



متن کاوی، ابزاری مؤثر در آینده پژوهی: معرفی روش و مرور ادبیات

مینا رضاییان^{۱*}، نگار جلیلیان^۲

یزد، دانشکده اقتصاد، مدیریت و حسابداری، دانشگاه یزد، mina.rezaian@ymail.com
یزد، دانشکده اقتصاد، مدیریت و حسابداری دانشگاه یزد، negar.jalilian@stu.yazd.ac.ir

چکیده

در محیط رقابتی امروز، شناسایی فرصتهایی که امیددهندهی رشد و پیشرفت هستند در گرو توجه به آینده می باشد که به واسطه‌ی آن می توان دید بلند مدتی نسبت به علوم و فناوری های تازه به دست آورد. آینده نگاری فرآیندی است که به واسطه‌ی آن فهم صحیح از نیروهای شکل دهنده آینده حاصل شده و سبب می شود تا این نیروها در کلیه‌ی مراحل تدوین راهبردها، برنامه ریزی و تصمیم گیری مورد توجه قرار گیرند. ابزارهای مختلفی برای تجزیه و تحلیل آینده و نیز پیش بینی آینده علم، تکنولوژی و نوآوری وجود دارد. در این میان متن کاوی یکی از روش های رایج و پرکاربرد در این حوزه محسوب می شود. در دنیای مدرن که در آن میزان انتشار متون علمی به حدی زیاد است که تقریباً برای هر محقق غیرممکن است که تمام اطلاعاتی را که با آن روبروست، دنبال کند؛ متن کاوی به عنوان ابزاری مفید در اختیار پژوهشگران قرار دارد. در این پژوهش با معرفی روش متن کاوی به عنوان روشی مؤثر در فرایند آینده پژوهی، مطالعاتی که در این زمینه صورت گرفته اند مرور شده و پس از بررسی زمینه استفاده از آن در حوزه های مختلف علمی - تحقیقاتی، زمینه کاربرد این روش در مقالات مختلف در پنج حوزه طبقه بندی متن، بازیابی اطلاعات، خوشه بندی، استخراج اطلاعات و پیش بینی و ارزیابی دسته بندی گردید.

واژه های کلیدی: آینده پژوهی، پیش بینی، متن کاوی

۱- مقدمه

امروزه پیشرفت های تکنولوژیکی جدید، مهمترین نیروی پیشران رشد و نوسازی در زمینه هایی مانند تکنولوژی اطلاعات، تامین انرژی و مهندسی ژنتیک است که پیش بینی می شود، کیفیت زندگی را در آینده افزایش دهند (Bjorn & Lubeck, 2003). بهره برداری درست از این تحولات علم و فناوری نوعی مزیت رقابتی برای حضور قدرتمند در عرصه بین المللی محسوب می گردد و این امر مستلزم آن است که بتوانیم سیاست های علم و فناوری خود را متناسب با تهدیدها و فرصت های مرتبط طراحی کنیم (نوری و ایوبی، ۱۳۸۸؛ زالی، ۱۳۹۰). از آنجا که آینده در بردارنده پدیده عدم قطعیت است، اطلاعات و واقعیات گذشته و حال می تواند رهنمون آینده باشد (زارعیان و ستارزاده، ۱۳۹۰) بنابراین برای داشتن دید وسیع نسبت به آینده (آینده نگری)، باید ابتدا آن را شناخت (آینده شناسی)، و سپس برای شناخت آن تحقیق کرد (آینده پژوهی). به عبارت دیگر، لازمه ی آینده نگری، آینده شناسی و آینده شناسی مستلزم آینده پژوهی است (سلامی و همکاران، ۱۳۹۱).

آینده پژوهی^۳ در اواخر قرن بیستم به یک رشته ی دانشگاهی ضروری تبدیل شد، چرا که آینده چالش های نمایان و سترگی را پیش روی ما قرار داده است و روبرو شدن با آن بستگی به این دارد که ما هرچه سریع تر آینده پژوهی را به عنوان وظیفه ای

۱ دانشجوی کارشناسی ارشد مدیریت اجرایی گرایش تولید و عملیات دانشکده اقتصاد، مدیریت و حسابداری، دانشگاه یزد

۲- دانشجوی دکترای مدیریت صنعتی گرایش تولید و عملیات، دانشکده اقتصاد، مدیریت و حسابداری، دانشگاه یزد

3 foresight



اساسی و محوری تلقی کنیم و زمینه بروز همه آنچه را که در عنوان تفکر آینده اندیشانه می‌گنجد، از سرچشمه‌های اختصاصی آن یعنی دانشگاه‌ها و موسسات تحقیقاتی فراهم آوریم (اسلاتر، ۱۳۸۶). در این راستا آینده‌پژوهی علم و فناوری به عنوان ابزاری برای تحقق این هدف به کار می‌رود (کریمی فرد، ۱۳۸۹). دو تعریف معروف از آینده‌پژوهی توسط محققان بریتانیایی ارائه شده است که دو جنبه از این علم را نشان می‌دهد. بن‌مارتین^۱ (۱۹۹۶) آینده‌پژوهی علم (پژوهش) را تلاشی سیستماتیک به منظور نگاه به آینده بلندمدت علم، فناوری، اقتصاد و جامعه با هدف شناسایی حوزه‌های پژوهشی استراتژیک و فناوری‌های عام نو ظهور معرفی می‌کند. وی بر این باور است که شناسایی چنین حوزه‌هایی می‌تواند به بیشترین منافع اقتصادی و اجتماعی منجر شود. همچنین، جورجیو^۲ (۱۹۹۶) آینده‌پژوهی فناوری را "ابزاری سیستماتیک برای ارزیابی پیشرفت‌های علمی و تکنولوژیکی که می‌تواند تاثیر زیادی بر رقابت صنعتی، ایجاد ثروت و کیفیت زندگی داشته باشد" می‌داند.

روش‌شناسی آینده‌نگاری فرآیندی است که به واسطه‌ی آن فهم صحیح از نیروهایی که آینده را شکل می‌دهند، حاصل شده و سبب می‌شود تا این نیروها در کلیه‌ی مراحل تدوین راهبردها، برنامه‌ریزی و تصمیم‌گیری مورد توجه قرار گیرند. هدف از این رویکرد آن است تا اولویت‌های نوآورانه، فناورانه و علمی جوامع با توجه به چشم‌انداز راهبردی برآورده شوند (Santo et al, 2006). انتخاب یک روش برای آینده‌پژوهی به عوامل مختلفی به ویژه زمان و منابع مالی در دسترس و همچنین اهداف ما از فعالیت آینده‌پژوهی بستگی دارد (Miles and Keenan, 2000). مدل‌ها و ابزارهای مختلفی برای تجزیه و تحلیل آینده و نیز پیش‌بینی آینده علم، تکنولوژی و نوآوری وجود دارد اما برای دستیابی به نتایج قابل قبول و معتبر باید تا حد امکان سعی شود از روش‌های کمی و کیفی به طور هم‌زمان استفاده شود (Santo et al, 2006; Singh et al, 2007).

بسیاری از ابزارهای جدید و بالقوه که در حال حاضر در مطالعات مربوط به آینده استفاده می‌شوند، از پیشرفت در فناوری اطلاعات و علم اطلاعات به دست آمده‌اند که در میان آن‌ها علم‌سنجی و کتاب‌سنجی ابزارهایی هستند که به طور سنتی توسط متخصصان علوم اطلاعات برای اندازه‌گیری بهره‌وری علمی و شناسایی شبکه‌های علم و تکنولوژی استفاده می‌شود. بنابراین، امروزه در پیش‌بینی‌های فناوری، تلاش شده است تا از ترکیبی از روش‌هایی نظیر یادگیری ماشینی، تجزیه و تحلیل آماری، تکنیک‌های مدل‌سازی و تکنولوژی‌های پایگاه داده، برای یافتن الگوها و روابط ظریف در داده‌ها و قواعد استنتاجی که شناسایی روندهای آینده را ممکن می‌سازد، استفاده شود (Santo et al, 2006).

متن‌کاوی^۳ یکی از تکنولوژی‌های نوظهوری است که تلاش می‌کند تا اطلاعات معناداری را از داده‌های متنی ساختار نیافته استخراج کند. این روش شکل توسعه یافته‌ی داده‌کاوی برای داده‌های متنی است که با تکیه بر داده‌های متنی ساختار نیافته، مانند فایل‌های متنی، فایل‌های HTML، پیام‌های متنی، مقالات و نامه‌های الکترونیکی به یافتن مدل‌های مفید، روندها، الگوها و یا مجموعه‌ای از قوانین برای آینده می‌پردازد. به عنوان یک تکنیک خودکار، متن‌کاوی می‌تواند برای شناسایی، استخراج، مدیریت، ادغام و بهره‌برداری دانش از متون موجود به گونه‌ای کارا و نظام‌مند، مورد استفاده قرار گیرد. برخلاف تجزیه و تحلیل محتوا به روش سنتی، متن‌کاوی عمدتاً از داده‌ها مشتق می‌شود و هدف اصلی آن شناسایی خودکار الگوها و روندهای پنهان در میان داده‌ها و پس از آن ایجاد مدل‌ها و تفسیر الگوها و روندهایی است که از داده‌های متنی حاصل می‌شود (He et al, 2013). با عنایت به اهمیت آینده‌پژوهی در توسعه علوم و فنون جدید و نیز نقش و اهمیت روش متن‌کاوی در محقق ساختن اهداف آینده‌پژوهانه، در پژوهش حاضر تلاش می‌شود تا رویکرد متن‌کاوی به عنوان ابزاری موثر در فرایند آینده‌پژوهی معرفی و شرح داده شود و زمینه استفاده از آن در حوزه‌های مختلف علمی - تحقیقاتی بررسی گردد.

¹ Ben Martin

² Luke Georghiou

³ Text mining



۲- مبانی نظری پژوهش

در این بخش از پژوهش حاضر تلاش می‌شود تا ضمن ارائه تعاریفی از آینده‌پژوهی و روش متن‌کاوی، به بررسی پیشینه‌ی پژوهش پرداخته شود.

۱-۲ آینده پژوهی

در محیط رقابتی امروز، شناسایی فرصت‌هایی که امیددهنده‌ی رشد و پیشرفت هستند در گرو توجه به آینده می‌باشد که به واسطه‌ی آن می‌توان دید بلند مدتی نسبت به علوم و فناوری‌های تازه به دست آورد (Yoon, 2012). هرچند مدت زمان زیادی نیست که "آینده‌پژوهی و آینده‌نگاری" به عنوان یک دانش نوپا به عرصه‌ی علم وارد شده است، اما از بدو ورود خود توانسته است تا در فرآیندهای تصمیم‌گیری، سیاست‌گذاری‌ها و تدوین راهبردهای کلان تاثیرگذار باشد. آینده‌پژوهی دانشی است که رویدادها، فرصت‌ها و مخاطرات احتمالی آینده را پیش‌بینی می‌کند و با کاهش میزان ابهام و تردید نسبت به آینده، تلاش می‌کند تا به افراد نشان دهد که به کجاها می‌توانند بروند؟ به کجاها باید بروند و از چه مسیرهایی می‌توانند به آینده مطلوب خود برسند؟ (مقیسه و همکاران، ۱۳۹۲). آینده‌پژوهی، اصول و روش‌های مطالعه و سپس تصمیم‌گیری، طرح‌ریزی و اقدام پیرامون علوم و فناوری مرتبط با آینده است. این علم، مبانی فلسفی، روش‌های علمی و مدل‌های مختلف را بررسی می‌کند و با مطالعه‌ی آینده، آینده‌های بدیل و احتمالی را ترسیم می‌کند. بنابراین می‌توان اذعان داشت که آینده‌پژوهی، ابزاری برای معماری و مهندسی هوشمندانه آینده است (فاتح راد و همکاران، ۱۳۸۹).

آینده‌پژوهی دانش تحلیل و برپایی هوشمندانه آینده است که چگونگی مواجهه شدن با پدیده "آینده" را تبیین می‌کند. این علم، دانش و معرفت شکل بخشیدن به آینده است که با تمرکز بر موضوعات بلند مدت و تعهد به نتایج، طی یک فرآیند نظام‌مند، به گونه‌ای آگاهانه عمل می‌کند. آینده‌پژوهی سبب می‌شود تا دید کلی نسبت به رویدادها و مخاطرات احتمالی به دست آمده و جامعه بتواند با علم به آینده ممکن، محتمل و مطلوب خود، انتخاب‌ها و تصمیمات هوشمندانه‌ای داشته باشد (بابا غیبی ازغندی، ۱۳۸۹). همواره انتخاب تاریخی مشخص به عنوان مبدأ رویدادی معین، اقدامی بحث برانگیز است؛ چرا که چنین انتخابی را به سادگی می‌توان با ارایه نمونه‌هایی قدیمی‌تر و ابتدایی‌تر، به چالش کشید. برای مثال، انتشار مقاله "پیش‌بینی‌هایی درباره‌ی تاثیر پیشرفت علمی و مکانیکی بر زندگی و تفکر انسان" اثر اچ جی ولز^۱ می‌تواند مبنای مناسبی برای تعیین تاریخ خواستگاه‌های جدید آینده‌پژوهی باشد. از جمله پیشنهادهای ولز، علم آینده^۲ بود. وی در سال ۱۹۳۲ در مصاحبه‌ای، تحت عنوان "نیاز ما به صاحب‌نظران آینده‌نگر" بیان می‌کند: "من به هیچ عنوان دیدگاه ضدتاریخی ندارم، ولی تعجب خود را هم از این بابت پنهان نمی‌کنم که دانشگاه‌های ما مجهز به تعداد زیادی دپارتمان‌های تاریخ و لشکری از تاریخ‌دان‌ها هستند که دائما به واری گذشته مشغولند، اما تعداد بسیار کمی "آینده‌پژوه" در آن‌ها کار می‌کنند" (اسلاتر، ۱۳۸۶).

آینده‌پژوهی علم و فناوری ریشه در دهه ۱۹۵۰ دارد، هنگامی که پیش‌بینی فناوری^۳ در بخش دفاعی ایالات متحده و در کار با مشاورانی مانند شرکت رند^۴ برجسته شد. هدف کلی از آینده‌پژوهی فناوری شناسایی فن‌آوری‌های در حال ظهوری بود که به احتمال زیاد به منافع اقتصادی و اجتماعی منجر می‌شد (Martin, 2001; Martin and Johnston, 1999). ژاپن در فعالیت‌های آینده‌پژوهی از سال ۱۹۷۰ درگیر شده بود و تعدادی از طرح‌های آینده‌پژوهی در فرانسه در اوایل دهه ۱۹۸۰ وجود داشت. بعد از آن دهه، کشورهای دیگری مانند سوئد، کانادا و استرالیا شروع به آزمایش با آینده‌پژوهی کردند. در طول دهه ۱۹۹۰، آینده‌پژوهی تکنولوژی کاربرد بسیار گسترده‌تری یافت. در سال ۱۹۹۰ هلند، ایالات متحده، استرالیا، آلمان، بریتانیا، فرانسه و دیگر کشورهای مختلف پروژه‌های آینده‌پژوهی بزرگی راه‌اندازی نمودند (Martin, 1996). امروزه محققان بسیاری به دنبال

¹ H. G. Wells

² Science of future

³ Technology Forecasting

⁴ RAND



کشف نقش ابعاد مختلف علم در زندگی بشر امروزی و پی بردن به ارتباط پدیده‌های جهان با زندگی انسان‌ها و آسان کردن کارها و راحتی او هستند.

۲-۲ متن کاوی

روش‌های متفاوت کمی و کیفی برای تجزیه و تحلیل تغییرات فنی- اجتماعی از طریق آینده‌پژوهی وجود دارد. روش‌هایی مانند تجزیه و تحلیل مورفولوژیک، متن کاوی^۱، برون‌یابی روند، تجزیه و تحلیل فناوری و SWOT برای شرح دقیق توسعه فناوری و ارائه ابزاری برای تصمیمات مدیریت استراتژیک به کار می‌روند (Zheng et al, 2012).

اسناد (برای مثال پتنت‌ها، مقالات، صفحات وب، گزارشات سالیانه و غیره) از مجموعه عظیمی از متون ساختار یافته و اعداد، به همراه اطلاعات متنی بدون ساختار تشکیل شده‌اند. این گونه اسناد اغلب حجیم و سرشار از اصطلاحات فنی هستند، که این مسئله تحلیل دستی آن‌ها را دشوار می‌سازد. ولی اگر به دقت مورد تحلیل قرار گیرند، می‌توانند جزئیات و روابط فنی را نشان داده، روندهای کسب و کار را آشکار ساخته و در تصمیم‌گیری‌های مرتبط با سیاست‌های سرمایه‌گذاری مورد استفاده قرار گیرند. متن کاوی می‌تواند در جهت درک بهتر اطلاعات موجود در این اسناد بکار گرفته شود. در واقع، عنصر اصلی متن کاوی مبتنی بر جمع‌آوری اسناد است. به زبان ساده، جمع‌آوری اسناد به معنای گروه‌بندی مستندات بر پایه‌ی متن است. در بسیاری از فرآیندهای متن کاوی، هدف، کشف الگوها در میان حجم وسیعی از مستندات است. تعداد اسناد در مجموعه‌ها می‌تواند هزاران و یا میلیون‌ها سند را در بر گیرد. جمع‌آوری اسناد به دو شکل ایستا و پویا صورت می‌گیرد. در جمع‌آوری اسناد به روش ایستا، اجزای اولیه‌ی سند تغییر نمی‌کند اما در روش پویا، اسناد پذیرنده‌ی مستندات جدیدی هستند و مستندات موجود نیز ممکن است طی بازه‌های زمانی مشخص به روزسانی شود.

متن کاوی، بازیابی و درک اطلاعات پنهان موجود در متون را تسهیل می‌کند (Tseng et al, 2007) و می‌توان آن را به عنوان یکی از شاخه‌های داده کاوی در نظر گرفت. این روش، نوعی اکتشاف و تحلیل داده‌های متنی توسط ابزارهای خودکار و نیمه خودکار به منظور کشف دانش جدید است. این مفهوم برای اولین بار توسط فلدمن و داگان در سال ۱۹۹۵ به عنوان روش تجزیه و تحلیل ماشینی متن، مطرح گردید. متن کاوی از روش‌هایی نظیر بازیابی اطلاعات، استخراج اطلاعات و همچنین فرآیند زبان طبیعی (NLP) استفاده کرده و سپس آن‌ها را با الگوریتم‌ها و روش‌های داده کاوی، یادگیری ماشینی و تجزیه و تحلیل‌های آماری مرتبط می‌کند (Hotho et al, 2003). طبق تعریفی که زهر در مورد متن کاوی ارائه می‌دهد، این روش به معنی استخراج اطلاعات ضمنی، از قبل نامعلوم و بالقوه مفید از حجم عظیمی از داده‌های متنی است (یوسف تبار، ۱۳۹۱). متن کاوی فرآیندی دانش محور است که کاربر می‌تواند با بهره‌گیری از این فرآیند و با استفاده از ابزارهای تحلیلی مناسب، با مجموعه‌ای از اسناد در تعامل قرار گرفته و برای پیش‌بینی آینده الگوها و روندهایی متناسب با داده‌های متنی موجود، ترسیم کند (Hotho et al, 2003).

۲-۳ پیشینه‌ی پژوهش

تاکنون، بسیاری از محققان تجربه موفقیت‌آمیزی در استفاده از روش‌های متن کاوی برای تجزیه و تحلیل حجم وسیعی از داده‌های متنی در حوزه‌ی کسب و کار، علوم پزشکی، آموزش و پرورش، مدیریت منابع انسانی، فناوری، مدیریت دانش و غیره داشته‌اند. از آن‌جا که هدف از این تحقیق، معرفی این روش و کاربرد آن در آینده‌پژوهی می‌باشد، در این قسمت از مرور تحقیقات مشابه به کاربرد این روش در مطالعات مختلف پرداخته می‌شود.

با عنایت به آن که، مقالات نمایه شده در موسسه اطلاعات علمی (ISI)، یکی از معتبرترین شاخص‌های سنجش علم و فناوری می‌باشد و طبقه‌بندی موضوعی آن‌ها یکی از چالش‌های بزرگ مدیریت فناوری است، تیمورپور و همکارانش تلاش کردند تا با

¹ Text mining



استفاده از یک روش نوین متن کاوی به نام (SUTC)، مقالات متخصصان ایرانی در حوزه‌ی فناوری نانو را دسته‌بندی کنند. در این راستا ابتدا استانداردهای معتبر در فناوری نانو با یکدیگر ادغام شد و طبقه‌بندی جامعی برای نانو مواد به دست آمد، سپس با استفاده از روش‌های بازبایی اطلاعات و متن کاوی، مقالات بدون دانش پیشین از برچسب داده‌ها، به طور هوشمند دسته‌بندی شدند (تیمورپور و همکاران، ۱۳۸۸).

روشن است که شناخت و درک هرچه بیشتر از آینده برای هر سازمانی در دنیای رقابتی امروز از اهمیت بالایی برخوردار است و آینده‌پژوه باید در زمینه‌ی شناسایی آینده مطلوب و توجه به چگونگی رسیدن به آن تلاش کند. صدر ارحامی و جولایی معتقدند که اطلاعات مورد نیاز برای آینده‌پژوهی اغلب به صورت متنی است و برای استخراج و پردازش این متون لازم است تا از روش‌های جدید استفاده شود. آن‌ها در پژوهش خود در صدد برآمدن تا با استفاده از روش متن کاوی در حوزه آنالیز معنایی به شناسایی معنایی که مشخص کننده‌ی مفید بودن یک مطلب است، بپردازند (صدرارحامی و جولایی، ۱۳۹۳).

زارع احمد آبادی و رضائیان با رویکرد آینده‌پژوهی علم و فناوری، بر این باورند که نگاه ویژه به فناوری بادگیر به عنوان یک تکنولوژی بومی ایران بسیار حائز اهمیت می‌باشد. در این راستا برای شناسایی روند حاکم بر توسعه این فناوری، تعداد ۷۰ مقاله بین‌المللی منتشر شده در زمینه بادگیر را شناسایی کرده و از رویکرد متن کاوی به عنوان ابزاری موثر در فرآیند آینده‌پژوهی بهره گرفتند. در این پژوهش با استفاده از نظر خبرگان، متن کاوی و تحلیل خوشه‌ای مقالات، حوزه‌های پژوهشی مد نظر پژوهشگران، شکاف‌های پژوهش و نیز بسترهای مطالعاتی جدید شناسایی شد. شناسایی فرصت‌های سرمایه‌گذاری در بخش تحقیق و توسعه در تهویه طبیعی و فناوری بادگیر از نتایج این پژوهش می‌باشد (زارع احمدآبادی و رضائیان، ۱۳۹۳).

بالوگلو و آسانته (۱۹۹۹) محتوای تحقیقات مربوط به مدیریت مهمانداری^۱ را با تجزیه و تحلیل نواحی موضوعی و روش تحقیق ۱۰۷۳ مقاله منتشر شده در پنج ژورنال هتلداری طی سال‌های ۱۹۹۶-۱۹۹۰ آزمایش کردند. آنها به این نکته دست یافتند که اکثر مقالات بر روی منابع انسانی و مسکن و صنایع خدماتی مواد غذایی متمرکز شده‌اند. ضمناً این مطالعه مشخص کرد که روش پرسشنامه، طرح تحقیقاتی اغلب مطالعات بوده و روش‌های میدانی و آزمایشگاهی کمترین استفاده را داشته‌اند.

در پژوهشی دیگر نقش متن کاوی به عنوان ابزاری موثر در مطالعات آینده‌پژوهی بررسی شده و این نکته مورد تأکید قرار گرفته است که برای افزایش اعتبار این روش، باید تلفیقی از روش‌های کمی (متن کاوی کامپیوتری) و روش‌های کیفی (فضاوت متخصصان) به کار گرفته شود. در این پژوهش با استفاده از متن کاوی مقالات منتشر شده در زمینه نانو تکنولوژی، کشورهای کلیدی پیشرو در این زمینه شناسایی شده و چشم‌اندازی کلی برای این کشورها رسم شده است (Santo et al, 2006).

مجله‌ی بین‌المللی مدیریت هتلداری در بازه‌ی زمانی سال ۱۹۹۴-۲۰۰۳ شاهد انتشار مقالات زیادی در حوزه‌ی مدیریت منابع انسانی بوده است. همین امر محققان را بر آن داشت تا زمینه‌های مطالعاتی پر کاربرد را در حیطه‌ی مدیریت منابع انسانی معرفی کنند. از این رو، داده‌های متنی با توجه به مقالات انتشار یافته جمع‌آوری شد و تجزیه و تحلیل محتوا با استفاده از روش متن کاوی صورت گرفت. نتایج حاصل از متن کاوی داده‌های موجود نشان داد که زمینه‌ی مطالعاتی مدیریت منابع انسانی در بازه‌ی زمانی مذکور، بیشتر بر موضوعاتی نظیر جایگاه شغلی هتلداری، آموزش، محیط کاری، رضایت شغلی، استخدام و گردش شغلی، مسائل قانونی و توسعه‌ی سازمانی تمرکز داشته است (Singh et al, 2007).

در پژوهشی تحت عنوان "محتوای مدیریت استراتژیک منابع انسانی در گزارشات سالانه شرکت‌ها: تجزیه و تحلیل از طریق متن کاوی" از رویکرد متن کاوی برای کشف الگو از یک مجموعه داده شامل کتب مربوط به مدیریت استراتژیک منابع انسانی استفاده کرده و سپس اهمیت این مفاهیم در گزارشات سالانه مورد بررسی قرار داده است. در حقیقت عبارات و مفاهیم مهم مورد استفاده در محیط آکادمیک شناسایی شده و با عباراتی که در محیط کسب و کار به عنوان گزارش سالانه مورد استفاده

¹ Hospitality Management



قرار می‌گیرند مقایسه شده‌اند. سپس این مفاهیم براساس حضورشان در گزارشات سالانه رتبه بندی شده و موضوعات مورد توجه در هر دو حوزه آکادمیک و کسب و کار مورد بررسی قرار گرفته‌اند (Kamal and Kumar, 2011).

العظمی (۲۰۱۳) بر این باور است که انقلاب فناوری‌های ارتباطی و اطلاعاتی سبب شده است تا عصر حاضر شاهد رشد انبوهی از اطلاعات دیجیتالی باشد. همین امر، واحدهای اقتصادی را بر آن داشته است تا برای بررسی و مطالعه‌ی داده‌های دیجیتالی و استخراج دانش و بینش مورد نیاز جهت توسعه‌ی علوم جدید، از فناوری‌هایی نظیر داده‌کاوی، متن‌کاوی و وب‌کاوی بهره‌گیرند. به طور کلی، ابزارهای کنکاش، نرم‌افزارهای خودکاری هستند که با یافتن روابط پنهان میان حجم گسترده‌ای از داده‌ها، وقایع آینده را پیش‌بینی می‌کنند. بنابراین العظمی پژوهشی را با هدف پاسخ به این پرسش که چگونه می‌توان با استفاده از ابزارهای کنکاش داده‌ها به مفهومی نظیر هوش تجاری دست یافت، انجام داد (Al-Azmi, 2013).

روش متن‌کاوی در زمینه‌ی علوم مرتبط با هنر نیز کاربرد دارد. در همین رابطه عباس و همکارانش با استفاده از روش متن‌کاوی، جایگاه هنر را در تجزیه و تحلیل پتنت مورد بررسی قرار دادند. آن‌ها بر این باور هستند که ابزارهای مورد نظر برای تجزیه و تحلیل پتنت باید قابلیت‌های زیادی را برای انجام دامنه‌ی وسیعی از عملیات نظیر تجزیه و تحلیل و پیش‌بینی روند برای فناوری، برنامه‌ریزی راهبردهای فناوری، شناسایی تخلف در ارائه و ثبت پتنت‌ها، تعیین کیفیت پتنت‌ها و شناسایی زمینه‌های تازه برای ارائه پتنت و خلاءهای موجود در این زمینه داشته باشند. بنابراین روش متن‌کاوی را برای انجام پژوهش خود برگزیدند (Abbas et al, 2014).

۳- روش شناسی متن‌کاوی

داده‌کاوی یا کشف دانش در پایگاه داده، رویکردی است که به واسطه‌ی آن، الگوهای جالب و مفهومی از مجموعه‌ی بزرگی از داده‌ها استخراج می‌شود. متن‌کاوی یکی از زمینه‌های داده‌کاوی می‌باشد که شباهت زیادی با داده‌کاوی دارد با این تفاوت که ابزارهای داده‌کاوی برای مدیریت داده‌های ساختاریافته طراحی شده است اما متن‌کاوی می‌تواند بر مجموعه‌ای از داده‌های نیمه ساختاریافته و ساختار نیافته نیز برای استخراج الگوها و روندهای جدید تمرکز کند (آقا‌کاردان و کیهانی نژاد، ۱۳۹۱). در حقیقت داده‌کاوی به دنبال استخراج اطلاعات مفید از اشکال مختلف داده است و به طور معمول در تجزیه و تحلیل داده‌های عددی کاربرد دارد، این در حالی است که روش متن‌کاوی بیشتر به پردازش و تحلیل داده‌های متنی می‌پردازد (Santo et al, 2006). در مقایسه با داده‌کاوی، متن‌کاوی به دنبال آن است که اطلاعات مفید را از منابع موجود استخراج کرده و الگوهای مورد نظر را شناسایی کند. از دیدگاه این روش، منبع داده‌های مورد نیاز مجموعه‌ای از اسناد هستند و الگوهای مورد نظر در میان پایگاه‌های داده‌ی رسمی یافت نمی‌شود بلکه این الگوها از داده‌های متنی ساختار نیافته قابل شناسایی هستند (Feldman & Sanger, 2006).

به طور کلی، متن‌کاوی حوزه‌ی جدید در علوم کامپیوتر است که تلاش می‌کند، مشکل همپوشانی اطلاعات را با استفاده از ترکیب تکنیک‌هایی نظیر یادگیری ماشینی، فرآیند پردازش زبان طبیعی، بازیابی اطلاعات و مدیریت دانش، بررسی و حل کند. این روش برای یادگیری و کشف اطلاعاتی که قبلاً نا شناخته بودند، مناسب است. اگرچه روش متن‌کاوی با هر نوع اطلاعاتی کار می‌کند، اما زمانی که اطلاعات مبتنی بر متن و از میان متون ارزشمند و دقیق انتخاب شود، نتایج بهتری ارائه می‌شود. در واقع این روش به کشف و سازماندهی روابط بین مفاهیم در داده‌های متنی کمک شایان توجهی می‌کند.

کاواینکاس و تئودوایدیس در سال ۲۰۰۲ بیان کردند که اهداف روش متن‌کاوی را می‌توان در نه گروه عملیاتی طبقه‌بندی کرد که شرح هر یک از این اهداف در جدول ۱ آورده شده است (Singh et al, 2007).



جدول 1: اهداف روش متن کاوی

اهداف	تعریف
استخراج ویژگی ^۱	* منظور از استخراج ویژگی، مشخص کردن تمایز میان عبارت اسمی، شخص، مکان و یا سازمان از سایر اشیا است. که عملیات مرتبط با این هدف، شامل استخراج کلمات و محاسبه‌ی تعداد دفعاتی است که هر کلمه در متن تحلیل شده است.
هدایت براساس آزمون ^۲ آزمون ^۲	* هدایت براساس آزمون، استخراج کننده متن را قادر می‌سازد تا کلمات مرتبط را در متن بیابد و روابط مهم بین آن‌ها را به هم ارتباط دهند.
جستجو و بازیابی ^۳	* جستجو و بازیابی، کاربر را بر آن می‌دارد تا بر اساس معیارهای از جستجوی از پیش تعیین شده به جستجو و بازیابی اطلاعات بپردازد.
خوشه‌بندی ^۴	* خوشه‌بندی، عملیات گروه‌بندی کلمات کلیدی براساس میزان شباهت‌ها و یا عدم تشابه می‌باشد. رایج‌ترین الگوریتم خوشه‌بندی براساس روش‌های طبقه‌بندی آماری است.
طبقه‌بندی ^۵	* طبقه‌بندی، عملیات تعریف مجموعه‌ای از کلمات مخصوص دامنه و روابط بین آن‌ها از طریق الگوریتم‌های طبقه‌بندی است.
تلخیص ^۶	* تلخیص عبارت است از عملیات کاهش مقدار داده‌های متنی با حفظ عناصر اصلی.
تجزیه و تحلیل روند ^۷	* تجزیه و تحلیل روند فرآیندی است که طی آن برای کشف روند از داده‌های متنی وابسته به سازمان استفاده می‌شود.
قواعد انجمنی ^۸	* هدف از پیوستگی، مرتبط ساختن الگوی استخراج شده با الگوهایی است که از قبل کشف و استخراج شده‌اند.
تجسم‌سازی ^۹	* در تجسم‌سازی از استخراج ویژگی و کلمات کلیدی فهرست شده برای ایجاد یک نمایش گرافیکی استفاده می‌شود که می‌تواند کاربر را برای شناسایی موضوعات اصلی با مفاهیم مهمشان یاری دهد.

۳-۱ کاربردهای متن کاوی

کاربردهای گوناگون روش متن کاوی، سبب شده است تا این روش در حوزه‌های مختلف علوم مورد استفاده قرار گیرد. هرچند تمرکز اصلی متن کاوی بر طبقه‌بندی داده‌ها و پیش‌بینی معطوف می‌باشد ولی می‌توان اذعان داشت که بازیابی اطلاعات، خوشه‌بندی اطلاعات و نیز استخراج اطلاعات از کاربردهای رایج این روش به شمار می‌رود. طبقه‌بندی متن^{۱۰}، جهت طبقه‌بندی اسناد استفاده می‌شود. در طی فرآیند طبقه‌بندی، اسناد در قالب پوشه‌ها سازماندهی می‌شوند و هر پوشه به یک موضوع خاص، اختصاص می‌یابد.

بازیابی اطلاعات^{۱۱} موضوعی است که با داده‌های آنلاین مرتبط است. اولین مسئله در انجام متن کاوی، جمع آوری مجموعه‌ای از داده‌های مرتبط با موضوع مورد بررسی است. این اسناد ممکن است از پیش آماده بوده و یا مستلزم انجام جستجوهای اولیه،

¹ Feature extraction

² Test base navigation

³ Search and retrieval

⁴ Clustering

⁵ Categorization

⁶ Summarization

⁷ Trend analysis

⁸ Associations

⁹ Visualization

¹⁰ Document Classification

¹¹ Information Retrieval



قبل از ادامه فرایند باشد. برای مثال برای عمل بازیابی اطلاعات در مورد یک صفحه وب، اسناد مرتبط همان صفحات خواهند بود و نیازی به جمع آوری آن‌ها نمی‌باشد.

خوشه بندی^۱، فرایند سازماندهی مفاهیم در قالب گروه‌های مختلف یا به عبارت ساده‌تر، گروه بندی داده‌ها (متون) در قالب زیر زیر گروه‌هایی به نام خوشه می‌باشد. خوشه بندی ابزار مهمی در متن کاوی، به منظور تشخیص توزیع داده‌ها و الگوهای موجود درون آن‌هاست. همان‌طور که هدف از طبقه‌بندی جایگذاری اسناد جدید به پوشه‌ی مناسب است، و این پوشه‌ها توسط کسی ایجاد می‌شود که دانش ساختار اسناد را داشته باشد، روش خوشه‌بندی زمانی مورد استفاده قرار می‌گیرد که مجموعه‌ای از داده‌ها با ساختار ناشناخته در دست باشد. بنابراین، با استفاده از خوشه‌بندی، می‌توان به طور مستقیم و بدون اتکا به هیچ پیش زمینه دانشی از داده‌ها، ساختارها و خوشه‌های جالب توجهی را استخراج کرد.

استخراج اطلاعات^۲ یکی از کاربردهای متن‌کاوی است که تلاش می‌کند پایه‌های متن‌کاوی را بر اسناد ساختار یافته (مانند داده‌کاوی) بنا نهد تا متن ساختار یافته شود. از آن جایی که در متن‌کاوی، اطلاعات مورد استفاده، کلمات هستند، تنها وجود یا عدم وجود کلمه در سند اندازه‌گیری نمی‌شود (فراوانی کلمه) و می‌توان اطلاعات عمیق و پایه‌ای را که به مراتب پیچیده‌تر داده‌کاوی عددی کلاسیک است، استخراج کرد.

هدف نهایی از متن‌کاوی، پیش‌بینی و ارزیابی^۳ می‌باشد. در پیش‌بینی هدف آن است تا طرح‌ریزی‌هایی که از نمونه‌های گذشته حاصل شده و بر پایه‌ی تجربیات قبلی بوده را به نمونه‌های جدید تعمیم داد و با ارزیابی اسناد و توجه به قوانین، الگوها و روندهای میان تجربیات قبلی، الگوها، قوانین و روندهایی را برای وقایع آتی پیش‌بینی نمود (Wiese et.al, 2010). در جدول شماره ۲، پژوهش‌های صورت گرفته در حوزه‌های مختلف علمی با استفاده از روش متن‌کاوی، به تفکیک کاربردهای آن نشان داده شده است.

جدول 2 طبقه‌بندی مطالعات صورت گرفته با توجه به کاربرد متن‌کاوی

پژوهشگر / سال پژوهش	پیش‌بینی و ارزیابی	استخراج اطلاعات	خوشه‌بندی متن	بازیابی اطلاعات	طبقه‌بندی متن
Tseng Yuen- Hsien, Lin Chi- Jen, Lin Yu-I (2007)			√		
Daim & Rueda (2006)	√				

¹ Clustering & Organizing document

² Information Extraction

³ Prediction & Evaluation



√					Denoyer & Gallinari (2004)
			√		Ventura Jao & Silva Joaquim (2012)
				√	Jun Sunghae & Lee Seung- Loo (2012)
√					Kaur Manmeet & Sapra Richa (2013)
	√				Tangil Guillermo, Tapiador Juan, Peris- Lopez Pedro, Blasco Jorge (2014)
	√				Hersh William (2005)
		√			Kogan (2001)
		√			Hutho, Staab & Stumme (2003)

۱-۳ فرآیند متن کاوی

فرآیند عمومی متن کاوی متشکل از مراحل سه گانه‌ی آماده سازی متن، پردازش متن و تحلیل متن می باشد که در هریک از این مراحل، مجموعه فعالیت های مرتبط با هدف مرحله انجام می شود. مراحل فرآیند متن کاوی در شکل ۱ نشان داده شده است.

شکل ۱: مراحل متن کاوی (Liddy, 2000)



- ✓ **آماده سازی متن:** انتخاب، پاک سازی و پردازش مقدماتی متن در مرحله‌ی آماده سازی متن صورت می گیرد. در این مرحله، منابعی که باید بر روی آن ها متن کاوی انجام شود انتخاب شده و به پردازش مقدماتی متن نظیر نشانه گذاری^۱، شناسایی جمله/ پاراگراف، برچسب گذاری نقش کلمه، ریشه یابی^۲ کلمه و ابهام زدایی^۳ پرداخته می شود.
- ✓ **پردازش متن:** در این مرحله، از الگوریتم داده کاوی جهت پردازش داده های آماده سازی شده استفاده می شود و مدل پژوهش بر پایه‌ی کلمات کلیدی استخراج شده و ویژگی های اسناد موجود، شکل می گیرد.
- ✓ **تحلیل متن:** در این مرحله، خروجی متن کاوی مورد ارزیابی قرار می گیرد تا مشخص شود که آیا دانشی به دست آمده است و این که آیا دانش کشف شده با اهمیت است یا خیر؟ در واقع در تحلیل متن، الگوها و مدل های ساخته شده در مرحله‌ی قبل، برای کشف مشابهت های بین اسناد بر پایه‌ی کلمات کلیدی، مورد تحلیل قرار می گیرد و با اجرای الگوریتم های، داده یا متن استخراج شده به فنون مختلفی تحویل داده می شود که امکان استفاده مستقیم از

¹ Tokenization

² Stemming

³ Word sense disambiguation



اطلاعات استخراج شده را از طریق ابزار کشف ارتباط یا مصورسازی فراهم می‌کنند (آقا کاردان و کیهانی نژاد، ۱۳۹۱؛ یوسف تبار، ۱۳۹۱).

با توجه به تشریح مراحل سه‌گانه‌ی فرآیند متن‌کاوی، مطلوب است تا در طی انجام فرآیند، هدف خاص متن‌کاوی، محدودیت‌های موجود برای داده‌ها یا متون و نیز نقاط ضعف و قوت الگوریتم‌های مورد نیاز برای تحلیل داده‌ها در نظر گرفته شود تا کشف دانش از داده‌ها محقق شده و آینده‌نگر ضمن دنبال کردن فرآیند متن‌کاوی به بهترین شکل به اهداف خود دست یابد.

۴- نتیجه‌گیری

آینده‌چالش‌های نمایان و سترگی را پیش روی ما قرار داده است و روبرو شدن با آن بستگی به این دارد که ما هرچه سریع‌تر آینده‌پژوهی را به عنوان وظیفه‌ای اساسی و محوری تلقی کنیم و زمینه بروز همه آنچه را که در عنوان تفکر آینده‌اندیشانه می‌گنجد، از سرچشمه‌های اختصاصی آن یعنی دانشگاه‌ها و موسسات تحقیقاتی فراهم آوریم. هرچند مدت زمان زیادی نیست که "آینده‌پژوهی و آینده‌نگاری" به عنوان یک دانش نوپا به عرصه‌ی علم وارد شده است، اما از بدو ورود خود توانسته است تا در فرآیندهای تصمیم‌گیری، سیاست‌گذاری‌ها و تدوین راهبردهای کلان تاثیرگذار باشد. آینده‌پژوهی دانشی است که رویدادها، فرصت‌ها و مخاطرات احتمالی آینده را پیش‌بینی می‌کند و با کاهش میزان ابهام و تردید نسبت به آینده، چگونگی مواجهه با پدیده "آینده" را تبیین می‌کند.

بسیاری از ابزارهای جدید یا بالقوه که در حال حاضر در مطالعات آینده استفاده می‌شوند، از پیشرفت فناوری اطلاعات و علم اطلاعات بدست آمده‌اند. در میان آن‌ها علم‌سنجی و کتاب‌سنجی ابزاری هستند که به طور سنتی توسط متخصصان علوم اطلاعات برای اندازه‌گیری بهره‌وری علمی و شناسایی شبکه‌های علم و تکنولوژی استفاده می‌شود. در حالی که ما به دهه سوم شبکه جهانی وب^۱ وارد می‌شویم، انقلاب متنی تغییرات گسترده‌ای را در دسترسی به اطلاعات آنلاین ایجاد کرده است. اگرچه پیدا کردن اطلاعات در هر زمینه‌ای تنها با یک کلیک امکان‌پذیر شده است و دیجیتال سازی و ایجاد داده‌های متنی به سرعت ادامه پیدا می‌کند، توانایی هدایت، کاوش و جستجوی اطلاعات در آن‌ها، فاصله زیادی با سرعت افزایش این داده‌ها دارد. می‌توان گفت در دنیای مدرن که در آن میزان انتشار متون علمی به حدی زیاد است که تقریباً برای هر محقق غیرممکن است که تمام اطلاعاتی را که با آن روبروست، دنبال کند؛ متن‌کاوی به عنوان ابزاری مفید در اختیار پژوهشگران قرار دارد.

در این پژوهش با معرفی روش متن‌کاوی به عنوان روشی موثر در فرایند آینده‌پژوهی، مطالعاتی که در این زمینه صورت گرفته‌اند مرور شده و پس از بررسی زمینه استفاده از آن در حوزه‌های مختلف علمی-تحقیقاتی، زمینه کاربرد این روش در مقالات مختلف در ۵ حوزه طبقه‌بندی متن، بازیابی اطلاعات، خوشه‌بندی، استخراج اطلاعات و پیش‌بینی و ارزیابی دسته‌بندی گردید. در نهایت با معرفی و شرح فرایند کلی متن‌کاوی روش استفاده از آن به صورت کلی توضیح داده شد. نتایج حاکی از آن است که کاربردهای گوناگون روش متن‌کاوی، سبب شده است تا این روش در حوزه‌های مختلف علوم مورد استفاده قرار گیرد.

مراجع

- اسلاتر، ر (۱۳۸۶)، نواندیشی برای هزاره نوین، ترجمه ملکی فر، ع، ابراهیمی، ا، وحیدی مطلق و، مرکز آینده‌پژوهی علوم و فناوری دفاعی، چاپ اول، تهران، ۱۳۸۶.
- آقا کاردان، ا. کیهانی‌نژاد، مینا. (۱۳۹۱)، ارائه‌ی مدلی برای استخراج اطلاعات از مستندات متنی، مبتنی بر متن‌کاوی در حوزه‌ی یادگیری الکترونیکی، فصلنامه علمی- پژوهشی فناوری اطلاعات و ارتباطات ایران، سال چهارم، شماره‌های ۱۱ و ۱۲، ۴۷-۵۴

¹ World Wide Web



۳. باباغیبی ازغندی، ع (۱۳۸۹) آینده پژوهی، رهیافتی نو در مدیریت جامع حمل و نقل شهری، فصلنامه مطالعات مدیریت ترافیک، سال پنجم، شماره ۱۶، ۷۷-۱۰۰
۴. تیمورپور، ب. سپهری، م. پزیشک، ل. (۱۳۸۸)، روشی نوین برای دسته‌بندی هوشمند متون علمی (مطالعه‌ی موردی مقالات فناوری نانو متخصصان ایران)، فصلنامه علمی- پژوهشی سیاست علم و فناوری، سال دوم، شماره ۲، ۱-۱۴
۵. زارع احمد آبادی، ح. رضاییان، م. (۱۳۹۳)، آینده پژوهی مطالعات انرژی باد در تهبویه طبیعی فضاهاى مسکونی: کاربرد متن‌کاوی و تحلیل خوشه‌ای، دهمین همایش بین‌المللی انرژی، تهران- ایران
۶. زارعیان، آ. ستارزاده، م. (۱۳۹۰)، رهیافتی در باب چگونگی ماهیت آینده پژوهی، مجله دانشکده پرستاری ارتش جمهوری اسلامی ایران، سال یازدهم، شماره مسلسل ۲۱، ۱-۵
۷. زالی، ن. (۱۳۹۰) " آینده نگاری راهبردی و سیاست‌گذاری منطقه ای با رویکرد سناریو نویسی " فصلنامه مطالعات راهبردی، سال چهاردهم، شماره چهارم، صص ۳۳-۵۴.
۸. سلامی، ر. خانی، م. صفاری دربرزی، ع. (۱۳۹۱)، بررسی تاثیر آینده‌پژوهی در حوزه تحقیقات ناجا، دو فصلنامه توسعه تکنولوژی صنعتی، شماره نوزدهم، ۷۳-۸۸
۹. صدرارحامی، ح. جولایی، ر. (۱۳۹۳)، استفاده از متن کاوی و آنالیز معنایی برای استخراج اطلاعات مورد نیاز آینده پژوهی از متون، همایش ملی مهندسی رایانه و مدیریت فناوری اطلاعات، تهران، شرکت علم و صنعت طلوع فرزین
۱۰. فاتح‌راد، م. جلیلود، م. مولایی، م. سمیعی، س. نصرالهی وسطی، ل. (۱۳۹۲)، مختصات روش‌شناختی فرا رشته آینده‌پژوهی به مثابه یک فرآیند یک پارچه، فصلنامه مطالعات میان رشته‌ای در علوم انسانی، دوره ششم، شماره یک، ۱۳۵-۱۶۱
۱۱. کریمی فرد، ح. (۱۳۸۹) "بررسی تکامل آینده نگاری علم و تکنولوژی در گذر زمان" مجله راهبرد یاس، شماره ۲۲، صص ۲۸۵-۲۶۷.
۱۲. مقیسه، ر. معلم، محمدحسین. علی‌دادی، علی. (۱۳۹۲)، فلسفه‌ی تاریخ رویکرد مناسب آینده‌پژوهی و آینده‌نگاری اسلامی، کنگره بین‌المللی فرهنگ و اندیشه دینی
۱۳. نوری، س. ایوبی، م. (۱۳۸۸)، توسعه مدل آینده‌پژوهی مبتنی بر پیمایش محیطی (مطالعه موردی در یک سازمان تحقیقاتی- دفاعی)، مدیریت فردا، شماره ۲۲، سال هشتم، ۵۵-۶۶
۱۴. یوسف‌تبار، ص. (۱۳۹۱)، پیش بینی تغییرات تکنولوژی کاشی و سرامیک بر اساس متن کاوی اختراعات ثبت شده، پایان‌نامه کارشناسی ارشد، رشته‌ی مدیریت صنعتی، دانشکده‌ی اقتصاد، مدیریت و حسابداری دانشگاه یزد.
15. Abbas, A. Zhang, L. Khan, S. (2014), A literature review on the state- of- the art in patent analysis, World patent information, 1-11
16. Al- Azmi, A.A. (2013), Data, Text and Webmining for business intelligence: A survey, International journal of Datamining & Knowledge management process, Vol.3, No.2, 1-21
17. Baloglu S, Assante L.M. (1999), A content analysis of subject areas and research methods used in five hospitality management journals, Journal of Hospitality and Tourism Research, 23, 53-70.
18. Bjorn L, Lubeck L. (2003), Foresight – a successful project, with many lessons learned, The second international conference on technology foresight, Tokyo, Swedish Technology.
19. Daim, T. Rueda, G. Martin, H. Gersdri P, (2006), Forecasting emerging technologies: Use of bibliometrics and patent analysis Technological Forecasting and Social Change 73 (8), 981-1012
20. Denoyer, L. Gallinari, P. (2004), Un modèle de mixture de modèles génératifs pour les documents structurés multimédias, Document numérique 8 (3), 35-54
21. Feldman, R. Sanger, J. (2006), The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data, Cambridge University Press
22. Georghiou, L, (1996), The UK Technology Foresight Program, Futures, 28, 359 – 377.
23. Hersh, W. (2005), Evaluation of biomedical text-mining systems: lessons learned from information retrieval, Briefings in bioinformatics 6 (4), 344-356
24. Hotho, A. Staab, S. Stumme, G. (2003). Ontologies improve Text document clustering. In Proc. Conf. On Datamining, 541-544
25. Jun, S. Lee, S (20120), Emerging technology forecasting using new patent information analysis, Software engineering and its application
26. Kamal Kumar, K., Kumar, K. (2011), Strategic Human Resource Management Content in the Annual Report of Companies: An Analysis through Text Mining, Indian Institute of Management Indore, 2, 15-26.



27. Kaur, M. Sapra, R. (2013), Classification of patents by using the Text mining approach based on PCA and logistics, International journal of engineering and advanced technology, Vol 2, 711-714
28. Kogan, J (2001), Means clustering for text data, Proceedings of the Workshop on Text Mining at the First SIAM International
29. Liddy, E. (2000), Text mining, Bulletin of American Society for information Sciences
30. Martin, B. (1996), Foresight in Science and Technology, Technology Analysis & Strategic Management, 7, 139-68.
31. Martin, B.R. (2001), Technology Foresight in a Rapidly Globalizing Economy, Technology Foresight for Central and Eastern Europe and the Newly Independent States', Vienna, Austria, 4-5.
32. Martin, B. R. and Johnston, R. (1999), Technology Foresight for Wiring up the National Innovation System: Experiences in Britain, Australia and New Zealand, Technological Forecasting and Social Change, 60, 37-54.
33. Miles I, Keenan, M (2000), Overview of methods used in Foresight, Institute of Innovation Research, University of Manchester, UK.
34. Santo, M, et al. (2006), Text mining as a valuable tool in foresight exercises: A study on nanotechnology, Technological Forecasting & Social Change, 73, 1013-1027.
35. Singh, N et al, (2007), "Text mining a decade of progress in hospitality human resource management research: Identifying emerging thematic development", Hospitality Management 26, 131-147.
36. Tangil, G; Tapiador, J.; Peris-López, P; Blasco, J. (2014), A text mining approach to analyzing and classifying code structures in Android malware families, Expert Systems with Applications, 41 (4), 1104-1117
37. Tseng, Y.H. Lin, C.J. Lin, Y.I. (2007), Textminig techniques for patent analysis, Information processing and management, 43, 1216-1247
38. Ventura, J. Silva, J. (2012), Mining Concepts from Texts, Procedia Computer Science, 27-36
39. Weiss, S. Indurkha, N. Zhang, T. (2010), Fundamentals of predictive Text mining, ISSN 1868-095X, Springer
40. Yoon, J. (2012), Detecting weak signals for long- term business opportunities using Textmining of Web news, Expert systems with application, 39, 12543-12550
41. Zheng, W, et al. (2012), Morphological Analysis of Technologies using Multidimensional Scaling, Journal of Business Chemistry, 9, 147-160.



Text mining, as an effective tool for future studies: Introduction and Review of Literature

Abstract

In today's competitive environment, identify opportunities which are promising for growth and development is dependent on attention to future trends. Through which, we can view the long term towards science and new technologies to be acquired. Foresight is a process by which a correct understanding of the forces shaping the future, can be achieved and allows these forces at all stages of policy formulation, planning and decision making be considered. Generally, there are many different tools for analyzing and forecasting the future of science, technology and innovation. Among these, text mining is a widely used method and is widely used in future study fields. In today's world, the publication of scientific literature is so high and following all the information, is almost impossible for any researcher, so for facilitating access to related information, Text mining is as a useful tool available for researchers. Regarding the importance of future studies, in this study, text mining method was introduced and investigated by studies carried out with this method. Then, the application of this method was tested in several articles and grouped in five areas of text classification, information retrieval, clustering, information extraction and forecasting and evaluation.

Keywords

Foresight, Forecasting, Textmining