

خوشه‌بندی تطبیقی چشم‌انداز، مأموریت و ارزش‌های شرکت‌های صنایع غذایی با رویکرد متن‌کاوی

مرتضی ملکی مین‌باش رزگانه*، امید افشاری‌زاده جعفری**

چکیده

بی‌شک درک مفاهیم تشکیل‌دهنده اسناد چشم‌انداز، مأموریت و ارزش‌های سازمانی به‌عنوان شاخص‌های کلیدی جهت‌گیری سازمان، کمک بزرگی در شناخت رویکرد سازمان‌ها به مشتریان دارد. در این پژوهش با به‌کارگیری تکنیک نوین متن‌کاوی و ترکیب آن با روش خوشه‌بندی K-means، متن سند چشم‌انداز، مأموریت سازمانی و ارزش‌های ۱۰۰ شرکت برتر صنایع غذایی ایالات متحده آمریکا و کانادا (رتبه‌بندی سال ۲۰۱۵) و ۲۴ شرکت صنایع غذایی داخلی بررسی و تحلیل شده است. نتایج پژوهش نشان می‌دهد که شرکت‌های بین‌المللی در متن چشم‌اندازشان با مفاهیمی چون غذا، رشد، ارزش، مشتری، محصول و بازار در قالب ۶ خوشه و شرکت‌های داخلی با مفاهیمی چون تولید، بازار، سلامت، محصول و کیفیت در قالب ۳ خوشه تقسیم‌بندی شده‌اند. درنهایت مشخص شد که شرکت‌های بین‌المللی در مواردی با شرکت‌های داخلی هم‌راستا بوده و در موارد فراوانی نوع دیدگاه شرکت‌های داخلی و بین‌المللی به این سندها و مفاهیم به‌کاررفته در آنها متفاوت است.

کلیدواژه‌ها: چشم‌انداز؛ مأموریت سازمانی؛ ارزش‌ها؛ خوشه‌بندی؛ متن‌کاوی.

تاریخ دریافت مقاله: ۱۳۹۵/۱۲/۰۶، تاریخ پذیرش مقاله: ۱۳۹۶/۰۹/۱۷.

* استادیار گروه مدیریت بازرگانی، دانشکده اقتصاد، مدیریت و علوم اداری، دانشگاه سمنان (نویسنده مسئول).

Email: mmaleki80@semnan.ac.ir

** دانشجوی دکتری مدیریت سیستم، دانشکده اقتصاد، مدیریت و علوم اداری، دانشگاه سمنان.

۱. مقدمه

بی‌شک آشنایی با چشم‌انداز، مأموریت و ارزش‌های بنیادی شرکت‌های برتر و مقایسه تطبیقی آن با شرکت‌های داخلی راهگشای مدیران در تحلیل و به‌روزرسانی بیانیه‌های شرکت‌های متبوع خواهد بود. بیشتر مدیران و رهبران بر این باور هستند که اهداف، راهبردها، قدرت عملیات درونی و مزایای کلیدی رقابتی سازمان‌ها فرصت ظهور و بروز آن‌ها را در سطح محیط خارجی مهیا کرده است و این مهم در پیوند با چشم‌انداز، مأموریت و موجودیت سازمان قرار دارد (گمبل و تامسون، ۲۰۱۱). بخش قابل‌توجهی از اطلاعات قابل‌دسترس به‌صورت متنی ذخیره شده‌اند و پایگاه داده‌های متنی به‌علت افزایش مقدار اطلاعات موجود به فرم الکترونیکی به‌سرعت رشد کرده‌اند. امروزه بیشتر اطلاعات در صنعت، کسب‌وکار^۱ و سازمان‌های دیگر به‌صورت الکترونیکی و به فرم پایگاه داده متنی ذخیره شده‌اند (ایمانی، ۱۳۹۱). بیش از ۸۰ درصد از اطلاعات به شکل متن ذخیره می‌شود (گوپتا و لهال، ۲۰۰۹). افزایش روزافزون این اطلاعات باعث تحریک پژوهشگران در استفاده از ابزارهای نوین و کشف دانش و الگوهای موجود در این اطلاعات شده است. یکی از مهم‌ترین این ابزارها متن‌کاوی است.

متن‌کاوی یکی از زمینه‌های است که به دنبال استخراج اطلاعات مفید از داده‌های متنی بدون ساختار، به‌وسیله شناسایی و اکتشاف الگوها است. متن‌کاوی، نخستین بار توسط فلدمن و داگان (۱۹۹۵) مطرح شد. متن‌کاوی، کشف اطلاعات جدید و از پیش ناشناخته به‌وسیله استخراج خودکار اطلاعات از منابع مختلف است. متن‌کاوی یک زمینه جوان میان‌رشته‌ای است (هوتهو، نرنبرگر و پاب، ۲۰۰۵). از خوشه‌بندی اسناد^۲ می‌توان به‌عنوان یکی از تکنیک‌های حوزه متن‌کاوی در هوش مصنوعی نام برد. خوشه‌بندی اسناد، به‌عنوان یکی از روش‌های یادگیری ماشین بدون ناظر^۳، در زمینه‌های مختلف پردازش زبان‌های طبیعی از قبیل بازیابی اطلاعات^۴، خلاصه‌سازی چندمتنی خودکار و غیره کاربرد گسترده‌ای دارد. برای مثال، در موتورهای جست‌وجو، خوشه‌بندی اسنادی که از نتایج موتور جست‌وجو به‌دست می‌آید تأثیر قابل‌ملاحظه‌ای در بهبود دقت بازیابی اطلاعات خواهد داشت (ایمانی، ۲۰۱۲).

با توجه به اهمیت تحلیل چشم‌انداز، مأموریت و ارزش‌های سازمان و از سویی دیگر کاربرد متن‌کاوی در کاوش این‌گونه اسناد، پژوهش حاضر بر آن است به بررسی عوامل بالا در ۱۰۰ شرکت برتر صنایع غذایی ایالات‌متحده و کانادا و مقایسه آن با ۲۴ شرکت داخلی با استفاده از متن‌کاوی در قالب روش خوشه‌بندی بپردازد، شرکت‌های مذکور را خوشه‌بندی کند و با

1. Business
2. Document clustering
3. Unsupervised machine learning
4. Information retrieval

استخراج واژه‌ها و مفاهیم کلیدی هر خوشه به تحلیل خوشه‌های شکل گرفته پردازد؛ بنابراین پژوهشگران در صدد پاسخ به پرسش‌های زیر در این پژوهش هستند:

- مفاهیم حاصل از خوشه‌بندی شرکت‌های بین‌المللی و داخلی چیست؟

- میزان تطابق این مفاهیم براساس چشم‌انداز، مأموریت و ارزش‌های شرکت‌ها چیست؟

۲. مبانی نظری و پیشینه پژوهش

متن‌کاوی. متن‌کاوی یعنی جستجوی الگوها در متون بدون ساختار. سه روش اساسی در مواجهه با این حجم وسیع از اطلاعات بدون ساختار وجود دارد که عبارت‌اند از: بازیابی اطلاعات؛ استخراج اطلاعات^۲ و پردازش زبان طبیعی^۳. در بازیابی اطلاعات با توجه به نیاز مطرح‌شده از سوی کاربر، نزدیک‌ترین متون و مستندات و یا کیسه کلمه از میان دیگر مستندات یک مجموعه بیرون کشیده می‌شود (فیضی درخشی و همکاران، ۲۰۱۳). این عمل یافتن دانش نیست؛ بلکه تنها فهرست کلماتی را که به نظر مرتبط‌تر به نیاز اطلاعاتی جست‌وجوگر است را به او تحویل می‌دهد. این روش هیچ دانشی به ارمغان نمی‌آورد. هدف کلی پردازش زبان طبیعی، رسیدن به درک بهتر از زبان طبیعی توسط رایانه‌ها است و روش‌های مستحکم و ساده‌ای را برای پردازش سری متن به کار می‌برد؛ همچنین از روش‌های تجزیه و تحلیل زبان‌شناسی نیز برای پردازش متن استفاده می‌کند. نقش پردازش زبان طبیعی در متن‌کاوی فراهم کردن یک سیستم در مرحله استخراج اطلاعات با داده‌های زبانی است. هدف روش‌های استخراج اطلاعات، استخراج اطلاعات خاص از سندهای متنی است. استخراج اطلاعات می‌تواند به‌عنوان یک فاز پیش‌پردازش در متن‌کاوی به کار برود. استخراج اطلاعات روزنامه‌ها، ایمیل‌ها، گزارش‌ها، مقاله‌های صفحات وب، هر پایگاه داده متنی و غیره به یک نمایش ساخت‌یافته و از پیش تعریف‌شده یا قالب‌هایی که وقتی پر می‌شوند، منتخبی از اطلاعات کلیدی از متن اصلی را نشان می‌دهند (ایمانی، ۲۰۱۲).

تحلیل خوشه‌ای. تحلیل خوشه‌ای یکی از تکنیک‌های مهم داده‌کاوی است و مزایای زیادی در شناسایی روندهای غیرقابل‌پیش‌بینی، همبستگی‌ها و الگوهای بین داده‌ها دارد (آنانیادو، ۲۰۰۸). خوشه‌بندی رایج‌ترین روش یادگیری بدون ناظر است و ابزار خوبی با گستره وسیعی از کاربردها در حوزه‌های متفاوت تجارت و علم است. از جمله این کارکردها یافتن سندهای متنی مشابه است (ایمانی، ۲۰۱۲). استفاده از خوشه‌بندی به‌عنوان مهم‌ترین مسئله در یادگیری بدون

1. Information Retrieval
 2. Information Extraction
 3. Natural Language Processing (NLP)

نظارت و ترکیب آن با رویکرد استخراج اطلاعات، در تولید مفاهیم از حجم زیادی از اسناد متنی بسیار ارزشمند خواهد بود.

پیشینه تجربی استفاده از متن‌کاوی در حوزه‌های مختلف نمایانگر حضور مؤثر این دانش نوین است. مطالعات فراوانی در حوزه متن‌کاوی و کاربردهای آن صورت گرفته است. متن‌کاوی در حوزه بهداشت، درمان و پزشکی به شکل برقراری ارتباط چند پرونده پزشکی با هم و یافتن روابط بین علائم بیماری و نسخه‌های تجویزی به کار رفته است (اسمیت، ۲۰۱۰). یک مطالعه موردی در زمینه استفاده از متن‌کاوی در تجزیه و تحلیل ثبت اختراع^۱ در سال ۲۰۰۳ انجام گرفت. در این پژوهش، نرم‌افزار PackMOLE به عنوان یک ابزار متن‌کاوی در زمینه اطلاعات ثبت اختراع‌های صنعت بسته‌بندی به کار رفت. پژوهشگران نشان دادند که این نرم‌افزار نسبت به تکنیک‌های تجزیه و تحلیل دستی ثبت اختراع‌ها مزایای بیشتری دارد (فاتوری، پدرازی و تورنا، ۲۰۰۳). متن‌کاوی در زمینه خوشه‌بندی اسناد آرشو شده به تعدادی خوشه و دسته‌بندی معنایی از پیش تعیین شده نیز به کار می‌رود (نتو و همکاران، ۲۰۰۰). اخیراً در زمینه‌های قضایی و قانونی به کار رفته و در خلاصه‌سازی متون مورد استفاده قرار گرفته است که به طور مؤثری کاراکترها، اسامی اصلی، افعال و بیشترین کلمات و رفرنس‌های مستندات حجیم را مشخص می‌کند (ال اظمی، ۲۰۱۳)

کولیشکینا و روین (۲۰۰۵)، نتایج تجزیه و تحلیل کاربرد متن‌کاوی در پایگاه خسارات بیمه را ارائه دادند. آن‌ها با استفاده از متن‌کاوی در فرم‌های شکایات، مفاهیمی را از این توضیحات استخراج کردند. فرانسیس (۲۰۰۶) و ویز و همکاران (۲۰۰۵) در به کارگیری خوشه‌بندی و سایر تکنیک‌های بدون ناظر در متن‌کاوی، مقاله مشروح و مفصلی ارائه کرده‌اند.

دیو و همکاران (۲۰۰۵)، استفاده از داده‌کاوی در گام‌های کلیدی تجزیه و تحلیل متن‌های غیرساختاریافته را نشان داده‌اند. گام‌های مورد اشاره عبارت‌اند از: الف) آماده‌سازی داده‌ها، شامل ورود داده و پاک‌سازی؛ ب) تجزیه و تحلیل روابط و ج) نمایش نتایج تجزیه و تحلیل، شامل نمایش گرافیکی نتایج که فهم نکات کلیدی متن را ساده‌تر می‌کند. الشامری و همکاران (۲۰۱۲)، با استفاده از رویکرد متن‌کاوی و نرم‌افزار CLUTO، چشم‌انداز و مأموریت ۷۷۲ شرکت حاضر در فورچون ۵۰۰ را خوشه‌بندی کردند و نشان دادند که چگونه می‌توان این خوشه‌ها را با مفاهیم جهانی شدن، نوآوری و ویژگی‌های استراتژی تبیین کرده و برای ارزیابی هم‌راستایی، تثبیت و جهت‌یابی استراتژیک استفاده کرد.

ویتن و همکاران (۲۰۰۳) از تکنیک‌های متن‌کاوی برای استخراج فراداده^۲ از اسناد موجود در یک کتابخانه و غنی‌سازی اسناد با پررنگ کردن آیتم‌های مناسب در متن، استفاده کردند و

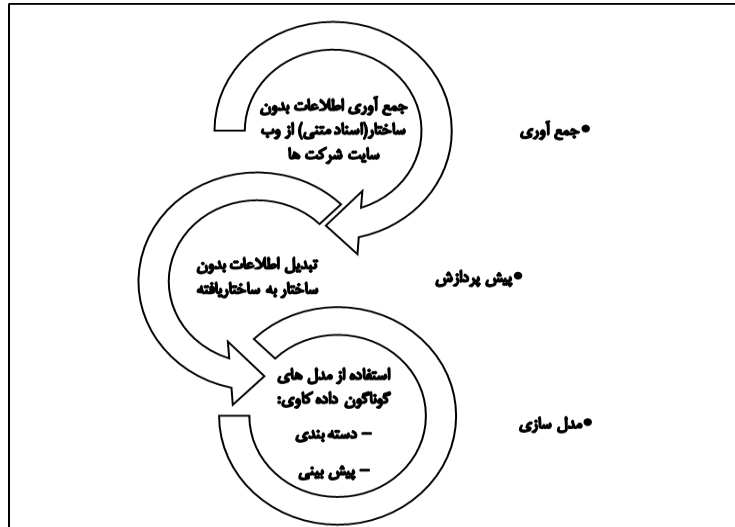
1. Patent
2. Meta Data

دریافتند که متن‌کاوی می‌تواند ارزش‌افزوده‌ای را به اسناد کتابخانه دیجیتال بیافزاید و تجربه کاربران از کتابخانه را غنی‌تر کند. فولر و همکاران (۲۰۱۱) از متن‌کاوی برای شناسایی فریب و دروغ در داده‌های دنیای واقعی استفاده کردند. نتایج آن‌ها نشان می‌دهد که تکنیک‌های متن‌کاوی توان بالقوه فراوانی در کمک به کسانی دارد که در پی کشف دروغ در متون هستند. هانگ (۲۰۱۲) با استفاده از تحلیل خوشه‌بندی به‌عنوان یک تکنیک اکتشافی، مبانی نظری آموزش الکترونیک را واکاوی کرد و به نمایش الگوهای موجود در آن بر اساس کلمات مشابه، ارزش صفات و قوانین گذشته پرداخت.

رسانه‌های اجتماعی یکی دیگر از حوزه‌های مهمی است که متن‌کاوی به آن ورود کرده است. با توجه به قابلیت‌های قدرتمند متن‌کاوی، اعتقاد پژوهشگران بر این است که استفاده از متن‌کاوی در داده‌های رسانه‌های اجتماعی می‌تواند به کشف یافته‌های جذابی در حوزه رفتار و تعامل انسان منجر شود (ابدوس و همکاران، ۲۰۱۲؛ باربیر و لیو، ۲۰۱۱؛ هی، ۲۰۱۳؛ پنگ و لی، ۲۰۰۸).

۳. روش پژوهش

پیشرفت الگوریتم‌های خوشه‌بندی متون، دستیابی به الگوهای موجود در داده‌ها را افزایش داده است. خوشه‌بندی و تشخیص مهم‌ترین ویژگی‌های مجموعه داده‌های متنی باعث بهبود تجزیه و تحلیل و در نتیجه تولید معنا از این داده‌های خام می‌شود. در پژوهش حاضر، پژوهشگران به بررسی متون چشم‌انداز، مأموریت سازمانی، ارزش‌ها و اهداف ۱۰۰ شرکت برتر صنایع غذایی ایالات متحده و کانادا و مقایسه تطبیقی آن با ۲۴ شرکت ملی پرداختند و در قالب خوشه‌بندی این متون و استخراج مهم‌ترین ویژگی‌های هر خوشه اهداف چهارگانه زیر را بررسی کرده‌اند: ۱. استخراج کلمات مهم و پرتکرار در چشم‌انداز، مأموریت و ارزش‌های شرکت‌های بین‌المللی و داخلی؛ ۲. خوشه‌بندی شرکت‌های بین‌المللی و ملی بر اساس کلمات استخراج‌شده؛ ۳. بررسی خوشه‌های ایجادشده و استخراج مفاهیم کلیدی هر خوشه و ۴. مقایسه تطبیقی بین شرکت‌های ملی و بین‌المللی با توجه به خوشه‌ها و مفاهیم کلیدی هر خوشه. در شکل ۱، مراحل اجرای پژوهش ارائه شده است.



شکل ۱. مراحل اجرای پژوهش

جمع آوری مجموعه داده‌های پژوهش. مجموعه داده مورد استفاده در این پژوهش شامل متن چشم‌انداز، مأموریت سازمانی و ارزش‌های شرکت‌های بین‌المللی و داخلی در حوزه صنایع غذایی است که به شرح زیر جمع‌آوری شده است:

– هر ساله فهرست ۱۰۰ شرکت برتر صنایع غذایی ایالات متحده آمریکا و کانادا توسط کارشناسان صنایع غذایی با لحاظ کردن عوامل مورد نظر از طریق تارنمای پردازش غذا^۱ اعلام می‌شود. فهرست سال ۲۰۱۵ این تارنما، مرجع پژوهش حاضر قرار گرفته است و با مراجعه به تارنمای شرکت‌ها و تارنماهای بیانیه‌های مأموریت^۲ و شعارها و بیانیه‌های شرکت‌ها^۳، متن چشم‌انداز، مأموریت سازمانی و ارزش‌های شرکت‌ها استخراج شده است.

– با مراجعه به سایت «بورس اوراق بهادار تهران»، فهرست ۵۰۰ شرکت برتر ایران در سال ۹۴ و تارنمای صنایع غذایی^۴، فهرست ۲۴ شرکت داخلی فعال در صنایع غذایی استخراج شده و با رجوع به تارنمای هر یک از این شرکت‌ها، متن چشم‌انداز، مأموریت و ارزش‌ها استخراج شد.

متن چشم‌انداز، مأموریت و ارزش‌های بنیادی شرکت‌ها به صورت رکوردهای مجزا در فایل اکسل وارد شده و در مرحله بعد به نرم‌افزار متن کاوی ریپیدماینر^۵ وارد شد.

1. www.foodprocessing.com
 2. http://www.missionstatements.com/company_mission_statements.html
 3. <http://www.company-statements-slogans.info/fortune-500-companies.htm>
 4. www.foodkeys.com
 5. Rapid Miner

نرم‌افزار متن‌کاوی مناسب باید قابلیت اطمینان بالا، ثبات و سرعت محاسباتی بالا داشته باشد. نرم‌افزارهای متن‌کاوی فراوانی در بازار موجود است. نوع اول، نرم‌افزارهای تجاری که راه‌حل گران‌قیمت و درعین‌حال توانمندی هستند و از جمله خصوصیات آن‌ها بدون محدودیت بودن و پشتیبانی از زبان‌های مختلف است؛ در مقابل این‌گونه نرم‌افزارها، نرم‌افزارهای متن‌باز قرار می‌گیرند که قوی‌ترین و رایج‌ترین آن‌ها «رپیدماینر» است.

این نرم‌افزار در سال ۲۰۰۴ در دانشگاه دورتموند تولید شد و پیوسته در حال به‌روزرسانی است. بر اساس گزارش تارنمای مربوطه^۱ تقریباً نیمی از متخصصان متن‌کاوی از این نرم‌افزار استفاده می‌کنند. این امر تصادفی نیست؛ زیرا شرکت‌های بزرگی چون «سیسکو»، «فورد»، «هوندا»، «اچ پی»، «آی بی ام» و «فیلیپس» نیز از قابلیت‌های فراوان داده‌کاوی و بیش از ۵۰۰ عملگر آن استفاده زیادی می‌برند.

در پژوهش حاضر از افزونه متن‌کاوی موجود در نرم‌افزار «رپیدماینر» برای متن‌کاوی متون و درنهایت خوشه‌بندی آن‌ها توسط الگوریتم K-means استفاده شده است.

پیش‌پردازش داده‌ها. اجرای این مرحله که با عنوان پیش‌پردازش داده‌ها شناخته می‌شود مستلزم گام‌های زیر است (شکل ۲):

- نخستین گام، تفکیک متون به اجزای تشکیل‌دهنده آن، یعنی کلمات است. در اصطلاح به این مرحله «تفکیک به کلمات»^۲ می‌گویند.

- گام دوم، حذف کلمات عمومی از جمله حروف اضافه، حروف ربط، قیود و غیره است که نقش چندانی در بار معنایی جملات ایفا نمی‌کنند و از آن به «حذف کلمات توقف»^۳ یاد می‌شود.

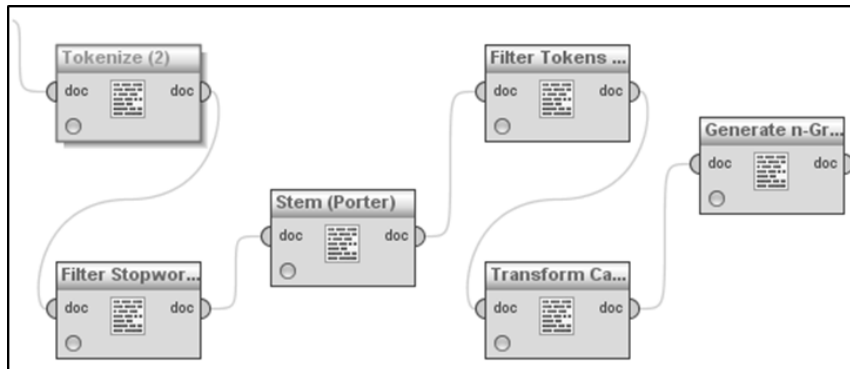
- گام سوم کلماتی که ریشه یکسان دارند به یک شکل در نظر گرفته می‌شوند. به این مرحله «ریشه‌یابی»^۴ می‌گویند.

- در گام بعد، حساسیت به کوچک و بزرگ بودن حروف^۵ نادیده گرفته می‌شود.

- بعضی از کلمات به صورت جفت یا حتی سه‌بخشی در متون ظاهر می‌شوند و معنا می‌یابند. در این مرحله که در اصطلاح «دنباله‌ای از یک یا چند کلمه»^۶ نامیده می‌شود، این مورد لحاظ می‌شود.

1. Rapid-i.com
2. Tokenize
3. Filter Stop words
4. Stemming
5. Transform Case
6. n-Grams

- فیلتر کلمات بر اساس معنا^۱: کلمات خاصی از متن به تشخیص پژوهشگر استخراج می‌شود و مورد مطالعه قرار می‌گیرد.



شکل ۲. پیش‌پردازش داده‌ها در رپیدماینر

مدل فضای برداری پایه‌ای‌ترین مدل در سیستم‌های بازیابی اطلاعات است. در مدل برداری، هر مستند به صورت برداری از کلمات در نظر گرفته شده و فضایی چندبُعدی که ابعاد آن را کلمات تشکیل می‌دهند، ایجاد می‌شود؛ سپس هر سند در این فضا به صورت یک بردار نمایش داده می‌شود. مؤلفه‌های این بردار سند، وزن‌هایی هستند که نشان می‌دهند هر یک از کلمات چقدر در متمایز کردن آن سند دخیل هستند؛ سپس بر اساس پارامترهای مختلفی مانند تعداد تکرار کلمه در متن، تعداد تکرار کلمه در اسناد مجموعه و مؤلفه‌های نرمال‌سازی وزنی به هر کلمه نسبت داده می‌شود.

یکی از پرکاربردترین روابط در حوزه بازیابی اطلاعات، پارامتر TF/IDF است (بایز - یاتس و ریپیرو - نتو، ۱۹۹۹) که از حاصل ضرب فراوانی کلمه در فراوانی معکوس سند به دست می‌آید. این روش یک روش مبتنی بر چند سند است که در آن منظور از فراوانی کلمه، فقط تعداد تکرار کلمه در یک سند خاص است؛ همچنین منظور از فراوانی معکوس سند، تعداد اسنادی است که این کلمه خاص در آن اسناد ظاهر شده است. دلیل مقبولیت این روش نسبت به سایر روش‌ها را می‌توان با توجه به سهولت در استفاده از این روش، محاسبات کم و نتایج قابل قبول دانست (تشکری، ۲۰۰۱). با توجه به مطالب بالا از پارامتر TF/IDF استفاده شده و در جدول‌های ۱ تا ۴، نمونه‌ای از خروجی نرم‌افزار شامل بردار TF/IDF^2 هر سند و تکرار کلمات در مجموعه متون چشم‌انداز، مأموریت و ارزش‌های شرکت‌های بین‌المللی و داخلی نمایش داده شده است.

1. Filter Tokens (by Content)
2. Term Frequency-Inverse Document Frequency

جدول ۱. نمونه‌ای از بردار TF-IDF چشم‌انداز شرکت‌های بین‌المللی

Row No.	custom	Food	Grow	Product	Value	World
۱	۰	۰/۴۴۷	۰	۰/۸۹۴	۰	۰
۲	۰	۰	۰	۰	۰/۷۲۵	۰/۶۸۸
۳	۰	۰	۰	۰/۷۲۶	۰/۶۸۸	۰

جدول ۲. کلمات پرتکرار در چشم‌انداز شرکت‌های بین‌المللی

کلمه	کل رخداد	تعداد سند
Custom	۲۶	۱۶
World	۱۶	۱۵
Valu	۱۵	۱۴
Grow	۱۴	۱۴
Product	۱۴	۱۳
Food	۱۳	۱۳

جدول ۳. نمونه‌ای از بردار TF-IDF چشم‌انداز شرکت‌های داخلی

ردیف	بازار	تولید	سالم	محصول	کیفیت
۱	۰/۸۱۶	۰/۴۰۸	۰/۴۰۸	۰	۰
۲	۰	۰	۰	۰	۰
۳	۰	۰	۰	۰	۰

جدول ۴. کلمات پرتکرار در چشم‌انداز شرکت‌های داخلی

کلمه	کل رخداد	تعداد سند
محصول	۱۳	۷
تولید	۱۱	۷
بازار	۸	۷
کیفیت	۸	۶
سالم	۷	۷

خوشه‌بندی. خروجی مرحله پیش‌پردازش داده‌ها به شکل فهرست کلمات و بردار اسناد است. فهرست کلمات بر اساس تعداد اسنادی که کلمه در آن رخ داده و تکرار کلمه در کل اسناد است. برای مثال، طبق جدول ۲، کلمه Custom در ۱۶ سند از ۵۴ سند موجود در مجموع ۲۶ بار تکرار شده است. در جدول ۴ که مربوط به متن چشم‌انداز شرکت‌های داخلی است کلمه محصول در ۷ سند از ۲۱ سند موجود در مجموع ۱۳ بار تکرار شده است. در جدول ۵، کلمات

پرتکرار در چشم‌انداز، بیانیه مأموریت و ارزش‌های شرکت‌های بین‌المللی و داخلی مشخص شده است.

جدول ۵. کلمات پرتکرار در چشم‌انداز، بیانیه مأموریت و ارزش‌های شرکت‌های بین‌المللی و داخلی

		تعداد سند	کل رخداد	کلمه
چشم‌انداز	بین‌المللی	۱۶	۲۶	Custom
		۱۵	۱۶	World
		۱۴	۱۵	Value
		۱۴	۱۴	Grow
		۱۳	۱۴	Product
	داخلی	۱۳	۱۳	Food
		۷	۱۳	محصول
		۷	۱۱	تولید
		۷	۸	بازار
		۶	۸	کیفیت
مأموریت سازمانی	بین‌المللی	۲۶	۳۶	Custom
		۲۴	۲۸	Product
		۱۴	۱۸	Employ
		۱۶	۱۸	Quality
		۱۶	۱۷	Value
	داخلی	۱۳	۲۸	تولید
		۱۶	۲۶	محصولات
		۱۲	۱۷	مشتریان
		۱۲	۱۲	کیفیت
		ارزش‌ها	بین‌المللی	۲۰
۳۱	۳۱			Integration
۲۰	۲۹			Custom
۲۲	۲۳			Respect
۱۹	۲۱			Quality
داخلی	۱۸		۱۹	Innovation
	۹		۱۲	مشتری
	۷		۷	صداقت
	۷		۷	نوآوری
	۷		۷	کارکن

در پژوهش حاضر از روش خوشه‌بندی K-means استفاده شده است. این الگوریتم از نوع افزایشی است و یک افزاز اولیه از داده‌ها به وجود می‌آید؛ سپس داده‌ها را به منظور دستیابی به یک افزاز بهتر بین خوشه‌ها جابه‌جا می‌کند. تعداد خوشه‌هایی (k) که باید داده‌ها در نهایت درون آن‌ها قرار بگیرد از ابتدای فرآیند مشخص می‌شود. پس از بررسی مقادیر مختلف k در بازه ۲ تا ۸، در نهایت تعداد خوشه‌ها در جدول ۶، بر اساس متن چشم‌انداز، مأموریت و ارزش‌های شرکت‌های بین‌المللی و داخلی تفکیک شده و نمایش داده شده است.

جدول ۶. تعداد خوشه‌ها (K)

متن ارزش‌ها	متن بیانیه مأموریت	متن چشم‌انداز
۶	۴	شرکت‌های بین‌المللی
۳	۴	شرکت‌های داخلی

می‌توان مفاهیم و کلماتی که بیشترین تفکیک را بین خوشه‌ها ایجاد می‌کنند مشخص کرد و به‌عنوان برچسب (مفهوم) خوشه به کار برد. در جدول ۷، به‌عنوان نمونه برچسب مربوط به خوشه‌های متون چشم‌انداز شرکت‌های بین‌المللی ارائه شده است. بر اساس این جدول، کلمه Food، خوشه اول را نسبت به سایر خوشه‌ها تفکیک کرده است و به‌عنوان برچسب این خوشه استفاده می‌شود و در خوشه دوم کلمه Grow به‌عنوان برچسب خوشه انتخاب می‌شود. به همین ترتیب، برچسب Value (ریشه لغت ارزش) برای خوشه سوم، کلمه Custom (ریشه لغت مشتری) برای خوشه چهارم، Product به‌عنوان برچسب خوشه پنجم و world به‌عنوان برچسب خوشه ششم معرفی می‌شوند.

جدول ۷. برچسب خوشه‌های چشم‌انداز شرکت‌های بین‌المللی

صفت	خوشه ۱	خوشه ۲	خوشه ۳	خوشه ۴	خوشه ۵	خوشه ۶
Custom	۰/۲۵	۰	۰	۰/۷۳۰	۰	۰
Food	۰/۷۷۶	۰/۰۹۱	۰	۰/۰۲۴	۰/۰۸۲	۰
Grow	۰/۰۶۴	۰/۸۸۸	۰	۰/۱۸۹	۰	۰
Product	۰	۰/۰۹۱	۰/۲۴۲	۰/۱۸۱	۰/۳۲۹	۰/۱۴۹
Valu	۰/۰۶۵	۰	۰/۸۵۰	۰/۲۸۷	۰	۰
World	۰/۲۶۰	۰/۰۸۶	۰/۱۱۵	۰/۱۳۸	۰	۰/۹۳۴

خروجی نرم‌افزار در خصوص مأموریت و ارزش‌های شرکت‌های بین‌المللی و داخلی همانند جدول ۷، است. در جدول ۸، خلاصه‌ای از برچسب‌های خوشه‌ها به تفکیک ارائه می‌شود.

جدول ۸. مقایسه برجسب خوشه‌ها

	خوشه ۱	خوشه ۲	خوشه ۳	خوشه ۴	خوشه ۵	خوشه ۶
	بین‌المللی	Food	Grow	Value	Custom	Product
چشم‌انداز	داخلی	محصول - کیفیت	سالم - بازار	تولید		
مأموریت	بین‌المللی	Custom	Value	Quality-Product	Employ	
سازمانی	داخلی	تولید	کیفیت	محصولات	مشتریان	
	بین‌المللی	Innovation	Integrity	Employ	Quality	Custom
ارزش‌ها	داخلی	نوآوری	صداقت	مشتری - کارکن		Respect

۴. تحلیل داده‌ها و یافته‌ها

پس از عملیات متن‌کاوی بر چشم‌انداز، مأموریت و ارزش‌های نمونه‌ای از شرکت‌های صنایع غذایی بین‌المللی و داخلی، مفاهیم قابل‌تأملی در قالب برجسب خوشه‌ها ارائه شد. در جدول ۹، این مفاهیم نمایش داده شده است. با بررسی جدول ۹ و مطالعه جداگانه سند چشم‌انداز، مأموریت سازمانی و ارزش‌های شرکت‌ها بر اساس مفاهیم استخراج‌شده از خوشه‌ها، نتایج در ادامه مرور خواهد شد.

مقایسه تطبیقی براساس چشم‌انداز. با توجه به جدول ۹ و بررسی متن چشم‌انداز شرکت‌های بین‌المللی و داخلی که در خوشه‌ها قرار گرفته‌اند در جدول ۱۰، مقایسه تطبیقی صورت گرفته است.

جدول ۱۰. مقایسه تطبیقی مفاهیم موجود در متن چشم‌انداز شرکت‌ها

Product	چگونگی تولید محصول، طعم محصول، محصولات کشاورزی، بهترین محصولات
محصول	محصولات متمایز، کیفیت محصول، محصولات متنوع، محصول جدید، محصول منجمد
Grow	رشد تغذیه، رشد فروش، رشد بازار جهانی، رشد بازگشت سرمایه، رشد پایدار جهانی تغذیه، رشد سریع، افزایش رشد تا رسیدن به رهبری بازار
Value	ارزش سهام، فرهنگ ارزش محور، ارزش‌های بنیادی شرکت
World	جهانی بهتر، بهترین بودن در جهان، افزایش کیفیت زندگی در جهان، به‌هم‌زدن معادلات جهانی
کیفیت	کیفیت محصول، کیفیت جهانی
سالم	غذای سالم، زندگی سالم
بازار	بازار ملی و فراملی
تولید	تولید ارزش، فناوری تولید، ظرفیت بهینه تولید، تولید محصول

با بررسی جدول ۱۰، می‌توان به نکته‌های قابل توجهی دست یافت. واژه‌های Product و محصول که به‌طور مشترک در متن چشم‌انداز شرکت‌های داخلی و بین‌المللی به‌کار رفته‌اند عموماً مفاهیم مشابهی را القا می‌کنند. واژه‌های Value, Grow و World در مجموع با اشاره به رشد بازارهای جهانی و رشد پایدار تغذیه در قالب ارزش‌های بنیادی شرکت‌ها در جست‌وجوی جهانی بهتر و افزایش کیفیت زندگی هستند و نشان از دیدگاهی فراتر از واژه‌های به‌کاررفته در متن چشم‌انداز شرکت‌های داخلی دارند. در شرکت‌های داخلی کماکان چالش اساسی فناوری تولید، ظرفیت بهینه تولید و تا حدودی کیفیت محصول در قالب غذایی سالم است. اگرچه تا حدی نگاه به بازارهای فراملی و کیفیت جهانی وجود دارد.

مقایسه تطبیقی بر اساس مأموریت. مقایسه تطبیقی مفاهیم خوشه‌های برآمده از متن کاوی مأموریت شرکت‌های داخلی و بین‌المللی در جدول ۱۱، ارائه شده است.

جدول ۱۱. مقایسه تطبیقی مفاهیم موجود در متن مأموریت شرکت‌ها

Customer	بهترین طعم برای مشتری، کاهش مستمر هزینه مشتریان، خدمت بهتر به مشتری، ارائه بهترین محصول به مشتری، افزایش ارزش افزوده به مشتری، ارائه غذای سالم به مشتری، افزایش توقع و انتظار مشتری، مشعوف کردن مشتری، کمک‌گرفتن از مشتریان در نوآوری
مشتریان	رشد مشتریان، علاقه مشتریان، بینش مشتریان، شیوه زندگی سالم، ساده و خوشمزه برای مشتری، تعامل دائم با مشتریان
Product	محصول مطابق با شیوه نوین زندگی، کیفیت محصول، محصول نوآورانه، محصول تازه، محصولات مغذی، محصولات ارگانیک، محصول فوق کیفی، باکیفیت‌ترین محصول
محصولات	محصول باکیفیت، محصول مطابق استاندارد جهانی، محصول متنوع، محصول با بهترین قیمت، محصول سالم، محصول با ارزش غذایی بالا، تکمیل سبد محصولات شرکت، محصول مطابق با نیاز مشتری
Quality	کیفیت قابل اعتماد، کیفیت بالا، کیفیت زندگی، کیفیت برتر، برترین کیفیت
کیفیت	تعریف نو از کیفیت با استفاده از فناوری‌های نوین
Value	عجین شدن با ارزش‌ها، خلق ارزش، به حداکثر رساندن ارزش سهام در بلندمدت، خلق ارزش‌های متمایز، بهینه کردن ارزش‌ها، ارزش برجسته، ایجاد ارزش بیشتر برای مشتریان
Employ	آینده بهتر برای کارکنان، افزایش توقع و انتظارات کارکنان، برقراری ارتباط با کارکنان در بالاترین استاندارد
تولید	تولید ثروت، اقتصادی بودن تولید، بهترین مواد اولیه برای تولید، افزایش پیوسته تولید، بهبود چرخه تولید، تولید طعم‌های جدید، تولید جهانی، تولید محصولات متنوع

بررسی جدول ۱۱، نشان می‌دهد که از جهاتی در مفهوم مشتری، بین شرکت‌های داخلی و بین‌المللی شباهت وجود دارد. در شرکت‌های داخلی، رشد تعداد مشتریان همراه با تعامل با

آن‌ها مورد توجه است؛ درحالی‌که به نظر می‌رسد شرکت‌های بین‌المللی پا را فراتر نهاده‌اند و در پی مشعوف کردن مشتری با ارائه بهترین محصول و استفاده از دیدگاه مشتریان در نوآوری هستند. شرکت‌های داخلی نیز با مدنظر قراردادن تغییر در بینش مشتریان به دنبال رسالتی نوین هستند. با نگاهی به واژه محصول، کاملاً مشخص است که شرکت‌های بین‌المللی با اشاره به شیوه نوین زندگی، محصولات ارگانیک و فوق‌کیفی، مسیر جدیدی را به‌عنوان مأموریت پذیرفته‌اند. در مأموریت شرکت‌های داخلی کماکان محدود بودن در استانداردها، تکمیل سبد محصولات و قیمت محصول به چشم می‌خورد. واژه «کیفیت» مفهوم بعدی در شرکت‌های داخلی در چالش تعریف نو با استفاده از فناوری‌های نوین است؛ درحالی‌که شرکت‌های پیشرو در حوزه صنایع غذایی در جست‌وجوی برترین کیفیت هستند. واژه‌های ارزش و کارکنان حلقه‌های مفقود در متن چشم‌انداز و مأموریت شرکت‌های داخلی می‌باشند؛ درحالی‌که شرکت‌های پیشرو به دنبال عجز کردن ارزش‌ها با کارکنان و فرآیند تولید و خدمت هستند و پیوسته سطح انتظارات کارکنان را از سازمان بالا می‌برند.

مقایسه تطبیقی بر اساس ارزش‌ها. با توجه به اینکه ارزش‌های بنیادی شرکت‌ها معمولاً در قالب تک‌واژه‌ها بیان می‌شوند می‌توان از مفاهیم جدول ۸، برای مقایسه تطبیقی استفاده کرد. بررسی این جدول نشان می‌دهد که مفاهیم Innovation و نوآوری، Integrity و صداقت، Employ و کارکنان، Customer و مشتری در ارزش‌های شرکت‌های داخلی و بین‌المللی تطابق کامل دارند. در ضمن مفاهیمی چون Quality و Respect در ارزش‌های شرکت‌های بین‌المللی مورد استفاده و دقت نظر قرار گرفته است.

۵. نتیجه‌گیری و پیشنهادها

با توجه به یافته‌های حاصل از متن کاوی چشم‌انداز شرکت‌های بین‌المللی می‌توان آشکارا به این نتیجه دست یافت که دیدگاه شرکت‌های مطرح صنایع غذایی به سمت جهانی شدن است و این دیدگاه با مأموریت و رسالت این شرکت‌ها که همانا مشعوف کردن مشتریان در بالاترین سطح و افزایش روزافزون توقعات آن‌ها است، قابل تلفیق است. در کنار این مسائل، ارزش‌هایی چون نوآوری، صداقت، احترام به مشتری و کیفیت عناصر لاینفک استراتژی این‌گونه سازمان‌ها است. در شرکت‌های صنایع غذایی داخلی با توجه به اینکه ارزش‌ها تا حدودی با شرکت‌های بین‌المللی تطابق دارد، مشخصاً هم‌راستایی بسیار کمتری با چشم‌انداز و رسالت آن‌ها دارد. باید این شکاف را در متون چشم‌انداز این شرکت‌ها جست‌وجو کرد که کماکان پا را از مرزها فراتر نهاده‌اند و مفاهیمی چون کیفیت، غذای سالم و تکنولوژی تولید را

مدنظر دارند. این مفاهیم به‌عنوان ابتدایی‌ترین اهداف شرکت‌های بین‌المللی در سند چشم‌انداز و رسالت آن‌ها، رنگ باخته‌اند و پا را فراتر نهاده و توقعات جدید در مشتری ایجاد می‌کنند. ازجمله محدودیت‌های پژوهش حاضر می‌توان به نبود پایگاه مناسبی که بر اساس استانداردهای خاص به رتبه‌بندی شرکت‌های صنایع غذایی بپردازد، اشاره کرد که به استفاده از چندین پایگاه و تارنما منجر شد. محدودیت دیگر پژوهش عدم‌درج متون چشم‌انداز، مأموریت و ارزش‌های شرکت‌های داخلی در تارنمای شرکت مربوطه بود؛ به‌طوری‌که بیش از ۷۰ درصد شرکت‌های موجود در فهرست ۱۰۰ شرکت برتر بین‌المللی در تارنمای شرکت به‌طور کاملاً مشخص این اسناد را گنجانده و کمتر از ۱۵ درصد شرکت‌های داخلی مورد‌مطالعه دارای این ویژگی بودند.

متن‌کاوی به‌عنوان یکی از ابزارهای نوین علمی در استخراج دانش از انبوه اسناد در صنایع مختلف کاربرد فراوان دارد. در متون علمی و صنایع داخلی از این ابزار استفاده شایسته‌ای نشده است. پژوهشگران معتقدند که به‌کارگیری این ابزار نوپا در صنایع مختلف می‌تواند به تصمیم‌گیری مدیران کمک زیادی کند. به‌عنوان پیشنهاد، بررسی و متن‌کاوی سخنان مدیران برتر صنایع، ازجمله صنایع غذایی، در کشف استراتژی‌ها و تغییر رفتار مدیران و کارکنان بسیار مؤثر خواهد بود و این مهم با انجام پژوهش‌های مستمر و ورود علوم جدیدی چون داده‌کاوی و به‌طور اخص، متن‌کاوی، دست‌یافتنی خواهد شد.

منابع

1. Abdous, M., He, W., & Yen, C. J. (2012). Using data mining for predicting relationships between online question theme and final grade. *Educational Technology & Society, 15*(3), 77-88.
2. Al-Azmi, A. (2013). Data, text, and web mining for Business Intelligence: a Survey. *International Journal of Data Mining & Knowledge Management Process (IJDKP)*.3(2).
3. Alshameri, F., Greene, G. R., & Srivastava, M. (2012). Categorizing top fortune company mission and vision statements via text mining. *International Journal of Management Information Systems. 16*(3), 227-238.
4. Ananiadou, S. (2008). National centre for text mining: *Introduction to tools for researchers*. Retrieved February, 8, 2009.
5. Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern information retrieval, (463)*. New York: ACM press.
6. Barbier, G., & Liu, H. (2011). Data mining in social media. *In Social network data analytics, 327-352*, Springer US.
7. Davi, A., Houghton, D., Nasr, N., Shah, G., Skaletsky, M., & Spack, R. (2005). A Review of Two Text Mining Packages: *SAS TextMining and WordStat. American Statistician 59*(1), 89-103.
8. Fattori, M., Pedrazzi, G., & Turra, R. (2003). Text mining applied to patent mapping: a practical business case. *World Patent Information 25*(4), 335-342.
9. Feizi darakhshi, M. R., Rashidi, Sh., & Mahmoodloo, F. (2013). Overview of text mining and methods. *Joint convention on computer engineering and mechanics*, Miandoab University of Applied Science and Technology. (In Persian)
10. Feldman, R., & Dagan, I. (1995). KDT-Knowledge Discovery in Texts, Proc. of the First Int. Conf. on Knowledge Discovery KDD, 112-117.
11. Francis, L., A. (2006). *Taming Text: An Introduction to Text Mining*, Casualty Actuarial Society Forum.
12. Fuller, C., Biros, D., & Delen, D. (2011). An investigation of data and text mining methods for real world deception detection. *Expert Systems with Applications, 38*(7), 8392-8398.
13. Gamble J. E., & Thompson, A. A. (2011). *Essentials of Strategic Management: The Quest for Competitive Advantage*. McGraw-Hill Erwin, New York, NY.
14. Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence 1*(1), 60-76.
15. He, W. (2013b). Improving user experience with case-based reasoning systems using text mining and web 2.0. *Expert System with Applications, 40*(2), 500-507.
16. Hotho, A., Nürnberger, A., & Pabb, G. (2005). A Brief Survey of Text Mining Export, *LDV Forum, Vol.20*, 19-62.
17. 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), 5-16.
18. Imani, M. (2012). Clustering of Persian texts. *Master's Thesis, Iran university of science and technolog* (In Persian).
19. Kolyshkina, I., & M. van Rooyen. (2005). Text Mining For Insurance Claim Cost Prediction. *Presented at the XVth General Insurance Seminar Institute of the Actuaries of Australia*.

20. Neto, J., Santos, A., Kaestner, 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.
21. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
22. Smith, D. (2010). *Using Data and Text Mining to Drive Innovation*. PhUSE UK.
23. Tashakori, M. (2001). Design of Automatic Indexing for persian Documents, *Computer & Information Technology Engineering Department*, Amirkabir university of technology (In Persian).
24. Weiss, S.M., Indurkha, N., Zhang, T., & Damerou, F. (2005). *Predictive Methods for Analyzing Unstructured Information*. Springer.
25. Witten, I., Don, K., Dewsnip, M., & Tablan, V. (2003). Text mining in a digital library. *International Journal of Digital Library*, 5, 1-4.