



Designing and Implementing an Emotion Analytic System (EAS) on Instagram Social Network Data

Seyed Faridoddin Kiaei ¹, Saeed Farzi ²

¹ Faculty of Engineering, University of Tehran, Tehran, Iran
sfd.kiaei@ut.ac.ir

² Faculty of Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran
saeedfarzi@kntu.ac.ir

Abstract

Being aware of people's attitudes and emotions of a person or event can have a huge impact on the decisions of individuals and organizations. With the rise of social networks and in particular Instagram because of its popularity, many people are sharing their attitudes on this network. Analyzing the emotions of users of this social network, which is a great example of community, can help managers make organizational decisions and predict important events such as elections. In this research, the EAS system was designed and implemented to extract emotion and visualize them. As a practical example, the Instagram users' feelings about the two main candidates for the 12th Iranian presidential election are also examined. The result shows a positive feeling about the word "روحانی" and a high level of confidence emotion along with anger and disgust emotions.

Keywords: Emotion Analysis, Visualization, Instagram, Election



طراحی و پیاده سازی سامانه تحلیل احساس در شبکه اجتماعی اینستاگرام

سید فریدالدین کیائی^۱، سعید فرضی^۲

^۱ پردیس دانشکده های فنی، دانشگاه تهران
sfd.kiaei@ut.ac.ir

^۲ دانشکده کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی
saeedfarzi@kntu.ac.ir

چکیده

آگاهی یافتن از نظر و احساس افراد نسبت به یک شخص یا واقعه می تواند در تصمیم گیری های افراد و سازمان ها تاثیر بسزایی داشته باشد. با گسترش شبکه های اجتماعی و به طور خاص اینستاگرام به دلیل محبوبیت بالای آن، بسیاری از افراد نظرات خود را در این شبکه به اشتراک می گذارند. تحلیل احساس کاربران این شبکه اجتماعی که نمونه بسیار خوبی از جامعه است، می تواند به مدیران در تصمیم گیری های سازمانی و پیش بینی وقایع مهم مانند انتخابات کمک کند. در این پژوهش سامانه EAS جهت استخراج احساسات موجود در عبارات و مصورسازی آن ها طراحی و پیاده سازی شده است. همچنین به عنوان یک مثال کاربردی، احساسات کاربران شبکه اجتماعی اینستاگرام نسبت به دو کاندید اصلی دوازدهمین دوره انتخابات ریاست جمهوری اسلامی ایران بررسی شده است. نتیجه حاصل نشان دهنده احساس مثبت نسبت به کلمه "روحانی" و بالا بودن میزان احساس اعتماد در کنار احساس خشم و انزجار است.

کلمات کلیدی

تحلیل احساس، مصورسازی، اینستاگرام، انتخابات

شبکه های اجتماعی بزرگی همچون فیسبوک تمایل دارند که بر چشم انداز شبکه های اجتماعی مسلط شوند. در این میان با فراگیر شدن تلفن همراه هوشمند استفاده از شبکه اجتماعی اینستاگرام که در دسترس تر از شبکه اجتماعی فیسبوک برای کاربران است، محبوبیت بیشتری یافته است [3]. تجزیه و تحلیل احساس^۱، که استخراج نظر^۲ نیز نامیده می شود، زمینه مطالعه ای است که نظرات، ارزیابی ها، نگرش ها و احساسات افراد را نسبت به موجودیت هایی از قبیل محصولات، خدمات، سازمان ها، افراد، مسائل، رویدادها، موضوعات و ویژگی های آن ها تحلیل می کند. عقاید تقریباً در تمام فعالیتهای انسانی بسیار مهم است زیرا آن ها تأثیرگذارترین رفتارهای ما هستند هر زمان که نیاز به تصمیم گیری داشته باشیم می خواهیم نظرات دیگران را بدانیم. در دنیای واقعی، مشاغل و سازمان ها همیشه می خواهند نظر مصرف کنندگان را درباره محصولات و خدمات خود بیابند. مصرف کنندگان نیز همچنین می خواهند قبل از تصمیم گیری از نظرات خریداران قبلی درباره یک محصول اطلاع پیدا کنند. در انتخابات رأی دهندگان می خواهند قبل از رأی

۱- مقدمه

اولین شبکه اجتماعی مجازی در ۱۹۹۷ ایجاد شد. این شبکه ها برای افراد امکان ساخت صفحه پروفایل عمومی، ساخت لیست دوستان و مشاهده ارتباط و لیست دوستان آن ها را فراهم می کرد. [1] امروزه مردم خواه ناخواه اطلاعات زیادی در مورد نظرات، تصمیمات و احساسات خود را در فضای مجازی به اشتراک می گذارند. جمع آوری و تحلیل این اطلاعات کاربرد های مختلفی از جمله پی بردن به احساسات جامعه نسبت به یک واقعه و یا شخصی خاص و پیش بینی وقایع پیش رو مانند انتخابات را دارد.

در [2] نشان داده شده است که چگونه رسانه های اجتماعی خرد جمعی را بیان می کنند، که در صورت استفاده صحیح، می توانند یک شاخص بسیار قدرتمند و دقیق از نتایج آینده را نشان دهند. با گسترش استفاده از شبکه های اجتماعی و افزایش تمایل مردم به اشتراک گذاری نظرات و دیدگاه های خود،



۳-۱- روند تحقیق

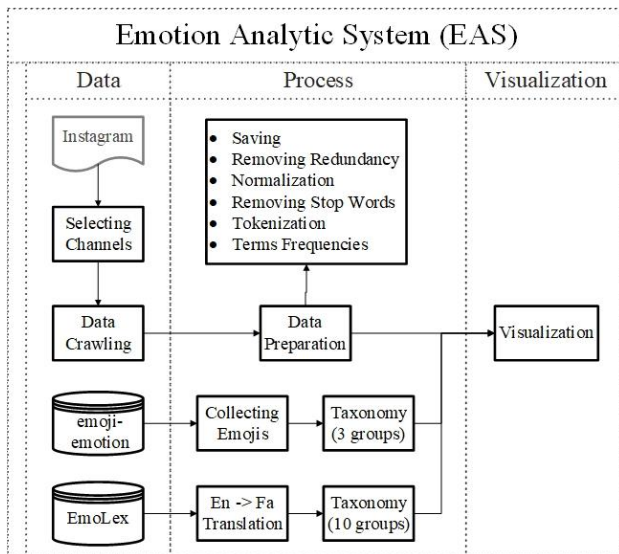
مرحله اول تحقیق (شکل ۱) شامل انتخاب صفحات مجازی اینستاگرام جهت جمع آوری داده، پیاده سازی خزشگر^۴، جمع آوری و ساخت پایگاه داده شکلک و واژه نامه است.

در مرحله دوم پیش پردازش داده‌های جمع آوری شده انجام می شود. شکلک‌ها به سه دسته مثبت، منفی و خنثی دسته بندی می شوند. پایگاه داده واژه نامه نیز ابتدا فارسی سازی شده سپس به ده گروه دسته بندی می شود. در آخرین مرحله، اطلاعات با استفاده از دو روش مبتنی بر شکلک و مبتنی بر واژه نامه مصورسازی می شوند.

۳-۲- خزش داده

ابتدا یک اکانت در اینستاگرام ساخته شده و ۴۰ کانال طرفدار مرتبط سیاسی دنبال^۴ گردید. در این مرحله سعی شد مجموع تعداد اعضای دو گروه سیاسی تقریباً با هم برابر باشند تا تحلیل اطلاعات دقیق تر انجام گیرد.

ساختار کامنت ها شامل متن، تاریخ و زمان، آی دی کامنت، آی دی پست و آی دی کاربر است. به دلیل محدودیت و سختی گرفتن api اینستاگرام و استفاده از Ajax برای نمایش همه کامنت ها در نسخه وب آن، با استفاده از reverse api از تاریخ ۱۲ April 2017 (1396/01/26 A.H.S) to 29 July 2017 (1396/05/07 A.H.S) که در محدوده انتخابات ریاست جمهوری ایران بود، همه کامنت ها به صورت json جمع آوری گردید. به دلیل سیاسی بودن پست ها و احتمال حذف آن ها، چند بار در روز داده ها جمع آوری شدند و نهایتاً تکراری ها حذف گردید. برای خزش داده ها از زبان جاوا و کتابخانه رایگان Jaunt که برای نوشتن برنامه‌های دریافت صفحات وب کاربرد دارد استفاده شد.



شکل (۱): روند نمای کلی تحقیق سه مرحله اصلی تحقیق عبارتند از جمع آوری داده، پردازش و مصورسازی.

دادن، نظرات دیگران را درباره نامزدهای انتخاباتی بدانند. در گذشته، وقتی فردی به عقاید نیاز داشت، از دوستان و خانواده می پرسید. هنگامی که یک سازمان یا یک کسب و کار به نظرات عموم مردم یا نظرات مصرف کنندگان نیاز داشت، از طراحی پرسشنامه یا نظرسنجی استفاده می کرد. به دست آوردن نظرات عمومی و مصرف کننده مدت هاست که برای تجارت و بازاریابی، روابط عمومی و شرکت های تبلیغاتی سیاسی یک تجارت بزرگ بوده است. با افزایش استفاده از شبکه های اجتماعی، افراد و سازمان ها به طور فزاینده از محتوای موجود در این شبکه ها برای تصمیم گیری استفاده می کنند. [4] در این پژوهش سامانه‌ای جهت تشخیص احساسات موجود در داده شبکه اجتماعی طراحی و پیاده سازی شده است. این سامانه از دو روش کلی تحلیل مبتنی بر شکلک و مبتنی بر واژه نامه جهت استخراج احساسات موجود در متون فارسی استفاده می کند. به منظور بررسی عملکرد این سامانه و اهمیت شبکه اجتماعی اینستاگرام در بین شبکه های اجتماعی، داده های این شبکه در بازه زمانی انتخابات ریاست جمهوری سال ۱۳۹۶ جمهوری اسلامی ایران جمع آوری شده و نتایج آن مورد بررسی قرار گرفت.

۲- پژوهش های مرتبط

در [5] سامانه‌ای جهت مصورسازی بلادرنگ میکرو بلاگ های توئیتر و تحلیل آن‌ها طراحی شده است. شبکه برجسب وزن دار جهت ارائه بینش معنایی پیشرفته در این پژوهش طراحی شده است. در [6] سامانه‌ای جهت یافتن الگو های معنادار از میان داده‌های ناهمگون و بدون ساختار شبکه‌های اجتماعی مانند اینستاگرام جهت تشخیص رویداد طراحی شده است. در [7] علاوه بر داده‌های کاربران اینستاگرام، از ترکیب داده‌های کاربران اینستاگرام و توئیتر جهت بهبود کیفیت تشخیص رویداد استفاده کرده است. [8] از روش های بدون نظارت جهت تشخیص رویداد با استفاده از داده‌های کاربران اینستاگرام و توئیتر استفاده کرده است. در [9] تعداد پیام‌های چهار دسته ترامپ، کلینتون و طرفداران، مخالفان آن‌ها در اینستاگرام مورد بررسی قرار گرفته است. تعیین دسته‌ها با استفاده از هشتگ‌ها انجام شده است. نتایج بررسی پیام‌های یک روز قبل از انتخابات ریاست جمهوری ۲۰۱۶ ایالات متحده آمریکا نشان می دهد که طرفداران کلینتون بیشتر بوده‌اند. اما از آنجایی که ۶۰٪ کاربران اینستاگرام در گروه سنی ۱۸-۳۵ سال قرار داشته‌اند، پیش بینی آن‌ها با نتایج واقعی متفاوت بوده است. این نمونه، به خوبی عدم امکان تعمیم نتایج برگرفته از اینستاگرام به کل جامعه را نشان می دهد.

۳- سامانه EAS

در این بخش نحوه جمع آوری، پیش پردازش، مصورسازی داده و معماری سامانه توضیح داده شده است.



۳-۳- پیش پردازش داده

سپس به دلیل حجم زیاد داده ها و دستیابی پویا، سریع و پایدار به اطلاعات، داده ها در Microsoft SQL Server ذخیره شد.

متن کامنت ها به صورت یونیکد دریافت شده بود که به دلیل فارسی بودنشان به UTF-8 و با دقت در حفظ شکلک ها تبدیل گردید. سپس به جهت جلوگیری از اشتباه کامپیوتر در تشخیص کلماتی که از نظر انسان ممکن است یکسان باشند اما از نظر کامپیوتر نه، با کمک کتابخانه هضم [10] در پایتون داده ها نرمال سازی شد. نرمال سازی شامل حذف نیمفاصله های تکراری، اصلاح فاصله گذاری پسوندها، تبدیل ارقام عربی و انگلیسی به معادل فارسی و... است. سپس کلمات توقف که علی رغم تکرار فراوان در متن، از لحاظ معنایی دارای اهمیت کمی هستند، جمع آوری شده [11]. کاستی های آن برطرف شده و به جهت بهبود نتایج پردازش و افزایش سرعت سامانه حذف گردید. پس از آن متن تقطیع شده و کلمات و شکلک ها جدا گردید.

برای ساخت گرافی که بتواند خلاصه ای از تمام داده های موجود را نمایش دهد، نیاز به دانستن بیشترین کلماتی که با هم تکرار شده اند است. بنابراین تعداد تکرار همه جفت واژه ها را یافته و تعداد تکرار مشترک و تعداد تکرار هر یک به تنهایی محاسبه و ذخیره گردید تا در پردازش های بعدی به راحتی بتوان بر تکرارترین جفت واژه ها را بدست آورد.

۳-۴- مصورسازی داده

داده ها مهم ترین چیزی هستند که با آن ها می توان مسائل را تحلیل و بررسی کرد. به همین دلیل انتخاب نمودار یا گراف درست و مناسب مهم است، چراکه ابزاری قوی برای درک بهتر داده ها را فراهم آورده و اگر درست انتخاب نشوند می توانند باعث گمراهی در فهمیدن داده ها شوند.

۳-۴-۱- تنظیمات

در بخش تنظیمات عبارات^۲ (شکل ۲) و (شکل ۳ - A) امکان جستجوی حداقل یک و حداکثر پنج عبارت وجود دارد. در سامانه Soda [5] فقط امکان جستجوی یک عبارت وجود داشت اما در سامانه EAS علاوه بر نمایش نتایج هر عبارت، امکان مقایسه تا پنج عبارت نیز وجود دارد. همچنین نمودار (شکل ۲) و (شکل ۳ - E) قابل رسم به صورت عادی و پشته ای است که تعیین نوع آن نیز در این قسمت انجام می شود. می توان برای آشنایی با داده ها و انتخاب عبارات مرتبط تر، از گراف موجود در سامانه استفاده کرد. همچنین امکان تعیین تاریخ و ساعت شروع و پایان کامنت ها جهت تحلیل دقیق تر نیز در تنظیمات زمان^۳ (شکل ۲) و (شکل ۳ - A) وجود دارد. پس از تعیین عبارات و بازه زمانی مورد نظر می توان عملیات مصورسازی را آغاز کرد.

۳-۴-۲- مدل سازی و مصورسازی کامنت ها

ابتدا خلاصه ای از کل داده های موجود در پایگاه داده توسط گراف (شکل ۲ - F) نمایش داده می شود. هر گره نماینده یک کلمه و اندازه آن نشان دهنده

تعداد تکرار کلمه و در نتیجه اهمیت آن است. گره ها با یال هایی به هم متصل می شوند که اندازه یال نشان دهنده تعداد تکرار مشترک هر دو کلمه ی دو سر یال با هم در یک کامنت است.

محاسبه گراف به این شکل است که ابتدا پنج کلمه از پر تکرارین کلمات انتخاب شده و پنج تا از بیشترین کلماتی که با آن ها تکرار شده اند نیز به آن ها اضافه می شوند. سپس با این اطلاعات گراف رسم می شود. گراف بدست آمده قادر است مهم ترین کلمات به کار رفته در متن را در یک نگاه به کاربر نشان داده و دیدی کلی از داده ها به کاربر بدهد. این گراف در ابتدای اجرای برنامه نمایش داده می شود تا بتواند به کاربر در انتخاب عبارات کمک کند.

۳-۴-۳- مدل سازی و مصورسازی توزیع احساسات

پس از انتخاب عبارت (عبارات) مورد نظر و آغاز مصورسازی، در بخش اول کامنت ها بر اساس شکلک های موجود در آن ها به سه دسته مثبت، منفی و خنثی دسته بندی می شوند. نتیجه در نمودار میله ای (شکل ۲) و (شکل ۳ - D) نمایش داده شده است که محور افقی نشان دهنده دسته و محور عمودی درصد کامنت ها به ازای هر عبارت ورودی را نشان می دهد. در این نمودار نتایج نهایی نرمال شده و مجموع دسته ها برای هر عبارت برابر ۱ است. در بخش دوم کامنت ها بر اساس کلمات احساسی به کار رفته در آن ها به ده دسته تقسیم می شوند. دسته بندی به این صورت انجام می شود که اگر کامنتی شامل کلمه موجود در هر یک از دسته ها بود، عضو آن دسته می شود. بنابراین هر کامنت می تواند عضو بیش از یک دسته باشد. سپس نتیجه در نمودار میله ای (شکل ۳ - D) نمایش داده می شود. محور افقی دسته ها و محور عمودی درصد کامنت های هر دسته را نشان می دهد.

۳-۴-۴- مدل سازی و مصورسازی زمان

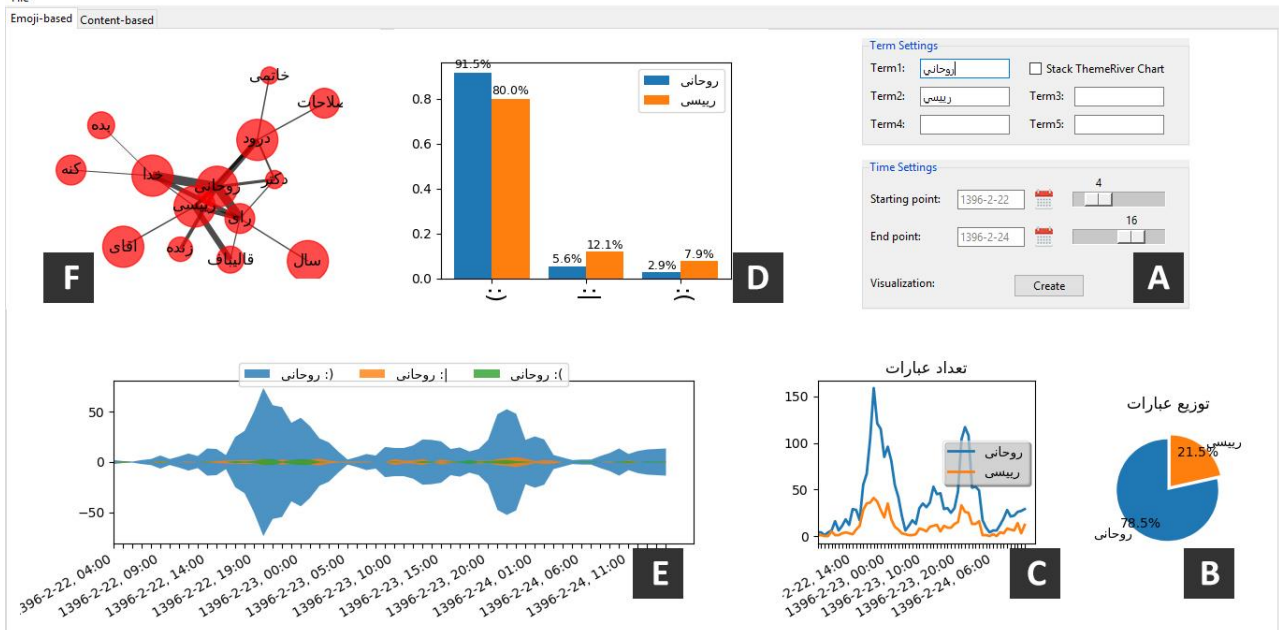
جهت بررسی دقیق تر عبارت اول که عبارت اصلی است، از نمودار روحانه ای^۴ که مناسب ترین نمودار جهت نمایش جریان داده در زمان است استفاده شد. به این صورت که در نمودار بخش اول (شکل ۲ - E) محور افقی زمان و محور عمودی تعداد کامنت های هر دسته را نشان می دهد. در این نمودار از حالت پشته ای استفاده نشده و مقایسه نسبت دسته ها آسان تر است. در نمودار بخش دوم (شکل ۲) و (شکل ۳ - E) محور مانند بخش اول محور افقی زمان و محور عمودی تعداد کامنت های هر دسته را نشان می دهد. در این شکل (شکل ۲) و (شکل ۳ - E) از حالت پشته ای این نمودار استفاده شده و هر دسته با رنگ خاصی مشخص شده است.

۳-۴-۵- مصورسازی تعداد و توزیع کامنت ها

جهت آگاهی از تعداد و توزیع عبارات مورد جستجو در بین تمامی کامنت ها و نسبت آنان، دو نمودار در برنامه قرار داده شده است. برای مقایسه کلی عبارات، توزیع کامنت ها به ازای هر عبارت مورد جستجو در نموداری دایره ای (شکل ۲) و (شکل ۳ - B) نمایش داده می شود. نمودار دایره ای برای بررسی داده های هم جنس در یک مقطع از زمان مناسب است. همچنین برای بررسی دقیق تر داده ها و اطلاع از تعداد دقیق کامنت های شامل هر یک از عبارت

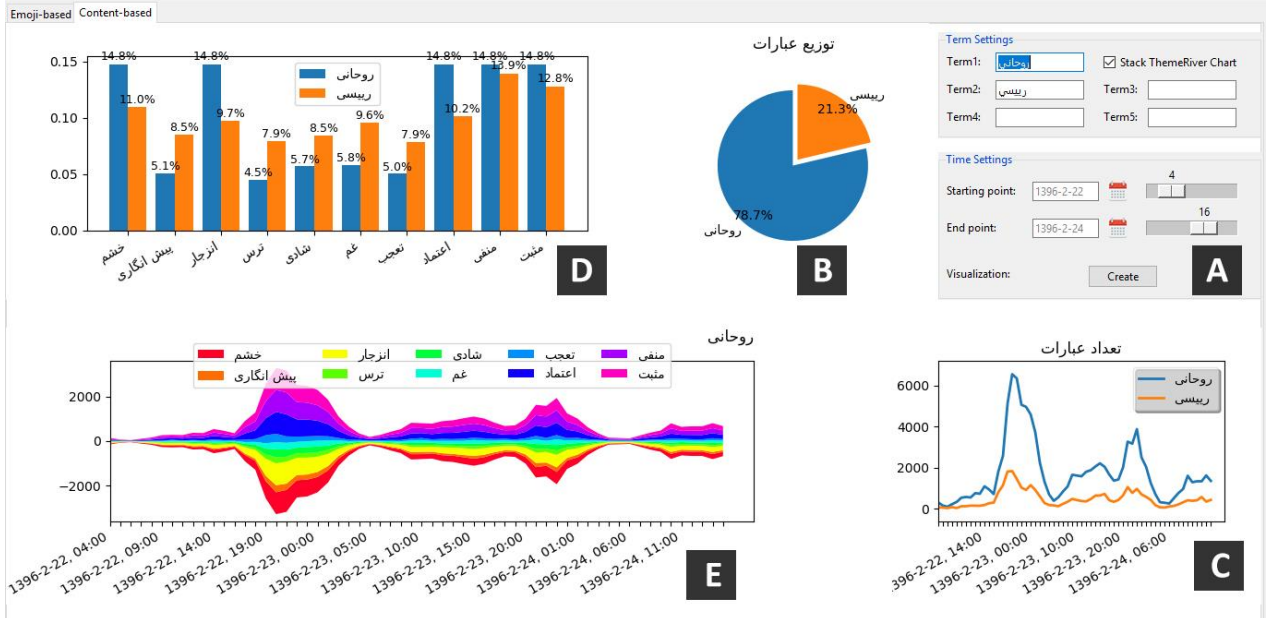


Emotions Analytic System (EAS)



شکل (۲): خروجی مبتنی بر شکلک سامانه تعاملی EAS برای عبارات "روحانی" و "رئیس" (چند روز پیش از دوازدهمین انتخابات ریاست جمهوری ایران) بر اساس تحلیل و مصورسازی کامنت های اینستاگرام دو بعدی (متن و زمان). گراف خلاصه اطلاعات (F) نمایشی از کلمات مهم داده ها که شامل "روحانی" و "رئیس" دو رقیب اصلی در انتخاب مذکور بودند را نشان می دهد. نمودار ستونی (D) نشان دهنده توزیع کلی کامنت ها به ازای هر عبارت در دسته های مثبت، خنثی و منفی است. این نمودار درصد محبوبیت بیشتری را برای "روحانی" که پس از این انتخابات به عنوان رئیس جمهور ایران انتخاب گردید نشان می دهد. نمودار رودخانه‌ای (E) توزیع نتایج عبارت اول را در زمان نشان می دهد.

Emotions Analytic System (EAS)



شکل (۳): خروجی مبتنی بر واژه نامه سامانه تعاملی EAS برای عبارات "روحانی" و "رئیس" (چند روز پیش از دوازدهمین انتخابات ریاست جمهوری ایران) بر اساس تحلیل و مصورسازی کامنت های اینستاگرام دو بعدی (متن و زمان). نمودار ستونی (D) نشان دهنده توزیع کلی کامنت ها به ازای هر عبارت در دسته های ده گانه است. این نمودار درصد اعتماد بیشتر و ترس کمتری را برای "روحانی" که پس از این انتخابات به عنوان رئیس جمهور ایران انتخاب گردید نشان می دهد.



- Negative (منفی)
- Positive (مثبت)

ابتدا به کمک مترجم گوگل آن ها را به زبان فارسی ترجمه کرده سپس کلمات تکراری که پس از ترجمه اضافه شده بودند حذف گردید. پس از آن مانند مرحله قبل تعداد و نسبت هر دسته به ازای عبارت (عبارات) مورد جستجو در بازه زمانی مورد نظر محاسبه شده و نمایش داده شد.

۴- نتایج

با استفاده از سامانه طراحی شده و کامنت های جمع آوری شده از اینستاگرام، دو کلمه کلیدی "روحانی" و "رئییسی" که دو کاندید اصلی انتخابات ریاست جمهوری سال ۱۳۹۶ ایران بودند، به سامانه داده شدند. نتیجه تحلیل مبتنی بر شکلک (جدول ۱) بدین صورت است که میزان خوشحالی نسبت به کلمه روحانی بیشتر از کلمه رئییسی، و میزان ناراحتی کمتر بوده است. نتیجه تحلیل مبتنی بر واژه نامه (جدول ۲) نیز نسبت به کلمه روحانی مثبت بیشتری را نشان می دهد. نکته جالب توجه در این جدول، بالا بودن احساس اعتماد نسبت به کلمه روحانی در عین بالا بودن احساس خشم و انزجار نسبت به آن است. بنابراین بر اساس شواهد می توان این گونه نتیجه گرفت که جناب آقای روحانی پیروز انتخابات هستند اما لزوما همه مردم از این موضوع خوشحال نیستند.

۵- چالش های تحقیقاتی

کشف احساس موجود در متن با کمک شکلک و یا واژگان احساسی به کار رفته در آن امکان پذیر است اما خطای زیادی دارد. برای دستیابی به نتایج

جدول (۱): نتایج تحلیل مبتنی بر شکلک

	☺	☹	☹	
روحانی	5.6%	2.9%	91.5%	
رئییسی	12.1%	7.9%	80%	

جدول (۲): نتایج تحلیل مبتنی بر واژه نامه

	رئییسی	روحانی	
خشم	11%	14.8%	
پیش انگاری	8.5%	5.1%	
انزجار	9.7%	14.8%	
ترس	7.9%	4.5%	
شادی	8.5%	5.7%	
غم	9.6%	5.8%	
تعجب	7.9%	5%	
اعتماد	10.2%	14.8%	
منفی	13.9%	14.8%	
مثبت	12.8%	14.8%	

(عبارات) مورد جستجو در طول زمان و مشاهده نرخ تغییرات آن ها، از نموداری خطی (شکل ۲) و شکل (۳ - C) در کنار نمودار قبلی استفاده شده است.

۳-۵- معماری سامانه

سامانه معرفی شده در بالا با موفقیت پیاده سازی گردید. پیاده سازی این سامانه با زبان پایتون و با کمک کتابخانه tkinter انجام شده است. برای کاربرپسند بودن بیشتر برنامه از ماژول تقویم برای انتخاب زمان شروع و پایان (شکل ۲) و شکل (۳ - A) استفاده شده است. همچنین همه نمودار ها دارای قابلیت زوم با غریبلک موس هستند. از آن جایی که این برنامه جهت بررسی صفحات فارسی طراحی شده است، همه تاریخ ها از میلادی به شمسی تبدیل شده اند.

۳-۶- تحلیل احساس

برای تشخیص احساس موجود در جمله از دو روش مختلف شکلک و واژه نامه استفاده شد.

۳-۶-۱- مبتنی بر شکلک

برای تشخیص احساس موجود در جمله توسط شکلک، ابتدا از یک مجموعه داده [12] که اکثر شکلک های مورد استفاده را دارا بوده و احساس موجود در آن ها را با قطبیت بین ۵- و ۵، نشان داده بود، استفاده شد. بر اساس این قطبیت، شکلک ها را به سه دسته مثبت، منفی و خنثی تقسیم گردید. سپس در هر کامنتی که شکلک وجود داشت، دسته آن شکلک تشخیص داده شده و به تعداد تکرار آن دسته یک واحد اضافه گردید. در آخر مجموع هر دسته برای عبارت (عبارات) مورد جستجو نمایش داده شد.

۳-۶-۲- مبتنی بر واژه نامه

این بار برای تشخیص احساس از واژه نامه ای [13] که کلمات انگلیسی را به هشت دسته احساسی و دو دسته مثبت و منفی تقسیم کرده بود استفاده گردید. در این مجموعه داده برای هر کلمه ده دسته وجود دارد که اگر هر کدام را داشته باشد مقدار یک و در غیر این صورت مقدار صفر جلوی آن نوشته شده است. این دسته ها به شرح زیر هستند:

- Anger (خشم)
- Anticipation (پیش انگاری)
- Disgust (انزجار)
- Fear (ترس)
- Joy (شادی)
- Sadness (غم)
- Surprise (تعجب)
- Trust (اعتماد)



- [6] K. Polous, A. Freitag, J. Krisp, L. Meng, and S. Singh, "A general framework for event detection from social media," in *Advances in Geographic Information Science*, 2015, vol. 19, pp. 85–105.
- [7] P. Giridhar and T. Abdelzaher, "Visualization of events using Twitter and Instagram," in *2017 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2017*, 2017, pp. 82–84.
- [8] S. Liu and P. Jansson, "City event detection from social media with neural embeddings and topic model visualization," in *Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017*, 2017, vol. 2018-Janua, pp. 4111–4116.
- [9] H. Schmidbauer, A. Rösch, and F. Stieler, "The 2016 US presidential election and media on Instagram: Who was in the lead?," *Comput. Human Behav.*, vol. 81, pp. 148–160, Apr. 2018.
- [10] HAZM, "Python library for digesting Persian text," *Sobhe*, 2014. [Online]. Available: <https://github.com/sobhe/hazm>. [Accessed: 14-Feb-2020].
- [11] "kharazi/persian-stopwords: Persian (Farsi) Stop Words List." [Online]. Available: <https://github.com/kharazi/persian-stopwords>. [Accessed: 14-Feb-2020].
- [12] T. Wormer, "List of emoji rated for valence in JSON." [Online]. Available: <https://github.com/words/emoji-emotion>. [Accessed: 14-Feb-2020].
- [13] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word-emotion association lexicon," in *Computational Intelligence*, 2013, vol. 29, no. 3, pp. 436–465.

زیر نویس ها

- ¹ Sentiment Analysis
- ² Opinion Mining
- ³ Crawler
- ⁴ Follow
- ⁵ Emoji
- ⁶ Tokenize
- ⁷ Term settings
- ⁸ Time settings
- ⁹ Theme River Chart

بتر می توان از روش های پیچیده تر مانند یادگیری ماشین و پردازش زبان طبیعی استفاده کرد.

از دیگر موارد قابل بررسی، تشخیص کاربران جعلی موجود در شبکه های اجتماعی است که در صورت انجام تحلیل های مبتنی بر اعضا، این مورد اهمیت بیشتری پیدا می کند.

مصورسازی بلادرنگ اطلاعات و نمایش آن ها در وب هم از مواردیست که به دلیل تغییر زیاد api اینستاگرام انجام نمی شود. این مسئله را نیز می توان در مراحل بعدی این پژوهش مدنظر قرار داد.

همچنین با بدست آوردن اطلاعات مکانی کاربران که درصد بسیار کمی آن را در پروفایل خود به اشتراک می گذارند، می توان تحلیل ها و نمودار های مبتنی بر مکان را نیز به این مجموعه اضافه کرد.

۶- نتیجه گیری

- سامانه تحلیل احساس که در دو حالت مبتنی بر شکلک و مبتنی بر واژه نامه عمل می کند، تحت عنوان EAS طراحی و پیاده سازی شد. گراف نمایش داده شده در این سامانه، مهم ترین کلمات به کار رفته در بین کل داده ها را همراه با نحوه و مقدار ارتباط بینشان به کاربر نمایش می دهد. کاربر با استفاده از این اطلاعات می تواند یک یا چند عبارت را در بازه زمانی دلخواه مورد بررسی قرار دهد.

- در این پژوهش سعی شد نتیجه انتخابات ریاست جمهوری سال ۱۳۹۶ جمهوری اسلامی ایران با بررسی کامنت های اینستاگرام نسبت به دو کاندید اصلی، با استفاده از سامانه طراحی شده پیش بینی گردد. همچنین میزان احساسات مختلف کاربران اینستاگرام نسبت به این دو کاندید نیز مورد بررسی قرار گرفت. نتیجه حاصل نشان دهنده احساس مثبت نسبت به کلمه "روحانی" و بالا بودن میزان احساس اعتماد در کنار احساس خشم و انزجار است.

مراجع

- [1] D. M. Boyd and N. B. Ellison, "Social network sites: Definition, history, and scholarship," *J. Comput. Commun.*, vol. 13, no. 1, pp. 210–230, Oct. 2007.
- [2] S. Asur and B. A. Huberman, "Predicting the future with social media," in *Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2010*, 2010, vol. 1, pp. 492–499.
- [3] R. Martinez-Pecino and M. Garcia-Gavilán, "Likes and problematic instagram use: The moderating role of self-esteem," *Cyberpsychology, Behav. Soc. Netw.*, vol. 22, no. 6, pp. 412–416, Jun. 2019.
- [4] B. Liu, *Sentiment Analysis: A Fascinating Problem*. 2012.
- [5] S. Hassan, J. Sängler, and G. Pernul, "SoDA: Dynamic visual analytics of big social data," in *2014 International Conference on Big Data and Smart Computing, BIGCOMP 2014*, 2014, pp. 183–188.