتماعی بیان شود[۸]. یک پست قصد صریح به‌طور مستقیم و به وضوع قصد خرید را نشان می‌دهد اما در پست‌های قصد ضمنی این موضوع به صورت غیرمستقیم است. استخراج پست‌های قصد ضمنی چالش برانگیزتر و ضروری‌تر است، زیرا بیشتر پست‌های قصد ماهیت ضمنی دارند[۹].

برای اولین بار در پژوهشی مشکل شناسایی آرزوها را در بررسی محصول مطالعه کردند و به شناسایی نظرات کاربرانی که محصولی را پیشنهاد می‌کنند و یا قصد خرید آن را دارند، پرداختند[۱۰].

در پژوهشی دیگر، یک روش یادگیری انتقال خاص برای طبقه‌بندی قصد پیشنهاد کردند و طبقه‌بندی قصد را به‌عنوان یک مسئله طبقه‌بندی دوتایی فرموله کردند. پست‌هایی که به‌صراحت قصد را بیان می‌کنند و پست‌های بدون قصد. نتایج حاصل عملکرد مناسب مدل را نسبت به روش‌های پایه مانند طبقه‌بندی‌کننده‌هایی با داده‌های برچسب‌گذاری شده و یک طبقه‌بندی‌کننده یادگیری انتقال پیشرفته نشان می‌دهد[۱۱].

در مقاله‌ای راه حلی با استفاده از مدل‌های طبقه‌بندی سنتی مانند ماشین بردار پشتیبان[4] برای شناسایی مقاصد تجاری در سطح توییت ارائه کردند. در مقاله‌ای دیگر به شناسایی قصد مصرف بر اساس یک شبکه عصبی پیچشی[5] پرداخته‌شد. یکی از کاستی‌های این کار عدم طبقه‌بندی محصولات در حوزه‌های مربوطه است.

یک تکنیک یادگیری عمیق با استفاده از مدل شبکه عصبی حافظه کوتاه مدت[6] برای پیش‌بینی رفتار خرید مشتری از داده‌های توییتر در پژوهشی پیشنهاد شد که سیستم‌های توصیه‌گر می‌توانند از آن استفاده کنند. این کار ثابت کرده است که استفاده از یک مدل یادگیری عمیق که اطلاعات متوالی را در بر می‌گیرد بهتر از نادیده‌گرفتن اطلاعات متوالی برای کار پیش‌بینی خرید عمل می‌کند[۱۲].

در یکی از پژوهش‌های اخیر، از داده‌های جمع‌آوری شده در مقالات پیشین برای شناسایی قصد خرید در نظرات کاربران در دامنه‌های مختلف تلفن همراه، لوازم الکترونیکی، دوربین، تلویزیون و .... استفاده شده و الگوریتم مورد استفاده آن‌ها، شبکه عصبی CNN-LSTM است[۱۳]. همچنین در پژوهش دیگری نیز از این داده‌های آماده و مدل شبکه عصبی CNN-BiLSTM جهت شناسایی قصد استفاده شده‌است[۱٤]. در این پژوهش‌ها، به بررسی توییت‌های کاربران در رابطه با یک محصول خاص از یک برند خاص پرداخته نشده‌است.

در پژوهشی به سنجش اثرگذاری احساسات کاربران بر قیمت بیتکوین با استفاده از شبکه عصبی پرداخته‌اند که نتایج به همبستگی میان قیمت واقعی و قیمت پیش‌بینی شده اشاره دارد[۱۵]. در پژوهش حاضر به بررسی اثرگذاری احساسات کاربران در پیش‌بینی قصد خرید پرداخته می‌شود که تا به حال انجام نگرفته‌است.

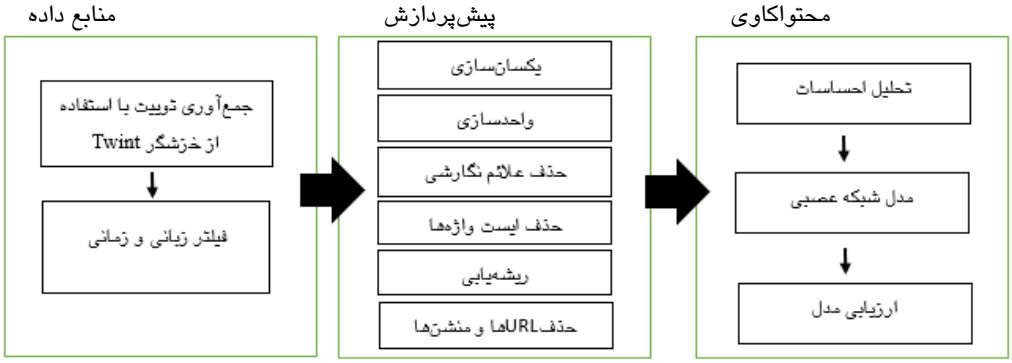
دو پژوهش به شناسایی اهداف پذیرش محصول خاص در توییتر می‌پردازند و آن را دسته کامل‌تری نسبت به قصد خرید معرفی می‌کنند؛ در پژوهش اول، از سه نمایش متن مختلف استفاده می‌کنند و با استفاده از مدل CNN-LSTM قصد پذیرش را از داده‌های توییتر استخراج می‌کنند. در پژوهش بعدی، علاوه بر داده‌های متنی ارتباطات اجتماعی کاربران جهت شناسایی قصد پذیرش محصول استفاده می‌شود. بر اساس مدل رفتاری یکپارچه یک مدل یادگیری عمیق ایجاد می‌شود که رویکرد پیشنهادی را تأیید می‌کند[۱٤]. پژوهش حاضر علاوه

بر پیش‌بینی قصد خرید کاربران توییتر به بررسی تأثیر نتایج تحلیل احساسات به عنوان یک ویژگی در این پیش‌بینی نیز می‌پردازد.

باتوجه‌به بررسی‌های انجام شده، در پژوهش‌های اخیر، جهت استخراج قصد خرید از مدل شبکه عصبی CNN-LSTM استفاده و اثربخشی این مدل تأیید شده است؛ در پژوهش حاضر داده‌های مورد استفاده از توییتر و پژوهش در رابطه با برند خاص آیفون و مدل ۱۳ آن است که جهت بازاریابی هدفمند استفاده از نوع خاص محصول به‌جای دسته‌های عمومی محصولات تأثیر بیشتری دارد. همچنین بررسی تأثیر تحلیل احساسات و ویژگی‌های متنی بر پیش‌بینی قصد خرید نیز در میان پژوهش‌ها دیده نمی‌شود. ازاین‌رو پژوهش حاضر با استفاده از شبکه عصبی عمیق CNN-LSTM قصد خرید توییت‌ها را استخراج کرده سپس با به‌کارگیری احساسات کاربران در حوزه آیفون ۱۳ و استخراج ویژگی‌های استاتیکی، سعی در ارائه مدلی جهت پیش‌بینی قصد خرید توسط مدل شبکه عصبی چندلایه پرسپترون [۷] دارد.

۳- روش‌شناسی پژوهش

در شکل ۱ نمای مراحل انجام پژوهش نمایش داده‌شده است. یکی از متداول‌ترین متدولوژی‌ها برای انجام پروژه‌های داده‌کاوی، فرآیندهای استاندارد صنعت متقابل برای داده‌کاوی نام دارد. این روش شامل ٦ مرحله است که در ادامه معرفی می‌شوند:



شکل ۱. نمای مدل‌سازی

۳-۱- منابع داده

در این پژوهش، توییت‌هایی از کاربران که حاوی عبارت iphone۱۳# بودند از تاریخ ۲۰۲۱/۰۷/۰۱ تا تاریخ ۲۰۲۲/۰۱/۳۰ میلادی جمع‌آوری شدند. داده‌های به‌دست‌آمده از نوع متنی هستند و توسط یک خزشگر به نام twint که در محیط پایتون توسعه داده شده است جمع‌آوری شدند. سپس تنها توییت‌های به زبان انگلیسی نگهداری شدند و ۲۳۰۰۰ سطر از آن‌ها جهت انجام پژوهش استفاده شد. ستون‌های اضافی داده‌ها حذف گردید و خروجی آماده پیش‌پردازش شد.

۳-۲- پیش‌پردازش داده‌ها

مرحله دوم، پیش‌پردازش توییت‌های ذخیره شده است. یک توییت ممکن است حاوی اطلاعات مبهم و غیرمرتبطی مانند برچسب‌های HTML، تبلیغات و داده‌های پرسروصدا باشد که تشخیص قطبیت آن را برای سیستم دشوار کند. عموماً پیش‌پردازش داده‌ها شامل مراحل ریشه‌یابی، بن‌واژه‌سازی، حذف ایست‌واژه‌ها، نرمال‌سازی، حذف علائم نگارشی، واحدسازی و برچسب‌گذاری است.

فرآیند ریشه‌یابی همه کلمات به شکل پایه یا ریشه تبدیل می‌شوند. در بن‌واژه‌سازی از قواعد زبان و اطلاعات بیشتری نسبت به ریشه‌یابی استفاده می‌شود. در این مرحله، شکل‌های مختلف یک کلمه در یک گروه‌بندی قرار می‌گیرند تا به‌عنوان یک کلمه تجزیه‌وتحلیل شوند. سپس ایست‌واژه‌ها که اطلاعات محتوایی زیادی ندارند حذف می‌شوند.

در مرحله بعد موارد دیگری مانند URLها، نام‌های کاربری، علائم نگارشی، ایموجی‌ها و ریتوییت‌ها نیز از توییت‌ها حذف می‌شوند. در این مرحله، برای نرمال‌سازی متون انگلیسی تمامی حروف متن به شکل حروف کوچک یا بزرگ درمی‌آیند تا برای تحلیل یکپارچه شوند. سپس با واحد سازی یا توکن‌سازی دنبالۀ کاراکترها به قطعات (کلمات/عبارات) به نام توکن یا نشانه، شکسته می‌شوند. [۱٦].

۳-۳- مدل‌سازی

۳-۳-۱- تحلیل احساسات

در این مرحله پس از انجام پیش‌پردازش داده‌ها با استفاده از تحلیل احساسات قطبیت توییت‌ها استخراج می‌شود و به سه دسته مثبت، منفی و خنثی تقسیم می‌شوند. در این پژوهش، برای تحلیل احساسات با روش مبتنی بر واژگان[۹] و واژه‌نامه VADER انجام شده است.

یکی از روش‌های تحلیل احساسات، رویکرد مبتنی بر واژگان است. این روش جهت‌گیری احساسی کل سند یا جمله را با استفاده از جهت‌گیری معنایی واژگان محاسبه می‌کند. جهت‌گیری معنایی می‌تواند مثبت، منفی یا خنثی باشد.

VADER یک واژه‌نامه و ابزار تجزیه‌وتحلیل احساسات مبتنی بر قانون است که می‌تواند کلمات، اختصارات، عبارات عامیانه، ایموجی‌ها و شکلک‌ها را که معمولاً در رسانه‌های اجتماعی یافت می‌شوند، مدیریت کند. این روش بسیار سریع‌تر از الگوریتم‌های یادگیری ماشین است زیرا نیازی به آموزش ندارد. این ابزار برای هر متن، بردار نمرات احساسات با قطبیت‌های منفی، خنثی، ترکیبی و مثبت تولید می‌کند[۱۷].

۳-۳-۲- استخراج قصد خرید

جهت استخراج قصد خرید، ابتدا توییت‌ها به دارای قصد خرید و بدون قصد خرید برچسب‌گذاری می‌شوند. برچسب‌گذاری یا حاشیه‌نویسی داده‌ها فرایند افزودن برچسب‌ها به داده‌های خام برای نشان‌دادن ویژگی‌های هدف یا پاسخ به مدل یادگیری ماشین است که انتظار می‌رود یاد بگیرد و پیش‌بینی کند.

داده‌های با صفر و یک برچسب‌گذاری شدند که نشانگر ۱ به داشتن قصد خرید و نشانگر ۰ به نداشتن قصد خرید اشاره دارد. ۱۳۰۲۹ توییت دارای قصد خرید و ۹۹۷۰ توییت بدون قصد خرید هستند. سپس توییت‌های پیش‌پردازش شده در مراحل قبل و برچسب‌ها برای استخراج قصد خرید آماده می‌شوند.

در این پژوهش، از مدل شبکه عصبی CNN-LSTM برای طبقه‌بندی توییت‌ها به دودسته دارای قصد خرید و بدون قصد خرید استفاده شده است. این مدل، در چهار مرحله اصلی عمل می‌کند: یک، جاسازی کلمه به‌عنوان ورودی به مدل داده می‌شود که یک نمایش برداری باارزش واقعی از یک کلمه است. دو، استخراج ویژگی با استفاده از لایه CNN. سه، به‌خاطر سپردن

وابستگی‌های طولانی‌مدت با استفاده از لایه LSTM و چهار، طبقه‌بندی متون به کلاس‌های باینری (دارای قصد/بدون قصد) انجام می‌گیرد. ازاین‌رو، پنج لایه در شبکه استفاده می‌شود: ورودی (جاسازی کلمه)، کانولوشن (نقشه ویژگی)، maxpooling (نقشه ویژگی ادغام شده)، LSTM (حفظ وابستگی‌های طولانی‌مدت) و خروجی (tanh).

۳-۳-۳- استخراج ویژگی

در این پژوهش از برخی ویژگی‌های آماری ساده به‌عنوان ویژگی ورودی استفاده شده است. میانگین تعداد کلمات در هر توییت، تعداد کاراکترها، نسبت تعداد کاراکتر بر کلمات و میانگین کلمات به‌عنوان ویژگی‌های مورد بررسی انتخاب شدند که در بهبود دقت مدل تأثیر دارند[۱۸]. همچنین در این پژوهش تأثیرگذاری احساسات کاربران بر پیش‌بینی قصد خرید نیز بررسی می‌گردد. یکی دیگر از ویژگی‌های استخراج شده تعداد تفاوت روز توییت تا روز رونمایی از محصول آیفون ۱۳ است که فرض می‌شود بر میزان قصد خرید محصول تأثیر دارد.

۳-۳-۴- انتخاب ویژگی

انتخاب ویژگی فرایند انتخاب حداقل ویژگی‌های مناسب از میان ویژگی‌های در دسترس است. این فرایند از طریق حذف ویژگی‌های نامرتبط و اضافی انجام می‌گیرد. [۱۹]. در این پژوهش، انتخاب ویژگی بر اساس همبستگی صورت گرفته است. ویژگی‌های استخراج شده اولیه تعداد کلمات، تعداد کاراکترها، میانگین کلمات، نسبت تعداد کاراکترها به تعداد کلمات، تحلیل احساسات و فاصله روز رونمایی مجصول تا روز توییت هستند که پس از بررسی همبستگی بین آن‌ها یکی از ویژگی‌هایی که بین آن‌ها همبستگی بالا بود کنار گذاشته شد. درنهایت، تعداد کلمات، نسبت تعداد کاراکتر به تعداد کلمات، فاصله روز توییت تا روز رونمایی محصول و تحلیل احساسات به عنوان ورودی وارد مدل پیش‌بینی شدند.

۳-۳-۵- شبکه عصبی پیش‌بینی قصد خرید

مدل شبکه عصبی پرسپترون چندلایه به‌منظور استخراج ویژگی و طبقه‌بندی توییت‌های کاربران به دارای قصد خرید و بدون قصد خرید استفاده می‌شود.

۴- ارزیابی

جهت ارزیابی مدل‌سازی از معیارهای زیر استفاده شد:

ماتریس درهم‌ریختگی: ماتریس درهم‌ریختگی ابزاری برای تحلیل پیش‌بینی در یادگیری ماشین است. ستون‌های این ماتریس مربعی مقادیر واقعی و ردیف‌های آن مقادیر پیش‌بینی شده توسط مدل را نمایش می‌دهد و بالعکس[۲۰].

معیار صحت: صحت معیاری برای تعداد پیش‌بینی صحیح مدل برای مجموعه‌داده آزمایشی است. این معیار، نشان می‌دهد که مدل طراحی شده چند درصد از کل مجموعه رکوردهای آزمایشی را به‌درستی پیش‌بینی کرده است.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$ (فرمول۱)

معیارهای دقت و فراخوانی: دقت و فراخوانی دو معیار مهم هستند که از مفاهیم ماتریس درهم‌ریختگی به دست می‌آیند. دقت بالا به این معنی است که بیشتر داده‌ها با برچسب مثبت، به‌درستی پیش‌بینی شده‌اند.

$$Precision = \frac{TP}{TP+TN}$$ (فرمول۲)

$$Recall = \frac{TP}{TP+FN}$$ (فرمول۳)

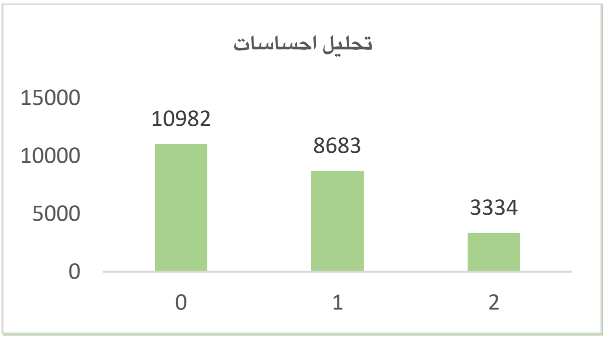
معیار سنجه F: با افزایش دقت مدل، فراخوانی کاهش می‌یابد و بالعکس. معیار سنجه F یک میانگین هارمونیک دقت و فراخوانی است و ترکیبی از این دو را نمایش می‌دهد.

$$F1\ Score = \frac{۱}{\frac{۱}{recall}+\frac{۱}{precision}}$$ (فرمول٤)

۵- یافته‌های کمی

در این بخش یافته‌های کمی هر بخش ارائه می‌شود:

۵-۱- تحلیل احساسات



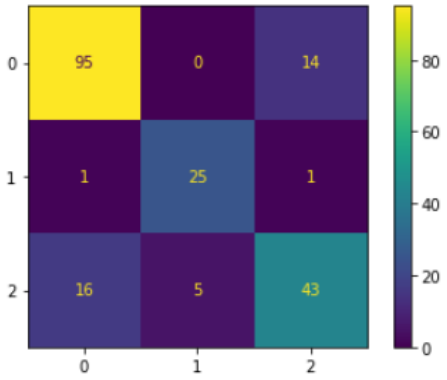
شکل ۲. نتایج تحلیل احساسات

در شکل ۲، دو (۲) نشانگر توییت‌های مثبت، یک (۱) نشانگر توییت‌های منفی و صفر (۰) نشانگر توییت‌های خنثی است. در مجموع در میان ۲۳۰۰۰ توییت ۳۳۳۴ توییت منفی، ۸۶۸۴ توییت مثبت و ۱۰۹۸۲ توییت خنثی ارزیابی شدند. میزان احساسات مثبت نسبت به احساسات منفی بیشتر است که نشان‌دهنده احساسات عمومی مثبت نسبت به آیفون ۱۳ است. به‌منظور بررسی علت پشت احساسات بیان شده کاربران در مرحله بعد به بررسی ابر کلمات مثبت و منفی توییت‌ها و کلمات پرتکرار پرداخته شد.

برای ارزیابی مدل تحلیل احساسات که رویکرد مبتنی بر واژگان است و از روش‌های نظارت نشده تحلیل احساسات است، ۲۰۰ داده انتخاب گردید و به‌صورت دستی برچسب احساسات زده شد. این داده‌های برچسب‌گذاری شده برای ارزیابی تحلیل احساسات پیش‌بینی شده مورداستفاده قرار گرفت و دقت مدل VADER، ۸۱ درصد ارزیابی شد. نتایج ارزیابی در جدول ۱ و شکل ۳ ارائه شده است.

جدول ۱. نتایج تحلیل احساسات

کلاس	صحت	دقت	بازخوانی	امتیاز F
توییت‌های خنثی	۰,۸۱	۰,۸۵	۰,۸۷	۰,۸۶
توییت‌های مثبت		۰,۸۳	۰,۹۳	۰,۸۸
توییت‌های منفی		۰,۷۴	۰,۶۷	۰,۷۰



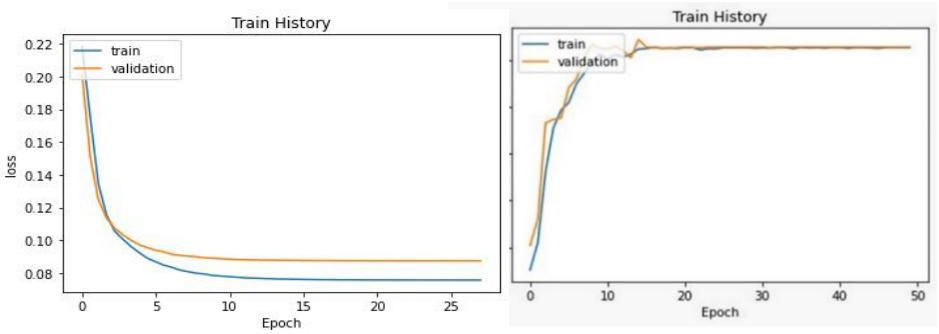
شکل ۳ – ماتریس درهم‌ریختگی تحلیل احساسات

۵-۲- استخراج قصد خرید

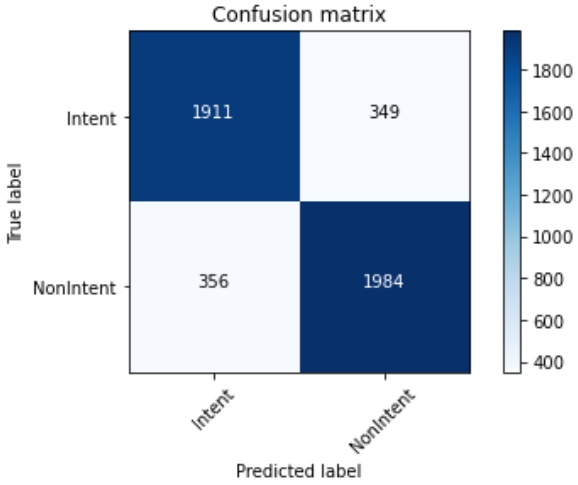
در این مرحله با استفاده از مدل نظارت شده و الگوریتم شبکه عصبی عمیق CNN-LSTM به استخراج قصد خرید توییت‌ها پرداخته می‌شود. داده‌ها پس از برچسب‌گذاری به ۳ دسته داده‌های آموزشی، اعتبارسنجی و آزمایشی تقسیم می‌شوند، این جداسازی توسط دستور (train_test_split) انجام می‌شود. که داده‌های آموزشی ٦٠ درصد مجموعهٔ داده، داده‌های آزمایشی ۲۰ درصد مجموعهٔ داده و داده‌های اعتبارسنجی نیز ۲۰ درصد آن را شامل می‌شود. نتایج استخراج قصد خرید در جدول ۲ و شکل ٤ و ٥ ارائه شده‌است.

جدول ۲ . نتایج استخراج قصد خرید

کلاس	صحت	دقت	بازخوانی	امتیاز F
توییت‌های دارای قصد خرید	۰,۸۵	۰,۸٤	۰,۸٥	۰,۸٤
توییت‌های بدون قصد خرید		۰,۸٥	۰,۸٥	۰,۸٥



شکل ۴. مقایسه دقت و هزینه در داده‌های آموزشی و تست در استخراج قصد خرید



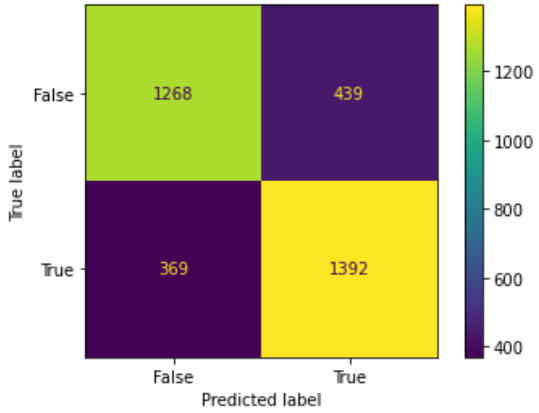
شکل ۵. ماتریس درهم‌ریختگی در استخراج قصد خرید

۵-۳- پیش‌بینی قصد خرید

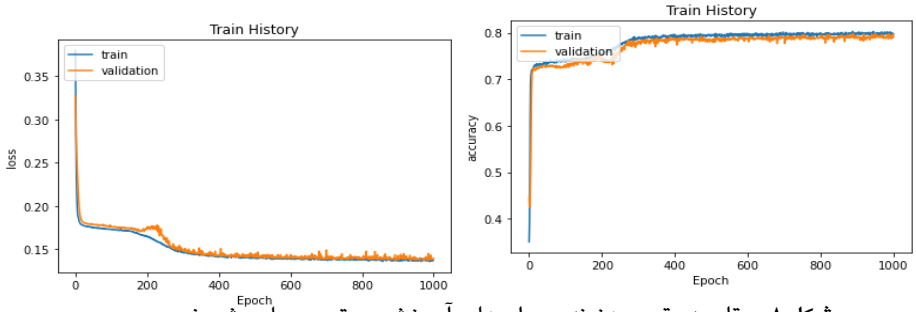
در این بخش نتایج پیش‌بینی قصد خرید در جدول ۳ و شکل ۷و ۸ ارائه شده‌اند.

جدول ۳ . نتایج پیش‌بینی قصد خرید

کلاس	صحت	دقت	بازخوانی	امتیاز F
توییت‌های دارای قصد خرید	۰,۸۰	۰,۷۷	۰,۸۵	۰,۸۱
توییت‌های بدون قصد خرید		۰,۸۳	۰,۷٤	۰,۷۸



شکل ۷- ماتریس درهم‌ریختگی پیش‌بینی قصد خرید

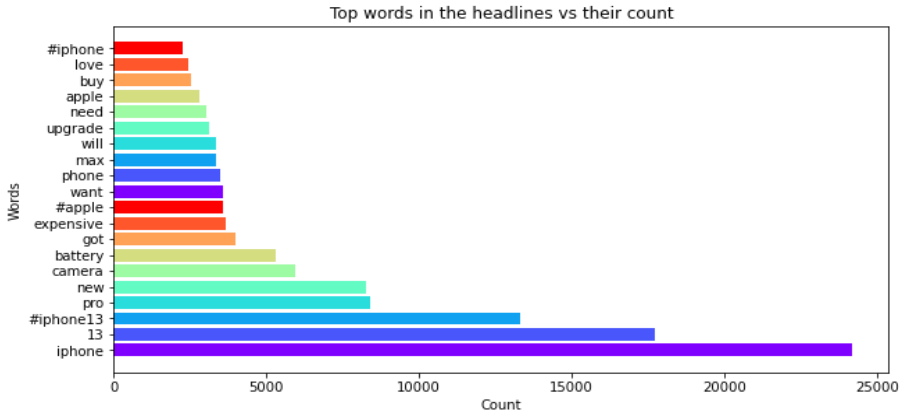


شکل ۸- مقایسه دقت و هزینه در داده‌های آموزشی و تست مدل پیش‌بینی

۶- یافته‌های کیفی

در این بخش یافته‌های کیفی هر بخش ارائه می‌شود:

۶-۱- ابر کلمات و کلمات پرتکرار

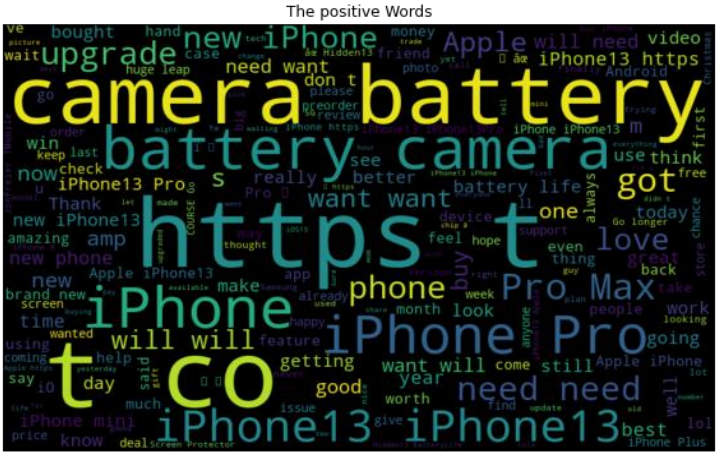


شکل ۹ . پرتکرارترین کلمات توییت‌ها

باتوجه‌به شکل۹ کلمات پرتکرار، پرتکرارترین کلمه آیفون ۱۳ است که باتوجه‌به کلمه کلیدی مورداستفاده برای جمع‌آوری داده (#iPhone) قابل پیش‌بینی است. همچنین کلماتی مانند علاقه (love)، نیاز (need)، ارتقا (upgrade)، خواستن (want)، گرفتن(got) و خریدن(buy) به چشم می‌خورد که علاقه به خرید و احساسات مثبت نسبت به آیفون را تأیید می‌کنند. از میان ویژگی‌های آیفون ۱۳ بیشترین موضوعاتی که در مورد آن‌ها صحبت شده است، باتری، دوربین و قیمت آن است که برای بررسی احساسات عمومی دررابطه‌با این ویژگی‌ها ابر کلمات توییت‌های مثبت و منفی استخراج شد. ابر کلمات یک نمایش بصری ساده و قدرتمند برای پردازش متن است که بیشترین کلمه را با حروف بزرگ‌تر و پررنگ‌تر و با رنگ‌های مختلف نشان می‌دهد. هر چه اندازه کلمه کوچک‌تر باشد اهمیت کمتری دارد.

پژوهش‌های نوین در تصمیم‌گیری ـــ

The positive Words



شکل ۱۰. ابر کلمات توییت‌های مثبت

با توجه به شکل ۱۰ در ابر کلمات توییت‌های مثبت ویژگی‌های دوربین و باتری دیده می‌شود که نشان می‌دهد نظرات کاربران نسبت به این ویژگی‌ها مثبت است و علت این امر می‌تواند عمر باتری بهبودیافته و سریع‌تر شارژ شدن این تلفن همراه باشد.

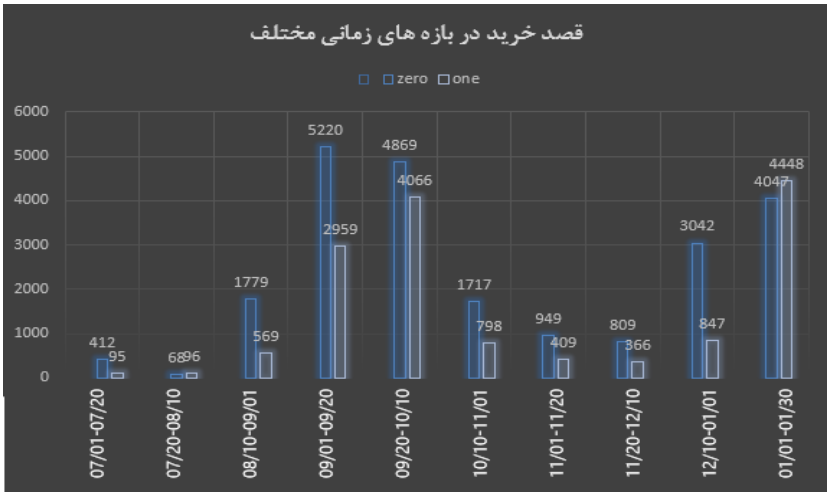
The negative Words

شکل ۱۱. ابر کلمات توییت‌های منفی

با توجه به شکل ۱۱ در ابر کلمات توییت‌های منفی کلمات گران (expensive)، شارژر (charger)، ارزش (worth)، بازکردن قفل (unlock) و هرگز (never) دیده می‌شود که نشان می‌دهد ویژگی‌های منفی آیفون باز شدن قفل صفحه، حذف شارژر از جعبه‌های آیفون و گران بودن بیش از حد این تلفن همراه است. همچنین طراحی (design) نیز در بین ویژگی‌های منفی دیده می‌شود که به عدم تغییر طراحی این تلفن همراه نسبت به مدل قبلی اشاره دارد.

۶-۲- قصد خرید در بازه‌های زمانی مختلف

همان‌طور که پیش‌تر گفته شد، توییت‌ها در بازه زمانی ۷ماهه از ۲۰۲۱/۰۷/۰۱ تا ۲۰۲۲/۰۱/۳۰ جمع‌آوری شدند و در این بخش به ده بازه زمانی تقسیم‌بندی شده‌اند و میزان قصد خرید به‌صورت جداگانه رسم شده است.



شکل ۱۲. قصد خرید در بازه زمانی‌های مختلف

همان‌طور که در نمودار شکل ۱۲ مشاهده می‌شود میزان قصد خرید در سه بازه زمانی افزایش یافته است. در مرحله بعد، بر اساس همبستگی بین ویژگی‌ها بررسی می‌شود که کدام یک از آن‌ها برای ورود به مدل مناسب هستند.

۶-۳- هم‌بستگی ویژگی‌ها

جدول ۴ . محاسبه همبستگی بین ویژگی‌ها

	تعداد کلمات	تعداد کاراکتر	نسبت تعداد کاراکتر به کلمه	تفاوت روز لانچ محصول تا روز توییت	میانگین کلمات	تحلیل احساسات	قصد خرید
تعداد کلمات	۱						
تعداد کاراکترها	۰,۹۱	۱					
نسبت تعداد کاراکتر به کلمه	-۰,۱۸	۰,۱۷	۱				
تفاوت روز لانچ محصول تا روز توییت	-۰,۰۴	-۰,۰۶	-۰,۱	۱			
میانگین کلمات	-۰,۱۸	۰,۱۶	۰,۹۹	-۰,۱۰	۱		
تحلیل احساسات	۰,۲۷	۰,۲۰	-۰,۱۵	۰,۰۴	-۰,۱۶	۱	
قصد خرید	-۰,۱۴	-۰,۲۰	-۰,۱۰	-۰,۱۴	-۰,۰۹	۰,۱۴	۱

دامنه تغییرات ضریب همبستگی از ۱+ تا ۱- نوسان دارد. یعنی در حالتی که همبستگی بین متغیرها وجود نداشته باشد ضریب همبستگی بین آن‌ها صفر و هر به سمت ۱+ نزدیک شود همبستگی مثبت شدیدتر می‌شود و همچنین هر چه به ۱- منفی نزدیک شود همبستگی منفی شدیدتر می‌شود.

باتوجه به جدول ٤، همبستگی بین ویژگی‌های تعداد کلمات و تعداد کاراکترها و همچنین بین میانگین کلمات و نسبت تعداد کاراکتر به تعداد کلمات بالا است و همبستگی بالا موجب کاهش دقت مدل و اختلال در پیش‌بینی است. ازاین‌رو، ویژگی‌های تعداد کاراکتر و میانگین کلمات کنار گذاشته شدند و ویژگی‌های تعداد کلمات، نسبت تعداد کاراکتر به کلمات، تفاوت روز رونمایی تا توییت و تحلیل احساسات برای پیش‌بینی قصد خرید انتخاب گردیدند. در میان آن‌ها تنها تحلیل احساسات و قصد خرید همبستگی مثبت دارند و افزایش یکی، افزایش دیگری را منجر می‌شود.

۷- بحث و نتیجه‌گیری

در این پژوهش سیستمی جهت استخراج قصد خرید کاربران توییتر ارائه شد که توییت‌ها را از منظر دارای قصد خرید بودن و بدون قصد خرید بودن تقسیم‌بندی می‌کند. سپس مدل پیش‌بینی قصد خرید مبتنی بر ویژگی‌های استخراج شده ارائه شد. هدف این سیستم ارائه سیستمی جهت بازاریابی هدفمند و ایجاد یک تبلیغات شخصی تنها برای مشتریان هدف محصول است که آگاهی آن‌ها را افزایش دهد، روابط را توسعه دهد و انرژی و هزینه برای کاربران خاص با قصد خرید بالا صرف شود.

معیارهای ارزیابی در جدول ۱ که نتایج تحلیل احساسات را نمایش می‌دهند، آورده شده است. در این پژوهش از واژه‌نامه VADER جهت تحلیل احساسات استفاده شد که صحت حاصل از مدل که نشان می‌دهد مدل درست آموزش‌دیده است یا خیر، ۰٫۸۱ است که عملکرد خوب در تحلیل احساسات را نشان می‌دهد. پایین بودن دقت در کلاس منفی به‌این‌علت است که تعداد کسانی که نظرات منفی در مورد آیفون ۱۳ داشته‌اند، کمتر بوده‌اند.

در میان توییت‌های جمع‌آوری شده، ۸٦۸٤ نظر مثبت و ۳۳۳٤ نظر منفی شناسایی شدند که می‌توان نتیجه گرفت احساسات عمومی دررابطه‌با آیفون ۱۳ مثبت بوده است. باتوجه‌به تحلیل ابر کلمات و کلمات پرتکرار، باتری و دوربین از ویژگی‌های مثبت این تلفن همراه و نبود شارژر و قیمت جنبه‌های منفی آن هستند. شناسایی جنبه‌های مثبت و منفی و بررسی احساسات عمومی جهت توسعه محصول، ارزیابی تصویر کلی برند، طراحی استراتژی‌های آینده و مقایسه با رقبا اهمیت زیادی دارد.

باتوجه‌به جدول ۲ نتایج استخراج قصد خرید از توییت‌ها، عملکرد خوبی را نشان می‌دهند و در میان توییت‌ها ۱۳۰۲۹ قصد خرید وجود دارد که نشان می‌دهد بیش از نیمی از توییت‌ها دارای قصد خرید هستند و مدل شبکه عصبی عمیق CNN-LSTM بادقت ۸۱ درصد آن‌ها را شناسایی می‌کند. درواقع، از میان ۱۰۰ نمونه، ۸۱ قصد خرید به‌درستی شناسایی شده است. پس از شناسایی قصد خرید ویژگی‌های ورودی به مدل پیش‌بینی استخراج می‌شوند.

همان‌طور که در شکل ۱۲ مشاهده می‌شود، میزان قصد خرید در زمان رونمایی محصول بالاترین میزان خود را دارد و بعد از آن در بازه زمانی (۱/۱-۱۲/۱۰) و (۱/۳۰-۱/۱) مجدداً افزایش‌یافته است که علت آن کمپین‌های تبلیغاتی آیفون ۱۳ است. این تبلیغات با برجسته‌کردن

ویژگی‌های مثبت آیفون ۱۳ به افزایش فروش این محصول کمک کرده‌اند. در این ویدئوهای تبلیغاتی به عمر باتری و دوام این تلفن همراه اشاره دارند. یکی از کمپین‌های آیفون که به دوربین آن اشاره دارد از دارندگان این تلفن همراه می‌خواهد که عکس‌هایی که با دوربین حالت سینماتیک آیفون ۱۳ خود انداخته‌اند را منتشر کنند. این تصاویر که کیفیت دوربین را نشان می‌دهند، در افزایش قصد خرید کاربران تأثیر به‌سزایی دارد. این مسئله به تأثیر کمپین‌هایی که کاربران نیز در آن مشارکت دارند بر قصد خرید اشاره دارد. که با استفاده از کمپین‌های تبلیغاتی میتوان قصد خرید کاربران را افزایش داد.

باتوجه‌به جدول همبستگی داده‌ها، قصد خرید تنها با تحلیل احساسات همبستگی مثبت دارد؛ اما همبستگی آن‌ها کم است. این همبستگی پایین، برخی کاربرانی هستند که با وجود احساسات مثبت توانایی خرید این محصول به علت هزینه بالای آن را ندارند و یا کاربرانی که تنها به علت لوکس بودن و برند آیفون آن را خریداری می‌کنند. همبستگی قصد خرید با دیگر ویژگی‌ها معکوس است.

درنهایت ویژگی‌هایی که با یکدیگر همبستگی بالایی دارند کنار گذاشته شدند و مدلی جهت پیش‌بینی قصد خرید از روی تعداد کلمات توییت، فاصله روز توییت تا روز رونمایی محصول، نسبت تعداد کاراکتر به تعداد کلمات و تحلیل احساسات ارائه شد. نتایج مدل شبکه عصبی عمیق پرسپترون چندلایه MLP در جدول۳ ارائه شده است و صحت ۸۰ درصد عملکرد خوبی را از مدل نشان می‌دهد و برای شناسایی بازار هدف و مشتریانی که می‌خواهند محصول را بخرند این سیستم به‌خوبی عمل می‌کند.

قصد خرید کاربران مهم‌ترین متغیر برای انجام اقدامات بازاریابی در دنیای دیجیتال است. مدیران بازاریابی با استفاده از پیش‌بینی قصد خرید می‌توانند تلاش‌های خود را محدود کنند و نتیجه بیشتری را دریافت کنند. درواقع آن‌ها می‌توانند زمان کمتری صرف مشتریان بدون قصد خرید کرده و زمان بیشتری را برای پرورش، درگیر و تبدیل مخاطبان دارای قصد خرید اختصاص دهند.

با توجه به نتایج حاصل، پیشنهاد می‌گردد سازمان‌ها و مدیران بازاریابی با شناسایی کاربرانی که قصد خرید محصول را دارند به جای بازاریابی عمومی از بازاریابی شخصی استفاده کنند. همچنین به علت اثرگذاری کمپین‌های تبلیغاتی بر قصد خرید کاربران برنامه‌ریزی دقیق‌تری

نسبت به اجرای کمپین‌ها داشته باشند تا هزینه‌ها مدیریت شده و بازگشت سرمایه سریع‌تر اتفاق بیفتد. در این پژوهش ویژگی‌های مثبت و منفی محصول از نگاه کاربران شناسایی شدند که نقش مهمی در مرحله تولید و توسعه محصول دارد و میتوان محصولی مطابق نظر کاربران ارائه داد. این سیستم برای محصول آیفون ۱۳ ارائه شد؛ ولی پیشنهاد می‌شود در پژوهش‌های آتی از محصولات و یا خدمات دیگر استفاده شود.

پیشنهاد می‌شود در پژوهش‌های آینده از مجموعه‌داده‌های دیگر در حوزه‌های متعدد مانند گردشگری، ورزش، مراقبت‌های بهداشتی، سلامت عمومی و رویدادهای غم‌انگیز مانند خودکشی جهت پیش‌بینی قصد استفاده شود. همچنین در این پژوهش از داده‌های توییتر برای پیش‌بینی قصد خرید استفاده شده است و استفاده از سایر شبکه‌های اجتماعی برای بررسی نیز پیشنهاد می‌شود. باتوجه‌به این مسئله که توسط سیستم ارائه شده مشتریان بالقوه محصول شناسایی می‌شوند، ارائه سیستم توصیه‌گر محصول که به آن‌ها محصول مورد نظرشان را توصیه می‌کند، نیز پیشنهاد می‌شود.

۸- پی‌نوشت‌ها

۱. Twitter
۲. Purchase Intent
۳. Sentiment Analysis
٤. Support Vector Machine(SVM)
٥. Convolutional Neural Network (CNN)
٦. Long Short-Term Memory (LSTM)
۷. Perceptron
۸. Cross-Industry Standard Process
۹. Lexicon Based Approach

۹- منابع

[1] Alikhani, Y., khadivar, A., & abbasi, F. (2022). Assessing the public perception of the fifth generation of mobile communication (5G) by sentiment analysis of Twitter users. Modern Research in Decision Making, 7(2), 111-135.

[2] Zarei, A., feiz, D., Taheri, G. (2021). Providing Social Market Intelligence Framework based on web 2.0 Using Text-Mining Technique on Social Media

Websites (Case Study: Competitive Analysis between Samsung and Emersun Brands). Management Research in Iran, 24(4), 98-125.

[3] Khani, M., saghaei, A., & Heidarzadeh Hanzaee, K. (2019). Modeling New Product Launch Strategies within agent-based simulation. Modern Research in Decision Making, 4(4), 176-196.

[4] T. Le Luong, T. H. Tran, Q. T. Truong, T. M. N. Truong, T. T. Phi, and X. H. Phan, "Learning to filter user explicit intents in online vietnamese social media texts," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016, vol. 9622, pp. 13–24, doi: 10.1007/978-3-662-49390-8_2.

[5] D. Gaikar and B. Marakarkandy, "Product Sales Prediction Based on Sentiment Analysis Using Twitter Data," vol. 6, no. 3, pp. 2303–2313, 2015.

[6] S. Kumar, A. K. Kar, and P. V. Ilavarasan, "Applications of text mining in services management: A systematic literature review," *Int. J. Inf. Manag. Data Insights*, vol. 1, no. 1, p. 100008, 2021, doi: 10.1016/j.jjimei.2021.100008.

[7] Afsharizadeh jafari, O., Maleki MinBashRazgah, M., Zarei, A., & Shafiei Nikabadi, M. (2021). Designing a ranking system for purchased products based on the consumer's and expert's opinions using an aspect-based sentiment analysis approach. Modern Research in Decision Making, 6(2), 20-47. [8]B. Hollerit, M. Kröll, and M. Strohmaier, "Towards linking buyers and sellers," no. May, pp. 629–632, 2013, doi: 10.1145/2487788.2488009.

[9] J. Guerreiro and P. Rita, "How to predict explicit recommendations in online reviews using text mining and sentiment analysis," *J. Hosp. Tour. Manag.*, vol. 43, no. November 2018, pp. 269–272, 2020, doi: 10.1016/j.jhtm.2019.07.001.

[10] A. B. Goldberg, N. Fillmore, D. Andrzejewski, Z. Xu, B. Gibson, and X. Zhu, "May all your wishes come true: A study of wishes and how to recognize them," *NAACL HLT 2009 - Hum. Lang. Technol. 2009 Annu. Conf. North Am. Chapter Assoc. Comput. Linguist. Proc. Conf.*, no. December 2013, pp. 263–271, 2009.

[11] Z. Chen, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh, "Identifying intention posts in discussion forums," *NAACL HLT 2013 - 2013 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. Proc. Main Conf.*, no. January, pp. 1041–1050, 2013.

[12] M. Korpusik, S. Sakaki, F. Chen, and Y. Y. Chen, "Recurrent neural networks for customer purchase prediction on Twitter," *CEUR Workshop Proc.*, vol. 1673, pp. 47–50, 2016, doi: 10.475/123.

[13] A. Khattak, A. Habib, M. Z. Asghar, F. Subhan, I. Razzak, and A. Habib, "Applying deep neural networks for user intention identification," *Soft Comput.*,

vol. 25, no. 3, pp. 2191–2220, 2021, doi: 10.1007/s00500-020-05290-z.

[14] A. Habib, N. Jelani, A. M. Khattak, S. Akbar, and M. Z. Asghar, "Exploiting Deep Neural Networks for Intention Mining," in *ACM International Conference Proceeding Series*, 2020, pp. 26–30, doi: 10.1145/3384544.3384607.

[15] Zolfaghar, P., khadivar, A., abbasi, F. (2022). Investigating the effect of perceived risk by social network users on the price of Bitcoin using sentiment analysis technique and NARX neural network. Management Research in Iran, 26(2), 18-41.

[16] M. Allahyari *et al.*, "A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques," 2017, [Online]. Available: http://arxiv.org/abs/1707.02919.

[17] A. Ibrahim, "Forecasting the early market movement in bitcoin using twitter's sentiment analysis: An ensemble-based prediction model," *2021 IEEE Int. IOT, Electron. Mechatronics Conf. IEMTRONICS 2021 - Proc.*, 2021, doi: 10.1109/IEMTRONICS52119.2021.9422647.

[18] Q. Han, J. Guo, and H. Schütze, "CodeX: Combining an SVM classifier and character N-gram language models for sentiment analysis on twitter text," **SEM 2013 - 2nd Jt. Conf. Lex. Comput. Semant.*, vol. 2, no. SemEval, pp. 520–524, 2013.

[19] B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A Survey on Evolutionary Computation Approaches to Feature Selection," *IEEE Trans. Evol. Comput.*, vol. 20, no. 4, pp. 606–626, Aug. 2016, doi: 10.1109/TEVC.2015.2504420.

[20] J. Sundström, "Sentiment analysis of Swedish reviews and transfer learning using Convolutional Neural Networks," *undefined*, 2018.