

آنالیز محتوای کمپین‌های انتخاباتی ۲۰۱۶ ریاست جمهوری ایالات متحده آمریکا در توییتر

پروین فخری *

مهدی حسین زاده **

چکیده

نتایج انتخابات ریاست جمهوری ۸ نوامبر ۲۰۱۶ در آمریکا بسیاری را شگفت‌زده نمود. اغلب کارشناسان و نظرسنجی‌های معمول و سنتی، هیلاری کلینتون را برنده نهایی انتخابات پیش‌بینی نمودند اما در واقعیت نتایج به گونه‌ای دیگر رقم خورد. در این پژوهش با استفاده از آنالیز شبکه‌های اجتماعی با رویکرد تحلیل محتوایی وضعیت کمپین‌های انتخاباتی دو نامزد اصلی انتخابات آمریکا بر اساس توییت‌های ارسالی طی بازه سه‌روزه قبل از ۸ نوامبر ۲۰۱۶ در شبکه اجتماعی توییتر بررسی شد. مشخص گردید در اغلب جنبه‌های تحلیلی از قبیل تحلیل توییت‌ها، ری توییت‌ها، تحلیل هشتک‌ها، تحلیل منشن‌ها و تحلیل پروفایل کاربران کمپین ترامپ دارای برتری محسوس و در میان کاربران با گرایش بی‌طرف نیز احتمالاً گرایش به کمپین ترامپ غالب است. بعلاوه در این پژوهش توانسته‌ایم برخی از استراتژی‌های انتخاباتی و جریان‌های سیاسی تأثیرگذار را شناسایی کنیم به عنوان مثال فهمیده‌ایم کمپین کلینتون به جای حمایت از کلینتون بیشتر به حمله به ترامپ پرداخته و بالعکس کمپین ترامپ به جای تاختن به کلینتون بیشتر به حمایت از ترامپ پرداخته است و یا کاربران بخصوص کاربران بدون گرایش و بی‌طرف، به سایت ویکی لیکس و ماجرای افشای ایمیل‌ها توجه خاص نشان داده‌اند.

کلیدواژه‌گان: شبکه‌های اجتماعی، توییتر، آنالیز محتوا، انتخابات ریاست جمهوری ایالات متحده آمریکا.

* دانشجوی دکتری، مهندسی سیستم‌های نرم‌افزاری، دانشگاه آزاد اسلامی، واحد علوم تحقیقات، تهران.

** دانشیار، عضو هیئت علمی، دانشگاه علوم پزشکی ایران، تهران. (نویسنده مسئول): Hosseinzadeh.m@iums.ac.ir

تاریخ پذیرش: ۱۳۹۶/۰۶/۲۵

تاریخ دریافت: ۱۳۹۵/۱۲/۲۳

مقدمه

نتایج انتخابات ریاست جمهوری ۸ نوامبر ۲۰۱۶ در آمریکا بسیاری را شگفت زده نمود. پس از انتخابات درون حزبی و تعیین نامزدهای نهایی در حزب دموکرات و جمهوری خواه به عنوان رقبای اصلی انتخابات ریاست جمهوری ۸ نوامبر ۲۰۱۶ آمریکا بسیاری تصور می کردند که هیلاری کلinton برای شکست ترامپ سختی چندانی در پیش رو ندارد. اغلب نظرسنجی های معمول و سنتی نیز هیلاری کلinton را برنده نهایی انتخابات پیش بینی نمودند. اما در واقعیت نتایج به گونه ای دیگر رقم خورد و بسیاری را شگفت زده نمود. سؤالی که مطرح است این است که آیا این شگفت زدگی به حق است و آیا ما نباید منتظر چنین نتایجی می بودیم؟

در این پژوهش ما از زاویه دیگری به قضیه نتایج انتخابات نگاه خواهیم کرد و آن زاویه یکی از بسترهای محبوب شبکه های اجتماعی در ایالات متحده یعنی توییتر است. ایده ما این است که با استفاده از یک تحلیل دوبعدی مکان-احساس برای محتوای کمپین های انتخاباتی دو نامزد اصلی انتخابات آمریکا بر روی توییتر بر اساس توییت های ارسالی طی سه روز قبل از ۸ نوامبر ۲۰۱۶ به صورت منطقی و مدلل بفهمیم افکار عمومی و خرد جمعی متمایل به چه کسی بوده است و پرسش اصلی این پژوهش این خواهد بود که با توجه به تحلیل های محتوایی داده های توییتر در نزدیک ترین بازه زمانی به انتخابات، آیا باید پیروزی ترامپ و کمپین انتخاباتی او را غیرمنتظره و شگفتی ساز دانست؟

ساختار این پژوهش در چهار بخش تدوین یافته است. در بخش دوم مباحث نظری مربوط به انتخابات در شبکه های اجتماعی، چشم انداز انتخابات ۲۰۱۶ ریاست جمهوری ایالات متحده آمریکا در شبکه اجتماعی توییتر، آنالیز محتوای شبکه های اجتماعی و پیشینه پژوهش پرداخته می شود. در بخش سوم به صورت خلاصه رویکرد تحلیل شبکه اجتماعی مورد استفاده ما به صورت گام به گام معرفی شده و در واقع گام های عملی آنالیز انجام گرفته معرفی و تشریح می شود. در بخش چهارم نتایج حاصل از آنالیز انجام گرفته طی مراحل قبل در قالب

شاخص‌ها و از زاویه و دیدگاه تحلیل توییت‌ها و ری توییت‌ها، تحلیل کاربران، تحلیل هشتک‌ها^۱ و ... امکان بحث و نتیجه‌گیری را فراهم می‌آورد.

مباحث نظری

شبکه‌های اجتماعی و انتخابات

به نظر می‌رسد پیش‌بینی‌های پیتر دراگر درباره این عصر خیلی زود محقق شده است. چراکه او گفته بود: «عصر آینده عصر شبکه‌ها خواهد بود نه شبکه‌های فناوری بلکه شبکه‌های انسانی در قالب فناوری‌های نوین. (جعفری و همکاران، ۱۳۹۵)

شبکه‌های اجتماعی جایگاه خود را در دنیای مدرن پیدا کرده و اغلب افراد هر روز با این شبکه‌های مجازی تعامل دارند. ایده‌ی اولیه شبکه‌های اجتماعی بسیار ساده است. یک شبکه‌ی اجتماعی، مجموعه‌ای از بازیگران (گره‌ها، نقطه‌ها، عامل‌ها و غیره) است که می‌توانند با یکدیگر رابطه داشته باشند.

بر اساس تئوری جامعه شبکه‌ای ولمن «فضای مجازی، امکان برقراری پیوندها و روابط ضعیف را به میزان نامحدود گسترش می‌دهد». او نتیجه می‌گیرد «افراد به صورت فزاینده تعاملاتشان را از فضای فیزیکی و عینی به فضای مجازی منتقل خواهند کرد. بر این اساس، ظهور واکنش‌های جمعی نسبت به یک رویداد در شبکه را می‌بایست برآیند حاصل از تعامل و برخورد نقش‌های حتی گاهی متضاد در چهارچوب روابط، در شبکه‌ی اعضا دانست (جیانگتاو و شن، ۲۰۱۴).

افراد به هنگام حضور در شبکه، کوله باری از دانش و تجربه‌های فرهنگی و اجتماعی خود را به همراه می‌آورند و درعین حال، تحت تأثیر اطلاعات و تعامل‌های درون آن قرار می‌گیرند. تعامل میان فردی آنلاین، به بعد مهمی از فعالیت‌های اجتماعی تبدیل شده است و تأثیرات آن

1. Hashtaq

2. Jiangtao & Shen

بر زندگی اجتماعی اشخاص، رو به فزونی است. این شبکه‌ها بر این واقعیت تمرکز دارد که کاربران، زندگی اجتماعی خود را فراتر از اینترنت نیز حفظ می‌کنند.

طبیعت پویا، متغیر و تعاملی شبکه‌های اجتماعی بر روی پلتفرم‌های رایگان باعث توانایی ایجاد دیدگاه‌های باورپذیر برای مخاطبان می‌گردد. از سوی دیگر یکی دیگر از مهم‌ترین آثار شبکه‌های مجازی، پدیدار شدن خرد جمعی است. خرد جمعی به جریان سیال و پویای افکار و اذهان کاربران در چنین محیط‌هایی اطلاق می‌شود که همانند یک پردازشگر عظیم، اطلاعات را پردازش می‌کند. همانند دنیای پردازنده‌های کامپیوتری که قدرت پردازشی کل از ترکیب چندین پردازشگر منفرد حاصل می‌شود، نیروی تفکر و ذهن کاربران با کمک تعاملات اجتماعی و بهره‌گیری از ابزارهای اینترنتی با یکدیگر ترکیب و همراه شده و نیروی عظیم با قدرت پردازشی بالا پدیدار می‌آورند.

شاید یکی از بهترین زمینه‌های نمود موارد ذکر شده مبحث انتخابات باشد. کاربران که همان رأی‌دهندگان آتی هستند به راحتی و بدون محدودیت خاص گرایش‌ها، تعاملات و فعالیت‌های انتخاباتی‌شان را از فضای فیزیکی به فضای مجازی منتقل می‌کنند. آن‌ها دانش، اطلاعات و تجربه خود را به سایرین ارائه خواهند نمود و در عین حال تحت تأثیر سایرین قرار خواهند گرفت. این دادوستدهای تعاملاتی منجر به ظهور خرد جمعی خواهد شد که البته در نهایت توسط اغلب افراد جامعه باور و در نهایت تبدیل به رفتار جمعی انتخاباتی خواهد شد.

انتخابات ۲۰۱۶ ریاست جمهوری ایالات متحده آمریکا در شبکه اجتماعی توییتر توییتر یکی از پر استفاده‌ترین و محبوب‌ترین شبکه‌های اجتماعی در ایالات متحده آمریکا محسوب می‌شود. توییتر در ۲۰۰۶ با شعار *What are you doing* با فلسفه به اشتراک گذاشتن وضعیت افراد به دوستان و اطرافیان پا به میدان گذاشت اما در ۲۰۰۹ این شعار به *What's happening* تغییر یافت و در خبررسانی به یک ابزار کاربردی مفید تبدیل شد.

تعداد زیاد کاربران، تنوع و طیف کاربران، به‌روز بودن مباحثات و توییت‌ها برحسب اتفاقات روز، محتوای مشارکتی و توزیعی، امکان پیام‌رسانی لحظه‌ای^۱، سرعت انتشار، آزادی ابراز عقاید و احساسات و امکان برجسته‌سازی با استفاده از هشتک‌ها، توییت‌ها را به یک ابزار ایده-آل برای انتخاب شونده‌گان جهت تبلیغات، جلب آرا و نظرها و برای انتخاب‌کنندگان جهت تصمیم‌سازی و برای رسانه‌ها و مراکز تحقیقاتی جهت پژوهش‌ها و نظرسنجی‌های مرتبط با عقاید و سنجش افکار عمومی نموده است.

طبق آمار سال ۲۰۱۵ توییت‌دارای بیش از ۲۸۲ میلیون کاربر فعال است. تعداد توییت‌های ارسالی توسط این کاربران در روز حدود ۵۰۰ میلیون توییت است. ۶۳٪ از برندهای دنیا حساب کاربری توییت‌گوناگون دارند و حدود ۹۲٪ کاربران فعال بیش از یک توییت در روز ارسال می‌کنند که نشان‌دهنده فعالیت خوب آن‌ها و موفقیت توییت‌ها در جذب بازاریابان و فعالان حوزه بازاریابی است. همچنین در سال ۲۰۱۵ توییت‌ها به‌عنوان دومین شبکه مورد اجتماعی‌پرکاربرد برای حوزه B2B شناخته شده است. در نهایت اثر نفوذ تبلیغات در توییت به سطحی رسیده است که طبق آمار اعلام شده، تبلیغات در این شبکه ۸ تا ۲۴ برابر فیس‌بوک، کلیک می‌شوند و این نکته به موفقیت توییت‌ها در زمینه تبلیغات آنلاین اشاره مستقیم دارد.

در انتخابات ریاست جمهوری ۲۰۱۶ آمریکا، مناظرات انتخاباتی ریاست جمهوری آمریکا با شکستن رکورد ۱۷ میلیون توییت در شبکه اجتماعی توییت تاریخ‌ساز شد. اگرچه هر دو نامزد انتخابات در طی این مناظرات هزاران فالوور کسب کردند و فالوورهای آن‌ها به ۱۳ و ۱۰ میلیون نفر در توییت رسید اما روشن است که کمپین انتخاباتی ترامپ در توییت بسیار قوی‌تر از کلینتون بوده است و وی از توییت‌ها به‌عنوان یکی از تریبون‌های اصلی خود استفاده نموده است. ترامپ بسیار بیشتر از کلینتون از توییت‌ها برای جلب نظر مردم و مطرح کردن نظراتش استفاده کرده است این برتری حتی در برخی موارد دو برابر رقیبش بوده است، وی در زمینه

جلب مشارکت و واکنش در شبکه‌های اجتماعی به خصوص توییتر بسیار فعال‌تر بوده است و بی‌شک همین استفاده مؤثر به جلب آرای ترامپ کمک بسیار نموده است.

آنالیز محتوای شبکه‌های اجتماعی

اگرچه بیشتر شبکه‌های اجتماعی امروزی‌اند و سنی کمتر از پانزده سال دارند اما ظهور و محبوبیت این شبکه‌های اجتماعی اینترنتی به‌طور ناگهانی موجب دسترسی به مقدار زیادی از داده‌های اجتماعی، تعاملات و رفتار انسان‌ها شد. این داده‌های عظیم یا بزرگ داده^۱، شامل عناصر مختلف است و می‌تواند به‌عنوان محتوای یک سیستم رابطه‌ای اجتماعی که توسط مجموعه‌ای از بازیگران و روابط اجتماعی بین آن‌ها تولید و مشخص شده است شناخته شود. اطلاعات عظیم شبکه‌های اجتماعی علاوه بر چالش‌های سنتی نظری 3V^۲، دیدگاه‌ها و چالش‌های جدیدی را به مطالعه و آنالیز شبکه‌های اجتماعی و نیز آنالیز محتوای شبکه‌های اجتماعی باز می‌کند. آنالیز محتوای شبکه‌های اجتماعی از علم ریاضی اجتماعی سرچشمه می‌گیرد و یکی از شاخه‌های آنالیز شبکه‌های اجتماعی محسوب می‌گردد و مجموعه‌ای از مفاهیم، اندازه‌ها و تکنیک‌ها برای آنالیز داده‌ها و رابطه‌هایی است که به‌طور تخصصی و بی‌نظیر برای درک مهم‌ترین ویژگی‌های ساختارها و وقایع اجتماعی استفاده می‌شود. در تجزیه و تحلیل داده‌های اجتماعی محققان شروع به استفاده و تجزیه و تحلیل آنچه شبکه‌های اجتماعی تولید کند، می‌کنند و خصوصاً تمرکز تجزیه و تحلیل داده‌های شبکه‌های اجتماعی در اصل بر محتوایی است که توسط کاربران و تعاملات و روابطشان تولید شده است. داده‌های تولیدشده در شبکه‌های اجتماعی ارزشمند، متنوع و فراوان‌اند که باعث می‌شود آن‌ها منبع مناسبی برای علم داده باشند. بسیاری از تحقیقات محاسباتی که داده‌های اجتماعی را به کار می‌گیرد از یادگیری ماشینی و پردازش زبان طبیعی (NLP^۳)، طبقه‌بندی و پیش‌بینی

-
1. Big data
 2. Natural Language Processing

استفاده می‌کند. آنالیز محتوای شبکه‌های اجتماعی در بسیاری از زمینه‌ها از جمله انتخابات بسیار مورد توجه قرار گرفته است.

پیشینه پژوهش

تحقیقات اولیه در آنالیز شبکه اجتماعی در زمینه‌های روانشناسی آموزشی و مطالعات رشد کودکان بوده است. همچنین آنالیز شبکه در زمینه‌هایی مانند جامعه‌شناسی و مردم‌شناسی توسعه بسیار یافته است. اخیراً آنالیز شبکه اجتماعی یک ابزار بسیار مهم برای درک اتصالات بین الگوها، تعاملات و خروجی‌های کسب‌وکار مانند بازدهی شده است (راپم^۱، ۲۰۱۳).

برای دانشمندان اجتماعی تئوری شبکه یک برگ برنده و یک معدن طلا برای توضیح رفتارهای پدیده‌های اجتماعی در یک طیف وسیع از رشته‌های روانشناسی تا اقتصاد محسوب می‌شود (بروگاتی و همکاران^۲، ۲۰۰۹).

احتمالاً بیشترین کار تأثیرگذار در تجزیه و تحلیل شبکه اجتماعی مقاله ابتدایی " تماس‌ها و اثرات " نوشته شده در دهه ۱۹۵۰ است.

اگرچه شبکه‌های اجتماعی در زمینه‌ی جامعه‌شناسی بیش از یک قرن است که استفاده می‌شود اما جمع‌آوری داده‌ها به دلیل کمبود مکانیسم‌های فناورانه‌ی مناسب، نسبتاً سخت است (آگاروال^۳، ۲۰۱۵).

در کل می‌توان مسائل بهره‌مند از آنالیز شبکه‌های اجتماعی را به صورت زیر خلاصه نمود:

- تشخیص باندهای نفوذ، مدل‌های مولد تشکیلات شبکه، تشخیص جوامع شبکه‌ای، نمایش شبکه‌ها، اندازه‌های توصیفی شبکه‌ها مانند خوشه‌بندی مرکزیت انجمن، بررسی مدل‌های نفوذ در شبکه، آنالیزهای رابطه‌ای (اومالی و همکاران^۴، ۲۰۱۴).

1. Rupam
2. Borgatti et al
3. Aggarwal
4. O'Malley et al.

- تصویرسازی، نمایش داده‌های شبکه‌ای، آنالیز احساسات، تشخیص و اکتشاف تمایلات، کشف اتفاقات واقعی برای بررسی وقایعی که در دنیای واقعی اتفاق می‌افتد از قبیل وقایع عمومی یا سیاسی، سیستم‌های توصیه اجتماعی، آنالیز تعاملات اجتماعی، پیشگویی آبخاری، انتشار شایعات، مطالعه تأثیر شبکه بر روی رفتار، خودسازمان‌دهی در موقعیت‌هایی که مرکز دستور دهی وجود ندارد و ارتباطات فاقد مرکزیت است، قوم‌نگاری و بررسی هدفمند از افراد و ترتیب‌های اجتماعی (کورکا و همکاران^۱، ۲۰۱۵).
 - آنالیز ساختاری (گارگ و همکاران^۲، ۲۰۰۸).
 - آنالیز توپولوژی شبکه‌ی اجتماعی، سنجش عملکرد، آنالیز داده‌های اجتماعی، انتشار اطلاعات و رفتارهای آبخاری در شبکه، پیشگویی ارتباط و تکامل شبکه اجتماعی، مسئله بازیگران کلیدی (لی هوی جیانگتائو و بینگ کین شن، ۲۰۱۴). آنالیز محتوای شبکه‌های اجتماعی در قالب شاخه‌ای قدرتمند و با استفاده از یک دید جایگزین که ویژگی‌های مختلف شبکه را به صورت سامانمند و بر مبنای علم آمار و ریاضیات مورد مطالعه قرار می‌دهد از طیف گسترده‌ای از علوم تحقیقات محاسباتی که داده‌های اجتماعی را به کار می‌گیرد از قبیل مدل‌سازی، آنالیز احساسات و عقیده کاوی، یادگیری ماشین، پردازش زبان طبیعی (NLP)، آزمون‌های آماری و... استفاده می‌کند. تمرکز تجزیه و تحلیل داده‌های اجتماعی در اصل بر محتوایی است که توسط کاربران در بسترهای شبکه‌های اجتماعی تولید شده است
- در زمینه آنالیز انتخابات در شبکه‌های اجتماعی پژوهش‌هایی با رویکردهای مختلف انجام شده است. اغلب این پژوهش‌ها در راستای پیش‌بینی نتایج و در بستر فیس‌بوک و توییتر صورت گرفته است به عنوان مثال:
- کارلیسه و همکاران^۳ مفهوم مشارکت سیاسی و فعالیت‌های سیاسی کاربران فیس‌بوک در

1. Kurka et al.

2. Garg et al.

3. Carlisle et al.

طول سال ۲۰۰۸ بررسی نموده‌اند و با استفاده از یک مدل منابع، مشارکت سیاسی در فیس‌بوک را توضیح می‌دهند و نیز بررسی می‌کند که آیا اندازه شبکه برای افزایش فعالیت‌های سیاسی مناسب است؟ (کارلیسه و همکاران، ۲۰۱۳).

آنستید و همکاران تجزیه و تحلیلی از پوشش رسانه‌ای از انتخابات ۲۰۱۰ UK نشان می‌دهند و نیز نشان می‌دهند چگونه مطالعه رسانه‌های اجتماعی برای تجزیه و تحلیل پویایی اجتماعی از طریق نظرات کارساز می‌شود (آنستید و همکاران^۱، ۲۰۱۵).

لارسون و همکاران با جمع‌آوری داده‌های توییتری و بهره‌گیری از تجزیه و تحلیل محتوای تولیدی کاربران توییتر در طول انتخابات سال ۲۰۱۰ سوئد فرصت‌های جالب مبارزات آنلاین و شور و مشورت انتخاباتی را بررسی می‌کنند (لارسون و همکاران^۲، ۲۰۱۲).

توماسجان^۳ با بررسی انتخابات فدرال آلمان در سال ۲۰۰۹ تحلیل می‌کنند که آیا پیام میکرو بلاگینگ در توییتر آینه معتبر چشم‌انداز سیاسی است و می‌تواند برای پیش‌بینی نتایج انتخابات استفاده شود یا خیر. (تامس جانگ و همکاران^۴، ۲۰۱۱).

وانگ هاو و همکاران، در سرویس میکرو بلاگینگ توییتر یک سیستم برای زمان واقعی تجزیه و تحلیل احساسات عمومی نامزدهای انتخابات ریاست جمهوری در سال ۲۰۱۲ ایالات متحده معرفی می‌کنند که در آن مردم به بیان نظرات و دیدگاه‌های احزاب سیاسی و نامزدهای خود و حوادث در حال ظهور و یا اخبار انتخاباتی می‌پردازند. علاوه بر این، تجزیه و تحلیل احساسات و کمک به کشف چگونگی این وقایع را مطرح می‌کند (وانگ هاو و همکاران^۵، ۲۰۱۲).

سانگ و همکاران به تحلیل احساسات و بر مبنای آن پیش‌بینی انتخابات سال ۲۰۱۱ مجلس سنا هلند پرداخته‌اند (سانگ و همکاران^۶، ۲۰۱۲).

1. Anstead et al.
2. Larsson et al
3. Tumasjan
- 4 Tumasjan et al.
5. Wang & Hao
6. Sang et al.

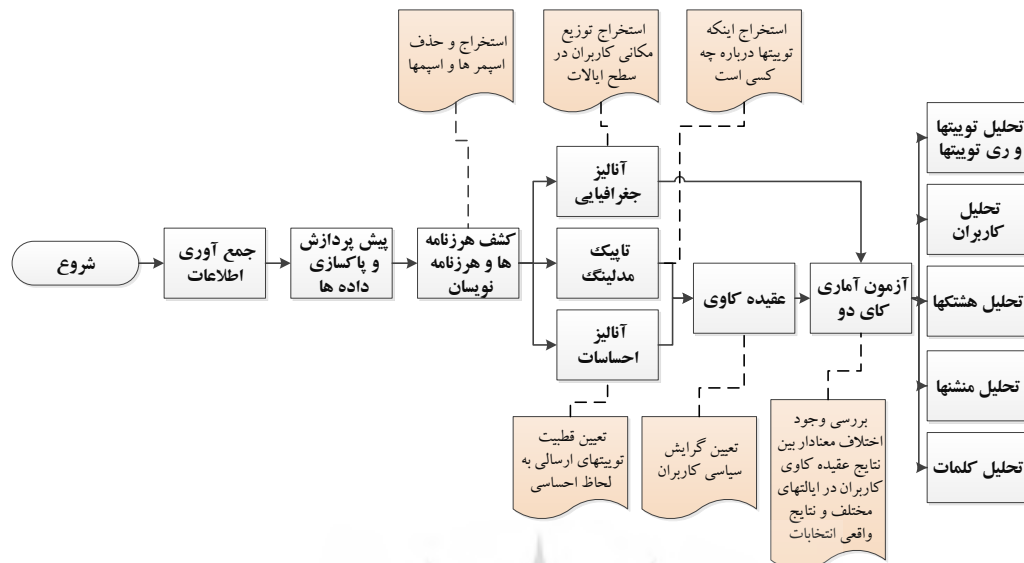
وولی و همکاران با استفاده از تجزیه و تحلیل محتوای کمی در انتخابات ریاست جمهوری سال ۲۰۰۸ به رقابت جان مک کین و باراک اوباما در فیس بوک پرداخته‌اند (وولی و همکاران^۱، ۲۰۱۰).

فرناندز و همکاران با استفاده از تجزیه و تحلیل محتوای گروه‌های دانشجویی به منظور تسهیل مشارکت سیاسی گفت‌وگوی مدنی در انتخابات ریاست جمهوری سال ۲۰۰۸ و رقابت جان مک کین و باراک اوباما در فیس بوک، بحث‌های سیاسی مربوط به روند مبارزات مدنی سیاسی، مسائل مربوط به سیاست، اطلاعات مبارزات انتخاباتی، مسائل نامزدها، رفتار رأی‌گیری و کسب محصولات مبارزاتی انتخاباتی بررسی کرده‌اند. (فرناندز و همکاران^۲، ۲۰۱۰).

رویکرد تحلیل شبکه اجتماعی استفاده شده

فرآیند آنالیز شبکه‌های اجتماعی با رویکرد آنالیز محتوایی که ما به صورت گام به گام انجام دادیم به صورت (شکل ۱) خلاصه می‌شود. لازم به ذکر است که آنالیزهای انجام شده بر روی داده‌ها هدفمند و بنا بر ماهیت موضوع و نیازمندی‌های آن است. به عنوان مثال ما برای اطلاع از توزیع طرفداران نامزدهای انتخابات، با توجه به سیستم رأی‌گیری الکترونال فدرالی در ایالات متحده نیازمند اطلاع از توزیع مکانی کاربران در سطح ایالات آمریکا هستیم بنابراین ناگزیر به استفاده از آنالیز جغرافیایی هستیم درحالی که ممکن است در سایر موضوعات نیازی به استفاده از این آنالیز نباشد. در خصوص تقدم و تأخر اجرای هر یک از گام‌های آنالیز نیز باید ذکر کرد برخی آنالیزها باید مقدم بر دیگر آنالیزها باشد یا ترتیب خاصی برای گام‌های آنالیزی مورد نیاز باشد و گاهی تقدم یا تأخر مطرح نیست به عنوان مثال آنالیز کاربران اسپم باید به دلیل نیاز به خروج آن‌ها از پژوهش مقدم بر سایر آنالیزها باشد درحالی که آنالیز جغرافیا و آنالیز احساسات در تقدم و تأخر ملزم به رعایت ترتیب خاصی نیست. (شکل ۱) مراحل انجام آنالیزها را با رعایت ترتیب گام‌ها نمایش می‌دهد:

1. Woolley et al.
2. Fernandes et al.



شکل ۱. فرآیند آنالیز محتوایی شبکه اجتماعی مورد استفاده در این پژوهش

۱- جمع آوری اطلاعات:

جمع آوری این داده‌ها از تاریخ ۵ لغایت ۷ نوامبر ۲۰۱۶ و با استفاده از API سایت توییتز بر اساس کلمات جستجوی #hillaryClinton، #hillary، #donaldTrump و #Trump انجام گرفته است لازم به ذکر است به دلیل حجم عظیم توییت‌ها تنها ۱۰٪ آن‌ها به صورت تصادفی انتخاب شده است که شامل حدود ۴۰۷۳۳۹ توییت است.

۲- پیش پردازش و پاکسازی داده‌ها:

در این گام اقدامات زیر صورت گرفت:

- استخراج توکن‌ها: تفکیک جمله به واژه، عبارت، نشانه و ...
- حذف اطلاعات زائد
- استانداردسازی زبان و حروف: حذف کاراکترهای غیر اسکی، یکسان‌سازی حروف به لحاظ بزرگی و کوچکی. حذف حروف تکرار شونده در توییت‌ها.

1. Tokenization

- حذف کلمات ایست^۱ و کاراکترهای خاص.

۳-مدل سازی و آنالیز

اقدامات مدل سازی و آنالیزی که بر روی توییت ها انجام گرفت شامل موارد زیر است:

- **کشف هرزنامه ها و هرزنامه نویسان:** برای شناسایی کاربران اسپم بر اساس

پژوهش های قبلی صورت گرفته از معیارهای زیر استفاده شده است.

۱. سن حساب کاربری
۲. نرخ ارسال پیام
۳. درصد تشابه توییت های ارسالی هر کاربر
۴. FF-ratio نسبت تعداد دنبال کنندگان^۲ به دوستان^۳
۵. نرخ توییت استاتوس^۴ S-ratio
۶. درصد توییت های حاوی URL

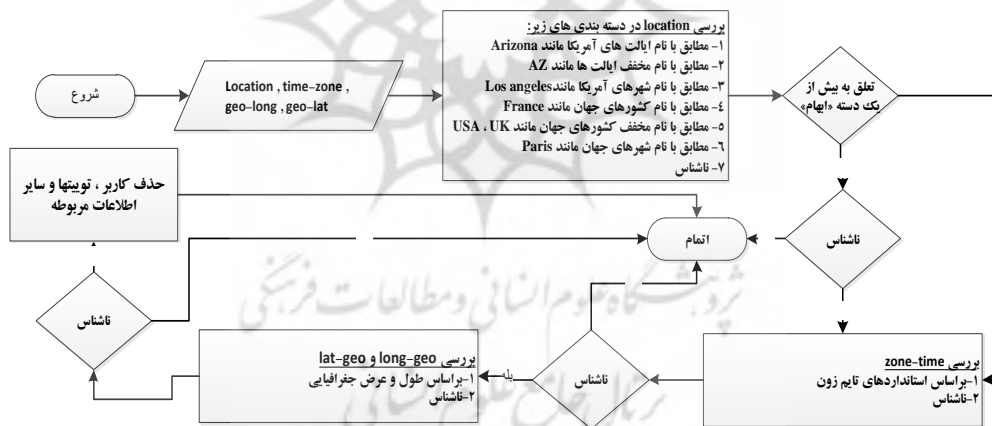
اغلب حساب های کاربری اسپم سن بالایی ندارند درعین حال نرخ ارسال پیام بسیار بالایی دارند توییت های ارسالی آنان نیز در بسیاری از موارد یکسان و یا از تشابه بالایی برخوردار است و پیام های آنها تقریباً همگی شامل یک URL است.

این حساب ها واقعی نیستند لذا تعداد دنبال کننده پایینی دارند ولی جهت ارسال پیام تعداد افراد زیادی را دنبال می نمایند. سن حساب کاربری و FF-ratio را به صورت مستقیم از اطلاعات پروفایل کاربران به دست آوردیم. برای محاسبه نرخ ارسال پیام از مجموع تعداد توییت های ارسال شده توانستیم نرخ ارسال پیام را بر حسب توییت بر روز محاسبه نماییم. برای محاسبه نرخ توییت های استاتوس. از نسبت تعداد استاتوس های ارسال شده به سن حساب کاربری نرخ ارسال استاتوس را بر حسب استاتوس بر سال محاسبه نمودیم. درصد

-
1. Stop words
 2. Follower
 3. Friends
 4. Status

توییت‌های حاوی URL نیز از نسبت تعداد توییت‌های حاوی URL بر تعداد کل توییت‌های ارسال شده یک کاربر به دست آمده است. برای هر یک از معیارها یک حد آستانه تعریف شده استفاده شده است به‌عنوان مثال حساب کاربری که بیش از ۷۰٪ پیام‌هایش مشابه باشند یا بیش از ۷۰٪ پیام‌هایش حاوی URL است مستعد اسپم بودن است. درعین حال ما تنها بر اساس یک معیار تصمیم‌گیری نکرده‌ایم زیرا هیچ کدام از معیارهای بالا به تنهایی نمی‌توانند نمایانگر کاربران اسپم باشد. هر کاربری که بیش از ۳ معیار اسپم برایش مثبت باشد به‌عنوان کاربر اسپم شناسایی می‌شود. از آنجاکه وجود چنین کاربران و یا توییت‌های آن‌ها ممکن است منجر به کج روی تحقیق شوند، حذف شدند.

• **آنالیز جغرافیایی:** با استفاده از فیلد location و فیلد time-zone و فیلدهای geo-long و geo-lat توزیع مکانی طرفداران کاندیدها مشخص می‌شود. برای پژوهش ما تفکیک کاربران به کاربران داخل و خارج آمریکا حیاتی است. بدین منظور ما الگوریتمی مطابق با (نمودار ۱) طراحی نموده و کد آن را پیاده‌سازی نمودیم.

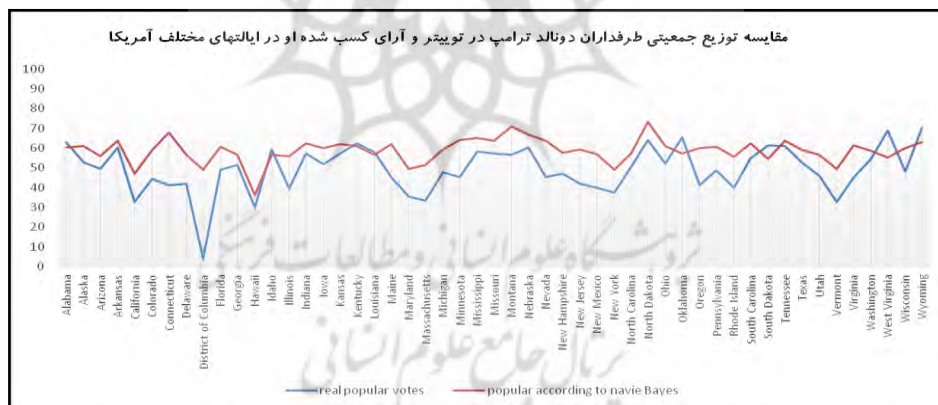


نمودار ۱. الگوریتم آنالیز جغرافیایی

• **تاپیک مدلینگ:** دسته‌بندی توییت‌ها با استفاده از تاپیک مدلینگ به دو صورت در این پژوهش مورد استفاده قرار گرفته است. ۱- برای تعیین موضوع مورد بحث در توییت ۲- برای تعیین اینکه توییت درباره چه کسی است. برای این منظور از نرم‌افزار Mallet^۱ استفاده شده که یک جعبه‌ابزار نرم‌افزاری متن‌باز است و از روش LDA برای تاپیک بندی استفاده می‌کند و خروجی این نرم‌افزار می‌تواند در پیدا کردن موضوع مورد بحث در توییت کمک‌کننده باشد.

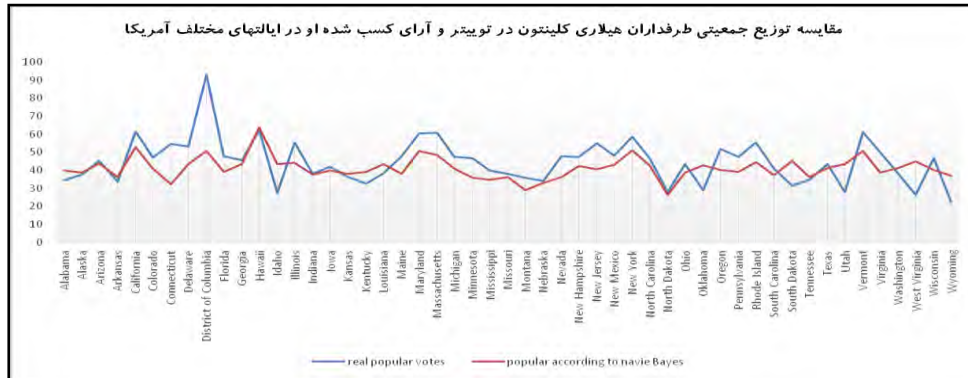
• **آنالیز احساسات:** الگوریتم مورد استفاده جهت آنالیز احساسات Naïve Bayes از نوع الگوریتم‌های یادگیری ماشین تحت نظارت و در سطح جمله است. با استفاده از سه فایل آموزشی^۲ positive training، negative training، بی‌طرف training سایر توییت‌های آزمایشی در قالب توییت‌های مثبت، منفی و بی‌طرف دسته‌بندی شده‌اند.

• **عقیده کاوی:** نوع گرایش و طرفداری کاربران بر اساس آنالیز برآیند نوع احساسات منعکس شده در توییت‌ها که در مرحله قبل حاصل شده، استخراج می‌گردد. بر این اساس توزیع جمعیتی طرفداران دونالد ترامپ در توییت و آرای کسب شده او در ایالت‌های مختلف آمریکا.



نمودار ۲. مقایسه توزیع جمعیتی طرفداران دونالد ترامپ در توییت و آرای کسب شده او در ایالت‌های مختلف آمریکا

1. <http://mallet.cs.umass.edu>
2. Training



نمودار ۳. مقایسه توزیع جمعیتی طرفداران هیلاری کلینتون در توییتر و آرای کسب شده او در ایالت‌های مختلف آمریکا

در مرحله‌های قبل در آنالیز احساسات به روش Naïve Bayes توانستیم قطبیت توییت‌ها را استخراج و با استفاده از تاپیک مدلینگ متوجه شدیم که هر توییت درباره کیست. در این حالت توییت‌های ارسالی به صورت زیر دسته‌بندی می‌گردد:

۱- طرفداری ترامپ ۲- ضد ترامپ ۳- طرفدار کلینتون ۴- ضد کلینتون ۵- خنثی
هر کاربر با توجه به برآیند ماهیت همه توییت‌های ارسالی در یکی از گرایش‌های زیر قرار می‌گیرد:

۱- طرفدار ترامپ ۲- طرفدار کلینتون ۳- بی طرف

در مرحله آنالیز جغرافیایی مکان کاربران نیز در سطح ایالت‌ها استخراج گردیده لذا می‌توان توزیع آمار جمعیتی کاربران در ایالت‌ها را به تفکیک گرایش‌های سیاسی استخراج نمود. بر این اساس توانسته‌ایم در (نمودار ۲) مقایسه‌ای از توزیع جمعیتی طرفداران دونالد ترامپ در توییتر و آرای واقعی کسب شده او در ایالت‌های مختلف آمریکا و در (نمودار ۳) مقایسه‌ای از توزیع جمعیتی طرفداران هیلاری کلینتون در توییتر و آرای واقعی کسب شده او در ایالت‌های مختلف آمریکا ارائه دهیم

• **آزمون آماری کای دو:** آزمون آماری کای دو یک آزمون نا پارامتریک است که برای بررسی معناداری تفاوت بین فراوانی‌های مشاهده شده و مورد انتظار مورد استفاده است و اصطلاحاً برای بررسی هم قواری متغیرها به کار می‌رود. برای بررسی صحت نتایج و بررسی وجود اختلاف معنادار بین نتایج عقیده کاوی کاربران در ایالت‌های مختلف و نتایج واقعی انتخابات در این ایالت‌ها در آزمون کای دو با درجه اطمینان ۰/۲۵/۰ روشن شد که تفاوت معناداری بین نتایج محاسبه شده و نتایج واقعی وجود ندارد و اصطلاحاً نتایج به دست آمده با نتایج واقعی هم قواره هستند جدول ۱ نتایج آزمون کای دو با درجه اطمینان ۰/۲۵/۰ را نشان می‌دهد.

جدول ۱. نتایج آزمون کای دو با درجه اطمینان ۰/۲۵

Trump-chiSq	Clinton-chiSq
۱۰۷E-۳/۴۸۴۹	۰۸E-۲/۲۱۹۲۶

- **تحلیل توییت‌ها و ری توییت‌ها:** توییت‌ها و ری توییت‌ها اصلی‌ترین دلیل حضور و تعامل در توییت‌هاست در این راستا در بخش ۴,۱ به تفصیل به تحلیل توییت‌ها و ری توییت‌ها پرداخته خواهد شد.
- **تحلیل کاربران:** به لحاظ اهمیت کاربران که مهم‌ترین موجودیت‌های شبکه‌های اجتماعی محسوب می‌شوند در بخش ۴,۲ به تفصیل به تحلیل کاربران پرداخته خواهد شد.
- **تحلیل هشتک‌ها:** هشتک‌ها به عنوان مراجع جریان ساز و ایجادکننده موج بسیار تعیین کننده هستند لذا در بخش ۴,۳ به تفصیل به تحلیل هشتک‌ها پرداخته خواهد شد.
- **تحلیل منشها:** منشها نمایانگر تعاملات و یکی از شاخص‌های بررسی مباحثات هستند که در بخش ۴,۴ به تفصیل به تحلیل منشها پرداخته خواهد شد.
- **تحلیل کلمات:** کلمات ساختار گفتمان‌ها را مشخص می‌کند و معیار بررسی گفتمان‌هاست. در بخش ۴,۵ به تفصیل به تحلیل کلمات پرداخته خواهد شد.

۴-بصری سازی و نمایش نتایج

نتایج با استفاده از نرم‌افزارهای آماری در قالب ۱۷ جدول و ۹ نمودار مقایسه‌ای در بخش‌های آتی نمایش داده شده است

بحث و آنالیز

تحلیل توییت‌ها و ری توییت‌ها

اصلی‌ترین ارزش در توییت داشتن تعامل و ارتباط گرفتن است اولین گام هم در این راستا ارسال توییت است. اگر توییتی مورد توجه کاربران قرار گیرد می‌تواند بازنشر یا اصطلاحاً ری توییت شود. ری توییت‌ها می‌توانند به سادگی زمینه ایجاد موج‌های اجتماعی، شایعات و یا ویروسی شدن را فراهم نمایند. توییت‌ها با توجه به محدود بودن اندازه می‌توانند در میان رسانه‌های اجتماعی بهترین و مستعدترین کاندیدا برای عقیده کاوی یا متن کاوی باشند. با توجه به اینکه در حین بازنشر یک توییت ممکن است تغییرات کوچکی در انتهای توییت صورت پذیرد مثلاً هنگام افزودن @Rt و نام نویسنده مرجع تعداد کاراکترها از ۱۴۰ بیشتر شده و به‌طور خودکار حروف آخر توییت که اضافه است حذف گردد لذا نمی‌توان برای پیدا کردن توییت‌های اصلی از شباهت کامل توییت و بازنشر آن استفاده شود. برای مقایسه توییت و بازنشر از یک روش ساده شباهت سنجی در مجموعه‌ها به نام Dice's t coefficient استفاده شده است که از شمارش تعداد عناصر مشابه مجموعه‌ها استفاده می‌کند. میزان شباهت دو مجموعه بر اساس این روش به‌صورت رابطه ۱ محاسبه می‌گردد (فریکس، ویلیام و یاتس، ۱۹۹۲)

$$S(x, y) = \frac{2 * |x * y|}{|x| + |y|} \quad \text{رابطه (1)}$$

$S(x, y)$ که مقداری در بازه صفر و یک و نشان‌دهنده میزان شباهت است. انتخاب حد آستانه ۰.۷۵٪ می‌تواند منجر به یافتن این شود که هر توییت بازنشر شده کدام توییت است و حتی گراف بازنشر قابل استخراج خواهد بود.

(جدول ۲) نشان می‌دهد درصد توییت‌های ضد ترامپ طرفداران کلیتون ۲/۱ برابر بیشتر از درصد توییت‌های طرفداری کلیتون است و می‌بینیم که کمپین کلیتون بیشتر به حمله به ترامپ پرداخته است تا حمایت از کلیتون. این موضوع برای کمپین ترامپ بالعکس است و درصد توییت‌های طرفداری ترامپ ۱/۲ برابر توییت‌های ضد کلیتون بوده است که به این معناست که کمپین ترامپ بیشتر به حمایت از ترامپ پرداخته است تا حمله به کلیتون. طرفداران کلیتون با فاصله زمانی کمتری توییت ارسال نموده‌اند. (جدول ۳) این موارد را تأیید می‌کند.

جدول ۲. شاخص‌های درصد توییت‌های ضدیت، طرفداری و بی‌طرف هر کاندید و نرخ ارسال

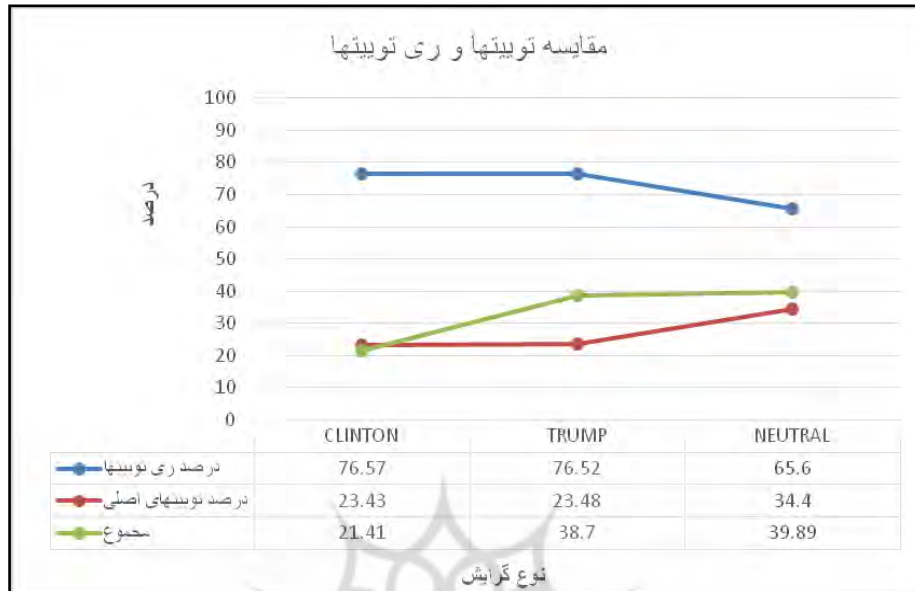
توییت‌ها برای هر گروه طرفداری

نرخ ارسال توییت‌ها	درصد توییت‌های بی‌طرف	درصد توییت‌های طرفداری کلیتون	درصد توییت‌های ضد کلیتون	درصد توییت‌های طرفدار ترامپ	درصد توییت‌های ضد ترامپ	نوع طرفداری
۰/۶۴	۳۴/۴۹	۱۹/۸۷	۱/۶۷	۲/۰۲	۴۱/۹۵	کلیتون
۰/۸۶	۴۲/۴۳	۲/۴۸	۲۳/۸۲	۲۹/۰۱	۲/۲۶	ترامپ
۰/۴۹	۹۰/۷۰	۲/۱۸	۲/۳۴	۲/۲۳	۲/۷۴	بی‌طرف

جدول ۳. شاخص‌های میانگین تعداد توییت‌های ضدیت، طرفداری و بی‌طرف هر کاندید برای هر

گروه طرفداری

میانگین تعداد توییت‌های بی‌طرف	میانگین تعداد توییت‌های طرفداری کلیتون	میانگین تعداد توییت‌های ضد کلیتون	میانگین تعداد توییت‌های طرفدار ترامپ	میانگین تعداد توییت‌های ضد ترامپ	نوع طرفداری
۰/۶۶	۰/۳۸	۰/۰۳	۰/۰۴	۰/۸۰	کلیتون
۱/۱۰	۰/۰۶	۰/۶۲	۰/۷۵	۰/۰۶	ترامپ
۱/۳۴	۰/۰۳	۰/۰۳	۰/۰۳	۰/۰۴	بی‌طرف



نمودار ۴. درصد توییت‌های اصلی و ری توییت برای هر گروه طرفداری

(نمودار ۴) درصد توییت‌های اصلی و ری توییت را مشخص نموده است که برای هر دو کمپین تقریباً برابر است و (جداول ۴ و ۵) به ترتیب پنج ری توییت برتر طرفداران کلینتون و پنج ری توییت برتر طرفداران ترامپ را نشان می‌دهد.

جدول ۴. پنج ری توییت برتر طرفداران کلینتون

ری توییت برتر طرفداران کلینتون	درصد
RT @nowthisnews: Donald Trump just straight up lied about how President Obama responded to a pro-Trump protester https://t.co/nd1Dhk3bwE	۶/۶۴
RT @BernieSanders: Donald Trump suffers from a very serious disease called amnesia. He seems to have forgotten 8 years ago we were losâ€¦	۴/۰۳
RT @Channel4News: "Don't boo, vote!" This is how Obama dealt with a pro-Trump protester who tried to disrupt his speech in support ofâ€¦	۲/۷۹
RT @carlajopetty: #Trump #america #hillaryClinton #iFunny https://t.co/D34x8XZ8F1 https://t.co/6hsSuOD5cP	۱/۷۳
RT @ddale8: Here is a complete list of Donald Trump's false claims over the last six weeks, sorted by category: https://t.co/h4BdWdZc6o	۰/۶۶

جدول ۵. پنج ری توییت برتر طرفداران ترامپ

ری توییت برتر طرفداران ترامپ	درصد
RT @TheSquidFace: https://t.co/1qp0MO7blG this video shows what DONALD TRUMP REPRESENTS FOR US as a leader.	۲/۳۴
RT @SandiBurkhart: It's ignorant to vote for Hillary Clinton without reading WikiLeaks https://t.co/K5rA3iMmmj via @denverpost	۱/۲۱
RT @StylishRentals: Check out "Donald Trump Your President" by Rick Poppe #Trump2016 #TrumpTrain - https://t.co/H5bowcbE5t https://t.co/bg8â€¦	۱/۰۷
RT @DrJillStein: If Saudi Arabia funded 9/11 and ISIS too, why does Obama protect them, Clinton arm them, & Podesta lobby for them? https://â€¦	۰/۶۳
RT @HillaryClinton: "Nobody respects women more than me." â€”Donald Trump\	۰/۵۹

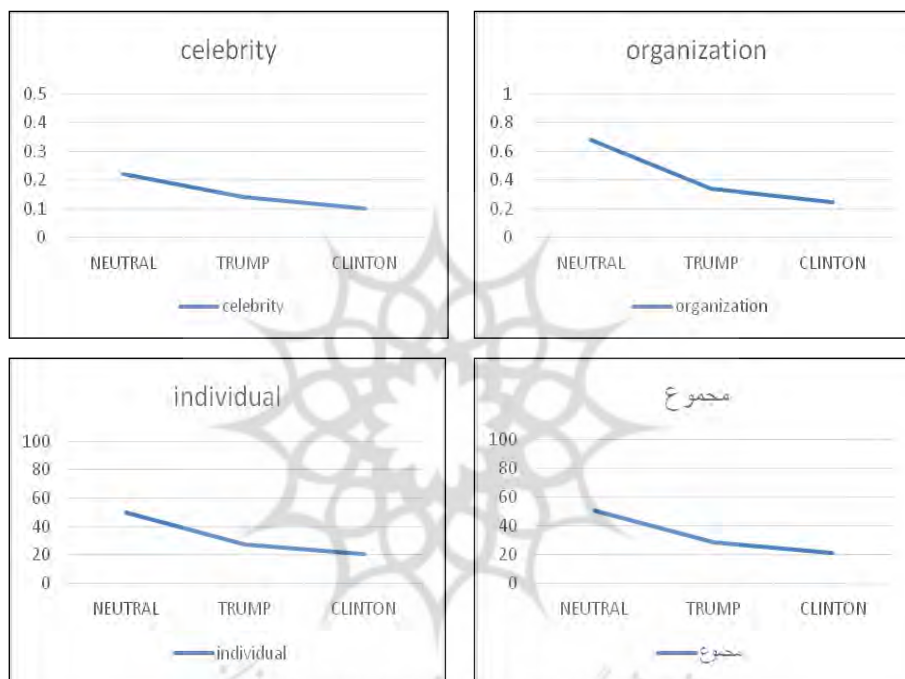
در پنج ری توییت برتر طرفداران ترامپ توییت‌های مورد توجه توییت‌هایی بوده‌اند که در راستای معرفی دستاورد او به‌عنوان یک رهبر برای آمریکا، تشویق افراد به مطالعه و یکی لیکس، اتهام پشتیبانی او با ما و کلینتون از داعش و تجهیز آن‌ها و پاسخ به اتهام عدم احترام به زنان پرداخته شده است. موضوعات این ری توییت‌ها بسیار هدفمند و متنوع است. ری توییت‌های برتر طرفداران کلینتون به اتهام زنی علیه ترامپ در خصوص رنج بردن از بیماری روانی و نحوه برخورد او با ما با یکی از طرفداران ترامپ پرداخته است که فاقد آن هدفمندی و تنوع و توجه به جریان‌های مهم انتخاباتی است.

در کمپین کلینتون به حواشی او با ما به‌عنوان یکی از چهره‌های شاخص طرفدار کلینتون پرداخته شده است اما کمپین ترامپ سعی در جا انداختن ترامپ در قواره یک رهبر و پاسخ به اتهامات درباره او دارد.

تحلیل کاربران

کاربران پرستیژ، قدرت و اهمیت یکسان ندارند (ناتلیانیس و همکاران، ۲۰۱۵). دامنه نفوذ و تأثیرگذاری کاربران بسته به نوع آن‌ها متفاوت است یک کاربر مشهور که در کانون توجه

است بسیار ساده‌تر می‌تواند کاربران و حتی کل شبکه را تحت تأثیر قرار دهد چنانکه این اتفاق تاکنون بارها اتفاق افتاده است. در انتخابات این تأثیرگذاری بسیار بیشتر به چشم می‌آید و حتی ممکن است یک کاربر مشهور نماد یک کمپین شناخته شود. با گروه‌بندی کاربران بر حسب سه دسته کاربران مشهور، سازمانی و شخص با محاسبه درصد جمعیتی برای هر گروه (نمودار ۵) را خواهیم داشت.



نمودار ۵. انواع کاربران بر حسب نوع طرفداری

مفهوم ارتباط بین کاربران اغلب با عنوان دنبال کنندگان پیاده‌سازی شده است. حوزه‌ی نفوذ کاربر را می‌توان تعداد دنبال کنندگان او دانست (لیو و بینگ، ۲۰۱۲). در کل میزان محبوب بودن کاربر به معنی تعداد دنبال کنندگانش است البته تأثیرگذاری مفهوم پیچیده‌تری است یعنی ممکن است همبستگی بین محبوبیت و تأثیرگذاری یک کاربر کمتر از حد انتظار باشد

وجود رابطه‌ی دنبال کردن، لزوماً به معنای تعامل واقعی بین آن‌ها نیست و یک شبکه‌ی اجتماعی مخفی واقعی در پس پرده‌ی روابط دنبال کردن شبکه‌های اجتماعی وجود دارد (جانگ، جین یی، هان و لی، ۲۰۱۵). تعداد دوستان اهمیت کمتری نسبت به دنبال کنندگان دارد و اغلب وابستگی مثبتی با چند مرجع مشخص مانند دبیرستان محل تحصیل، شهر،... دارد (جانگ، جین یی، هان و لی، ۲۰۱۵)

اگر شاخص FF-ratio نسبت تعداد دنبال کنندگان به دوستان فرض شود و شاخص S-ratio را نسبت تعداد پیام‌های استاتوس نسبت به سن حساب کاربری فرض کنیم در (جدول ۶) داریم:

جدول ۶. میانگین FF-ratio، میانگین سنی حساب کاربری، میانگین S-ratio بر حسب نوع طرفداری

نوع طرفداری	میانگین FF-ratio	میانگین سنی حساب کاربری	میانگین S-ratio
کلینتون	۹۲/۰۵	۴/۰۶	۶۶۹۸/۰۹
ترامپ	۵۳/۵۴	۳/۸۶	۷۳۸۲/۶۶
بی طرف	۶۳/۷۲	۳/۹۶	۵۵۵۶/۸۵

با توجه به میانگین سنی حساب کاربری، طرفداران کلینتون در توییتر به طور میانگین دارای سابقه بیشتری نسبت به طرفداران ترامپ هستند و با توجه به شاخص FF-ratio محبوبیت بیشتری دارند اما طرفداران ترامپ با توجه به شاخص S-ratio فعال ترند. مقایسه کاربران بر حسب ملیت توزیع جمعیتی (نمودار ۶) را برای کمپین‌های طرفداری رقم می‌زنند. کاملاً مشخص است که کلینتون چه در بین کاربران آمریکایی و چه کاربران غیر آمریکایی محبوبیت کمتری نسبت به ترامپ دارد. البته درصد نسبتاً بالایی از کاربران چه آمریکایی و چه غیر آمریکایی دارای گرایش‌های بی طرف بوده‌اند.



نمودار ۶. مقایسه توزیع جمعیتی طرفداران بر حسب ملیت

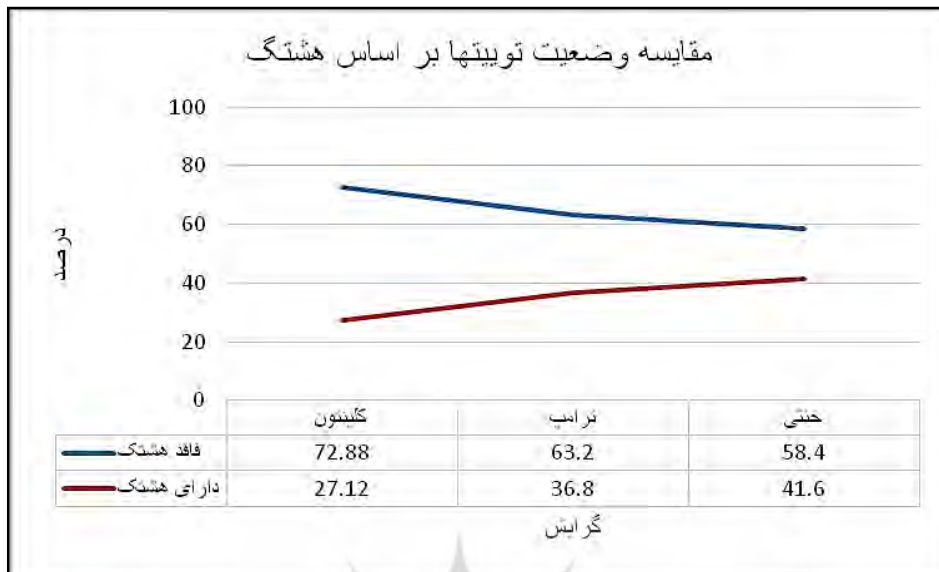
تحلیل هشتک‌ها

هشتک گذاری که نوعی نمایه‌گذاری انسانی توسط کاربران است ابداع شد تا کاربران در شبکه‌های اجتماعی در بین هشتک‌ها که همان کلیدواژگان درون‌متنی است به جستجو بپردازد. ماهیت هشتک از چیزی به نام کلیدواژه بسیار فراتر رفت و موج‌سازی اجتماعی و انقلاب‌های توییتری را رقم زد. کاربران می‌توانند در قبال یک پدیده اجتماعی یا سیاسی به‌طور دست‌جمعی اعتراض نموده یا رضایت خود را با درج هشتک‌های واحد در توییت‌های شبکه‌های اجتماعی خود اعلام کنند. اگرچه تگ، هشتک و منشن در ابتدا در توییت‌ها ظاهر شدند، اما اکنون در بسترهای مختلفی از رسانه‌های اجتماعی گسترده شده‌اند. لازم به ذکر است تگ‌های اجتماعی با هشتک تفاوت دارند. تگ‌های اجتماعی در فهرست ثابت شامل مجموعه‌ای از تگ‌های از پیش تعریف‌شده سازمان‌دهی شده است درحالی‌که هشتک نه تنها به‌صورت ثابت توسط دسته‌بندی موضوعی گروه‌بندی نشده است بلکه می‌تواند شامل

تعاملات ارتباطی پویا مانند اتفاقات زنده و بی‌درنگ و ... باشد. (جدول ۷) میانگین تعداد هشتک در هر توییت و درصد سهم از کل هشتک‌ها برحسب نوع طرفداری هر گروه نشان می‌دهد. سهم کل کمپین ترامپ از هشتک‌ها بیش از ۲/۵ برابر کمپین کلینتون بوده است و میانگین تعداد هشتک در توییت‌ها برای آنان ۱/۵ برابر طرفداران کلینتون بوده است. این یعنی انتشار سازمان‌دهی شده تر و هدفمندتر مطالب. آمار هشتک‌ها به‌طور واضح به نفع کمپین ترامپ است و این مسئله بسیار مهم است زیرا جوامعی که توسط هشتک‌ها به‌صورت ضمنی تشکیل می‌شوند ممکن است در یک بازه‌ی زمانی طولانی وجود نداشته باشند اما می‌توانند به‌طور بی‌نهایت کارا باشد حتی اگر برای یک‌زمان کوتاه وجود داشته باشند. (نمودار ۷) درصد توییت‌های دارای و فاقد هشتک را برحسب نوع طرفداری هر گروه نشان می‌دهد. (جدول ۸) ده هشتک برتر طرفداران را همراه با درصد نسبت به تعداد کل هشتک‌های هر گروه نشان می‌دهد.

جدول ۷. میانگین تعداد هشتک در هر توییت و درصد سهم از کل هشتک‌ها برحسب نوع طرفداری

نوع طرفداری	درصد سهم از کل هشتک‌ها	میانگین تعداد هشتک در هر توییت
کلینتون	۲۶/۱۵	۰/۶۳
ترامپ	۳۸/۷۱	۰/۸۹
بی‌طرف	۴۶/۰۳	۱/۰۲



نمودار ۷. درصد توییت‌های دارای و فاقد هشتک بر حسب نوع طرفداری

جدول ۸. ده هشتک برتر طرفداران را همراه با درصد نسبت به تعداد کل هشتک‌های هر گروه

افراد بی طرف		طرفداران ترامپ		طرفداران کلینتون	
برترین هشتک‌ها	درصد	برترین هشتک‌ها	درصد	برترین هشتک‌ها	درصد
#Trump	۱۳/۲۰	#Trump	۱۰/۹۹	#Trump	۱۲/۱۷
#hillary	۵/۶۰	#hillary	۵/۷۴	#hillary	۵/۴۷
#hillaryClinton	۴/۹۲	#maga	۴/۰۹	#hillaryClinton	۳/۵۳
#Clinton	۳/۶۷	#hillaryClinton	۳/۷۳	#imwithher	۳/۱۵
#imwithher	۳/۴۹	#Trumptrain	۲/۳۳	#Clinton	۲/۵۸
#maga	۲/۲۵	#Clinton	۲/۳۱	#maga	۲/۵۲
#donaldTrump	۱/۸۳	#imwithher	۲/۲۲	#wikileaks	۱/۹۱
#podestaemails	۱/۰۵	#Trump2016	۲/۰۱	#podestaemails31	۱/۷۵
#podestaemails31	۱/۰۲	#draintheswamp	۱/۶۷	#Trumpcheck	۱/۵۵
#spiritcooking	۰/۹۶	#podestaemails31	۱/۵۴	#draintheswamp	۱/۱۹

طرفداران کلیتون در توییت‌های خود از هشتک #Trump در رتبه اول و به نسبت ۲/۲۲ برابر بیشتر از هشتک #hillary در رتبه دوم استفاده نموده‌اند درحالی که طرفداران ترامپ از هشتک در توییت‌های خود از هشتک #Trump در رتبه اول و به نسبت ۱/۸۴ برابر بیشتر نسبت به #hillary استفاده نموده‌اند. افراد بی طرف نیز این نسبت را با اندازه ۲/۵ برابر حفظ نموده‌اند. در هر حال #TRUMP با برتری میانگین ۲/۱۳ برابر نسبت به #hillary در هر سه گروه هشتک برتر است و برتری حدود دو برابری آن بر هشتک #hillary که در رتبه بعدی است مشهود است. چیزی که قابل توجه است این است که در رتبه‌بندی هشتک‌های افراد بی طرف #podestaemails و #podestaemails31 و #spiritcooking به چشم می‌خورد و آن‌ها به این هشتک‌ها توجه نشان داده‌اند. این هشتک‌ها یادآور ماجرای Podesta email است. (جدول ۹) الگوهای مکرر هشتک‌های هر گروه طرفداران را نشان می‌دهد. الگوهای مکرر به مجموعه‌ای کلمات، آیتم‌ها، زیررشته‌ها و زیر گراف‌ها و یا دیگر واحدهای داده گفته می‌شود که به صورت مکرر در مجموعه داده‌ها تکرار شده‌اند و برای پیش‌بینی روند کلی داده‌ها قابل استفاده هستند (جیاوی، هان و کامبر ۲۰۰۱) و در داده کاوی، گراف کاوی، عقیده کاوی و ... استفاده دارند.

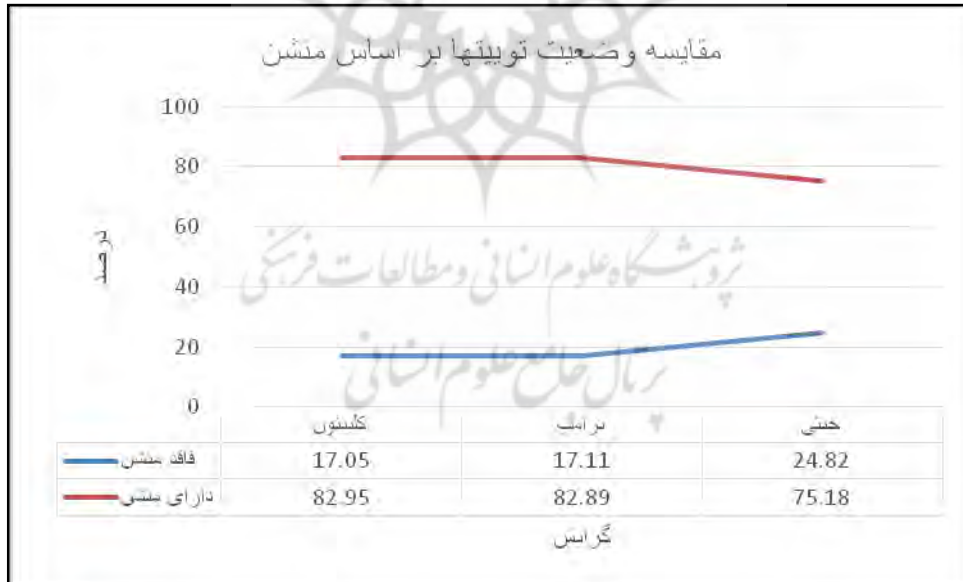
جدول ۹. الگوهای مکرر هشتک‌های هر گروه طرفداران

الگوهای هشتک افراد بی طرف	درصد	الگوهای هشتک طرفدار ترامپ	درصد	الگوهای هشتک طرفدار کلیتون	درصد
#Trump	۳/۸۹	#Trump	۳/۰۴	#Trump	۲/۲۵
#hillary	۱/۳۲	#hillary	۱/۱۷	#Trumpcheck	۰/۹۲
#imwithher	۱/۳۲	#Trump2016 #Trumptrain	۱/۰۷	#hillary	۰/۸۳
#hillaryClinton	۰/۹۷	#hillaryClinton	۰/۵۴	#wikileaks	۰/۵۶
#decision2016	۰/۹۰	#Clinton	۰/۵۱	#dnleak2	۰/۴۳
#americafirst	۰/۵۸	#maga	۰/۴۹	#hillaryClinton	۰/۴۰
#Clinton	۰/۵۳	#podestaemails31	۰/۴۷	#podestaemails31	۰/۳۸
#donaldTrump	۰/۴۶	#spiritcooking	۰/۴۵	#Clinton	۰/۳۴

۰/۳۲	#imwithher	۰/۳۲	#americafirst	۰/۳۹	#podestaemails #podestaemails31 #hillaryClinton #imwithher
۰/۲۳	#spiritcooking	۰/۲۹	#donaldTrump	۰/۳۸	#emabiggestfanslad ygaga

تحلیل منشها

منشن که با فرمت @username درون متون تعیبه می‌شود، افرادی که نام کاربری آن‌ها در متن با @ ذکر می‌شد را خطاب می‌کرد. منشن در جنجال‌ها توانایی کمتری نسبت به هشتک ندارد. کاربران اگر تصمیم می‌گرفتند اعتراضی را علیه یک شخص خاص که در توییت به راه‌اندازند و او را مورد خطاب قرار دهند، کافی بود در متن خود با @ نام کاربری آن فرد را درج کنند تا هم پست درج شده به آن فرد Notify و اعلان شود و هم دیگران ببینند که این پست خطاب به چه فرد یا افرادی بوده است. آن فرد مورد خطاب حتی می‌تواند یک رئیس‌جمهور یا یک فرد مشهور باشد. (جدول ۱۰) میانگین تعداد منشن در هر توییت و درصد سهم از کل منشها برحسب نوع طرفداری هر گروه نشان می‌دهد



نمودار ۸. درصد توییت‌های دارای و فاقد منشن را برحسب نوع طرفداری هر گروه

جدول ۱۰. میانگین تعداد منشن در هر توییت و درصد سهم از کل منشنها برحسب نوع طرفداری

نوع طرفداری	میانگین تعداد منشن در هر توییت	درصد سهم از کل منشنها
کلینتون	۰/۹۹	۲۰/۸۹
ترامپ	۱/۰۳	۳۹/۳۶
بی طرف	۱/۰۱	۳۹/۷۵

(نمودار ۸) درصد توییت‌های دارای و فاقد منشن را برحسب نوع طرفداری هر گروه نشان می‌دهد. (جدول ۱۱) ده منشن برتر طرفداران را همراه با درصد نسبت به تعداد کل منشنهای هر گروه نشان می‌دهد. بیشترین منشنهای طرفداران ترامپ خطاب به اول-ویکی لیکس، دوم-دونالد ترامپ و سوم-هیلاری کلینتون بوده است و یکی لیکس سایت افشاگری است که در جریان انتخابات برخی اطلاعات تخریب‌کننده درباره هیلاری کلینتون فاش نمود. طرفداران کلینتون بیشترین منشنهای خود را خطاب به اول- nowthisnews، دوم- berniesanders و سوم- channel4news انجام داده‌اند. nowthisnews و channel4news از خبر گذاری‌های شبکه‌های اجتماعی هستند و berniesanders یکی از نامزدهای ریاست جمهوری در دور مقدماتی است. افراد بی طرف نیز اول- دونالد ترامپ، دوم-هیلاری کلینتون و سوم-ویکی لیکس را منشن نموده‌اند. بنابراین می‌توان نتیجه گرفت طرفداران ترامپ در زمینه منشن هدفمندتر از طرفداران کلینتون عمل نموده‌اند و اصطلاحاً در خط مقدم فعالیت کرده‌اند نه پشت جبهه. افراد بی طرف نیز با منشن نمودن بیشتر دونالد ترامپ و ویکی لیکس بیشتر در جبهه مقابل کلینتون تعامل داشته‌اند. مسئله‌ای که باز هم دیده می‌شود تمرکز افراد بی طرف بر روی سایت ویکی لیکس است.

(جدول ۱۲) الگوهای مکرر منشنهای هر گروه طرفداران را نشان می‌دهد. الگوهای پرتکرار برتر طرفداران کلینتون دقیقاً همانند منشنهای منفرد برتر آنهاست اما طرفداران ترامپ - darren32895836 را در الگوهای پرتکرار جانشین - realdonaldTrump نموده‌اند که احتمالاً حساب کاربری پرنفوذ است. افراد بی طرف نیز در الگوهای پرتکرار، شبکه خبری cnn را جایگزین hillaryClinton نموده‌اند؛ اما رتبه ویکی لیکس به رتبه نخست تغییر

یافته که حاکی از توجه خاص آنان به سایت افشاگر ویکی لیکس است که علیه هیلاری کلینتون فعالیت می‌نمود.

جدول ۱۱. منشهای برتر گروه طرفداران

افراد بی طرف		طرفداران ترامپ		طرفداران کلینتون	
برترین منشها	درصد	برترین منشها	درصد	برترین منشها	درصد
@realdonaldTrump	۳/۰۹	@wikileaks	۴/۲۸	@nowthisnews	۶/۰۱
@hillaryClinton	۲/۴۸	@realdonaldTrump	۲/۴۳	@berniesanders	۴/۳۶
@wikileaks	۲/۲۶	@hillaryClinton	۲/۱۲	@channel4news	۲/۹۶
@cnn	۱/۲۸	@darren32895836	۱/۵۰	@ddale8	۲/۰۴
@youtube	۱/۱۸	@wdfx2eu8	۱/۴۰	@realdonaldTrump	۱/۸۲
@wdfx2eu8	۱/۰۱	@youtube	۱/۲۶	@hillaryClinton	۱/۴۰
@nbcphiladelphia	۰/۸۹	@berniesanders	۱/۰۸	@wikileaks	۱/۱۷
@syeddoha	۰/۸۹	@drjillstein	۱/۰۴	@cnn	۰/۹۵
@darren32895836	۰/۸۰	@lindasuhler	۱/۰۳	@mitchellvii	۰/۹۲
@kellyannpolls	۰/۶۵	@stylishrentals	۰/۹۰	@wdfx2eu8	۰/۸۷

جدول ۱۲. الگوهای مکرر منشهای گروه طرفداران

درصد	الگوهای منش طرفدار کلینتون	درصد	الگوهای منش طرفدار ترامپ	درصد	الگوهای منش افراد بی طرف
۵/۸۹	@nowthisnews	۳/۳۷	@wikileaks	۱/۷۲	@wikileaks
۴/۲۲	@berniesanders	۱/۳۴	@darren32895836	۱/۱۸	@realdonaldTrump
۲/۹۱	@channel4news	۱/۲۲	@hillaryClinton	۰/۹۳	@cnn
۱/۹۶	@ddale8	۱/۰۳	@drjillstein	۰/۹۰	@syeddoha @nbcphiladelphia
۰/۸۴	@mitchellvii	۰/۹۸	@berniesanders	۰/۷۷	@youtube
۰/۸۱	@wikileaks	۰/۹۷	@lindasuhler	۰/۶۹	@darren32895836

تحلیل کلمات

با توجه به تعداد محدود و ثابت کاراکترها در یک توییت هر کاراکتر در یک توییت ارزشمند است. کلمات بیشتر و متن طولانی تر به معنای فرصت بیشتر برای استفاده از کلمات مفیدتر و کلیدی تر است در نتیجه از نظر چگالی و رتبه بندی در میان توییت ها، معمولاً توییت های طولانی تر شانس بیشتری برای مورد توجه قرار گرفتن و ری توییت شدن دارد. (جدول ۱۳) ده کلمه با بیشترین استفاده در توییت های ارسالی در بازه زمانی مطالعه همراه با درصد آنان نسبت به تعداد کل کلمات نشان می دهد.

جدول ۱۳. ده کلمه با بیشترین استفاده در توییت های مطالعه همراه با درصد آنان نسبت به تعداد کل

کلمات

افراد بی طرف		طرفداران ترامپ		طرفداران کلینتون	
برترین کلمات	درصد	برترین کلمات	درصد	برترین کلمات	درصد
hillary	۳/۹۶	Trump	۴/۸۱	Trump	۶/۴۱
Trump	۳/۶۲	donald	۴/۲۹	donald	۴/۸۳
Clinton	۳/۴۵	hillary	۴/۲۶	hillary	۳/۲۶
donald	۲/۹۶	Clinton	۳/۸۳	Clinton	۳/۰۴
vote	۰/۹۴	vote	۰/۸۷	obama	۱/۴۴
president	۰/۶۵	fbi	۰/۸۲	protester	۱/۲۶
fbi	۰/۴۹	rally	۰/۷۲	pro	۱/۱۷
rally	۰/۴۷	emails	۰/۶۹	president	۱/۱۴
jay	۰/۴۶	president	۰/۵۱	vote	۱/۰۱
election	۰/۴۶	Clintons	۰/۴۹	lied	۰/۸۳

در (جدول ۱۴) بار احساسی کلمات برتر مشخص شده است. کاربران اغلب به طور مستقیم احساس خود را بروز نمی دهند و ممکن است به شیوه های متعدد نظر خود را بروز دهد و یکی از این شیوه ها استفاده از کلمات دارای بار احساسی است. بازی با کلمات در کمپین های

انتخاباتی یکی از بازی‌های جنگ روانی است. اگرچه استفاده از کلمات مثبت لزوماً به معنای مثبت صحبت کردن نیست اما می‌تواند شاخصی از مثبت صحبت کردن ارائه بدهد. طرفداران ترامپ کلمات مثبت بیشتری نسبت به طرفداران کلینتون بکار برده‌اند.

جدول ۱۴. بار احساسی کلمات برتر در کمپین‌های طرفداری بر اساس دیکشنری (تان، سانگو و همکاران، ۲۰۰۹)

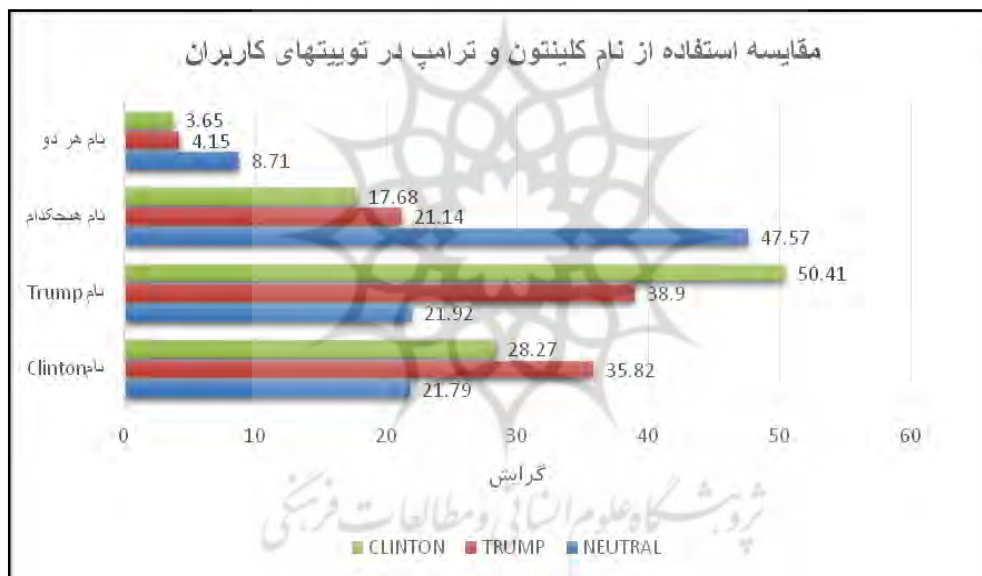
افراد بی طرف		طرفداران ترامپ		طرفداران کلینتون	
برترین کلمات	بار احساسی	برترین کلمات	بار احساسی	برترین کلمات	بار احساسی
hillary	.	Trump	.	Trump	.
Trump	.	donald	.	donald	.
Clinton	.	hillary	.	hillary	.
donald	.	Clinton	.	Clinton	.
vote	+	vote	+	obama	.
president	.	fbi	.	protester	0
fbi	.	rally	+	pro	0
rally	+	emails	+	president	.
jay	-	president	.	vote	+
election	+	Clintons	.	lied	-

(جدول ۱۵) میانگین تعداد کلمه در هر توییت و درصد سهم از کل کلمات برحسب نوع طرفداری هر گروه نشان می‌دهد. طرفداران کلینتون تعداد کمتری توییت ارسال نموده‌اند و توییت‌های آن‌ها به لحاظ تعداد کلمات کوتاه‌تر بوده است. سهم طرفداران ترامپ از کل کلمات ۱/۶ برابر طرفداران کلینتون و حتی بیشتر از افراد بی طرف بوده است

جدول ۱۵. میانگین تعداد کلمه در هر توییت و درصد سهم از کل کلمات برحسب نوع طرفداری

نوع طرفداری	درصد سهم از کل کلمات	میانگین تعداد کلمه در هر توییت
کلینتون	۲۳/۸۶	۸/۱۹
ترامپ	۴۰/۳۳	۷/۶۶
بی طرف	۳۵/۸۱	۶/۶۰

(نمودار ۹) درصد استفاده از نام کاندیداها در توییت‌ها برحسب نوع طرفداری هر گروه نشان می‌دهد. طرفداران کلینتون از نام ترامپ بیشتر در توییت‌هایشان استفاده کرده‌اند و حتی این نسبت به ۱/۷ برابر می‌رسد درحالی‌که طرفداران ترامپ بالعکس نام ترامپ را استفاده کرده‌اند. این نیز می‌تواند بدان معنا باشد که کمپین کلینتون به‌جای پرداختن به حمایت از کلینتون به حمله به ترامپ پرداخته‌اند و بالعکس کمپین ترامپ به‌جای تاختن به کلینتون بیشتر به حمایت از ترامپ پرداخته‌اند. کاربران بی‌طرف با درصد تقریباً برابر از نام ترامپ و کلینتون استفاده کرده‌اند و با درصد خیلی بیشتر (بیش از دو برابر) نام هر دو را به‌کاربرده‌اند یا اصلاً نامی از هیچ‌یک ذکر نکرده‌اند که کاملاً منطقی و قابل‌قبول است و گواهی بر تشخیص صحیح نوع طرفداری کاربران توسط الگوریتم ماست.



نمودار ۹. مقایسه استفاده از نام کلینتون و ترامپ در توییت‌های کاربران

نتیجه‌گیری

اثرات شبکه‌های اجتماعی در روابط شخصی، فرهنگی و زندگی، تعلیم و تربیت ثابت شده و قابل مشاهده است اما اثبات توانایی این شبکه‌ها در برخی زمینه‌ها مانند انعکاس گرایش سیاسی

غالب در انتخابات یک مسئله چالش‌برانگیز و قابل تأمل است. استفاده از فرآیند آنالیز شبکه‌های اجتماعی با رویکرد آنالیز محتوا به‌عنوان ابزار استخراج دانش از داده‌های بدون ساختار و یا شبه ساخت یافته شبکه‌های اجتماعی مانند توییتر می‌تواند نتایج جالب و حیرت‌انگیزی تولید نماید. در انتخابات ۸ نوامبر ریاست جمهوری ایالات متحده ۲۰۱۶ در تصور همگانی، نتایج نظرسنجی‌های سنتی و تحلیل‌های کارشناسی هیلاری کلینتون برنده انتخابات تصور می‌شد در حالی که با بررسی توییت‌های سه روز منتهی به انتخابات به‌طور کاملاً شفاف و واضح می‌توان دید که کمپین انتخاباتی دونالد ترامپ در بسیاری از زمینه‌ها بسیار موفق‌تر عمل کرده است و این برتری حتی در مواردی به بیش از دو برابر می‌رسد. در میان افراد بی‌طرف نیز گرایش به ترامپ غالب است این بدین معناست که احتمالاً آرا درصد بالایی از این افراد در نهایت به صندوق ترامپ ریخته خواهد شد. در این پژوهش علاوه بر موارد ذکر شده ما حتی توانسته‌ایم استراتژی‌های انتخاباتی را تشخیص دهیم به‌عنوان مثال فهمیده‌ایم کمپین کلینتون به‌جای حمایت از کلینتون بیشتر به حمله به ترامپ پرداخته و بالعکس کمپین ترامپ به‌جای تاختن به کلینتون بیشتر به حمایت از ترامپ پرداخته است. ما توانسته‌ایم جریان‌های سیاسی تأثیرگذار را شناسایی کنیم به‌عنوان مثال کاربران بخصوص کاربران بدون گرایش و بی‌طرف، به سایت ویکی لیکس و ماجرای افشای ایمیل‌ها توجه خاص نشان داده‌اند. در کل می‌توان نتیجه گرفت در زمینه انتخابات در صورتی که شبکه اجتماعی به اندازه کافی محبوب باشد که مشکل دموگرافیک وجود نداشته باشد می‌توان بهره‌برداری خوبی از محتوای تولیدی شبکه‌های اجتماعی محبوب مانند توییتر در راستای بررسی جریان‌های سیاسی، پیش‌بینی برنده نهایی انتخابات، آنالیز رفتاری کمپین‌های انتخاباتی و... انجام داد.

منابع

سید محمدباقر جعفری، آصف کریمی و زاهده ابرقوی زاده، (۱۳۹۵)، عوامل تأثیرگذار بر تمایل به ادامه استفاده از وبسایت شبکه‌های اجتماعی. **فصلنامه مطالعات مدیریت فناوری اطلاعات**، دانشگاه علامه طباطبائی، دوره ۵، شماره ۱۷، پاییز ۱۳۹۵، صفحه ۱۴۷-۱۸۲

Aggarwal, Charu C. "Outlier analysis." *Data mining*. Springer International Publishing, 2015.

Anstead, Nick, and Ben O'Loughlin. "Social media analysis and public opinion: The 2010 UK general election." *Journal of Computer-Mediated Communication* 20.2 (2015): 204-220.

Baccianella, Stefano, Andrea Esuli, and Fabrizio Sebastiani. "SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining." *LREC*. Vol. 10. 2010

Benevenuto, F. et al. (2010). Detecting spammers on twitter. Collaboration, electronic messaging, **anti-abuse and spam conference** (CEAS).

Borgatti, Stephen P., et al. "Network analysis in the social sciences." *science* 323.5916 (2009): 892-895.

Carlisle, Juliet E., and Robert C. Patton. "Is social media changing how we understand political engagement? An analysis of Facebook and the 2008 presidential election." *Political Research Quarterly* 66.4 (2013): 883-895.

Fernandes, Juliana, et al. "The writing on the wall: A content analysis of college students' Facebook groups for the 2008 presidential election." *Mass Communication and Society* 13.5 (2010): 653-675.

Frakes, William B., and Ricardo Baeza-Yates. "**Information retrieval: data structures and algorithms**." (1992).

Garg, Neha P., et al. "Role recognition for meeting participants: an approach based on lexical information and social network analysis." **Proceedings of the 16th ACM international conference on Multimedia**. ACM, 2008.

Jang, Jin Yea, Kyungsik Han, and Dongwon Lee. "No Reciprocity in Liking Photos: Analyzing Like Activities in Instagram." **Proceedings of the 26th ACM Conference on Hypertext & Social Media**. ACM, 2015.

Jiawei, Han, and Micheline Kamber. "Data mining: concepts and techniques." San Francisco, CA, itd: **Morgan Kaufmann** 5 (2001).

Kurka, David Burth, Alan Godoy, and Fernando J. Von Zuben. "**Online social network analysis: A survey of research applications in computer science.**" 1504.05655 (2015).

Larsson, Anders Olof, and Hallvard Moe. "Studying political microblogging: Twitter users in the 2010 Swedish election campaign." **New Media & Society** 14.5 (2012): 729-747.

Li, Hui, Jiangtao Cui, and Bingqing Shen. "A Study towards Application-driven Social Network Analysis." **JCP** 9.1 (2014): 134-145.

Liu, Bing. "Sentiment analysis and opinion mining." **Synthesis lectures on human language technologies** 5.1 (2012): 1-167.

Ntalianis, Klimis, Abdel-Badeeh M. Salem, and Ibrahiem El Emary. "Social media content ranking based on social computing and user influence." **Procedia Computer Science** 65 (2015): 148-157.

O'Malley, A. James, and Jukka-Pekka Onnela. "**Topics in social network analysis and network science.**" 1404.0067 (2014).

Sang, Erik Tjong Kim, and Johan Bos. "Predicting the 2011 dutch senate election results with twitter." Proceedings of the workshop on semantic analysis in social media. **Association for Computational Linguistics**, 2012.

Some, Rupam. "A Survey on Social Network Analysis and its Future Trends." **International Journal of Advanced Research in Computer and Communication Engineering** 2.6 (2013).

Stringhini, G., et al. (2010). Detecting spammers on social networks. **Proceedings of the 26th Annual Computer Security Applications Conference**, ACM.

Tan, Songbo, et al. "Adapting naive bayes to domain adaptation for sentiment analysis." **European Conference on Information Retrieval**. Springer Berlin Heidelberg, 2009.

Tumasjan, Andranik, et al. "Election forecasts with Twitter: How 140 characters reflect the political landscape." **Social science computer review** 29.4 (2011): 402-418.

Wang, Hao, et al. "A system for real-time twitter sentiment analysis of 2012 us presidential election cycle." **Proceedings of the ACL 2012 System Demonstrations**. Association for Computational Linguistics, 2012.

Woolley, Julia K., Anthony M. Limperos, and Mary Beth Oliver. "The 2008 presidential election, 2.0: A content analysis of user-generated political Facebook groups." **Mass Communication and Society** 13.5 (2010): 631-652.

Yardi, S., et al. (2009). "**Detecting spam in a twitter network.**" First Monday 15(1)

