



ORIGINAL RESEARCH PAPER

Clustering and comparative review of the vision document, organizational mission, and goals of international and domestic insurance companies with a text mining approach

M. Shafiei Nikabadi^{1,*}, O. Afsharizadeh Jafari²

¹ Department of Industrial Management, School of Economics, Management and Administrative Sciences, Semnan University, Semnan, Iran

² Department of System Management, School of Economics, Management and Administrative Sciences, Semnan University, Semnan, Iran

ARTICLE INFO

Article History

Received: 22 November 2015

Revised: 31 December 2015

Accepted: 05 September 2016

Keywords

Vision; Organizational Mission; Clustering; Text Mining; K-Means.

ABSTRACT

Organizational vision, organizational mission and organization's goal are the main indicators of the organization's direction. In this research, by using the new text mining technique and combining it with the k-means clustering method, the text of the vision document, organizational mission, and goals of 39 international insurance companies in the list of the top 500 companies in the world and 18 domestic insurance companies Analysis has been done. The results of the research show that international companies with concepts such as financial security, value, service, product-service, and quality in the form of 4 clusters and domestic companies with concepts such as professionalism, peace, profitability-financial, and service delivery in the form of 4 clusters are divided. Finally, international companies are aligned with domestic companies in some cases, and in many cases, the views of domestic and international companies on these documents and the words used in them are different.

***Corresponding Author:**

Email: shafiei@profs.semnan.ac.ir

DOI: 10.22056/ijir.2016.03.01



خوشه‌بندی و بررسی تطبیقی سند چشم‌انداز، مأموریت سازمانی، و اهداف شرکت‌های بیمه بین‌المللی و داخلی با رویکرد متن‌کاوی

محسن شفیعی نیک‌آبادی^{۱*}، امید افشاری‌زاده جعفری^۲

^۱گروه مدیریت صنعتی، دانشکده اقتصاد، مدیریت و علوم اداری، دانشگاه سمنان، سمنان، ایران

^۲گروه مدیریت سیستم، دانشکده اقتصاد، مدیریت و علوم اداری، دانشگاه سمنان، سمنان، ایران

اطلاعات مقاله

تاریخ دریافت: ۰۱ آذر ۱۳۹۴

تاریخ داوری: ۱۰ دی ۱۳۹۴

تاریخ پذیرش: ۱۵ شهریور ۱۳۹۵

کلمات کلیدی

چشم‌انداز

مأموریت سازمانی

خوشه‌بندی

متن‌کاوی

k- میانگین

چکیده:

چشم‌انداز سازمانی، مأموریت سازمانی و هدف سازمان، شاخص‌های عمده جهت‌گیری سازمان محسوب می‌شوند. در این تحقیق با به‌کارگیری تکنیک نوین متن‌کاوی و ترکیب آن با روش خوشه‌بندی k- میانگین، متن سند چشم‌انداز، مأموریت سازمانی، و اهداف ۳۹ شرکت بیمه بین‌المللی موجود در لیست ۵۰۰ شرکت برتر دنیا و ۱۸ شرکت بیمه داخلی مورد تحلیل قرار گرفته است. نتایج تحقیق نشان می‌دهد که شرکت‌های بین‌المللی با مفاهیمی چون امنیت مالی، ارزش، خدمات، محصول- خدمت، و کیفیت در قالب ۴ خوشه و شرکت‌های داخلی با مفاهیمی چون حرفه‌ای، آرامش، سودآوری- مالی، و ارائه خدمات در قالب ۴ خوشه تقسیم‌بندی می‌شوند. درنهایت شرکت‌های بین‌المللی در مواردی با شرکت‌های داخلی هم- راستا هستند و در موارد فراوانی نوع دیدگاه شرکت‌های داخلی و بین‌المللی به این سندها و واژه‌های به- کاررفته در آن‌ها متفاوت است.

*نویسنده مسئول:

ایمیل: shafiei@profs.semnan.ac.ir

DOI: 10.22056/ijir.2016.03.01

روشهای داده کاوی ارائه دهنده تکنیکهای گوناگون محاسباتی هستند که به طور فزاینده‌ای در حال تبدیل شدن به رویکردهای مهم و حیاتی-اند و به طور کارا و مؤثری اطلاعات جدید را از مجموعه داده‌های عظیم استخراج می‌کنند (Rasmuseen and Karypis, 2004). پایگاه داده‌های متنی به علت افزایش مقدار اطلاعات موجود به فرم الکترونیکی سریع رشد می‌کنند. امروزه بیشتر اطلاعات در صنعت، کسب و کار^۱ و سازمانهای دیگر به صورت الکترونیکی و به فرم پایگاه داده متنی ذخیره شده‌اند (ایمانی، ۱۳۹۱). بیش از ۸۰ درصد از اطلاعات به شکل متن ذخیره می‌شود (Gupta and Lehal, 2009). افزایش روزافزون این اطلاعات باعث ایجاد انگیزه برای محققان در استفاده از ابزارهای نوین و کشف دانش و الگوهای موجود در این اطلاعات شده است.

متن کاوی^۲ علمی میان‌رشته‌ای است که شامل زبان‌شناسی محاسباتی^۳، آمار، و یادگیری ماشین^۴ است. فرایند کاوش در اسناد متنی شامل تحلیل زبان‌شناسی و معنایی متن و در نهایت ساختار آن است. در نتیجه این فرایند، صفات پنهان متن آشکار می‌شود (Al-Azmi, 2013). از خوشه‌بندی اسناد^۵ می‌توان به عنوان یکی از تکنیکهای حوزه متن کاوی در هوش مصنوعی نام برد. خوشه‌بندی اسناد، به عنوان یکی از روشهای یادگیری ماشین بدون ناظر^۶، در زمینه‌های مختلف پردازش زبانهای طبیعی از قبیل بازیابی اطلاعات^۷، خلاصه‌سازی چندمتنی خودکار و ... کاربرد گسترده‌ای دارد (ایمانی، ۱۳۹۱). چشم‌انداز و مأموریت، عصاره جهت‌گیریهای سازمان هستند و چشم‌انداز بر آینده و مأموریت، به حال حاضر شرکت تمرکز دارد. اکثریت مدیران و رهبران بر این باورند که اهداف، راهبردها، قدرت عملیات درونی، و مزایای کلیدی رقابتی سازمانها فرصت ظهور و بروز آنها را در سطح محیط خارجی مهیا می‌کند و این مهم در پیوند با چشم‌انداز، مأموریت و موجودیت سازمان قرار دارد (Gamble and Thompson, 2011).

این تحقیق بر آن است که با استفاده از متن کاوی و در قالب روش خوشه‌بندی به بررسی چشم‌انداز، بیانیه مأموریت سازمانی و اهداف ۳۹ شرکت بیمه بین‌المللی موجود در لیست ۵۰۰ شرکت برتر دنیا و ۱۸ شرکت بیمه داخلی بپردازد و در پی پاسخگویی به این است که آیا مفاهیم به کاررفته در متن اسناد شرکتهای بیمه داخلی با شرکتهای پیشرو بیمه جهانی هم‌راستایی دارد؟ شرکتهای بیمه تا چه میزان مفاهیم موردنظر چشم‌انداز و مأموریت صنعت بیمه کشور در افق بیست‌ساله را مدنظر قرار داده‌اند.

مروری بر پیشینه پژوهش

چشم‌انداز و مأموریت سازمانی

بیانیه مأموریت به عنوان یک ابزار مدیریتی قوی که قادر به برانگیختن کارکنان و متمرکز کردن آنها بر اهداف سازمانی است، شناخته شده است (Brown and Yoshioka, 2003). بیانیه مأموریت می‌تواند یک ابزار مهم برای برقراری ارتباط با ذی‌نفعان داخلی و خارجی باشد (Bartkus, et al., 2000). گاهی اوقات مأموریت و چشم‌انداز با یکدیگر اشتباه گرفته می‌شوند (Dess et al., 2008). حال آنکه بیانیه مأموریت سندی است که علت و فلسفه وجودی سازمان را بیان می‌کند (David, 2007). اما چشم‌انداز به عنوان آنچه که سازمان قصد دست یابی به آن را در آینده دارد تعریف می‌شود (Grant, 2002). چشم‌انداز و مأموریت دو مقوله بهم وابسته و در هم تنیده‌اند. محتوای این اسناد به دلیل تأثیر بر رفتار افراد و تخصیص منابع از اهمیت فراوانی برخوردار است (Verna, 2010). تحقیقات انجام‌گرفته در حوزه اسناد چشم‌انداز و مأموریت نشان‌دهنده ارزش چشم‌انداز و اصول راهنما در کمک به شکل‌دادن به روابط بین کارکنان درون سازمان و افراد خارج سازمان، به‌ویژه مشتریان و ذی‌نفعان است (Morphew and Hartley, 2006; Fox, 2006; Verna, 2010).

در تحقیقات فراوانی متن مأموریت و چشم‌انداز شرکتها مورد مطالعه قرار گرفته است. برژنیک^۸ (۲۰۱۲) به بررسی محتوای اسناد مأموریت شرکتهای مطرح اسلونیایی پرداخت. ابعاد استخراج‌شده از تحلیل شبکه واژه‌ها در چهار مفهوم، "کیفیت زندگی"، "رشد"، "همکاری و

1. Business

2. Text Mining

3. Computational Linguistics

4. Machine Learning

5. Document Clustering

6. Unsupervised Machine Learning

7. Information Retrieval

8. Breznik

نوآوری، و "مشتري" ارائه شد. دیوید^۱ (۲۰۰۷ و ۱۹۸۹) و پیرس^۲ و دیوید (۱۹۸۷) با بررسی ۵۰۰ شرکت منتخب مجله فورچون^۳ به لحاظ عملکرد بالا و پایین، به این نتیجه رسیدند که این شرکتها به لحاظ وجود سه عنصر از ۹ عنصر در بیانیه مأموریت خود که شامل فلسفه سازمان، برداشت از خود، و توجه به تصور مردم است، نسبت به سایر شرکتها تفاوت زیادی دارند. مطالعه اخیر برزنيک و درمول^۴ (۲۰۱۲) بیانگر ۵ مفهوم کلیدی شامل کیفیت زندگی، کارایی شرکت، همکاری و نوآوری، جهت‌گیری رشد، و مشتری‌مداری است. حسین (۲۰۰۴) با بررسی بیانیه مأموریت ۱۵ شرکت بنگلادشی براساس مدل دیوید به این نتیجه رسید که این شرکتها در بیانیه مأموریت خود به ویژگیهای اساسی این سند توجهی نمی‌کنند. احمد و همکاران (۲۰۱۳) با بررسی بیانیه مأموریت ۱۶۹ بانک اسلامی و استفاده از تکنیک شمارش ساده^۵ و تحلیل خوانایی^۶ به این نتیجه دست یافتند که خوانایی و درک مأموریت این بانکها برای ذی‌نفعان بسیار سخت است و بایستی مورد بازنگری جدی قرار گیرد. ویلیامز^۷ (۲۰۰۸) در مطالعه‌ای، مأموریت شرکتها برتر (موجود در لیست فورچون ۱۰۰۰) را با شرکتهایی که کارآمدی پایین داشتند، با روشهای تحلیل محتوا و آنالیز متن مورد مقایسه قرار داد. یافته‌های تحلیل محتوای این مطالعه نشان می‌دهد که محتوای ویژگیهای مشخص شده توسط دیوید (۱۹۸۹) هنوز در مأموریت فعلی شرکتها وجود داشته و برای انتقال مهم‌ترین مفاهیم به ذی‌نفعان داخلی و خارجی استفاده شده و در دو دهه اخیر به طرز چشمگیری تغییر نکرده است.

متن‌کاوی

متن‌کاوی به معنی جستجوی الگوها در متون غیرساخت‌یافته است. سه روش اساسی در مواجهه با این حجم وسیع از اطلاعات غیرساخت یافته وجود دارد که عبارت‌اند از: بازیابی اطلاعات^۸، استخراج اطلاعات^۹ و پردازش زبان طبیعی^{۱۰}. بازیابی اطلاعات که بیشتر بازیابی مستندات و مدارک، با توجه به نیاز مطر شده از سوی کاربر است، مرتبط‌ترین متون و مستندات و یا در واقع لیست کلمه‌ها را از میان مستندات یک مجموعه بیرون می‌کشد. پردازش زبان طبیعی یک درک بهتر از زبان طبیعی توسط کامپیوترهاست. هدف کلی آن استفاده از تکنیکهای آنالیز زبان‌شناسی برای پردازش کردن متن است. و درنهایت هدف روشهای استخراج اطلاعات، استخراج اطلاعات خاص از سندهای متنی است. استخراج اطلاعات می‌تواند به‌عنوان یک فاز پیش‌پردازش در متن‌کاوی به کار برود (ایمانی، ۱۳۹۱).

در خوشه‌بندی سعی می‌شود تا داده‌ها به خوشه‌هایی تقسیم شوند که شباهت بین داده‌های درون هر خوشه حداکثر و شباهت بین داده‌های درون خوشه‌های متفاوت حداقل شود. از جمله کارکردهای خوشه‌بندی یافتن سندهای متنی مشابه است (ایمانی، ۱۳۹۱). مطالعات فراوانی در حوزه متن‌کاوی و کاربردهای آن صورت گرفته است. یک مطالعه موردی در زمینه استفاده از متن‌کاوی در تحلیل ثبت اختراع^{۱۱} در سال ۲۰۰۳ انجام گرفته است. نویسندگان مقاله نشان دادند که نرم‌افزار PackMOLE نسبت به تکنیکهای تحلیل دستی ثبت اختراعها مزایای بیشتری دارد (Fattori et al., 2003). متن‌کاوی در حوزه بهداشت و درمان و پزشکی به شکل برقراری ارتباط چند صد پرونده پزشکی با هم و پیدا کردن روابط بین علائم بیماری و نسخه‌های تجویزی به کار رفته است (Smith, 2010). متن‌کاوی در مسائل مربوط به رسانه به‌ویژه از لحاظ جنبه‌های سیاسی و مسائل بحث‌برانگیز کاربرد فراوانی دارد. در زمینه خوشه‌بندی، اسناد آ‌شیشوده به تعدادی خوشه و دسته‌بندی معنایی از پیش تعیین‌شده نیز به کار می‌رود (Neto et al., 2000). متن‌کاوی اخیراً در زمینه‌های قضایی و قانونی مورد استفاده قرار گرفته و در خلاصه‌سازی متون استفاده شده است که به طور مؤثری کاراکترها، اسامی اصلی، افعال و بیشترین کلمات و منابع مستندات حجیم را مشخص می‌کند (Al-Azmi, 2013).

1. David

2. Pearce

3. 500 Top Fortune Companies

4. Breznik and Dermol

5. Simple Counting Techniques

6. Readability Analyses

7. Williams

8. Information Retrieval

9. Information Extraction

10. Natural Language Processing (NLP)

11. Patent

کولیشکینا و روین^۱ (۲۰۰۵) نتایج تحلیل کاربرد متن‌کاوی در پایگاه خسارات بیمه را ارائه کردند. آنها با استفاده از متن‌کاوی در فرمهای شکایات، مفاهیمی را از این توضیحات استخراج کردند. فرانسسیس^۲ (۲۰۰۶) و ویز^۳ و همکاران (۲۰۱۰) در به‌کارگیری خوشه‌بندی و سایر تکنیکهای بدون ناظر در متن‌کاوی، مقاله مشروح و مفصلی ارائه کرده‌اند. داوی^۴ و همکاران (۲۰۰۵) استفاده از داده‌کاوی در گامهای کلیدی تحلیل متنهای غیرساختاریافته را نشان داده‌اند.

محققین از متن‌کاوی نیز در حوزه‌های علوم پزشکی، (Li et al. 2012)، حوزه‌های آموزشی (Abdous and He, 2011; Hung, 2012) و کسب‌وکار (Ingvaldsen and Gulla, 2012) استفاده کرده‌اند. ویتن^۵ و همکاران (۲۰۰۳) از تکنیکهای متن‌کاوی برای استخراج فراداده^۶ از اسناد موجود در یک کتابخانه و کمک به غنی‌سازی اسناد با پررنگ کردن آیتمهای مناسب در متن استفاده کردند و دریافتند که متن‌کاوی می‌تواند ارزش افزوده‌ای را به اسناد کتابخانه دیجیتال اضافه و تجربه کاربران از کتابخانه را غنی‌تر کند. فولر^۷ و همکاران (۲۰۱۱) از متن‌کاوی برای شناسایی فریب و دروغ در داده‌های دنیای واقعی استفاده کردند. هانگ^۸ (۲۰۱۲) با استفاده از تحلیل خوشه‌بندی به‌عنوان یک تکنیک اکتشافی، ادبیات آموزش الکترونیک را واکاوی کردند و به نمایش الگوهای موجود در آن براساس کلمات مشابه، ارزش صفات، و قوانین-کدشده پرداخت.

به طور ویژه، تحلیل خوشه‌ای یکی از تکنیکهای مهم داده‌کاوی است و مزایای زیادی در شناسایی روندهای غیرقابل پیش‌بینی، همبستگیها و الگوهای بین داده‌ها دارد (Ananiadou, 2008). یکی دیگر از حوزه‌های مهمی که متن‌کاوی به آن ورود کرده است، رسانه‌های اجتماعی است. با توجه به قابلیت‌های قدرتمند متن‌کاوی، اعتقاد محققین بر این است که استفاده از متن‌کاوی در داده‌های رسانه‌های اجتماعی می‌تواند منجر به کشف یافته‌های جذابی در حوزه رفتار و تعامل انسان شود (Pang and Lee, 2008; Barbier and Liu, 2011; Abdous et al., 2012; He, 2013)

اگرچه متن‌کاوی در حوزه‌های مختلفی به کار رفته است اما در زمینه بررسی متن چشم‌انداز و مأموریت سازمانی، تنها مطالعه صورت‌گرفته، الشامری^۹ و همکاران (۲۰۱۲) است. این محققان با استفاده از رویکرد متن‌کاوی چشم‌انداز و مأموریت، ۷۷۲ شرکت حاضر در فورچون ۵۰۰ را خوشه‌بندی کرده و نشان دادند که چگونه این خوشه‌ها را می‌توان با مفاهیم جهانی شدن، نوآوری، و ویژگیهای استراتژی تبیین و برای ارزیابی هم‌راستایی، تثبیت و جهت‌یابی استراتژیک استفاده کرد.

تحلیل

چارچوب تحقیق

در تحقیق حاضر، محققین برآنند تا اطلاعات معنی‌داری از چشم‌انداز، مأموریت سازمانی، و اهداف شرکتهای مطرح بیمه بین‌المللی و ملی در قالب خوشه‌بندی این متون و استخراج مهم‌ترین ویژگیهای هر خوشه ارائه کنند. اهداف این تحقیق را می‌توان بدین شکل مطرح کرد: ۱. استخراج کلمات مهم و پرتکرار در چشم‌انداز، مأموریت، و اهداف شرکتهای بیمه ملی و بین‌المللی، ۲. خوشه‌بندی شرکتهای بیمه ملی و بین‌المللی براساس 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 تحقیق ارائه شده است.

1. Kolyshkina and Rooyen

2. Francis

3. Weiss

4. Davi

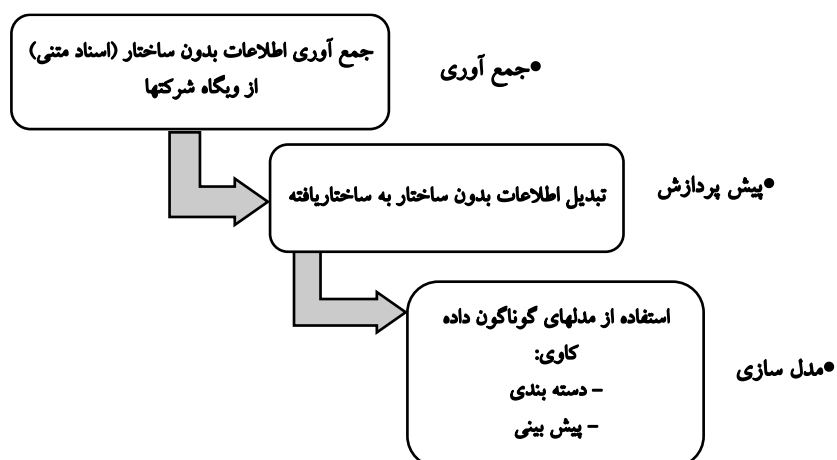
5. Witten

6. Meta Data

7. Fuller

8. Hung

9. Alshameri



شکل ۱: مراحل اجرای تحقیق

جمع‌آوری داده‌ها

مجموعه داده‌های مورد استفاده در این تحقیق شامل متن چشم‌انداز و مأموریت، و اهداف سازمانی تعدادی از شرکتهای بیمه بین‌المللی و داخلی است که به شرح ذیل جمع‌آوری شده است:

شرکت بیمه بین‌المللی که در لیست فورچون ۵۰۰ (۵۰۰ شرکت برتر دنیا) قرار گرفته‌اند و داده از وبگاههای مربوطه استخراج شده است. وبگاه اول^۱ شامل متن سند مأموریت و چشم‌انداز شرکتهاست که به دسته‌های مختلف تقسیم شده‌اند، مانند ساخت‌وساز، حقوقی، تولید، املاک و مستغلات، فن‌آوری، و غیره. وبگاه دوم^۲ شامل سند مأموریت سازمانی ۵۰۰ شرکت برتر جهانی است که به ترتیب حروف الفبا ذکر شده و اطلاعات گسترده‌تری مربوط به هر شرکت فراهم می‌کند.

۱۸ شرکت بیمه داخلی که نام آنها از وبگاه مربوط به بیمه مرکزی^۳ استخراج شده و سپس به وبگاه هر شرکت مراجعه و اطلاعات مورد نیاز به دست آمده است.

چشم‌انداز و مأموریت شرکتهای به صورت رکوردهای مجزا در فایل اکسل وارد شده و در مرحله بعد در نرم‌افزار متن‌کاوی ریپیدماینر^۴ وارد شده است. براساس گزارش وبگاه^۵ مربوطه تقریباً نیمی از متخصصان متن‌کاوی از این نرم‌افزار استفاده می‌کنند. در تحقیق حاضر از افزونه متن-کاوی موجود در نرم‌افزار ریپیدماینر برای متن‌کاوی متون و در نهایت خوشه‌بندی آنها توسط الگوریتم k- میانگین استفاده شده و نرم-افزار WORDLE در نمایش گرافیکی خوشه‌ها و مفاهیم به کار رفته است. نرم‌افزار WORDLE برای تولید "اِبْرهای کلمه"^۶ از داده‌های متنی فراهم شده توسط محقق استفاده می‌کند. اِبْرها، کلماتی را که در متن پرتکرارتر هستند، برجسته‌تر می‌کند و با فونتهای مختلف، طرح‌بندی و رنگهای گوناگون نمایش می‌دهند (Alshameri et al., 2012).

پیش‌پردازش داده‌ها

قبل از اعمال هر گونه مدلی بایستی فرمت متنها که غیرساختاریافته است به ساختاریافته تبدیل شود. به عبارتی متن به داده تبدیل شود تا بتوان انواع مدل‌های داده‌کاوی را بر آن اجرا کرد. مرحله پیش‌پردازش داده‌ها دارای چندین گام است:

1. http://www.missionstatements.com/company_mission_statements.html

2. <http://www.company-statements - slogans.info/fortune-500-companies.htm>

3. <http://www.centinsur.ir/>

4. Rapid Miner

5. www.rapid-i.com

6. Word Clouds

معمولاً اولین گام در ساختاردهی به متون، تفکیک آن به اجزای تشکیل دهنده‌اش، یعنی کلمات است. اصطلاحاً به این مرحله تفکیک به کلمات^۱ می‌گویند.

گام دوم پیش‌پردازش، حذف کلمات عمومی از جمله حروف اضافه، حروف ربط، قیود و... است که نقش چندانی در بار معنایی جملات ایفا نمی‌کنند و از آن به حذف کلمات توقف^۲ یاد می‌شود.

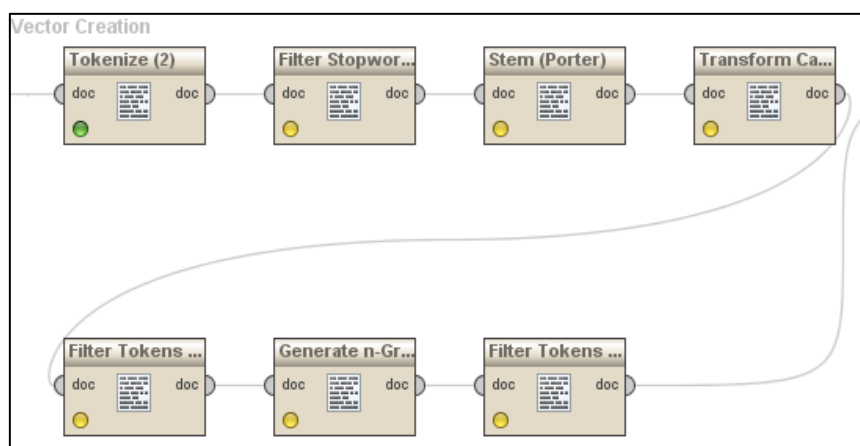
در گام بعد کلماتی که ریشه یکسان دارند به یک شکل در نظر گرفته می‌شوند. مثلاً کلمات مسئولیت‌پذیری و مسئولیت به ریشه خود که مسئول است برگردانده و یکسان در نظر گرفته می‌شوند. به این مرحله ریشه‌یابی^۳ می‌گویند.

در گام بعد، حساسیت به کوچک و بزرگ‌بودن حروف^۴ حذف می‌شود.

بعضی از کلمات به صورت جفت یا حتی سه‌بخشی در متون ظاهر شده و معنی می‌یابند. در این مرحله که اصطلاحاً دنباله‌ای از یک یا چند کلمه^۵ نامیده می‌شود، این مورد لحاظ می‌شود.

گام آخر فیلتر کلمات براساس معنی^۶ است که در آن کلمات خاصی از متن به تشخیص محقق استخراج می‌شود و مورد مطالعه قرار می‌گیرد.

شکل ۲ گام‌های مرحله پیش‌پردازش داده‌ها را نشان می‌دهد.



شکل ۲: پیش‌پردازش داده‌ها در ریپدماینر

پس از اجرای مراحل فوق، متون موردنظر به کلمات تبدیل شده و در ساختاری جدول‌گونه که هر رکورد نشان‌دهنده یکی از اسناد و هر ستون یکی از کلمات است، ارائه می‌شود. در جدولهای ۱ و ۲ نمونه‌ای از خروجی نرم‌افزار شامل بردار فراوانی کلمه-عکس فراوانی سند (TF/IDF^۷) در مجموعه متنها نمایش داده شده است.

1. Tokenize
2. Filter Stop Words
3. Stemming
4. Transform Case
5. n-Grams
6. Filter Tokens (by Content)
7. Term Frequency-Inverse Document Frequency

جدول ۱: بردار TF-IDF و کلمات پرتکرار در اسناد شرکتهای بیمه بین‌المللی

ردیف	financi	Product_servic	qualiti	secur	Servic	Valu
۱	۰/۵۰۹	۰	۰/۵۸۶	۰/۶۳	۰	۰
۲	۰	۰/۵۵۳	۰/۵۹۴	۰	۰/۱۹	۰/۵۵۳
۳	۰	۰	۱	۰	۰	۰

ادامه جدول ۱:

کلمه	کل رخداد	تعداد سند
Servic	۳۵	۲۶
Financi	۳۱	۱۳
Product_servic	۱۳	۱۲
Valu	۱۵	۱۲
Qualiti	۱۱	۱۱
Secur	۱۰	۱۰

جدول ۲: بردار TF-IDF و کلمات پرتکرار در اسناد شرکتهای بیمه داخلی

ردیف	آرامش	ارائه خدمات	حرفه‌ای	سودآوری	مالی
۱	۰	۰/۳۲۶	۰/۹۴۵	۰	۰
۲	۰	۰/۴۱	۰/۵۹۵	۰/۶۹۲	۰
۳	۰	۱	۰	۰	۰

کلمه	کل رخداد	تعداد سند
ارائه خدمات	۱۶	۱۳
مالی	۱۰	۱۰
حرفه‌ای	۷	۷
آرامش	۷	۶
سودآوری	۹	۶

خوشه‌بندی

خروجی مرحله پیش‌پردازش داده‌ها، به شکل لیست کلمات و بردار اسناد است. لیست کلمات براساس تعداد اسنادی که کلمه در آن رخ داده و تکرار کلمه در کل اسناد است. به طور مثال بر طبق جدول ۱، کلمه خدمات در ۲۶ سند از ۳۹ سند موجود در مجموع ۳۵ بار تکرار شده است. با توجه به اینکه بعضی از واژه‌ها در تعداد کمی از اسناد رخ داده‌اند و در سوی مقابل تعدادی نیز به شکل عمومی در بیشتر اسناد تکرار شده‌اند با عمل هرس کردن^۱، در اسناد مربوط به شرکتهای بین‌المللی واژه‌هایی که در بیش از ۱۰ سند تکرار شده‌اند لحاظ شده و در شرکتهای داخلی کلمات با تکرار بیش از ۶ سند مورد مطالعه قرار گرفته است.

¹. Prune

در تحقیق حاضر از روش خوشه‌بندی k- میانگین استفاده شده است. در این روش تعداد خوشه‌هایی (k) که باید داده‌ها در نهایت درون آنها قرار بگیرد از ابتدای فرایند مشخص می‌شود. در پژوهش حاضر پس از بررسی مقادیر مختلف k در بازه ۲ تا ۸، در نهایت تعداد ۴ خوشه برای مجموعه شرکت‌های بین‌المللی (جدول ۳) و تعداد ۴ خوشه برای شرکت‌های داخلی در نظر گرفته شد (جدول ۴).

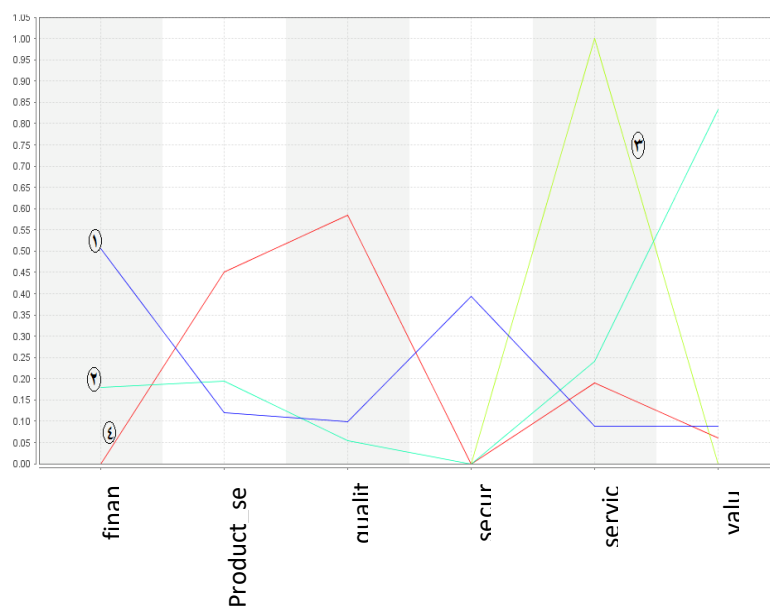
جدول ۳: خوشه‌بندی شرکت‌های بیمه بین‌المللی

خوشه ۱	خوشه ۲	خوشه ۳	خوشه ۴	صفت
۰/۵۰۵	۰/۱۸۰	۰	۰	Financi
۰/۱۲۱	۰/۱۹۵	۰	۰/۴۵۰	Product_servic
۰/۰۹۸	۰/۰۵۵	۰	۰/۵۸۴	Qualiti
۰/۳۹۳	۰	۰	۰	Secur
۰/۰۸۷	۰/۲۴۲	۱	۰/۱۸۹	Servic
۰/۰۸۸	۰/۸۳۲	۰	۰/۰۶۱	Valu

جدول ۴: خوشه‌بندی شرکت‌های بیمه داخلی

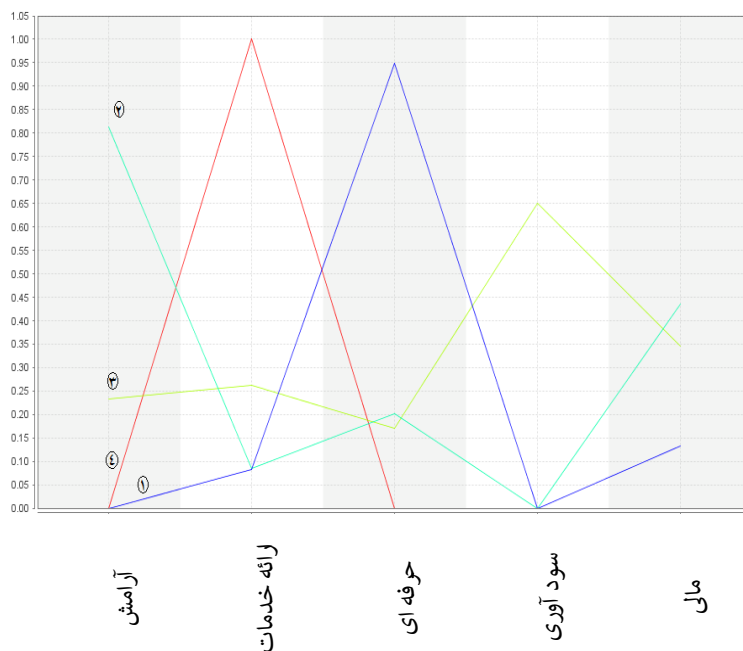
خوشه ۱	خوشه ۲	خوشه ۳	خوشه ۴	صفت
۰	۰/۸۱۳	۰/۲۳۲	۰	ارائه
۰/۰۸۱	۰/۰۸۴	۰/۲۶۲	۱	جامعه
۰/۹۴۹	۰/۲۰۱	۰/۱۶۹	۰	خدمات
۰	۰	۰/۶۵۱	۰	سرمایه
۰/۱۳۲	۰/۴۳۵	۰/۳۴۶	۰	مالی

در ادامه، رابطه بین کلمات و خوشه‌ها در قالب نمودار به نمایش درآمده که به این وسیله می‌توان مفاهیم و کلماتی که بیشترین تفکیک را بین خوشه‌ها ایجاد می‌کنند، استخراج کرده و به‌عنوان برچسب خوشه مورد استفاده قرار داد. شکل ۲ نمودار مربوط به شرکت‌های بیمه بین‌المللی است. براساس این نمودار کلمه Financi (ریشه لغت مالی) و Secur (ریشه لغت امنیت)، خوشه اول را نسبت به سایر خوشه‌ها تفکیک کرده و به‌عنوان برچسب این خوشه انتخاب می‌شوند. در خوشه دوم کلمه Valu (ریشه لغت ارزش) به‌عنوان برچسب خوشه انتخاب می‌شود. به همین شکل، خوشه سوم برچسب Servic (ریشه لغت خدمت) و برای خوشه چهارم، کلمات Qualiti (ریشه لغت کیفیت) و Product_Servic به‌عنوان برچسب خوشه و مهم‌ترین کلمات تفکیک‌کننده خوشه از سایر خوشه‌ها انتخاب می‌شوند.



شکل ۲: نمودار رابطه بین کلمات و خوشه‌ها در شرکت‌های بیمه بین‌المللی

شکل ۳ نمودار مربوط به شرکت‌های بیمه داخلی است. براساس استدلال فوق به ترتیب در خوشه اول کلمه حرفه‌ای به‌عنوان برچسب انتخاب شده و در خوشه دوم کلمه آرامش به‌عنوان برچسب خوشه انتخاب شده و در خوشه سوم، کلمه‌های سودآوری و مالی و در خوشه چهارم کلمه ارائه خدمات به‌عنوان برچسب خوشه انتخاب شده است.



شکل ۳: نمودار رابطه بین کلمات و خوشه‌ها در شرکت‌های بیمه داخلی

در شکل‌های ۴ و ۵ خوشه‌های شکل‌گرفته، برچسب هر خوشه و شرکت‌های مربوطه نمایش داده شده است.



شکل ۴: برچسب خوشه‌های شرکت‌های بیمه بین‌المللی و نام شرکت‌ها



شکل ۵: برچسب خوشه‌های شرکت‌های بیمه داخلی و نام شرکت‌ها

نتایج و بحث

پس از عملیات متن‌کاوی بر چشم‌انداز، مأموریت، و اهداف نمونه‌ای از شرکتهای بیمه بین‌المللی و داخلی، مفاهیم قابل تأملی در قالب برچسب خوشه‌ها ارائه شد. در جدول ۵، این مفاهیم نمایش داده شده است. با بررسی جدول ۵ و مطالعه جداگانه سند چشم‌انداز، مأموریت سازمانی و اهداف شرکتهای براساس مفاهیم استخراج‌شده از خوشه‌ها، نتایج ذیل حاصل شد:

مقایسه خوشه ۱ (شرکت‌های بین‌المللی) و خوشه ۳ (شرکت‌های داخلی):

شرکتهای بین‌المللی در ۵ مورد به طور خاص واژه امنیت مالی^۱ را استفاده کرده‌اند و در سایر موارد، کلمه مالی در متن اسناد به طور دقیق جهت‌دهی شده است و مفاهیمی چون قدرت مالی^۲، برنامه‌ریزی مالی^۳، نیازهای مالی^۴، شبکه مالی^۵، راه‌حل مالی^۶، محصولات رقابتی مالی^۷ و ... به کار رفته است. در مقابل شرکتهای داخلی به طور عمومی از واژه‌های سودآوری و مالی استفاده کرده‌اند و در متن اسناد، با اشاره به افزایش توانگری مالی شرکت، به افزایش سودآوری سهامداران و مشتریان پرداخته‌اند.

مقایسه ترکیب خوشه‌های ۳ و ۴ (شرکتهای بین‌المللی) و خوشه ۴ (شرکتهای داخلی):

شرکتهای بین‌المللی که متعلق به خوشه ۳ هستند، مفهوم خدمات را به شکل عمومی در اسناد خود به کار برده‌اند در حالی که در خوشه ۴ این مفاهیم جهت‌دهی خاص دارد و مفاهیمی چون خدمت با کیفیت و خدمت‌نوآورانه^۸ دیده می‌شود. در خوشه ۴ (شرکتهای داخلی)، بجز یک مورد که توسعه خدمات را منوط به توسعه شبکه فروش دانسته در بقیه موارد ارائه خدمات با کیفیت، نوآورانه و خلاقانه مورد تأکید قرار گرفته است که با شرکتهای بین‌المللی مشترک است.

بررسی خوشه ۲ (شرکتهای بین‌المللی) نشان می‌دهد که واژه ارزش در اسناد این شرکتها تبیین شده و در قالب کلماتی چون صداقت^۹، تنوع^{۱۰}، نوآوری، برتری^{۱۱}، نظم^{۱۲}، کارایی^{۱۳} و اعتماد متقابل^{۱۴} نگاشته شده است. بدین شکل شرکتها درون‌مایه واژه ارزش را به مشتریهای خود نشان داده‌اند.

بررسی خوشه ۱ (شرکتهای داخلی) که برچسب حرفه‌ای را به خود اختصاص داده‌اند گویای این مطلب است که استفاده از تخصص حرفه‌ای، شبکه فروش حرفه‌ای و ابزارهای حرفه‌ای سرلوحه اهداف شرکتهای این خوشه است. البته در مورد حرفه‌ای عمل کردن این شرکتها روشنگری خاصی انجام نگرفته است.

بررسی خوشه ۲ (شرکتهای داخلی) نشان می‌دهد شرکتهای حاضر در این خوشه واژه آرامش را بدون هیچ‌گونه روشنگری و تعریف از آرامش به‌عنوان رویکردی در چشم‌انداز و مأموریت و اهداف خود ذکر کرده‌اند.

جدول ۵: مفاهیم و برچسب خوشه‌ها

برچسب خوشه (شرکتهای بین‌المللی)	برچسب خوشه (شرکتهای بیمه داخلی)
امنیت- مالی	حرفه‌ای
ارزش	آرامش
خدمات	سودآوری- مالی
محصول- خدمت و کیفیت	ارائه خدمات

با بررسیهای فوق و در مجموع می‌توان به این نتیجه دست یافت که شرکتهای بیمه بین‌المللی در اسناد چشم‌انداز و مأموریت خود از واژه‌های دقیق و تعریف‌شده استفاده کرده‌اند و رویکرد شرکت را به طور واضح برای مشتری مشخص کرده‌اند. علت این امر این است که این شرکتها به دنبال فعالیت در کلاس جهانی بوده و بایستی به طور دقیق جهت‌گیری اقتصادی مبتنی بر ارزش افزوده در بیانیه مأموریت و چشم‌انداز آنها لحاظ شود و سطح مشترک اقتصادی ذی‌نفعان برایشان مهمتر از فرهنگ یک‌بعدی مختص به یک کشور است. در مقابل

1. Financial Security
2. Financial Strength
3. Financial Planning
4. Financial Needs
5. Financial Network
6. Financial Solution
7. Competitive Financial Products
8. Innovative
9. Integrity
10. Diversity
11. Excellence
12. Clarity
13. Performance
14. Mutual Trust

شرکتهای ایرانی، بجز مفهوم خدمات و محصولات باکیفیت و نوآورانه که جزء لاینفک مأموریت و چشم‌انداز اکثر شرکتهاست، در سایر مفاهیم چندان خود را موظف به ارائه جزئیات ندانسته و به واژه‌های کلی، آرامش و سودآوری قناعت کرده‌اند. بدون شک حضور در سطح بین‌المللی نیازمند تجدیدنظر در مأموریت و چشم‌انداز این شرکتها و تطابق با دیدگاه‌های جهانی است.

سند چشم‌انداز صنعت بیمه کشور در افق بیست ساله صنعت بیمه ایران، برگرفته از سایت بیمه مرکزی ج.ا.ا. مؤید این است که "بیمه در افق ۱۴۰۴ صنعتی است اقتصادی، عدالت‌محور، پایدار، سالم و قابل اعتماد، برخوردار از اخلاق حرفه‌ای و عجزین شده با جامعه، که اطمینان را برای روند طبیعی زندگی آحاد مردم ایران و چرخه فعالیتهای اقتصادی (کشاورزی، صنعتی و خدماتی) کشور به شیوه‌ای آسان، سریع، مشتری-مدار و با کمترین هزینه و به شکلی همه جانبه تأمین کرده و ضمن تعامل سازنده و مؤثر در سطح بین‌الملل، در جایگاه اول صنعت بیمه بین‌بیمه‌گران منطقه آسیای جنوب‌غربی (شامل آسیای میانه، قفقاز، خاورمیانه و کشورهای همسایه) قرار دارد".

بیانیه مأموریت صنعت بیمه کشور "تولید و عرضه اطمینان برای چرخه فعالیتهای اقتصادی در بخش کشاورزی، صنعتی و خدماتی کشور و جریان زندگی آحاد مردم با حداکثر کیفیت و قیمت متناسب با آن، به صورتی همه‌جانبه، مکفی و پایدار" است.

بررسی دو متن فوق و هم‌راستایی آن با نتایج حاصل از پژوهش محققین حاکی از این واقعیت است که مفاهیمی چون اخلاق حرفه‌ای، اطمینان، و کاهش هزینه که در متن سند چشم‌انداز بیست ساله صنعت بیمه ذکر شده است در مفاهیم مستخرج از خوشه‌بندی شرکتهای بیمه داخلی در قالب واژه‌های حرفه‌ای‌بودن، آرامش، و سودآوری-مالی مشاهده شده است. در مقابل، مفاهیمی چون مشتری‌مداری، عدالت-محوربودن، حوزه فعالیت (کشاورزی، صنعتی و خدماتی) و محدوده جغرافیایی فعالیت شرکتهای مورد تأکید قرار نگرفته است. در متن بیانیه مأموریت صنعت بیمه کشور به کیفیت و پایداربودن خدمات توجه ویژه شده اما در متن اسناد شرکتهای داخلی به کیفیت توجه خاصی نشده است در حالی که شرکتهای بیمه بین‌المللی به این مقوله توجه کامل داشته‌اند.

جمع‌بندی و پیشنهادها

مدلهای زیادی در حوزه مأموریت سازمانی و مفاهیم مؤثر در تدوین آن وجود دارد، از جمله این مدلها، مدل ۹ شاخصه دیوید است که می‌توان به بررسی متن مأموریت سازمانی شرکتهای پرداخته و مفاهیم حاصل از متن کاوی را با مفاهیم ارائه‌شده در مدل دیوید (۲۰۰۷) مقایسه کرد و میزان هم‌راستایی بیانیه مأموریت شرکتهای و مدل ذکرشده را مورد تحلیل قرار داد. در ضمن پیشنهاد می‌شود مهمترین ابعاد حاصل از خوشه‌بندی تحقیق حاضر را با ابعاد کارت امتیازی متوازن سازمانها و نحوه عملکردشان مقایسه و در پژوهشی دیگر مورد مطالعه قرار داد. بررسی و متن کاوی گزارش شکایات بیمه‌گذاران در تغییر رفتار و استراتژی شرکتهای بیمه بسیار مؤثر است و این مهم با انجام تحقیقات مستمر و ورود علوم جدیدی چون داده‌کاوی و به طور اخص، متن‌کاوی، قابل حصول است.

منابع و ماخذ

ایمانی، م.، (۱۳۹۱). خوشه‌بندی متون فارسی. پایان‌نامه کارشناسی ارشد. تهران: دانشگاه علم و صنعت.

Abdous, M.; He, W., (2011). Using text mining to uncover student's technology related problems in live video streaming. *British Journal of Educational Technology*, 40(5), pp. 40–49.

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), pp. 77–88.

Ahmed, I.; Shaukat, M.Z.; Islam, T., (2013). Mission statements readability: an insights into islamic banks. *Journal of Islamic Accounting and Business Research*, 4 (2), pp.132-150.

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).

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 (Online)*, 16(3), p.227.

Ananiadou, S., (2008). National centre for text mining: introduction to tools for researchers. Retrieved February, 8, 2009.

- Barbier, G.; Liu, H., (2011). Data mining in social media. *Social Network Data Analytics*, pp.327-352.
- Bartkus, B.R.; Glassman, M.; McAfee, R.B., (2000). Mission statements: are they smoke and mirrors?. *Business Horizons*, 43(6), pp.23-28.
- Breznik, k., (2012). Ranking and visualizing the keywords in mission statements of Slovenian companies. *knowledge and learning: global empowerment. Proceedings of the Management, Knowledge and Learning International Conference 2012.*
- Breznik, K.; Dermol, V., (2012). Identifying the mission statement's clusters in Slovenian economy. *TIIM 2012, Lublin, Poland.*
- Brown, W.A.; Yoshioka, C.F., (2003). Mission attachment and satisfaction as factors in employee retention. *Nonprofit Management & Leadership*, 14 (1), pp.5-18.
- Davi, A.; Haughton, 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), pp.89-103.
- David, F.R., (1989). How companies define their mission. *Long Range Planning*, 22(1), pp.90-97.
- David, F.R., (2007). *Strategic management: concepts and cases.* Upper Saddle, N.J: Pearson Practice.
- Dess, G.G.; Lumpkin, G.T.; Eisner, A.B., (2008). *Strategic management: text and cases.* Boston: McGraw Hill, Irwin.
- Fattori, M.; Pedrazzi, G.; Turra, R., (2003). Text mining applied to patent mapping: a practical business case. *World Patent Information*, 25(4).
- Fox, R., (2006). Using corpus linguistics to describe corporations' ideologies. *Tourism and Hospitality Management*, 12(2), pp.15-24.
- Francis, L.A., (2006). Taming Text: An introduction to text mining. In *Casualty Actuarial Society Forum*, pp.51-88.
- 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), pp.8392-8398.
- Gamble J.E.; Thompson, A.A., (2011). *Essentials of strategic management: the quest for competitive advantage.* McGraw-Hill Erwin, New York.
- Grant, R.M., (2002). *Contemporary strategy analysis: concepts, techniques, applications.* Malden, A, USA: Black Well Publishers.
- Gupta, V.; Lehal, G.S., (2009). A survey of text mining techniques and applications. *Journal of Emerging Technologies in Web Intelligence*, 1(1).
- He, W., (2013). Improving user experience with case-based reasoning systems using text mining and web 2.0. *Expert System with Applications*, 40(2), pp.500-507.
- Hossain, D.M., (2004). An analysis of the mission statements of selected Bangladeshi companies. *Journal of Business Administration*, 6, pp.1-7.
- 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), pp.5-16.
- Ingvaldsen, J.E.; Gulla, J.A., (2012). Industrial application of semantic process mining. *Enterprise Information Systems*, 6(2), pp.139-163.
- Kolyshkina, I.; van Rooyen, M., (2005). Text mining for insurance claim cost prediction. Presented at the XVth General Insurance Seminar Institute of the Actuaries of Australia.
- Li, L.; Ge, R.; Zhou, S.; Valerdi, R., (2012). Guest editorial integrated healthcare information systems. *IEEE Transactions on Information Technology in Biomedicine*, 16(4), pp.515-517.
- Morphew, C.C.; Hartley, M., (2006). Mission statements: a thematic analysis of rhetoric across institutional type. *The Journal of Higher Education*, 77(3), pp.456-471.
- 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.*

- Pang, B.; Lee, L., (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), pp.1-135.
- Pearce, J.A.; David, F., (1987). Corporate mission statements: the bottom line. *Academy of Management Executive*, 1(2), pp.109-116.
- Rasmussen, M.; Karypis, G., (2004). CLUTO-an interactive clustering, visualization, and analysis system. Technical Report TR#04-021, Department of Computer Science, University of Minnesota, Minneapolis, MN.
- Smith, D., (2010). Using data and text mining to drive innovation. PhUSE UK.
- Verna., (2010). The mission statement is a strategic tool: when used properly. *Management Decision*, 40(5), pp.448-455.
- Weiss, S.M.; Indurkha, N.; Zhang, T.; Damerau, F., (2010). *Text mining: predictive methods for analyzing unstructured information*. Springer Science & Business Media.
- Williams, L.S., (2008). The mission statement a corporate reporting tool with a past, present, and future. *Journal of business communication*, 45(2), pp.94-119.
- Witten, I.; Don, K.; Dewsnip, M.; Tablan, V., (2003). Text mining in a digital library. *International Journal of Digital Library*, 5, pp.1-4.