



## تحلیل احساسی مبتنی بر جنبه بر روی داده های شبکه اجتماعی توئیت درمورد گیاهخواری

فاطمه صلاحی<sup>۱</sup>

<sup>۱</sup> کارشناسی ارشد، گروه کامپیوتر، دانشگاه آزاد اسلامی، بیرجند، ایران

Salahi.fateme@gmail.com

### چکیده

گیاهخواری از جریان‌هایی است که بازخوردهای زیادی در شبکه‌های اجتماعی داشته است. مطالب منتشر شده توسط کاربران نشان دهنده احساسات و نظرات آنان نسبت به این جریان و جنبه‌های مختلف آن می‌باشد. در همین راستا، مجموعه داده‌ای شامل بیش از شصت هزار توئیت<sup>۱</sup> منتشر شده در سال ۲۰۲۳ در مورد گیاهخواری جمع آوری شده است. این مجموعه برای استخراج احساسات کاربران نسبت به جنبه‌های مختلف گیاهخواری استفاده شده است. ابتدا روشی مبتنی بر مدل زبانی RoBERTa برای تحلیل احساسات ضمنی نهفته در توئیت‌ها ارائه می‌شود. سپس با استفاده از مدل سازی موضوعی LDA<sup>۲</sup>، تعدادی جنبه و موضوع مرتبط با گیاهخواری استخراج می‌شود. در مرحله بعد با استفاده از روشی مبتنی بر مدل زبانی DeBERTa به تحلیل احساسات توئیت‌ها نسبت به جنبه‌های مختلف استخراج شده، پرداخته می‌شود. نمودارهای مختلف فراوانی و توزیع احساسات برای جنبه های مختلف در حیطه ی گیاهخواری مورد بررسی قرار می‌گیرد. با نمودارهایی نتایج حاصل از تحلیل احساسات مبتنی بر RoBERTa با نتایج حاصل از DeBERTa در کنار هم، مورد بحث و بررسی قرار می‌گیرد. تجزیه و تحلیل داده‌ها با استفاده از مدل مبتنی بر DeBERTa نشان می‌دهد که کاربران در مورد جنبه‌های plant و lifestyle توئیت‌هایی اکثراً با جهتگیری مثبت منتشر کرده‌اند. در مورد جنبه Animal غالباً با احساسی منفی مطالبی منتشر کرده‌اند. برای هر یک از جنبه های Diet و Co<sup>۳</sup> با مقادیری نزدیک به هم، اکثر توئیت‌ها مثبت و یا خنثی هستند. در میان بحث، تعدادی دانش ضمنی در رابطه با این موضوع مورد بررسی قرار می‌گیرد.

**واژه‌های کلیدی:** تحلیل احساسات مبتنی بر جنبه، مدل سازی موضوعی، پردازش زبان طبیعی، پردازش متن

<sup>۱</sup> Tweet

<sup>۲</sup> Latent Dirichlet allocation

<sup>۳</sup> company

## ۱. مقدمه

گياهخواری یکی از جریان هایی است که از سالیان پیش موافقان و معترضان داشته است. این جریان که به نوعی سبک زندگی تبدیل شده بر جنبه های مختلف زندگی تاثیرگذار بوده است .

افراد نسبت به جنبه های مختلف این سبک زندگی دیدگاه های متفاوتی دارند. گاهی نسبت به بخشی از آن دیدگاهی مثبت و نسبت به بخشی دیگر دیدگاهی منفی دارند. بررسی جزئیات دیدگاه افراد ، دید ژرفتری نسبت به این جریان فراهم می کند.

روزانه افراد مختلف در تایید این سبک زندگی و یا رد آن ، در شبکه های اجتماعی مطالبی را منتشر می کنند. سایت های رسانه های اجتماعی منبع اصلی اطلاعات تولید شده توسط کاربران در مورد سیاست، محصولات، ایده ها و خدمات هستند . (Ahmed et al., 2017) مطالب منتشر شده از جانب افراد ، بیانگر فکر و عقیده ی آنها در مورد موضوع مورد بحث است. بررسی این مطالب یاری دهنده ای برای شناختن سمت و سوی این جریان و زوایای آن خواهد بود.

یکی از شبکه های اجتماعی گسترده و مورد اقبال در جامعه جهانی ، توئیتر<sup>۴</sup> می باشد. که روزانه بستر خیل عظیمی از پیام های منتشر شده است. میلیون ها نفر از توئیتر استفاده می کنند و این سایت به عنوان یکی از پربازدیدترین سایت ها در سراسر جهان شناخته می شود (Zainuddin et al., 2020). افراد در مورد موضوعات مختلف از جمله گياهخواری در این شبکه ی اجتماعی نظراتی را بیان می کنند. لذا توئیتر به عنوان یک منبع عظیم از داده هاست که امکان مطالعه ی احساسات و عقاید افراد را فراهم می کند. بنابراین ما برای بررسی دیدگاه ها و احساسات مختلف افراد در مورد جنبه های مختلف این سبک زندگی، پلتفرم عظیم توئیتر را برگزیدیم.

تجزیه و تحلیل احساسات مبتنی بر جنبه<sup>۵</sup> (ABSA) با هدف شناسایی احساسات موجودیت های هدف و جنبه های آنها انجام می شود. (Ding et al., 2022) این نوع تحلیل که برای شناخت احساسات افراد نسبت به جنبه های مختلف یک موضوع نیز ، یک زمینه ی تحقیقاتی است که در حیطه ی آن پژوهش هایی انجام شده است. یکی از بسترهای پرکاربرد در تحلیل احساسات مبتنی بر جنبه ، تحلیل نظر افراد در مورد محصولات مختلف تجاری است ، چرا که جزئیات دیدگاه افراد را نسبت به جنبه های مختلف مانند، قیمت، پارامترهای کیفیت و ... ارائه می دهد. کار تحلیل احساسات مبتنی بر جنبه (ABSA) با هدف شناسایی نظرات مصرف کنندگان در مورد جنبه های مختلف محصولات یا خدمات انجام می شود. (Silva and Marcacini, 2021)

با توجه به تاثیرات این سبک زندگی بر جامعه ی بشری و زندگی افراد، این پژوهش به مطالعه و تحلیل نظرات افراد و احساسات آنان نسبت به جنبه های مختلف سبک زندگی گياهخواری که در شبکه ی اجتماعی توئیتر منتشر شده است می پردازد تا دید عمیق تری در مورد این جریان و زوایای مختلف آن فراهم کند. بدین منظور بیش از شصت هزار توئیت انگلیسی مرتبط با این سبک زندگی و انواع آن، که در سال ۲۰۲۳ منتشر شده است جمع آوری شده است. این توئیت ها به طور تصادفی از آغاز سال ۲۰۲۳ تا آوریل همان سال گردآوری شده و مورد تحلیل قرار گرفته است. در گام اول توئیت ها جمع آوری

<sup>۴</sup> <https://twitter.com>

<sup>۵</sup> Aspect-Based Sentiment Analysis

شده و سپس مورد پیش پردازش و پاکسازی قرار گرفته اند. در گام بعد با استفاده از مدل زبانی RoBERTa (Liu et al., 2019) احساسات نهفته در توئیت ها مورد تحلیل و بررسی قرار گرفته است. در گامی دیگر با استفاده از LDA به مدلسازی موضوعی<sup>۶</sup> مجموعه ی توئیت ها پرداختیم. در این روش کلمات موجود در داده ها جمع آوری می شود و کلمات به یکی از موضوعات سند اختصاص می یابد. مجموعه ی داده به تعداد کمی موضوع تقسیم می شود. با استفاده از کلمات یافته شده در مدل سازی موضوعی، گامی دیگر در جهت تحلیل توئیت ها برداشتیم. در این مرحله روشی برای تحلیل احساسات مبتنی بر جنبه را ارائه دادیم. با استفاده از کلمات مختلف، احساسات افراد نسبت به آن موضوع را در مجموعه ی توئیت ها سنجیدیم. این کلمات از بین کلمات مدلسازی موضوعی بر اساس وزن کلمات و در بین آنها به صورت دلخواه انتخاب شده اند. روشی مبتنی بر مدل DeBERTa برای تحلیل مبتنی بر جنبه را در این گام به کار گرفتیم. در نهایت نتایج حاصل از تحلیل احساسات مبتنی بر جنبه را در نمودارهای مختلف به تصویر کشیدیم. و در مجموع نتایج حاصل از آن را در کنار نتایج حاصل از RoBERTa (Liu et al., 2019) قرار دادیم.

ادامه ی این پژوهش در چهار بخش سازماندهی شده است: در بخش ۲، روش تحقیق ارائه شده است. بخش ۳، به بیان یافته ها و بررسی آنها می پردازد و در بخش ۴ نتایج پژوهش ارائه شده است.

## ۲. پیشینه پژوهش

تحلیل احساسات بر روی داده های توئیت در حوزه های مختلف انجام شده است. مثلاً بر روی داده های مرتبط با بیماری ها مانند کرونا (Nezhad and Deihimi, 2022) و یا بر روی توئیت ها مرتبط با واکسن های کووید-۱۹ (Marcec and Likic, 2022)، بررسی عقاید در مورد ارزشهای دیجیتالی (Hassan et al., 2022)، بر روی داده های مرتبط با صنایع مختلف مانند صنعت هوایی (Aljedaani et al., 2022) و در بسیاری از زمینه های دیگر تحلیل احساسات با استفاده از داده های توئیت انجام شده است.

در حوزه ی گیاه خواری نیز کارهای تحقیقاتی جهت تحلیل توئیت ها انجام شده است. روند احساسات نسبت به گیاهخواری روی حدود ۱۱۳ هزار توئیت از سال ۲۰۱۰ تا ۲۰۲۲ بررسی شده است. که روند مثبتی را در طی این دوازده سال کشف کرده است. (Shamoi, Elvina, et al., 2022) متن کاوی روی توئیت های سال ۲۰۱۹ و ۲۰۲۰ در رابطه با غذای سالم با جستجوی #healthyfood انجام شد؛ که غذاهای مرتبط با این هشتگ، گیاه خواری، غذای خانگی و ارگانیک کشف شدند، و پر استفاده ترین هشتگ همراه با آن، vegan بوده است (Pilař et al., 2021) در برخی از تحقیقات مربوط به این حوزه از تحلیل احساسات مبتنی بر جنبه<sup>۷</sup> و مدلسازی موضوعی استفاده شده است. (Jang et al., 2021) و یا در پژوهشی با استفاده از تکنیک تحلیل احساسات مبتنی بر جنبه تاثیر بیماری کووید ۱۹ را بر روی صنعت هوایی سنجیده اند. (Chang et al., 2022) همچنین در زمینه ی تحلیل نظرات مشتریان در شبکه های اجتماعی، پژوهش هایی (Banjar et al., 2021) و (Prakash and Sharma, 2023) انجام شده است. در پژوهشی از تحلیل احساسات مبتنی بر جنبه برای تشخیص تنفر در

<sup>۶</sup> Topic modeling

<sup>۷</sup> Aspect-Based Sentiment Analysis

توئیتر استفاده شده است. (Zainuddin et al., 2020) در پژوهشی دیگر، نظرات دانش آموزان با تحلیل احساسات مبتنی بر جنبه تحلیل شده است. (Sivakumar and Reddy, 2017)

در داده های مرتبط با سلامتی در توئیتر و با استفاده از مدل سازی موضوعی و همبستگی کاوی، کشف همبستگی بین گیاه خواری و یوگا انجام شد (Islam, 2019) ارزیابی دیدگاه کلی نسبت به گیاه خواری که در قالب های مختلف داده های آن جمع آوری شد؛ از جمله جمع آوری توئیتهایی که در آنها به کلمه ی *vegan* اشاره شده بود و با استفاده از الگوریتم Hedonometer میزان شادی را ارزیابی کرد. (Jennings, Laura, et al., 2019)

### ۳. روش تحقیق

#### ۳-۱- جمع آوری داده

داده های مورد استفاده در این پژوهش حاصل کاوش شبکه ی اجتماعی توئیتر با استفاده از *snsrape* می باشد. در طی این مرحله ، از میان توئیتهای منتشر شده در سال ۲۰۲۳ در مورد گیاهخواری ، بیش از شصت هزار توئیته انگلیسی به طور تصادفی در طی چهار ماه اول این سال جمع آوری شده است. با استفاده از اعمال محدودیت برای تعداد جمع آوری برای هر کلیدواژه در هر ماه، توزیع مجموعه ی توئیته ها در بین چهار ماه، تقریباً یکسان می باشد.

#### ۳-۲- پیش پردازش

برای انجام پیش پردازش و پاکسازی آنها، اقداماتی صورت گرفت. تمامی لینک ها از مجموعه ی توئیته ها حذف شد. همچنین نام های کاربری، اعداد، علائم و آنچه غیر از کلمات انگلیسی می باشد از مجموعه ی توئیته ها پاکسازی شد . همچنین تمام کلمات زائد<sup>۸</sup> از این مجموعه حذف شد .

#### ۳-۳- تحلیل احساسات

یکی از اقدامات مهم و قابل انجام بر روی توئیته ها، تحلیل احساسات آنها می باشد. با تحلیل احساسات می توان، احساسات کلی افراد را نسبت به موضوعات مختلف سنجید و از نتایج آن در کاربردهای دیگر بهره برد. برای انجام تحلیل احساسات در این مرحله از روشی مبتنی بر مدل زبانی *RoBERTa* (Liu et al., 2019) استفاده شده است. *RoBERTa* یک مدل زبانی پیشرفته که در زمینه ی پردازش زبان طبیعی<sup>۹</sup> بسیار پر کاربرد می باشد. *RoBERTa* با اعمال بر روی هر توئیته، سه جهت گیری احساسی مختلف شامل مثبت، منفی و خنثی را مورد سنجش قرار می دهد. برای هر توئیته به ازای هر یک از این مولفه های احساسی ، عددی در بازه ی ۰ و ۱ نسبت می دهد که نشان دهنده ی قدرت یا ضعف آن جهت گیری احساسی در آن توئیته می باشد. پس از محاسبه ی هر یک از این سه مولفه برای هر توئیته وارد مرحله ی بعدی تحلیل می شویم.

#### ۳-۴- مدل سازی موضوعی

<sup>8</sup> Stop word

<sup>9</sup> Natural language processing

مدل سازی موضوعی نوعی مدل سازی آماری برای کشف «موضوعات» انتزاعی است که در مجموعه ای از اسناد وجود دارد. مدل سازی موضوعی یک ابزار متن کاوی پرکاربرد برای کشف ساختارهای معنایی پنهان در یک متن است. LDA مدلی برای مدل سازی موضوعی است که برای طبقه بندی متن در یک سند به یک موضوع خاص استفاده می شود. LDA یکی از رایج ترین رویکردها برای تشخیص جنبه (Schouten and Frasincar, 2015) در میان روش های یادگیری ماشینی بدون نظارت است (Banjar et al., 2021). LDA سند را موضوع بندی کرده و واژگان را به موضوع نسبت می دهد. در این مرحله با استفاده از LDA، موضوعات مختلف در بین مجموعه ی توئیت ها را کشف و لیست واژگان هر موضوع را استخراج کردیم. ما استفاده از LDA را انتخاب کردیم زیرا ساده و محبوب است. (Jang et al., 2021) ما ۱۵ موضوع مختلف در این مجموعه داده دسته بندی کردیم. این تعداد دسته بعد از تست تعداد دسته بندی های متفاوت انتخاب شد. با توجه به لیست واژگان، از میان واژگان با امتیاز بالاتر، که بعضا در تعدادی از این دسته ها تکرار شده بود، ۵ جنبه را به دلخواه برگزیدیم.

### ۳-۵- تحلیل احساسات مبتنی بر جنبه

تجزیه و تحلیل احساسات مبتنی بر جنبه یک موضوع حیاتی در ارزیابی احساسات در سطح جزئیات است که قصد دارد یک محاسبه خودکار از قطبیت احساسات را با توجه به جنبه خاصی در زمینه آن ارائه دهد. ABSA جنبه ها را استخراج می کند و قطبیت احساسات جنبه را با درک نظرات نویسنده طبقه بندی می کند (Scaria et al., 2023). تحلیل احساسات بر مبنای جنبه یا ویژگی، به تعیین احساسات بیان شده در مورد ویژگی ها یا جنبه های مختلف موجودیت ها اشاره دارد. ویژگی های مختلف، می توانند واکنش های احساسی متفاوت ایجاد کنند. شناخت احساسات در مورد جنبه های متفاوت یک موضوع، می تواند به درک بهتر عقاید نسبت به آن موضوع کمک کند. برای انجام تحلیل احساسات مبتنی بر جنبه که به اختصار ABSA نامیده می شود، روشی مبتنی بر مدل زبانی DeBERTa (He et al., 2020) به کار گرفتیم. با استفاده از این مدل زبانی، قادر به تشخیص احساسات نهفته در متن نسبت به یک جنبه ی خاص از موضوع هستیم. مدل های زبان مبتنی بر BERT در برنامه هایی که نیاز به درک عمیق زبان دارند، مانند تجزیه و تحلیل احساسات، با موفقیت مورد استفاده قرار گرفته اند. با به کارگیری روش مبتنی بر DeBERTa (He et al., 2020)، می توانیم سوگیری احساسی هر متن را در قالب سه مقدار برای مولفه های احساس مثبت، منفی و خنثی، استخراج کنیم.

### ۴- یافته ها

بیش از شصت هزار توئیت انگلیسی مرتبط با انواع سبک های گیاهخواری در چهار ماه آغازین سال ۲۰۲۳ جمع آوری کردیم. پس از پیش پردازش و پاکسازی توئیت ها، مدل زبانی RoBERTa (Liu et al., 2019) را برای کشف سوگیری احساسی هر توئیت به کار گرفتیم.

سپس با استفاده از LDA، به مدل سازی موضوعی پرداختیم. پس از تست تعداد متفاوت موضوعات، با در نظر گرفتن تعداد موضوعات برابر ۱۵ به گروه بندی موضوعات متن پرداختیم. از میان گروه های مختلف موضوعی، در بین واژگان با امتیاز بالا که گاهی در چندین گروه تکرار شده بودند، تعدادی جنبه را به دلخواه انتخاب کردیم. جنبه های انتخاب شده به قرار جدول ۱ است.

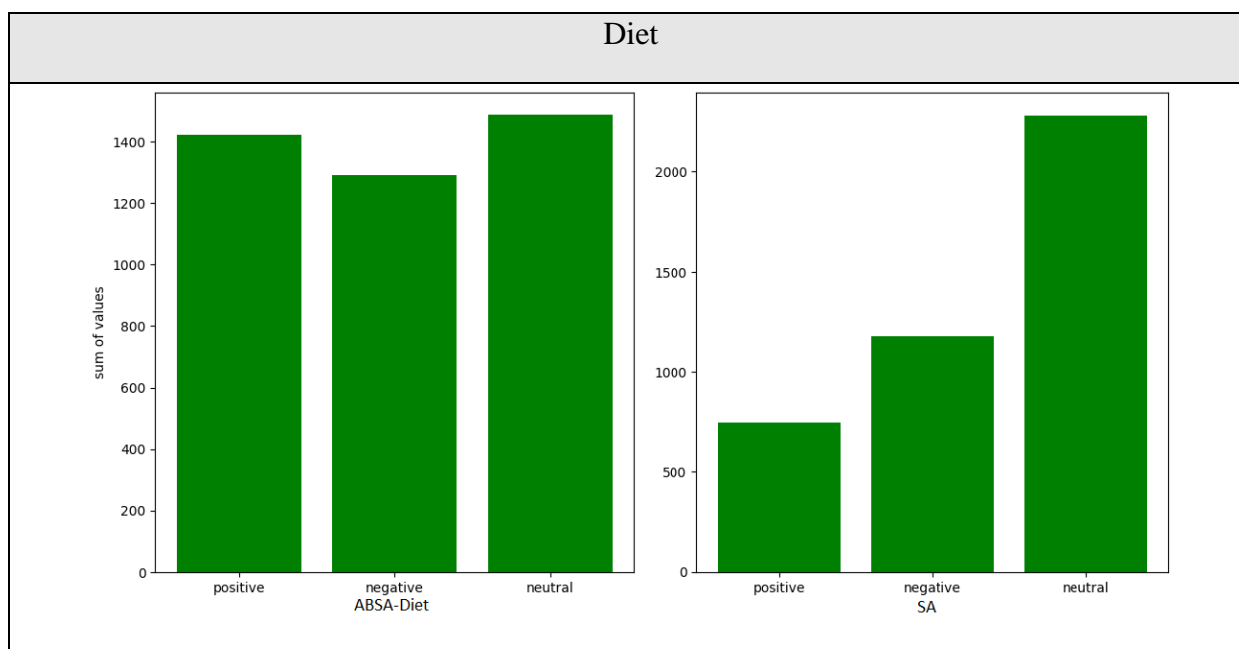
جدول ۱. جنبه های انتخاب شده مرتبط با موضوع گیاهخواری

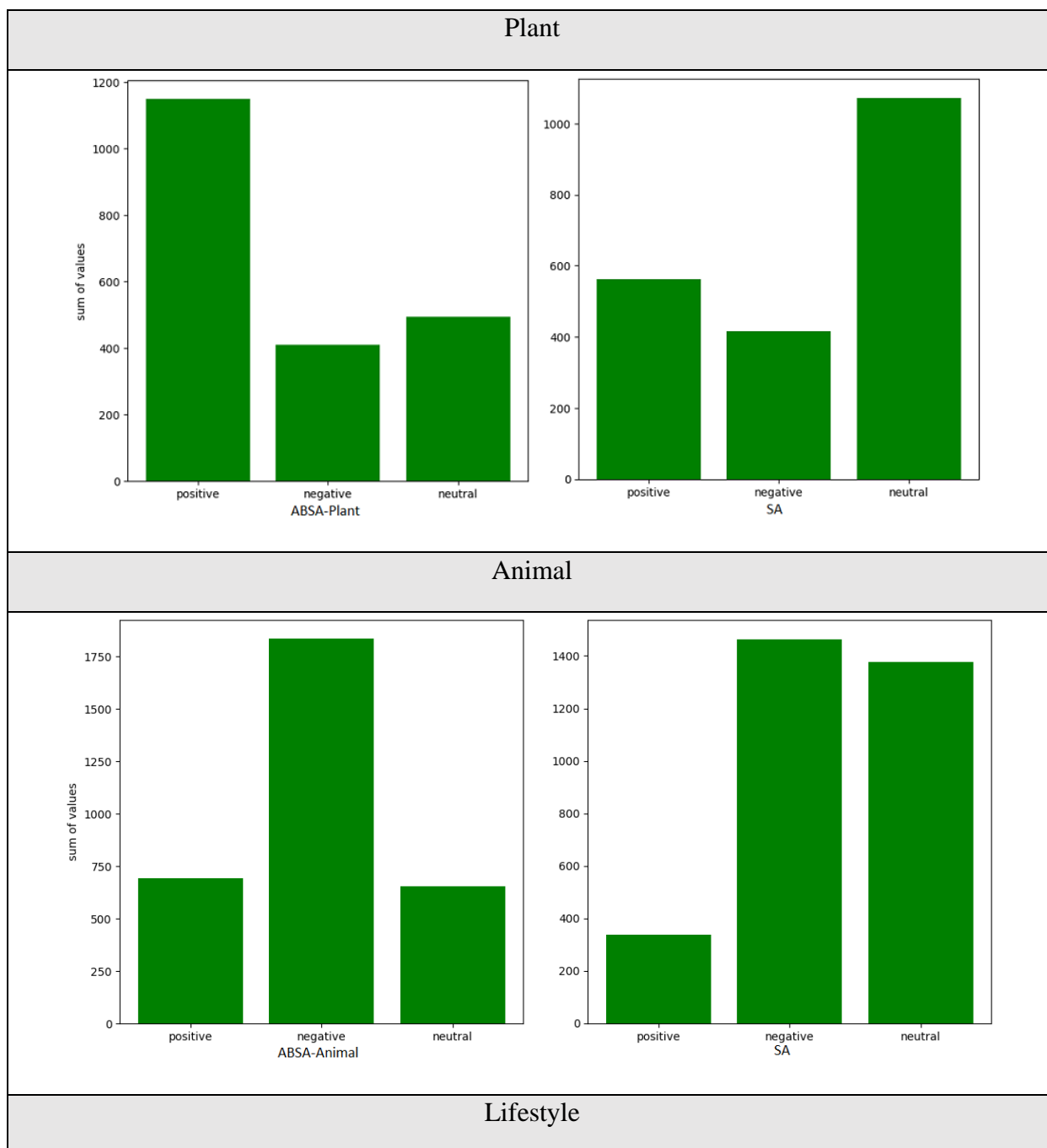
diet	animal	co
plant	lifestyle	

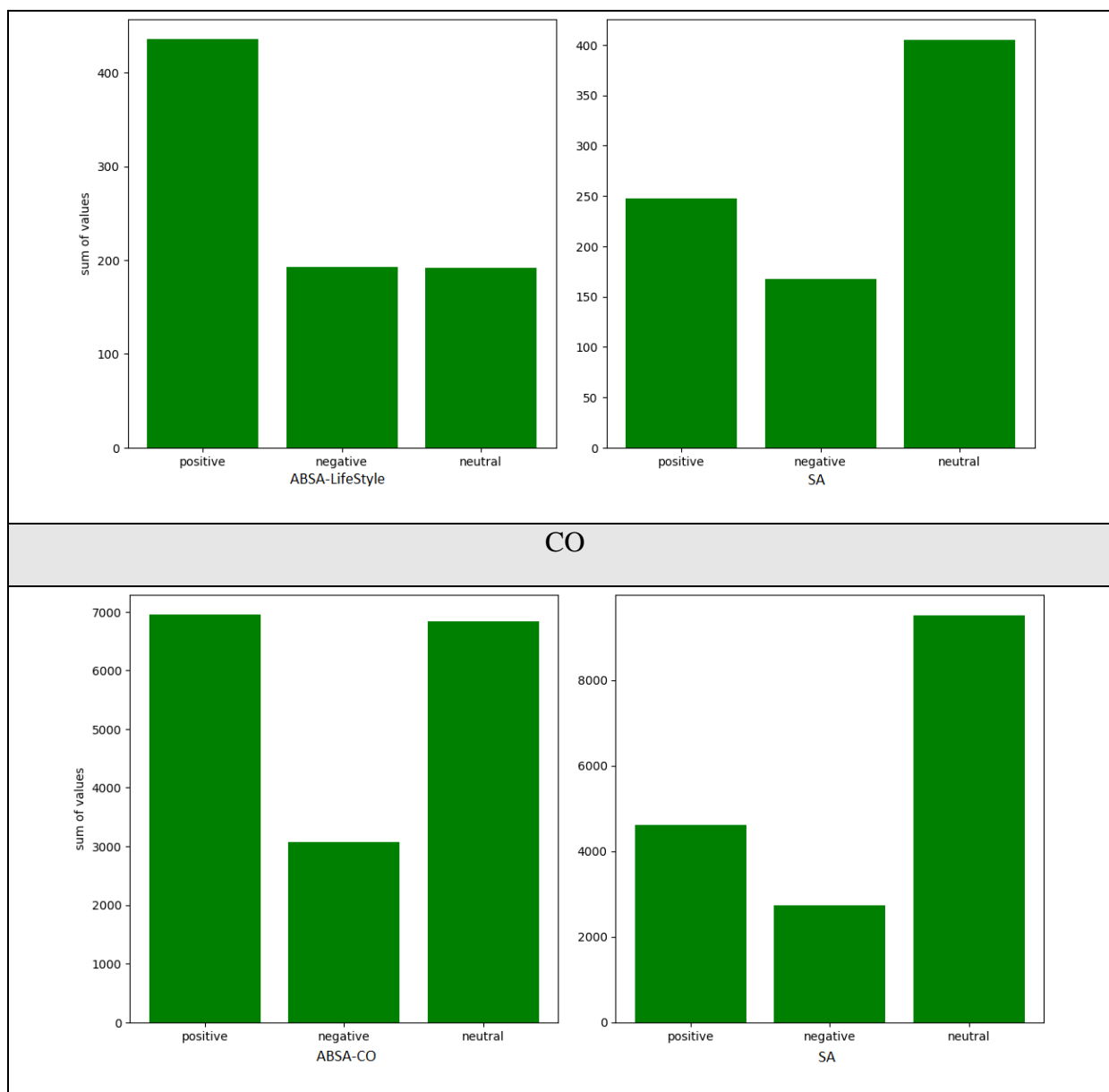
در گام بعدی، روشی مبتنی بر مدل زبانی DeBERTa (He et al., 2020) برای تحلیل احساسات مبتنی بر جنبه (Yang et al., 2021) را به کار گرفتیم. برای هر یک از جنبه های ذکر شده در جدول ۱، احساسات نهفته در توئیت ها را استخراج کردیم. و نمودارهای حاصل از محاسبه ی میزان احساسات مختلف را برای هر جنبه به صورت مجزا در ادامه آورده ایم.

در ابتدا مجموع وزنی احساسات مختلف مثبت، منفی و یا خنثی را، با توجه به نتایج حاصل از تحلیل احساسات مبتنی بر جنبه برای هر یک از جنبه های در نظر گرفته شده محاسبه کردیم که در نمودار ۱۰ قابل مشاهده است. سپس مجموع وزنی نتایج حاصل از تحلیل احساسات مبتنی بر RoBERTa را برای این مجموعه توئیت به ازای هر یک از جهتگیری های احساسی مثبت، منفی و خنثی به نمایش درآوردیم. که در نمودار ۱ مشهود است. همان طور که در نمودار ۱ برای هر یک از جنبه ها مشاهده می شود، تفاوت مجموع وزنی احساسات مختلف برای کل توئیت، با مجموع وزنی احساسات نسبت به جنبه های ذکر شده قابل درک است. این نشان دهنده ی آن است که کاربران با وجود آنکه در جمله ای جهتگیری احساسی ای را پیش گرفته اند اما نسبت به جنبه های ذکر شده در کلامشان ممکن است احساس متفاوتی داشته باشند.

از نمودار ۱ پیداست که نظرات قریب به اتفاق کاربران نسبت به جنبه های "سبک زندگی" و "گیاه" سوگیری مثبت دارد. اما در مقابل، برای موضوع "حیوان" با شدت بیشتر نسبت به سایر احساسات سوگیری منفی دارند. همچنین برای موضوعات "رژیم" و "شرکت" با میزان تقریباً یکسان، اکثر مجموع وزنی مثبت و خنثی در مجموع توئیت ها وجود دارد.





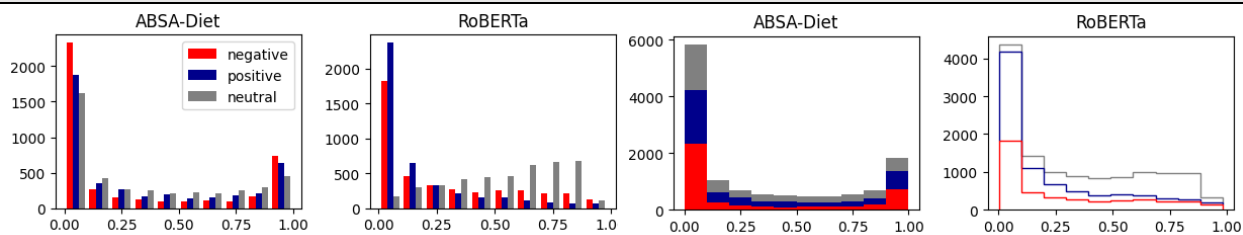


نمودار ۱. مجموع وزنی نتایج حاصل از SA و ABSA برای هر جنبه

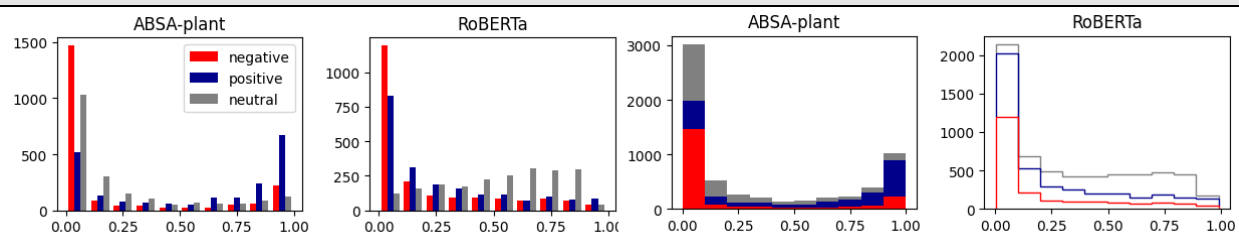
در گام بعد توزیع مقادیر احساسات مختلف در تحلیل احساسات مبتنی بر جنبه (ABSA) و همچنین توزیع احساسات مختلف در تحلیل احساسات کلی توثیت ها را در قالب دو دسته نمودار ۲ به تصویر کشیدیم. این مجموعه نمودار مقایسه ای بین توزیع مقادیر حاصل از دو تکنیک <sup>۱۰</sup>SA و ABSA برای هر یک از جنبه های بیان شده در جدول ۱ می باشد.

<sup>10</sup> Sentiment analysis

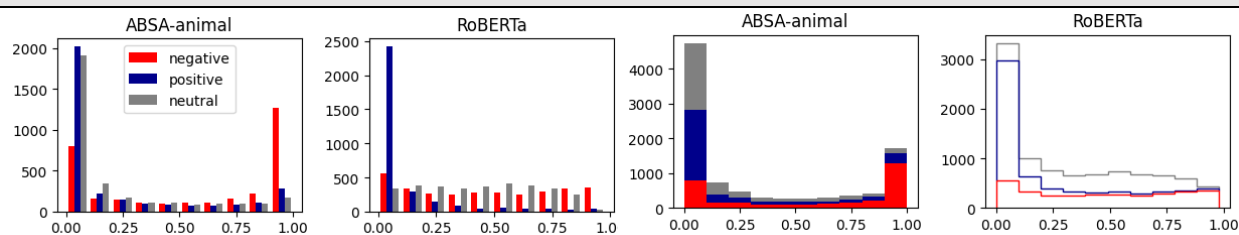
## Diet



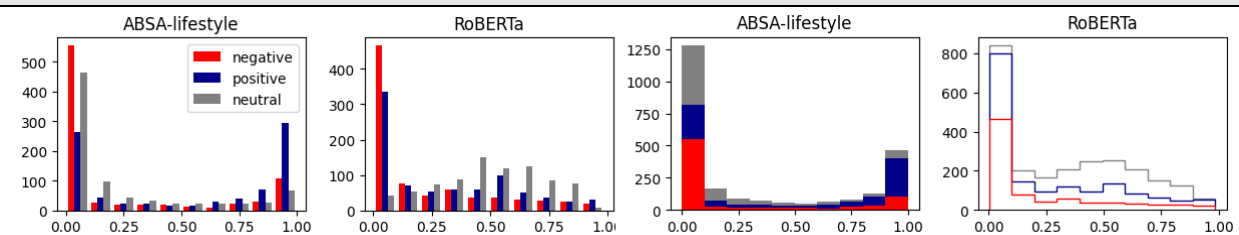
## Plant



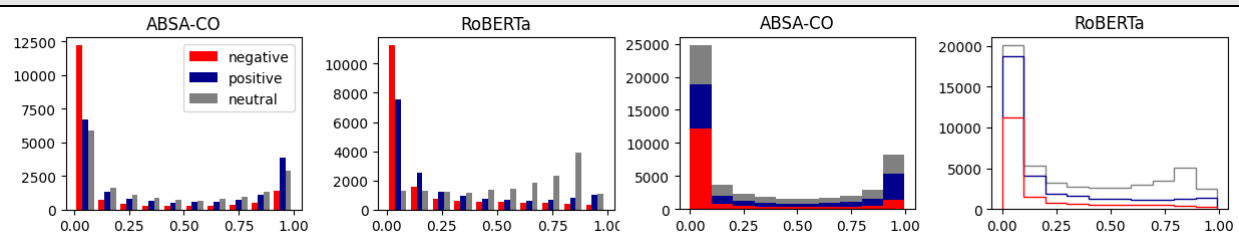
## Animal



## Lifestyle

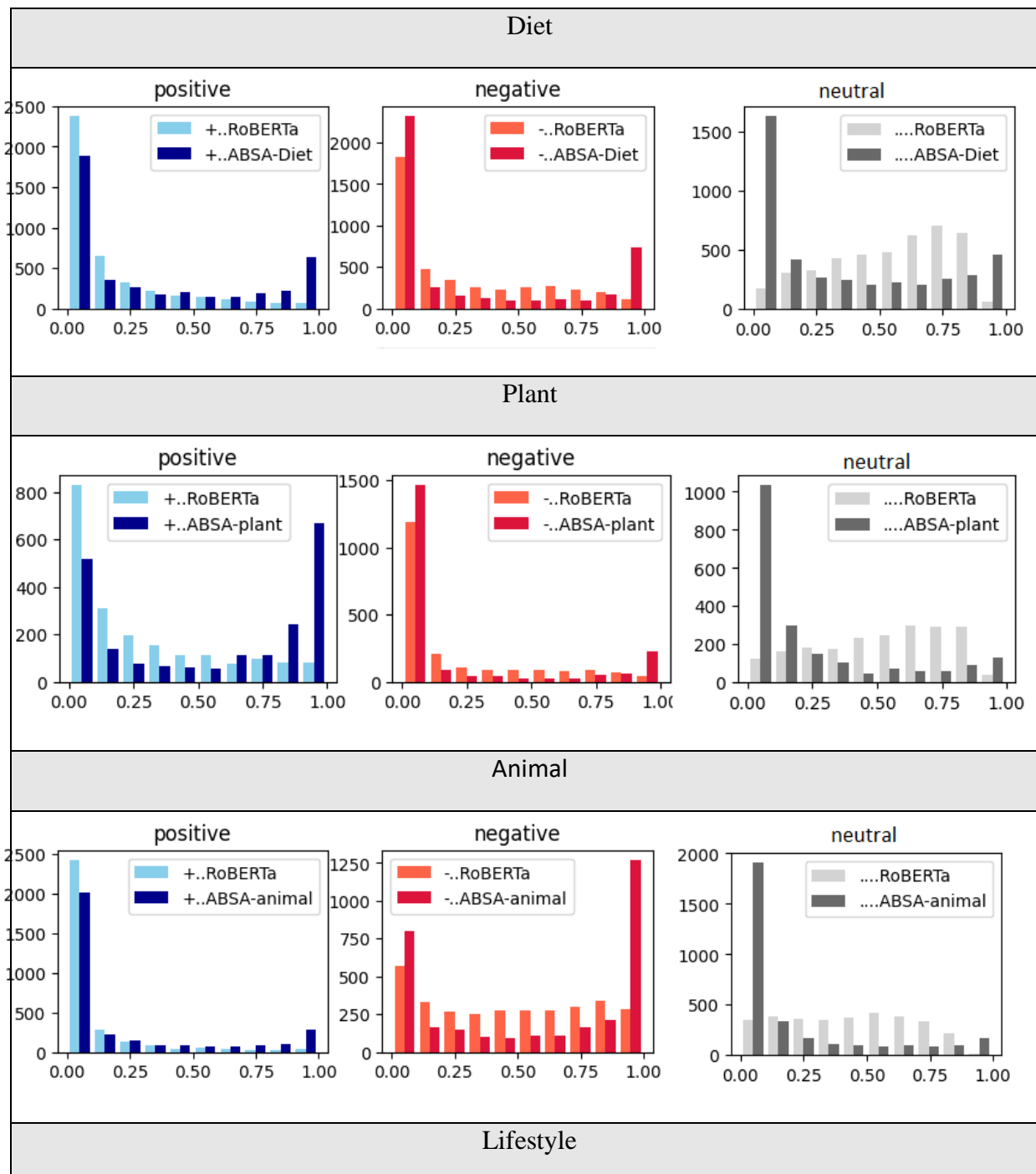


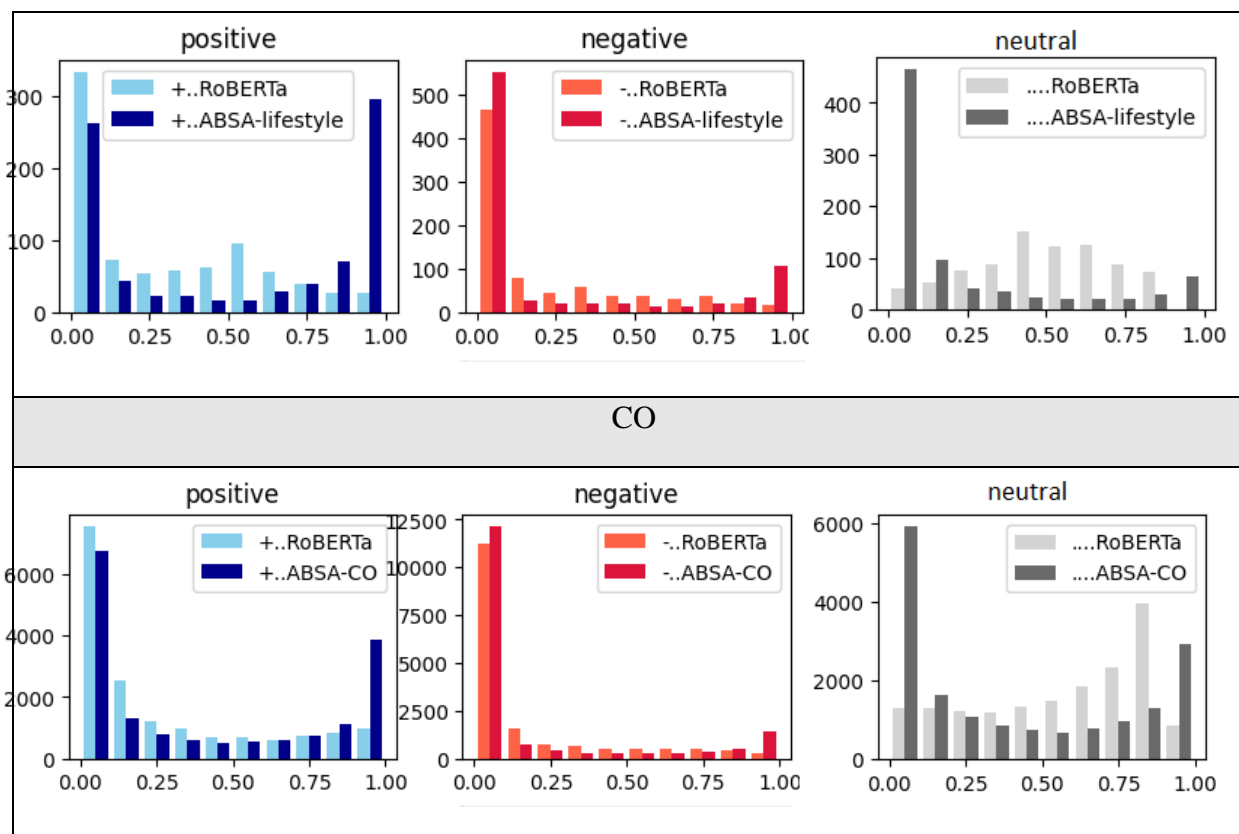
## CO



نمودار ۲. توزیع مقادیر حاصل از SA و ABSA

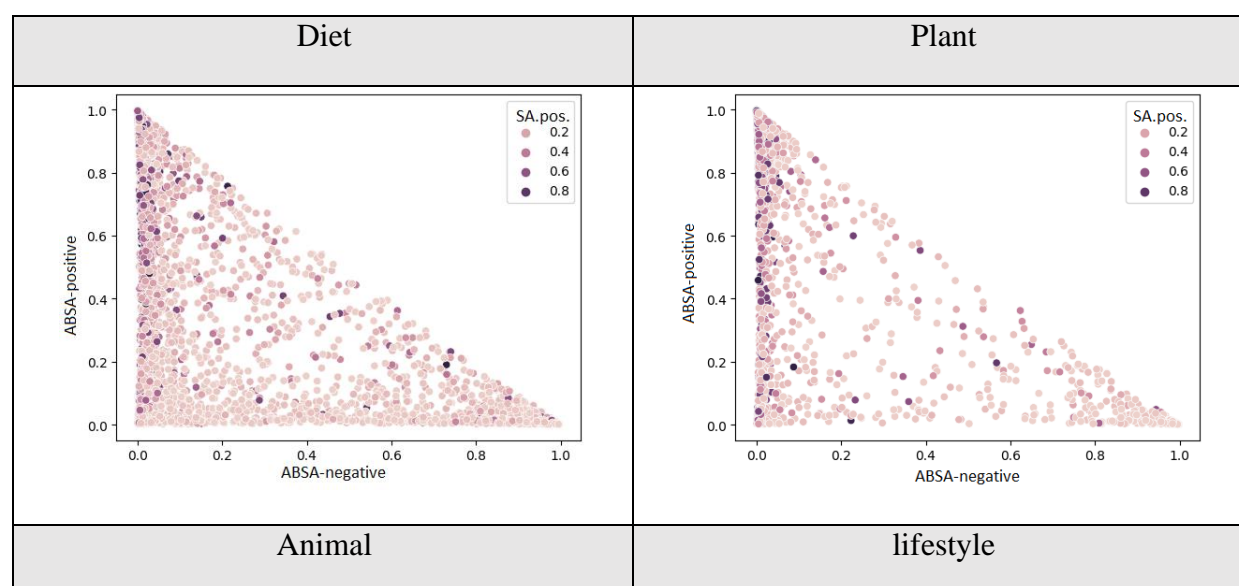
در مرحله ی بعد، برای هر یک از جنبه ها، به تفکیک سه جهت گیری احساسی مثبت، منفی و خنثی، نمودارهای مقایسه ای را بین توزیع مقادیر حاصل از تحلیل احساسات مبتنی بر RoBERTa بر روی هر مجموعه تئیت و نتایج حاصل از تحلیل احساسات مبتنی بر جنبه مبتنی بر DeBERTa، در نمودار ۳ به تصویر کشیدیم.

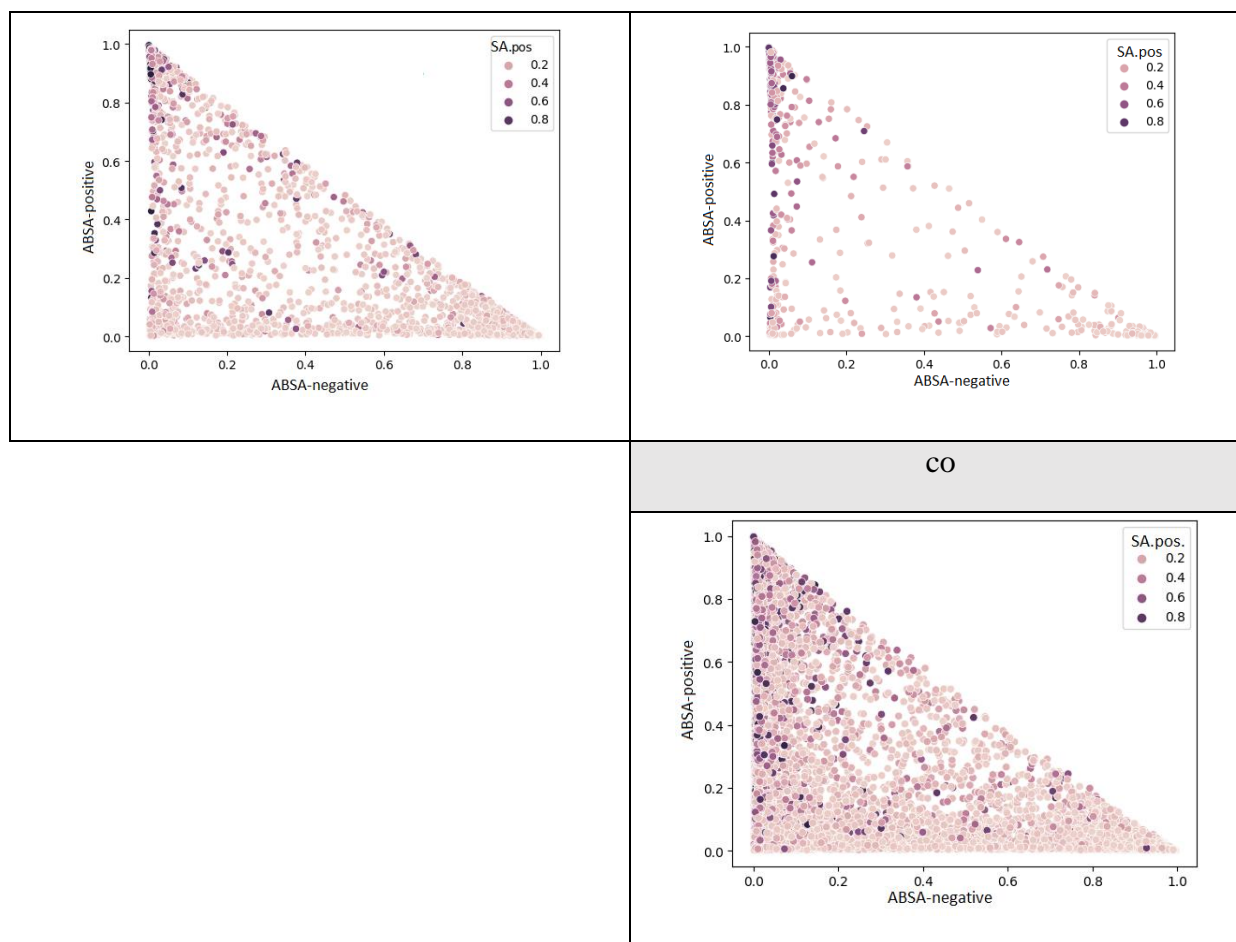




نمودار ۳. توزیع مقادیر حاصل از SA و ABSA به تفکیک هر احساس در هر جنبه

برای درک چگونگی توزیع احساسات افراد نسبت به هر جنبه، نمودارهای پراکندگی مقادیر محاسبه شده در تحلیل احساسات مبتنی بر جنبه را بر حسب مقدار منفی محاسبه شده (در محور X) و مقدار مثبت محاسبه شده (در محور Y) برای هر توثیت در هر دسته، در نمودار ۴ ترسیم کردیم. رنگ هر نقطه در این نمودارها نشان دهنده ی میزان مثبت بودن کلیت آن توثیت با استفاده از روش مبتنی بر RoBERTa می باشد.

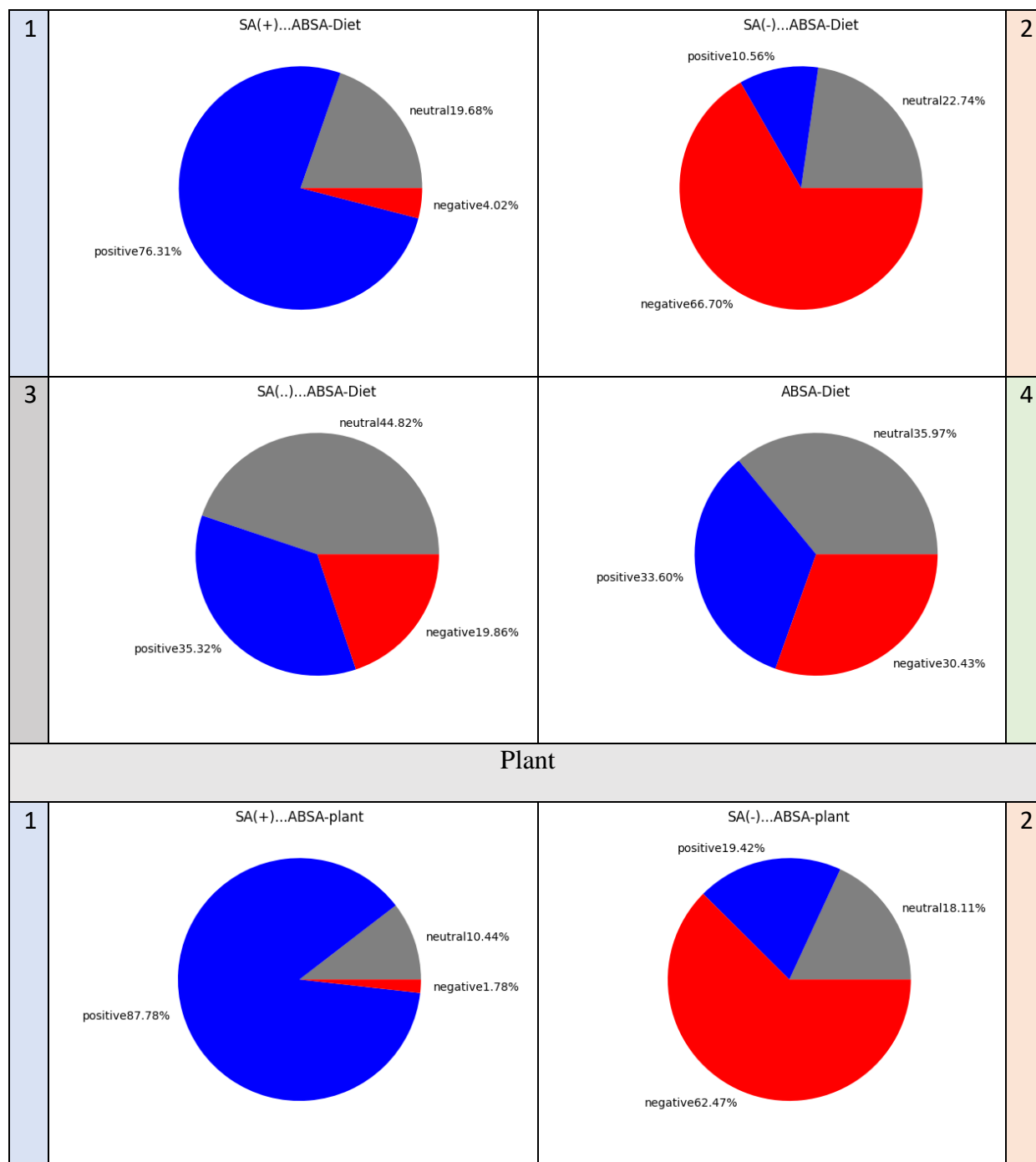


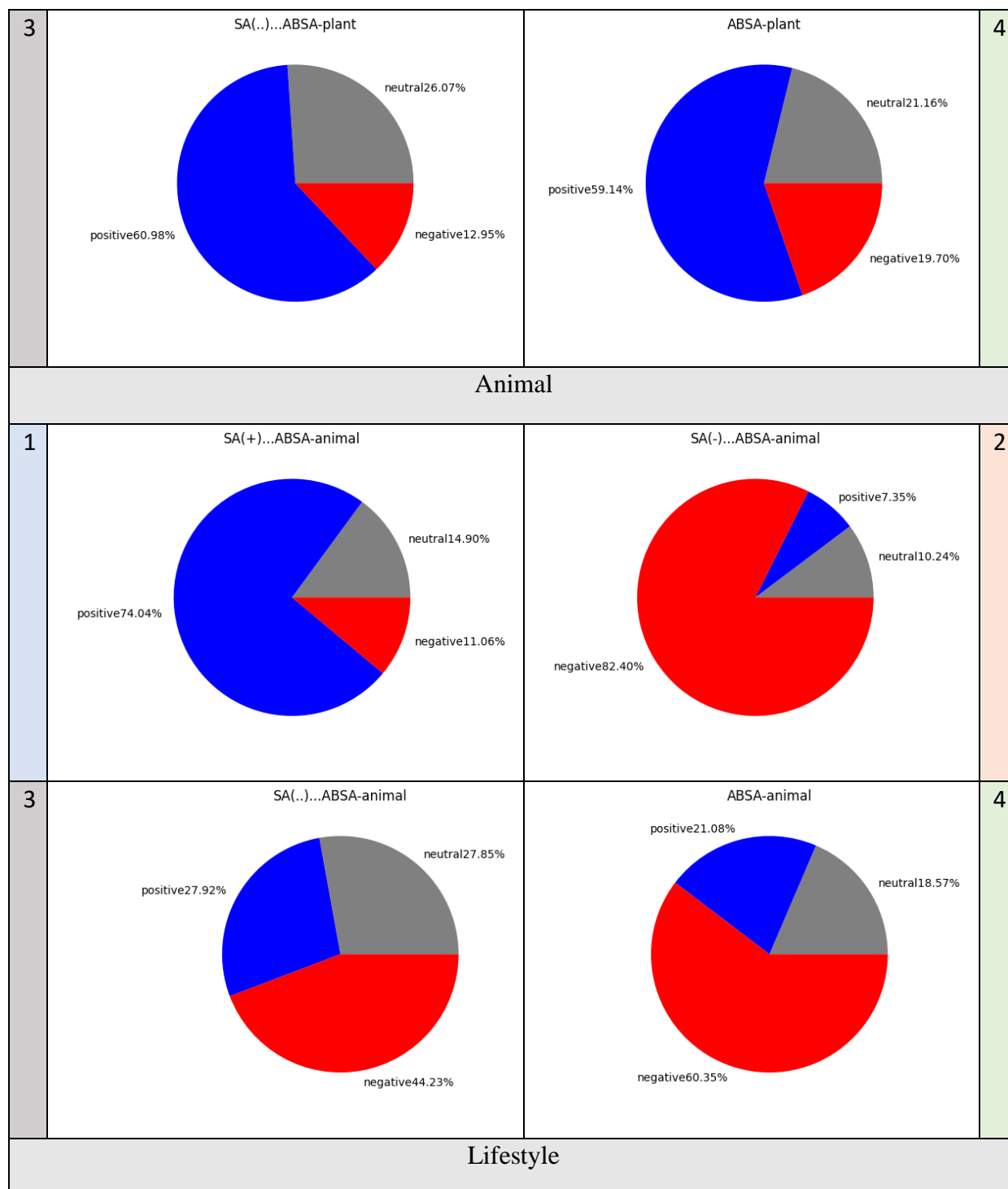


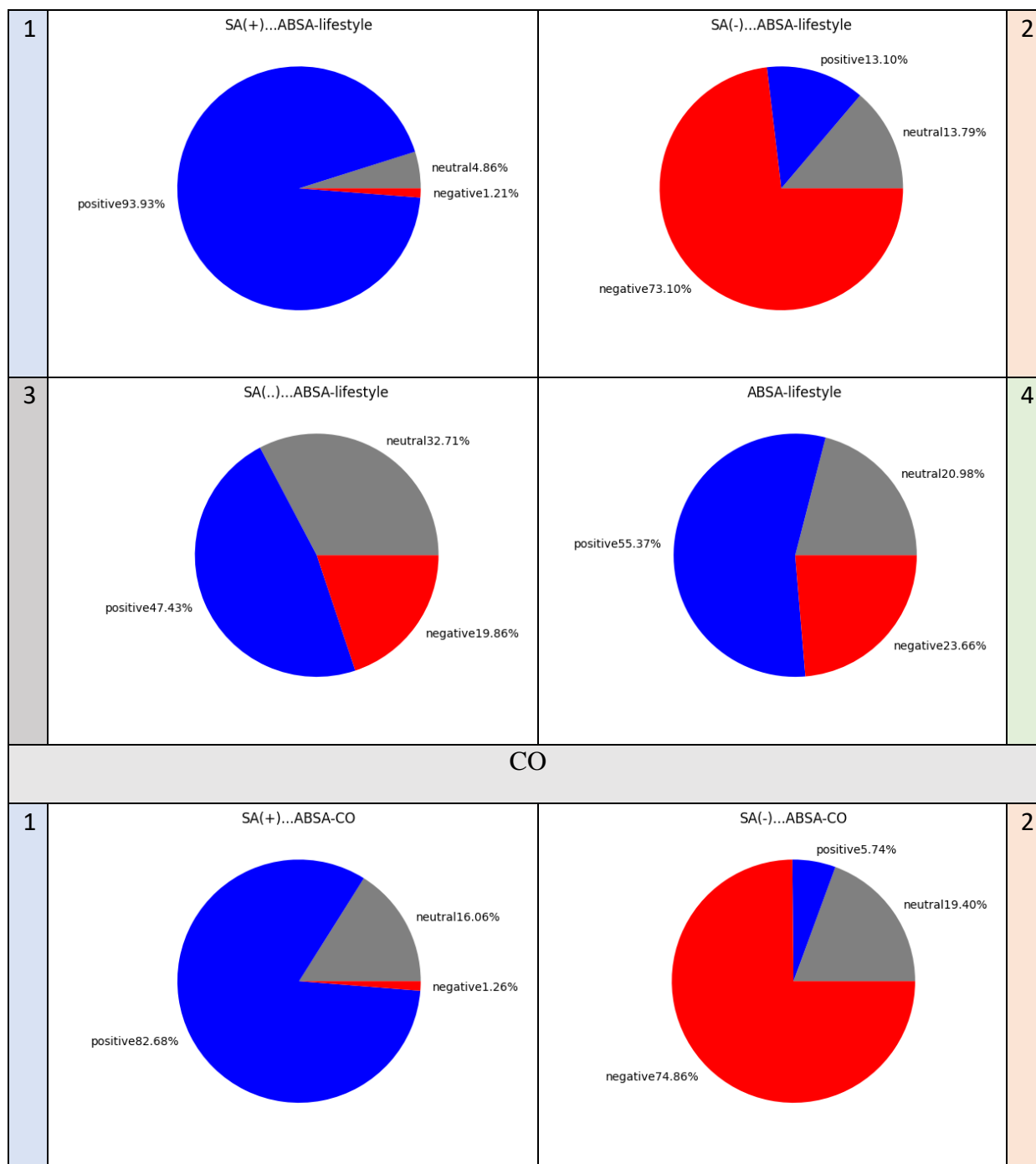
نمودار ۴. توزیع مقادیر منفی بر حسب مقادیر مثبت حاصل از ABSA

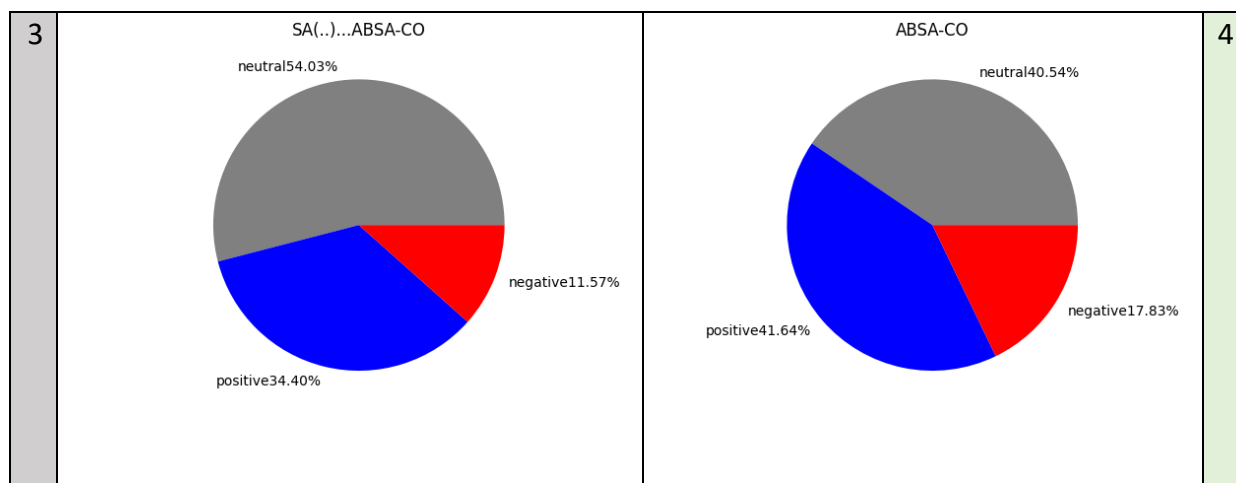
در مرحله ی بعد برای هر توثیت در هر دسته، احساس غالب آن توثیت را تشخیص دادیم. ( احساس غالب برای هر توثیت احساسی است که مقدارش از مقدار دو احساس دیگر برای آن توثیت بیشتر باشد). برای توثیت هایی که احساس غالبشان با استفاده از روش مبتنی بر RoBERTa، مثبت بوده است. فراوانی احساسات مبتنی بر جنبه را محاسبه کرده (برای هر توثیت نسبت به هر جنبه، احساس غالب را احساس اصلی در نظر گرفته و به شماره آوردیم) و در نمودارهای ۱-۵ به تصویر درآوردیم. در ادامه برای هر جنبه، در میان توثیت هایی که احساس غالبشان، منفی می باشد، درصد فراوانی احساسات غالب مبتنی بر جنبه را در نمودارهای ۲-۵ ترسیم کردیم. برای توثیت های خنثی نیز همین عمل را در هر یک از جنبه ها انجام دادیم و نمودارهای ۳-۵ به وجود آمدند. نمودار های ۴-۵ درصد فراوانی احساسات مختلف، بر حسب مقادیر محاسبه شده تحلیل احساسات مبتنی بر جنبه در هر دسته توثیت می باشد. نمودار های ۱، ۲ و ۳ در دسته نمودارهای ۴-۵ نشان دهنده ی تمایز جهتگیری احساسی کلی هر توثیت با جهتگیری احساسی هر توثیت نسبت به جنبه های معین شده، می باشد. همچنین نمودارهای ۴ در دسته نمودار ۵ نشان دهنده ی درصد فراوانی احساسات مختلف افراد نسبت به هر یک از جنبه ها می باشد.

Diet









نمودار ۵. درصد فراوانی احساسات مختلف در توئیت ها برای هر جنبه در چهار دسته (۱- بین توئیت های مثبت (۲) بین توئیت های منفی (۳) بین توئیت های خنثی (۴) کلیه ی توئیت ها

## ۵- بحث و نتیجه گیری

این پژوهش با هدف شناسایی احساسات مختلف افراد نسبت به جنبه های متفاوت از جریان گیاهخواری در چهار ماه اول سال ۲۰۲۳ با استفاده از مطالب منتشر شده توسط کاربران در این مدت، انجام پذیرفت. در این مطالعه ضمن مدلسازی موضوعی، روشی برای تحلیل احساسات کاربران نسبت به جنبه های مختلف از جریان گیاهخواری ارائه داده است. از روش مبتنی بر مدل **ROBERTa**، برای تحلیل سوگیری احساسی توئیت های مرتبط با جریان گیاهخواری استفاده کردیم. سپس با استفاده از مدل مبتنی بر **DeBERTa** برای تحلیل احساسات مبتنی بر جنبه، احساسات کاربران نسبت به جنبه های گزینش شده از مرحله ی مدلسازی موضوعی را استخراج کردیم. با نمودار های مختلف توزیع و فراوانی، سوگیری عمومی احساسی مختلف کاربران و همچنین جهتگیری احساسی آنان نسبت به جنبه های مختلف را به تصویر کشیدیم. برای پژوهش های آینده، بهتر آن است که از سایر منابع عقاید کاربران مانند انواع دیگر شبکه های اجتماعی بهره ببریم، تا دید وسیع تری نسبت به این موضوع به دست آوریم. همچنین، برای درک بهتر روند گیاهخواری می توان، احساسات افراد را نسبت به جنبه هایی دیگر از گیاهخواری نیز سنجید. برای سنجش دقیق تر احساسات کلی افراد در هر توئیت می توان از مدل های مبتنی بر **ROBERTa** که بر اساس داده های توئیت آموزش دیده اند استفاده کرد.

## ۶-منابع

- Zainuddin, N., Selamat, A., & Ibrahim, R. (2020). Discovering hate sentiment within twitter data through aspect-based sentiment analysis. In *Journal of physics: conference series* (Vol. 1447, No. 1, p. 012056). IOP Publishing.
- Ding, H., Huang, S., Jin, W., Shan, Y., & Yu, H. (2022). A novel cascade model for end-to-end aspect-based social comment sentiment analysis. *Electronics*, 11(12), 1810.
- Jang, H., Rempel, E., Roth, D., Carenini, G., & Janjua, N. Z. (2021). Tracking COVID-19 discourse on twitter in North America: Infodemiology study using topic modeling and aspect-based sentiment analysis. *Journal of medical Internet research*, 23(2), e25431.
- Schouten, K., & Frasincar, F. (2015). Survey on aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*, 28(3), 813-830.
- Banjar, A., Ahmed, Z., Daud, A., Abbasi, R. A., & Dawood, H. (2021). Aspect-based sentiment analysis for polarity estimation of customer reviews on Twitter. *Computers, Materials & Continua*, 67(2), 2203-2225.
- Ahmed, S., Haman, S., Atwell, E., & Ahmed, F. (2017). Aspect based sentiment analysis framework using data from social media network. *IJCSNS Int. J. Comput. Sci. Netw. Secur*, 17, 100-105.
- Silva, E. H. D., & Marcacini, R. M. (2021). Aspect-based sentiment analysis using BERT with disentangled attention. In *Proceedings*.
- He, P., Liu, X., Gao, J., & Chen, W. (2020). DeBERTa: Decoding-enhanced bert with disentangled attention. *arXiv preprint arXiv:2006.03654*.
- Yang, H., Zeng, B., Xu, M., & Wang, T. (2021). Back to Reality: Leveraging Pattern-driven Modeling to Enable Affordable Sentiment Dependency Learning. *arXiv preprint arXiv:2110.08604*.
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.
- Scaria, K., Gupta, H., Sawant, S. A., Mishra, S., & Baral, C. (2023). InstructABSA: Instruction Learning for Aspect Based Sentiment Analysis. *arXiv preprint arXiv:2302.08624*.
- Nezhad, Z. B., & Deihimi, M. A. (2022). Twitter sentiment analysis from Iran about COVID 19 vaccine. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 16(1), 102367.
- Prakash, Y., & Sharma, D. K. (2023, March). A Survey on Aspect-based Sentiment Analysis for Online Purchased Product Feedback. In *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1342-1347). IEEE.
- Sivakumar, M., & Reddy, U. S. (2017, November). Aspect based sentiment analysis of students opinion using machine learning techniques. In *2017 international conference on inventive computing and informatics (ICICI)* (pp. 726-731). IEEE.
- Chang, Y. C., Ku, C. H., & Le Nguyen, D. D. (2022). Predicting aspect-based sentiment using deep learning and information visualization: The impact of COVID-19 on the airline industry. *Information & Management*, 59(2), 103587.

## Aspect-based sentiment analysis on Twitter social network data about vegetarianism

Fatemeh Salahi

Department of Electrical and Computer Engineering, Islamic Azad University, Birjand, Iran

[Salahi.fateme@gmail.com](mailto:Salahi.fateme@gmail.com)

### Abstract

Vegetarianism is one of the trends that has received a lot of feedback on social networks. The content published by users reflects their feelings and opinions towards this trend and its various aspects. In this regard, a dataset containing more than sixty thousand tweets published in 2023 about vegetarianism was collected. This dataset was used to extract user sentiment towards different aspects of vegetarianism. First, a method based on RoBERTa language model was proposed to analyze the implicit sentiment hidden in tweets. Then, using the Latent Dirichlet Allocation topic modeling approach, some relevant aspects and topics related to vegetarianism were extracted. In the next step, a method based on DeBERTa language model was used to analyze tweet sentiment towards different aspects that had been extracted. Various frequency and sentiment distribution charts for different aspects in the field of vegetarianism were examined. The results of emotional analysis based on RoBERTa and DeBERTa models were compared side by side. Data analysis using the DeBERTa model showed that users had mostly tweeted positive sentiments regarding the plant and lifestyle aspects. However, for the Animal aspect, most tweets were negative. For both Diet and Company aspects, most tweets were positive or neutral with values close to each other. During the discussion, some implicit knowledge related to this topic was also examined.

**Keywords:** Aspect Based Sentiment Analysis, Topic Modeling, Natural Language Processing, Text Processing