

بررسی تطبیقی متن شعارهای تبلیغاتی فروشگاه‌های اینترنتی صنایع غذایی داخلی و خارجی با استفاده از تکنیک متن کاوی و خوشه‌بندی

تاریخ دریافت: ۹۷/۰۷/۲۸

تاریخ پذیرش: ۹۷/۰۹/۱۳

نوشته

سبا هدایتی فر *

علیرضا عباسی گرجی **

چکیده

از آنجایی که امروزه اکثر کسب و کارها، برای مدیریت کارآمدتر و ساده‌تر مشتریانشان از فروشگاه‌های اینترنتی برای عرضه محصولات خود استفاده می‌کنند، تحلیل متن شعارهای تبلیغاتی آنها که به نوعی تجلی بیانیه مأموریتشان نیز هست، اهمیت بسزایی دارد. در این پژوهش، متن شعارهای تبلیغاتی شرکت‌های برتر ایرانی و خارجی، شامل ۲۹ شعار جذاب فروشگاه‌های اینترنتی داخلی و ۲۴ شعار جذاب فروشگاه‌های اینترنتی خارجی مورد بررسی قرار گرفتند. جهت پیش‌پردازش متون و ترکیب آن با روش خوشه‌بندی و الگوریتم میانگین K-means از روش متن کاوی استفاده شده، همچنین برای فرارگیری در خوشه‌های جدا و تشخیص مهم‌ترین ریشه‌ها در هر خوشه، تکنیک تصمیم‌گیری چند شاخصه MADM به کار گرفته شده است؛ تا علاوه بر مشخص کردن محتوای شعارها، تفاوت‌ها و شباهت‌های بین این دو جهت‌گیری نیز آشکار شود. با تطبیق خوشه‌های داخلی و خارجی معلوم شد که توجه شرکت‌های داخلی بیشتر بر مفاهیمی چون "محصول" و "کلی‌گویی" است، اما شرکت‌های خارجی بر مفاهیمی "همچون توجه به مشتری" متمرکزند.

کلیدواژه: شعارهای تبلیغاتی، فروشگاه‌های اینترنتی، متن کاوی، خوشه‌بندی، میانگین K.

مقدمه

طبق نظرسنجی‌های گسترده‌ای که در مؤسسه‌های ComScore و UPS، انجام شده خرید آنلاین به میزان قابل توجهی از خرید حضوری پیشی گرفته است. این سبقت فروشگاه‌های اینترنتی، همزمان با روی گرداندن از کسب و کارهای فیزیکی، تغییراتی را در نظام اقتصادی و تجاری مانند: یکپارچگی قیمت‌ها، که با آن مشتریان می‌توانند قیمت محصول را در فروشگاه‌های مختلف مقایسه کنند؛ امکان دسترسی به خدمات و محصولات در تمام ساعات شبانه روز؛ حل مشکلات حمل و نقل؛ امکان مقایسه جزئیات محصولات مختلف و عدم وجود مشکل فضا و ... به وجود آورده است. بنابراین رویکرد افراد به خرید از فروشگاه‌های اینترنتی رشد بسیاری داشته است، لذا با توجه به بازار رقابتی بین فروشگاه‌ها، توجه به شعارهای تبلیغاتی از اهمیت ویژه‌ای برخوردار است.

نخستین بار فلدمن و همکارانش متن کاوی را مورد استفاده قرار دادند (شهرابی، ۱۳۹۲). متن کاوی، علمی میان رشته‌ای و شامل آمار، یادگیری ماشین و زبان‌شناسی است. در فرایند کاوش متن، تحلیل‌ها از نظر زبان‌شناسی، معنای متن و نیز ساختار متن انجام می‌شوند. این فرایند، زوایای پنهانی از متن را آشکار می‌کند (Al-Azmi, 2013). خوشه‌بندی نیز به معنای دسته‌بندی مجموعه‌ای از اشیاء است که در زمینه‌های مختلف، به‌خصوص در داده‌کاوی، پردازش زبان‌های طبیعی نظیر بازیابی اطلاعات و خلاصه‌سازی چندمتنی خود، کاربرد دارد. (ایمانی، ۱۳۹۱)

با در دست داشتن چنین ابزارهایی، قادریم شعارهای تبلیغاتی فروشگاه‌های اینترنتی را مقایسه و آنالیز متنی کرده و با شناسایی گروه‌های مشابه و قرار دادن آنها در یک شاخه، رویکردهای اصلی استفاده شده در این تبلیغات را شناسایی کنیم. قرار دادن حجم انبوهی از داده‌ها و دسته‌بندی جهت‌گیری‌های انبوهی از شرکت‌ها در تعداد کمی از خوشه‌ها، باعث می‌شود بررسی آنها راحت‌تر انجام گیرد. در این پژوهش سعی بر این است تا مفاد اصلی شعارهای تبلیغاتی فروشگاه‌های اینترنتی، در داخل و خارج از ایران بررسی و همچنین شباهت‌ها و تفاوت‌های آنها در نظر گرفته شود. از جمله نوآوری‌های این پژوهش می‌توان به استفاده از تکنیک متن کاوی در بررسی شعارهای تبلیغاتی، که تا به حال مشابه آن در ایران انجام نگرفته اشاره کرد؛ همچنین استفاده از تکنیک خوشه‌بندی جهت گروه‌بندی اسناد، نیز به نوبه خود نگرشی جدید است که در جهت انتخاب مهم‌ترین ریشه به بهبود نتایج کمک شایانی می‌کند.

اهداف و سؤال‌های تحقیق

هدف اصلی در این پژوهش، خوشه‌بندی متن شعارهای تبلیغاتی در ایران و خارج از ایران است. همچنین در ادامه به دنبال بررسی و مقایسه تطبیقی خوشه‌های حاصل از متن شعارهای تبلیغاتی فروشگاه‌های اینترنتی هستیم. بنابراین، پژوهش به دنبال پاسخ به این پرسش است که خوشه‌بندی فروشگاه‌های اینترنتی در ایران و خارج از ایران، بر اساس شعارهای تبلیغاتی به چه صورت خواهد

بود؟ همچنین به بیان شباهت‌ها و تفاوت‌های شعارهای تبلیغاتی در حوزه صنایع غذایی در ایران با سایر نقاط دنیا می‌پردازد.

چارچوب نظری

فروشگاه‌های اینترنتی، پل ارتباطی بین شما و مشتریان از طریق شبکه اینترنت هستند. فروشگاه‌های اینترنتی مزایای بسیاری نسبت به خرید حضوری دارند از جمله خدمات بهتری به مشتریان خود ارائه می‌کنند و می‌توانند بیشتر در کنار مشتریان خود باشند، حتی سوال‌های از پیش پاسخ داده شده‌ای برای آنان در نظر گرفته و در صورت لزوم به صفحه مورد نظر، در قالب فروم (forum) هدایت می‌شوند. دسترسی در تمام طول شبانه روز و در هر مکانی به محصولات و خدمات آنها از دیگر مزایای فروشگاه‌های اینترنتی هستند.

دلایل بسیار زیادی برای برقراری این فروشگاه‌ها وجود دارد از جمله، به هزینه بسیار کمتری برای راه‌اندازی نیاز دارند و می‌توانند با تعداد کمتری کارمند کار کنند و فقط از طریق اینترنت با هزینه خیلی کمی مشغول به کار شوند. در واقع نیازی به خرید ملک و لوازم اداری و ... نیست. تنها لازم است سایتی راه‌اندازی و محصولات و خدمات معرفی شوند. حتی برای برقراری این سایت‌ها سکوها آماده‌ای، نظیر وردپرس وجود دارند که سکوی پرتاب خواهند بود. از دیگر مزایای فروشگاه اینترنتی می‌توان به کوتاه شدن فاصله بین مشتری و خریدار اشاره کرد که تنها با یک کلیک، نیاز محقق می‌شود. هزینه‌های عملیاتی شرکت نیز کاهش می‌یابد؛ برای مثال با ثبت آنلاین سفارش‌ها، دیگر نیازی به کارکنان بخش خدمات مشتریان نیست. با توجه به پایه‌ای بودن بحث اینترنت، در این فروشگاه‌ها سرعت پاسخ‌دهی به مشتریان بسیار بالا و امکان مقایسه دقیق محصولات و مشاهده ابعاد مختلف آنها برایشان فراهم شده است. کاهش آلودگی و ترافیک، پرداخت در محل پس از تحویل گرفتن کالا، امکان استفاده از خدمات پس از فروش آنلاین و حضوری، دسترسی به توضیحات کامل در مورد محصول، چه به صورت عکس و یا ویدئو و ... از جمله مزایای فروشگاه‌های اینترنتی است که می‌توان به آن اشاره کرد.

فروشگاه‌های سنتی فقط می‌توانستند تعداد محدودی مشتری را در مکان‌های محدود بپذیرند، اما با استفاده از فروشگاه‌های اینترنتی می‌توان بازار جهانی را هدف گرفت، به‌خصوص زمانی که محصولات و خدمات با تبلیغات بهینه به جهانیان معرفی شوند و ترافیک وب‌سایت فروشگاه افزایش یابد.

مفهوم متن کاوی پیدا کردن الگو در بین متن‌های غیرساخت یافته است. برای رویارویی با حجم وسیعی از اطلاعات غیرساخت یافته سه روش بازیابی اطلاعات، استخراج اطلاعات و پردازش زبان طبیعی وجود دارد. بازیابی اطلاعات، در واقع بازیابی مستندات با توجه به نیاز کاربر و بیرون کشیدن فهرست کلمه‌های مرتبط از بین مجموعه‌ای از آنها است. پردازش زبان طبیعی فهم بهتری از بیان طبیعی رایانه‌ها است و هدف آن بهره بردن از تکنیک‌های آنالیز زبان‌شناسی برای پردازش

متون است. امکان استفاده از استخراج اطلاعات، یک مرحله پیش‌پردازش در متن‌کاوی است، که هدف آن استخراج اطلاعاتی خاص از سندهای متنی است. (ایمانی، ۱۳۹۱)

اغلب دو مرحله اصلی برای متن‌کاوی در نظر گرفته می‌شود؛ مرحله نخست: پیش‌پردازش مستندات، مرحله دوم: استخراج دانش. خروجی مرحله نخست می‌تواند دو شکل مبتنی بر سند و مبتنی بر مفهوم، داشته باشد.

شکل نخست (مبتنی بر سند) شیوه بهتری برای نمایش مستندات است؛ مثل تبدیل اسناد به یک فرمت میانی و نیمه ساخت‌یافته و یا هر نوع نمایشی که کار کردن با اسناد را راحت‌تر می‌کند. هر موجودیت در این نمایش باز هم یک سند خواهد بود.

در شکل دوم (مبتنی بر مفهوم) نمایش اسناد بهبود بخشیده می‌شوند. تمامی اطلاعات قابل فهم مثل کلیه مفاهیم و معانی موجود در سند و همچنین ارتباط مفهومی بین آنها از متن استخراج می‌شود. در این فرم با مفاهیمی که از این مستندات به دست آمده، روبه‌رو هستیم. قدم بعد، استخراج دانش از فرم‌های میانی است که تفاوت‌هایی بر اساس نحوه نمایش هر سند دارد. نمایش مبتنی بر سند، برای گروه‌بندی، طبقه‌بندی و تجسم استفاده می‌شود؛ اما نمایش مبتنی بر مفهوم برای یافتن روابط میان مفاهیم، ساختن اتوماتیک آنالوژی و ... به کار می‌رود.

یکی از فاکتورهای مهم برای اثربخشی تبلیغات، شعارهای تبلیغاتی است. با در نظر گرفتن فاکتور زمان، شعار تبلیغاتی می‌تواند گام مؤثری در جهت عرضه کالا به مشتری باشد.

یکی از مهم‌ترین عناصر آمیخته بازاریابی عنصر ترفیع است، که تبلیغات در زیرمجموعه آن قرار می‌گیرد. کارکرد اصلی آن ارتباط با بازار برای رسیدن به سه مقوله ۱. اطلاع‌رسانی در مورد ویژگی‌های محصول و معرفی آن؛ ۲. تشویق مشتریان به خرید محصول و یا استفاده از خدمات؛ ۳. یادآوری مداوم محصولات و خدمات به بازار است. (Katler & Keller, 2012) شعار تبلیغاتی، بخشی کلیدی در مبارزه‌های تبلیغاتی بین شرکت‌ها و هدف اصلی آن ذخیره شدن در ذهن مشتریان است؛ به‌وضوح عبارتی است که با هدف تثبیت در حافظه مدام تکرار می‌شود (Wells et al., 1989) برای شناخت نام تجاری، دو عامل "تصویر نام تجاری" و "آگاهی از نام تجاری" بسیار مهم هستند. معروف‌ترین شاخص‌های آگاهی از نام تجاری، به خاطر آوردن و شناخت آن است. به خاطر آوردن نام تجاری، در واقع همان به خاطر آوردن نام تجاری بدون هیچ گونه تلاش ذهنی است. اما شناخت نام تجاری، نیازمند کمی تلاش ذهنی است. در کل، نقش شعارهای تبلیغاتی، ارتقای آگاهی از نام تجاری و تصویر نام تجاری و در نهایت شناخت نام تجاری است. (Kohli et al., 2007)

برای یک شعار تبلیغاتی صفات مختلفی بیان شده است؛ برای مثال: فوستر (۲۰۰۱) ده شاخص یک شعار تبلیغاتی مناسب را، به خاطر آوردن، به یادآورنده نام تجاری، بیانگر یک مزیت کلیدی، متمایزکننده نام تجاری، منعکس‌کننده شخصیت نام تجاری، باورپذیری، راهبردی بودن، رقابتی، اصیل و منحصر به فرد بودن (عدم استفاده توسط دیگران) معرفی می‌کند.

پیشینه موضوع در حوزه پژوهش‌های مرتبط

مطالعات فراوانی در حوزه متن‌کاوی صورت گرفته که به برخی از آنها اشاره می‌شود: در سال ۲۰۰۳ مطالعه‌ای موردی در زمینه استفاده از متن‌کاوی در تحلیل ثبت اختراعات انجام شده است. نویسندگان نشان دادند که نرم افزار Packmole، نسبت به تکنیک‌های تحلیل دستی ثبت اختراعات از مزایای بیشتری برخوردار است. (fattori et al., 2003)

یکی دیگر از کاربردهای متن‌کاوی در حوزه رسانه است و در جنبه‌های سیاسی کاربردهای فراوانی دارد. همچنین، در زمینه خوشه‌بندی، اسناد بایگانی‌شده به تعدادی خوشه و دسته بندی معنایی از پیش تعیین شده مورد استفاده قرار می‌گیرد. (Neto et al., 2000)

در تحقیقی که کاسمساپ با نام کاربردهای چندگانه داده‌کاوی، هوش تجاری و مدیریت دانش انجام داد، به کاربردهای عملی داده‌کاوی از نظر بازاریابی، کسب و کار و منابع انسانی پرداخته است. (Kasemsap, 2018)

روانو و همکاران (۲۰۱۸) پژوهشی را روی پیشرفت بیماری‌های قلبی و عروقی با استفاده از تکنیک داده‌کاوی انجام دادند و نتایج مثبتی در خصوص آمار پرونده‌های بیماران کسب کردند. داده‌کاوی همچنین در حوزه‌های دیگری نظیر کسب و کار (Ingvaldsen and Gulla, 2012)، استخراج فراداده از اسناد کتابخانه‌ای و کمک به غنی‌سازی آن (Witten et al., 2003)، در حوزه‌های علوم پزشکی (Li et al., 2012)، به جهت شناسایی دروغ در داده‌های دنیای واقعی (Fuller et al., 2011) نیز استفاده شده است.

گروهی دیگر از محققان با استفاده از تحلیل خوشه‌بندی به عنوان یک تکنیک اکتشافی، ادبیات آموزش الکترونیک را واکاوی و الگوهای موجود در آن را بر اساس کلمه‌های مشابه، ارزش صفت‌ها و قوانین گذشته واکاوی کردند. (Hang, 2012)

روش پژوهش

این پژوهش از لحاظ زمانی، مقطعی و از نظر هدف، جزء پژوهش‌های کاربردی محسوب می‌شود. متغیر این پژوهش، کیفی و نوع تحقیق توصیفی است، که با استفاده از ابزار متن‌کاوی، تکنیک‌های خوشه‌بندی و تصمیم‌گیری چندشاخصه انجام شده است. داده‌ها با روش کتابخانه‌ای گردآوری شده‌اند که ادبیات تحقیق همچنین متن شعارهای تبلیغاتی فروشگاه‌های اینترنتی در ایران و خارج از ایران را در بر می‌گیرد و از تارنمای آنها به دست آمده‌اند.

پژوهش در چند مرحله انجام شده، در ابتدا متن شعارهای تبلیغاتی فروشگاه‌های اینترنتی، در حوزه صنایع غذایی در ایران و سایر نقاط دنیا از طریق تارنماهای رسمی آنها جمع‌آوری شده است. به منظور یافتن شرکت‌های برتر صنایع غذایی در دنیا از فهرستی که مجله فورچون منتشر می‌کند، بهره بردیم. همچنین برای بررسی شرکت‌های برتر صنایع غذایی در ایران از شرکت‌های منتخب

وزارت صنعت و معدن با جست‌وجو در تارنماهای آنها استفاده شده است؛ که در نهایت ۲۴ مورد شرکت خارجی و ۲۹ مورد شرکت داخلی مورد بررسی قرار گرفتند.

با توجه به اینکه هدف، استفاده از تکنیک متن‌کاوی و خوشه‌بندی بود، لذا صرفاً از شرکت‌هایی استفاده شد که متون شعارهای تبلیغاتی آنها به زبان انگلیسی در تارنمایشان قرار داشت.

پس از گردآوری اسناد مورد نیاز، برای تجزیه و تحلیل در ورودی رپیدماینر، اسناد در یک فایل متنی به صورت جدا برای شرکت‌های داخلی و خارجی قرار داده شدند. بنابراین پایگاه داده پژوهش، به شکل پوشه‌هایی شامل اسناد مربوط به هر شرکت بودند.

در مرحله بعدی، برای انجام عملیات متن‌کاوی متون جمع‌آوری شده، لازم است که آنها را از متون غیرساخت یافته به ساختاریافته تبدیل کنیم. (Weiss et al., 2010) برای این منظور از نرم‌افزار "رپیدماینر"، که در سال ۲۰۰۴ در دانشگاه دورتموند تولید شده، بهره گرفته شده است. ویمر و پاول، در سال ۲۰۱۶ این نرم‌افزار را یکی از پنج ابزار قدرتمند داده‌کاوی معرفی کرده‌اند. پیش‌پردازش متن، ابتدا جهت دسترسی به عملگرهای مختلف مربوط به متن‌کاوی در رپیدماینر، بایستی افزونۀ مربوط به پردازش متن اضافه شود. در این پژوهش از خروجی بردار کلمه‌ها TF-IDF استفاده می‌شود؛ که بیانگر میزان اهمیت یک کلمه در متن است. همچنین با در نظر گرفتن عدم حذف کلمه‌های مهم، حداقل تکرار هر کلمه در کلیه اسناد ۸ مرتبه و حداکثر عددی به نسبت بزرگ در نظر گرفته شده، تا حتی المقدور کلمه‌های مهم باقی بمانند. در مرحله بعدی، به منظور کشف کلمه‌های موجود در متن، آن را به شکل عبارت‌ها، کلمه‌ها، علامت‌ها و یا دیگر عناصر معنادار تجزیه می‌کنند. داده‌های متنی به صورت رشته‌ای از کاراکترها دیده می‌شوند، در حالی که در بازیابی اطلاعات، کلمه‌ها مورد نیاز هستند. بنابراین به تجزیه‌کننده‌ای نیاز است که عملیات تفکیک کلمه‌ها را انجام دهد. (Weiss et al., 2010) در مرحله بعدی، برای از بین بردن حساسیت به حروف کوچک و بزرگ، تمام کاراکترها به حروف کوچک تبدیل می‌شوند. سپس کلمه‌های بی‌اثر در متن حذف می‌شوند؛ مواردی همچون حروف اضافه، حروف ربط و ... کلمه‌های بی‌اثر در نظر گرفته می‌شوند. (Verma et al., 2014)

جهت کاهش تعداد واژه‌ها و پرهیز از بررسی اقسام گوناگون آن، از عملگر ریشه‌یاب استفاده می‌شود. در واقع طی این فرایند، کلمه‌هایی که مشتق از یک ریشه هستند، فقط یکبار در نظر گرفته می‌شوند؛ برای مثال Production, Product, Producing, Produce. همگی از ریشه Produc هستند که منطقی‌تر این است که به فرم ریشه خود بازگردند (شفیعی نیک‌آبادی و کرباسی خیر، ۱۳۹۶). در قسمت بعد، عبارت‌های n کلمه‌ای تشکیل خواهند شد. در برخی موارد عبارات یا واژه‌هایی در متن وجود دارند که با قرار گرفتن به دنبال هم بار معنایی مشخصی را به وجود می‌آورند، بنابراین به وسیله این فیلتر، کلمه‌ها یا ریشه‌هایی به دست می‌آید. در اینجا می‌توان پیشینه تعداد واژه‌هایی را به دنبال هم می‌آیند مشخص کرد. در این پژوهش با توجه به کلمه‌ها و

بار معنایی متن مقدار n عدد ۲ در نظر گرفته شده است. در مرحله بعدی کار، طول تکه‌های جدا شده فیلتر خواهد شد. بدین ترتیب حداقل و حداکثر طول تکه‌ها در پژوهش مشخص می‌شود. در این پژوهش حداقل طول، چهار کاراکتر و حداکثر طول، بیست و پنج کاراکتر قرار داده شده است. تمام تنظیم‌های فوق در نرم‌افزار رپیدمایز انجام شده است؛ در شکل زیر این مراحل مشاهده می‌شود:



مرحل پیش پردازش متون

در مرحله بعد ماتریس TF-IDF، که شامل دو بخش است تشکیل خواهد شد. نخستین جزء (TF) مقدار نرمال شده تکرار کلمه‌ها در متن است که با تعداد دفعاتی که یک کلمه در متن ظاهر می‌شود تقسیم بر تعداد کل کلمه‌های موجود در متن برابر است. دلیل نرمال کردن این است که یک کلمه شاید به تعداد بیشتری در متون بلند نسبت به متون کوتاه ظاهر شود.

قسمت دوم (IDF) میزان اهمیت یک کلمه را می‌سنجد. در نهایت با ضرب این دو مقدار، مقدار نهایی TF-IDF برای هر کلمه، و با کنار هم قرار دادن این مقادیر برای کلمه‌ها و اسناد مختلف، ماتریس مورد نظر به دست خواهد آمد. (WU et al., 2008)

ماتریس TF-IDF، که در مرحله قبل به دست آمده، به عنوان ورودی برای خوشه‌بندی مورد استفاده قرار می‌گیرد. در اینجا با داشتن یک ماتریس عددی می‌توان داده‌ها را با هم مقایسه کرد و هر سند را بُرداری از کلمه‌ها در نظر گرفت. خوشه‌بندی به صورتی انجام می‌گیرد که داده‌هایی با بیشترین تشابه به گونه‌ای در یک خوشه قرار گیرند که کمترین میزان شباهت را با سایر خوشه‌ها داشته باشند (Han et al., 2011). در این مرحله می‌توانیم متون مشابه شعارهای تبلیغاتی مورد بررسی را در هر خوشه داشته باشیم. از خوشه‌بندی، برای استخراج داده‌های بدون برجسب و از الگوریتم میانگین K-means برای تخصیص هر شی تنها به یک خوشه بهره

می‌گیریم؛ بنابراین هر متن تنها به یک خوشه تخصیص می‌یابد. در اینجا مقدار K -means جهت شروع حلقه‌های تکرار به حالت رندوم، انتخاب شده که نشان‌دهنده تعداد مراکز خوشه‌هاست و بدین منظور با استفاده از روش آرتور و واسیلویتسکی (Arthur & vassitvitskii, 2007) داده‌ها به نزدیک‌ترین مرکز تخصیص می‌یابند. در مرحله نهایی با تغییر جای مراکز، آنها را برابر با میانگین مقدار داده‌ها می‌کنیم. این عمل تا رسیدن به بهترین تقریب همگرا، انجام می‌شود. با توجه به این نکته که هر چه انتخاب بهتری برای مراکز صورت بگیرد، عملکرد میانگین K بالاتر می‌رود، بهتر است تعداد تکرار سیکل بالاتر رود. در این پژوهش مقدار k در بازه ۳ تا ۱۰ در نظر گرفته شده است و برای هر دو دسته ایرانی و خارجی مقدار عدد خوشه ۵ به دست آمده است؛ همچنین به منظور افزایش دقت، تعداد تکرار برای محاسبات بسیار بالا بوده است.

پس از مشخص شدن خوشه‌ها، پژوهش در پی یافتن موضوعات مختلف هر خوشه است تا بتواند اهداف کلیدی شرکت‌ها را شناسایی کند. بنابراین ریشه‌های به دست آمده با معیارهایی رتبه‌بندی شدند. این معیارها: میزان رخداد کلمه در کلیه متون، میانگین مقادیر TF-IDF در هر خوشه برای هر کلمه و مقادیر مراکز خوشه برای هر ریشه است که نرم‌افزار محاسبه کرده. استفاده از مقادیر TF-IDF به تنهایی کافی نبود، زیرا بیشتر ریشه‌هایی را نشان می‌دهند که در متن اسناد تمایز ایجاد می‌کنند (کلینبرگ، ۲۰۱۳)، بنابراین به منظور رتبه‌بندی ریشه‌ها از روش تصمیم‌گیری چندشاخصه MADM بهره گرفته شده است.

یکی از پرکاربردترین تکنیک‌ها از زیرمجموعه روش‌های مربوط به MADM تکنیک مجموع ساده وزین یا SAW است، در این روش سعی بر برآورد تابع مطلوبیتی به ازای هر گزینه است تا به گزینه‌ای با بیشترین مطلوبیت برسد. در این شیوه می‌توان با به دست آوردن وزن اهمیت شاخه‌ها به برتری گزینه‌ها دست یافت (Ismanto, 2015). در این پژوهش با فرض برابری وزن معیارها، ماتریس تصمیم‌گیری تشکیل شده و مقادیر بی‌مقیاس یا به اصطلاح "نرمالیزه" شده‌اند و در نهایت با ضرب ماتریسی وزن‌ها در سطرهای ماتریس حاصل ماتریس ستونی به دست آمده که با استفاده از مرتب‌سازی واژه‌ها بر اساس بزرگ‌ترین مقدار به کوچک‌ترین، مقدار ریشه‌ها رتبه‌بندی می‌شود. (Abdullah & Ada wiyah, 2014)

در این مرحله، نتایج حاصل از انتخاب واژه‌های مهم در هر خوشه با استفاده از تصمیم‌گیری چندمعیاره مشخص می‌شود. برای انتخاب مهم‌ترین واژه‌ها صرفاً به نتایج ریاضی حاصل اکتفا نشده و نظر چند تن از خبرگان در حوزه تبلیغات و صنایع غذایی نیز افزوده می‌شود.

مهم‌ترین ریشه‌های مربوط به شرکت‌های ایرانی و خارجی، که به تأیید خبرگان و متخصصان امر نیز رسیده و با استفاده از چهار معیار میزان رخداد کلی، رخداد در سند، میانگین TF-IDF و مقادیر مراکز خوشه‌ها، رتبه‌بندی شده‌اند و به ترتیب در جدول‌های ۱ و ۲ آمده است.

جدول ۱. مهم‌ترین ریشه‌های یافته شده در هر خوشه از شرکت‌های خارجی

خوشه‌ها	مهم‌ترین ریشه‌ها
خوشه اول	Quality-good-life-family-change-commit-competit-world
خوشه دوم	Employe-you-safe-sustain-local-ethic-more
خوشه سوم	Develop-quality-live-research-innov-nutrition-life-custom
خوشه چهارم	People-social-brand-finance-growth-response-great
خوشه پنجم	Serv-nutrit-fresh-valu-ethic-people- integ- consum-taste-good seconds-eat-more

جدول ۲. مهم‌ترین ریشه‌های یافته شده در هر خوشه از شرکت‌های ایرانی

خوشه‌ها	مهم‌ترین ریشه‌ها
خوشه اول	Taste-develop-innov-satisfact-iran-respons-fresh
خوشه دوم	Brand-product-group-world-creat-society-standard-growth-local
خوشه سوم	Establish-team-export-market-product-area-standard
خوشه چهارم	Standard-group-improve-safety-iran-produc-plan-continu-taste-quaility
خوشه پنجم	Human-leader-quality-standard-satisfact-resource-share-competit-brand

نتیجه‌گیری

شرکت‌های صنایع غذایی برتر در خارج از ایران، به ۵ خوشه و در داخل ایران نیز به ۵ خوشه تقسیم شدند. در جدول ۳، مواردی که بیشتر تأکید خوشه‌ها بر آنهاست آورده شده.

جدول ۳. جدول مقایسه بین وجوه مشترک در خوشه‌های ایرانی و خارجی

خوشه‌ها	تمرکز عمده شرکت‌های خارجی	تمرکز عمده شرکت‌های داخلی	وجه اشتراک
خوشه اول	تأکید بر کیفیت توجه به سلامت خانواده توجه به توسعه محصول اطمینان دادن به مشتری و القای حس برتری نگاه جهانی به محصول	توجه به محصول خلاقیت در ارائه محصول القای حس تازگی تأکید بر ایرانی بودن توجه به توسعه محصول	توجه به توسعه محصول
خوشه دوم	اشاره به قلمرو جهانی محصول اشاره به قلمرو فعالیت محلی تشویق به مصرف بیشتر	اشاره به قلمرو فعالیت محلی تمرکز بر معرفی محصول	اشاره به قلمرو فعالیت محلی

خوشه سوم	تأکید بر طبیعی بودن محصول تأکید بر استاندارد بودن محصول اشاره به محلی بودن محصول تأکید بر رضایت مشتری	اشاره به استاندارد بودن محصول تأکید بر سلامت محصول اشاره به استاندارد بودن محصول	اشاره به استاندارد بودن محصول
خوشه چهارم	تأکید بر عالی بودن توجه به مزه تأکید بر خدمت رسانی	تأکید بر کیفیت محصول تأکید بر مصرف مداوم تأکید بر طعم و مزه	توجه به مزه
خوشه پنجم	توجه به پیوند با لحظات خوش تأکید بر مشتری توجه به مزه تأکید بر مصرف بیشتر محصول	تأکید بر رضایت مشتری تشریح چگونگی تولید محصول	تأکید بر مشتری

با توجه به جدول‌ها، در حالت کلی مشاهده می‌شود شرکت‌های خارجی دیدگاه کلی‌تری نسبت به شرکت‌های ایرانی دارند؛ مثل بیان کلمه‌های نظیر همیشه، هر موقع، همیشگی و ...؛ که در مورد شرکت‌های ایرانی بیان کلمه‌هایی مانند: لحظه‌ها، ثانیه‌ها نشان می‌دهد که دیدگاه لحظه‌ای و آنی وجود دارد شعارهای تبلیغاتی خارجی بیشتر پیام همراهی و کمک به مصرف‌کننده دارند؛ مثل پیشنهاد یک شکلات در زمان گرسنگی که باعث می‌شود مشتری در موقع نیاز، پیام شعار را به خاطر بیاورد.

شعارهای تبلیغاتی خارجی، بیان عباراتی همچون "شکلات را تکه تکه کن و از هر قطعه آن لذت ببر" لذت مصرف را تداعی می‌کنند.

در هر دو مورد شعارهای تبلیغاتی ایرانی و خارجی، تلاش برای استفاده از عبارات‌های زیبا (استفاده از واج‌آرایی) مشهود بود. هر دو مورد تلاش می‌کردند مفهوم عشق را در بیان کلمه‌های همچون رویا، آرزو و ... برسانند. در شعارهای هر دو دسته، جنبه‌های حفاظتی در کلمه‌های همچون زندگی، امنیت، پیشرفت و ... وجود داشت.

نکته‌ای که در خوشه‌بندی‌ها وجود داشت توجه به بحث جهانی شدن بود که تمایل شرکت‌های خارجی بیشتر از شرکت‌های داخلی بود. شرکت‌های داخلی بیشتر سعی می‌کنند خدمات و محصولات خود را مطابق سلیقه مشتریانشان عرضه کنند.

شرکت‌های داخلی تأکید فراوانی بر استاندارد محصول داشتند در حالی که این تأکید در شعارهای شرکت‌های خارجی زیاد مشاهده نشد.

شعارهای تبلیغاتی خارجی، عباراتی همچون صبحانه قهرمانان نوعی احساس برتری در مصرف‌کنندگان محصولاتشان ایجاد می‌کردند و همین موجب وفادار ماندن مشتریانشان می‌شود. شعارهای تبلیغاتی غالباً عبارات کوتاهی هستند که اطلاعات تشویقی در مورد برند به مشتری منتقل می‌کنند و می‌توانند نقش مهمی در جنبه‌های بازاریابی محصول داشته باشند. شعارها، همچنین

در ارتقای ارزش ویژه برند مؤثرند؛ در واقع، آنها در درک بهتر مفهوم برند به یاری مصرف‌کننده می‌شتابند و به نوعی بیانیه مأموریت و چشم‌انداز شرکت را در قالب چند کلمه به مشتری ارائه می‌دهند.

در موارد بسیاری به مقایسه تطبیقی بین چشم‌انداز و بیانیه مأموریت شرکت‌ها پرداخته شده و نتایج بیشتر تحقیق‌ها با پژوهش حاضر همراستا است؛ برای نمونه: لین و همکاران (Line et al., 2016) شرکت‌های چینی را با شرکت‌های بین‌المللی مقایسه کرده و اذعان داشتند شرکت‌های چینی بر مواردی چون رشد، بقا و سودآوری تمرکز دارند؛ در حالی که شرکت‌های غیر چینی سعی داشتند تصویری مثبت از سازمان و هویتی شخصی ایجاد کنند.

با عنایت به اینکه شعارهای تبلیغاتی و آگهی‌های تجاری که بین برنامه‌ها پخش می‌شوند، تنها ترفندی برای سوق دادن بینندگان به خرید کالاها می‌باشند، بلکه از هر جهت بیانگر ارزش‌های فرهنگی غالب در جامعه‌اند (پاینده، ۱۳۸۵)، و مخاطبان را به داشتن الگوهای رفتاری و ارزش‌های معین ترغیب می‌کنند؛ بنابراین بایستی با توجه بیشتر به متن این شعارها، علاوه بر اطلاع‌رسانی و بیان وجوه مختلف کالاها و خدمات، سبک‌های زندگی، ارزش‌ها و باورهای یک جامعه نیز تقویت شود.

منابع

- ایمانی، م (۱۳۹۱)، خوشه‌بندی متون فارسی، پایان‌نامه کارشناسی ارشد، دانشگاه علم و صنعت، تهران.
 پاینده، حسین (۱۳۸۵)، "بررسی یک آگهی تلویزیونی از منظر مطالعات فرهنگی"، فصلنامه مطالعات فرهنگی و ارتباطات، شماره ۵.
 شفیعی نیک‌آبادی، محسن و عطا کرباسی‌خیر (۱۳۹۶)، "بررسی تطبیقی متن چشم‌انداز و بیانیه مأموریت شرکت‌های برتر و شرکت‌های ایرانی در صنایع غذایی با استفاده از تکنیک متن‌کاوی و خوشه‌بندی"، فصلنامه آینده پژوهی مدیریت، شماره ۱۱۱.
 شهرابی، جمال (۱۳۹۲)، واده‌کاوی، جهاد دانشگاهی، تهران.

Abdullah, L. & Adawiyah, C. R. (2014), "Simple additive weighting methods of multi criteria decision making and applications: A decade review", *International Journal of Information Processing and Management*, 5(1), 39.

Al-Azmi, A. A. R. (2013), "Data, text and web mining for business intelligence: a survey", *arXiv preprint arXiv:1304.3563*.

Arthur, D. & Vassilvitskii, S. (2007), "k-means The advantages of careful seeding. In Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms", *Society for Industrial and Applied Mathematics*.

Fattori, M., Pedrazzi, G. and Turra, R. (2003), "Text mining applied to patent mapping: a practical business case", *World Patent Information*, 25(4).

Foster, T. R. V. (2001), "The Art & Science of Advertising Slogan", [http://www.adslogans.co.uk/ans/adslogans_artsci-ence.pdf].

- Fuller, C., Biros, D. and Delen, D. (2011), "An investigation of data and text mining methods for real world deception detection", *Expert Systems with Applications*, 38(7).
- Han, J., Pei, J., & Kamber, M. (2011), *Data mining: concepts and techniques*, Elsevier.
<https://www.thebalancesmb.com/putting-offline-business-online-2531853>
- Hung, J (2012), "Trends of e-learning research from 2000 to 2008: use of text mining and bibliometrics", *British Journal of Educational Technology*, 43(1).
- Ingvaldsen, J.E. and Gulla, J.A. (2012), "Industrial application of semantic process mining", *Enterprise Information Systems*, 6(2).
- Ismanto, H. (2015), "Multi Attribute Decision Making Method for Determination of Regional Development Plan Priority in Papua Province based on the Indicator of Gross Regional Domestic Product (GRDP)", *International Journal of Computer Applications*, 117(7).
- Kasemsap, K. (2018), "Multifaceted Applications of Data Mining, Business Intelligence, and Knowledge Management, In Intelligent Systems: Concepts", *Methodologies, Tools, and Applications IGI Global*.
- Kelinkenberg, R. (Ed). (2013), *Rapidminer: Data mining use. Cases and business analytics applications*, Chapman and Hall.
- Kohli, C., Leuthesser, L., Suri, R. (2007), "Got Slogan? Guideline for Creating Effective Slogans", *Business Horizon*, 50.
- Kotler, P., and Keller, K.L. (2012), *Marketing Management*. 14 Edition, New Jersey: Prentice Hall.
- Liddy, E. D. (2000), "Text mining", *Bulletin of the American Society for Information Science and Technology*, 27(1).
- Lin, Q., Zhu, R., Zhang, Y., Huang, Y., Zhang Y. (2016), "Comparative analysis of mission statements in Chinese and non-Chinese companies: A study from the perspective of Linguistics", *The Journal of Business and Management, proceeding*.
- Neto, L., Santos, A., Kaeßner, C., Freitas, A. (2000), Document clustering and text summarization, In the proceeding of the 4th international conference practical Applications of knowledge Discovery and Data Mining PADD- 2000, London, UK.
- Ruano, M.G., Almeida, G.P., Palma, F., Raposo, J.F., R. T. (2018), "Ribeiro Reliability of medical databases for the use of real word data and data mining techniques for cardiovascular diseases progression in diabetic patients", *IEEE Xplore*.
- Verma, T., Renu, D. G., & Gaur, D. (2014), Tokenization and filtering process in RapidMiner, *International Journal of Applied Information Systems*, 7(2).
- Weiss, S. M., Indurkha, N., & Zhang, T. (2010), *Fundamentals of predictive text mining* London, Springer.
- Wells, W., Burnett, J., & Moriarty, S. (1989), *Advertising: Principles and Practice*, Englewood Cliffs, New Jersey, Prentice Hall.
- Witten, I., Don, K., Dewsnip, M. and Tablan, V., (2003), "Text mining in a digital library", *International Journal of Digital Library*, 5.
- Wu, H. C., Luk, R. W. P., Wong, K. F., & Kwok, K. L. (2008), "Interpreting tf-idf term weights as making relevance decisions", *ACM Transactions on Information Systems (TOIS)*, 26(3), 13.