

نوآوری در مدل کسب و کار با استفاده از تحلیل کلان داده‌ها

منیره حسینی

دانشیار، گروه مهندسی فناوری اطلاعات، دانشکده مهندسی صنایع، دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران
hosseini@kntu.ac.ir

تینا اخلاقی مسکین

دانشجوی کارشناسی ارشد، گروه مهندسی فناوری اطلاعات، دانشکده مهندسی صنایع، دانشگاه خواجه نصیرالدین طوسی، تهران، ایران
tina.akhlaghimeskin@email.kntu.ac.ir

چکیده

با پیشرفت روزافزون تکنولوژی اینترنت و دستگاه‌های قابل حملی که به آن متصل می‌شود، ارتباط مشتریان با کسب و کارها نسبت به گذشته تغییر کرده است. علاوه بر شبکه‌های اجتماعی، کاربران نظرات خود را در وبسایت‌های خرید و فروش و فروم‌های تخصصی با یکدیگر به اشتراک می‌گذارند و این نظرات در خریدهای آتی دیگران تأثیرگذار است. در نتیجه این تعاملات حجم عظیمی از داده ایجاد می‌شود تحلیل متن به صورت سنتی و دستی هزینه زمانی، مالی و نیروی انسانی زیادی را بر شرکت‌ها متحمل می‌کند. با استفاده از متن کاوی^۱ و یادگیری ماشین^۲ می‌توان با سرعت بیشتر و هزینه مالی و انسانی کمتر تحلیل را انجام داد و از نتیجه آن در تدوین استراتژی و برنامه‌ریزی‌های آتی و نوآوری در مدل کسب و کار^۳ شرکت بهره برد. در این تحقیق مدلی پیشنهادی برای ترکیب نتایج مدل‌سازی موضوعی^۴ با استفاده از الگوریتم انتساب نهفته دیریکله^۵ و تحلیل احساس^۶ برای دستیابی به بینش عمیق‌تر درباره داده‌های متنی و استفاده از آن در نوآوری مدل کسب و کار پیشنهاد شده است.

واژگان کلیدی: کلان داده، مدل کسب و کار، مدل‌سازی موضوعی، تحلیل احساس

مقدمه

کاربران شبکه‌های اجتماعی با سرعت بالا داده‌هایی حجیم و متنوع در شبکه‌های اجتماعی، وبسایت‌ها و فروم‌ها تولید می‌کنند. (Hallikainen et al, 2020؛ Chaudhary et al, 2015؛ Chapman et al, 2019) این داده‌ها به علت خود اظهاری توسط کاربران از سوگیری‌های تجاری جلوگیری می‌کنند. همچنین به علت حجم بالا دید کلی‌تری درباره مشتریان، تقاضای آن‌ها و علاقیشان به کسب و کارها در طول زمان ارائه می‌دهند. متن کاوی^۸ به عنوان زیرشاخه‌ای از داده کاوی امکان بررسی داده‌های متنی تولیدشده توسط کاربران را با استفاده از یادگیری ماشین ممکن می‌سازد. (Ray et al, 2021) با کمک متن کاوی این داده‌های بدون

^۱ Forum

^۲ Data Mining

^۳ Machine Learning

^۴ Business Model

^۵ Topic Modeling

^۶ Latent Dirichlet Allocation (LDA)

^۷ Sentiment Analysis

^۸ Text Mining

ساختار برای کشف الگوها، ویژگی‌ها و روابط بررسی می‌شوند. (He et al, 2013) با استفاده از متن کاوی می‌توان طبقه‌بندی اسناد، مدل‌سازی موضوعی، ترجمه و تشخیص زبان متن، تشخیص اخبار جعلی و تحلیل احساس انجام داد. (Chatterjee et al, 2020) مدل موضوعی به دنبال کشف روندها و تم‌های پنهان در داده‌های بدون ساختار بدون نیاز به نظارت و از طریق روش‌های آماری و احتمالی با استفاده از رویکردهای پردازش زبان است. (Kwayu et al, 2022؛ Kushwaha et al, 2021؛ Wang and Goh, 2020) یکی از الگوریتم‌های مدل‌سازی موضوعی LDA است که ساختارهای معنایی در متون را با استفاده از تقسیم آن‌ها به موضوعات مشخص می‌کند. (Bastani et al, 2018) نوآوری مدل کسب‌وکار به معنی ایجاد تغییراتی در مدل کسب‌وکار شرکت به منظور بهبود عملکرد، رشد درآمد و حاشیه سود مدل کسب‌وکار است. (Acciarini et al, 2023) در این تحقیق به بررسی تحقیقات پیشین در حوزه استفاده از کلان داده‌ها در نوآوری مدل کسب‌وکار خواهیم پرداخت و سپس مدلی عملی برای استفاده از متن کاوی جهت تحلیل داده‌های موجود در شبکه‌های اجتماعی به منظور نوآوری در مدل کسب‌وکار ارائه خواهیم داد.

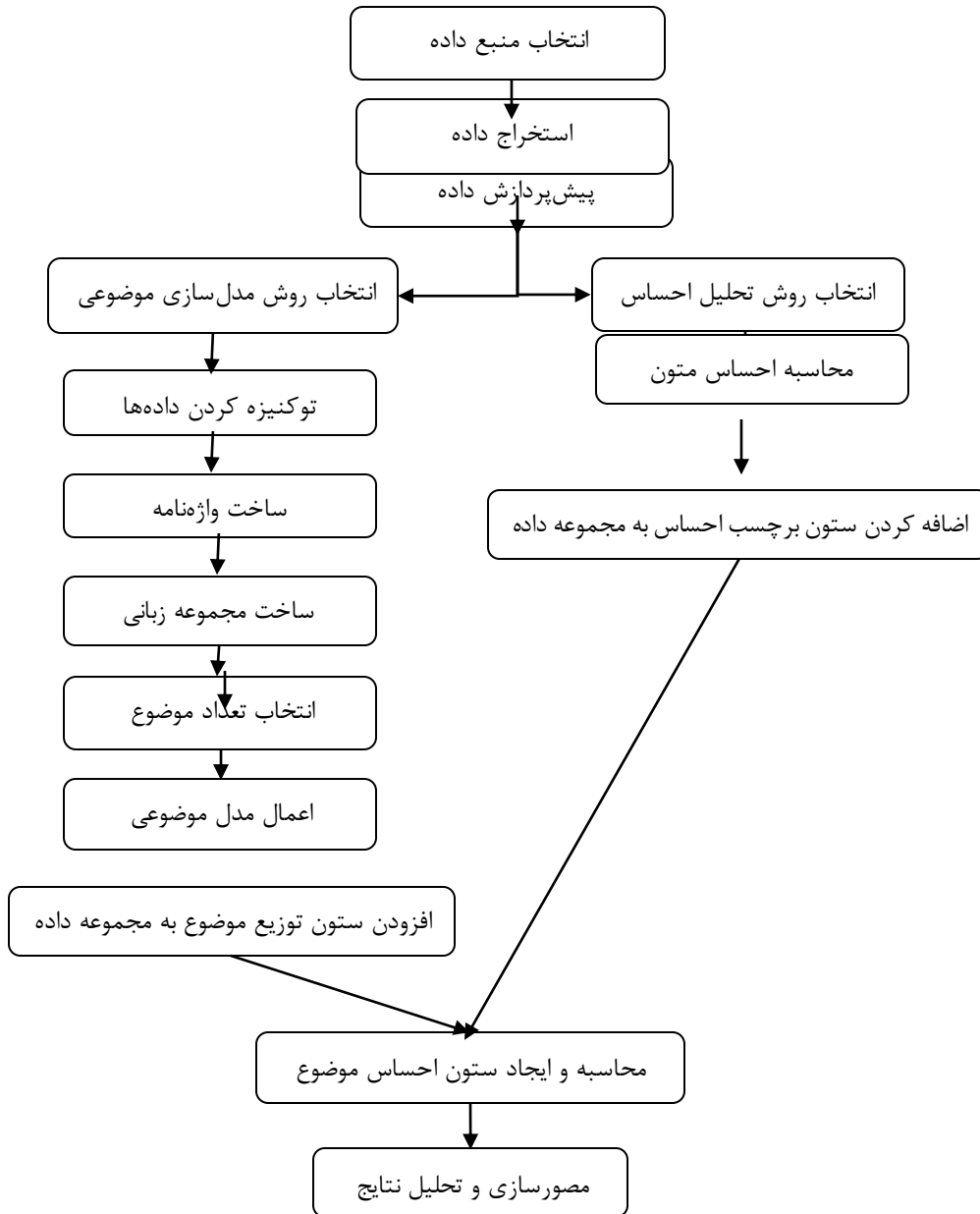
استفاده از کلان داده‌ها در نوآوری مدل کسب‌وکار

با توجه به حرکت سازمان‌ها از حالت دستی به دیجیتال، کلان داده‌ها^۹ با سرعت و حجمی بالا از منابع درونی و بیرونی در دسترس سازمان‌ها قرار گرفته‌اند که سازمان‌ها در تلاش‌اند تا از آن داده‌ها به بینش و اطلاعات مفیدی در جهت بهبود مدل کسب‌وکار خود جهت ایجاد نوآوری، حفظ موقعیت خود در بازار پویا و پیشی گرفتن از رقبای خود به منظور کسب ارزش بیشتر استفاده کنند. کلان داده‌ها منجر به افزایش سرعت نیاز به تغییر و بازپیکربندی مدل‌های کسب‌وکار شده است. (Wang, 2022) از کلان داده‌ها می‌توان برای شناسایی مهم‌ترین مشتریان، شناسایی محصولات و خدمات مورد تقاضای بیشتر، شناسایی نوآوری‌های سودآورتر در گذشته و شناسایی بخش‌های سودآورتر به صورت خودکار استفاده کرد. (Chaudhary et al, 2015) نوآوری در مدل کسب‌وکار می‌تواند تغییری در ارزش ارائه‌شده به مشتری توسط سازمان باشد. (Sorescu, 2017) کسب‌وکارها می‌توانند با استفاده از کلان داده‌ها فرآیند تولید محصول را بهبود دهند. (Dai and Liang, 2022) همچنین به کمک تحلیل کلان داده‌ها می‌توانند خدمات و محصولات اضافه‌تری به مشتریان ارائه دهند و تصمیمات بهتر در حوزه کسب‌وکار خود اتخاذ کنند. (Acciarini et al, 2023) استفاده از کلان داده‌ها علاوه بر مزایای عنوان‌شده شامل چالش‌هایی مانند هزینه ساخت پایگاه داده، مدیریت و تحلیل داده‌ها، حفظ امنیت، کاهش کیفیت تصمیم‌ها به علت زیاد بودن داده و تهدیدهای ایجادشده در صورت درز اطلاعات به بیرون و نیاز به محرمانگی است. (Acciarini et al, 2023)

مدل پیشنهادی برای استفاده از کلان داده‌ها در نوآوری مدل کسب‌وکار

مقالات بررسی‌شده از روش‌ها و ابزار گوناگون و گسترده‌ای استفاده کرده بودند اما فرآیند متن کاوی به‌طور کلی شامل سه مرحله پیش‌پردازش متن، تحلیل متن و تحلیل نتایج به دست آمده (کسب بینش از متن) می‌شود. شکل ۱ مراحل کلی مدل پیشنهادی این تحقیق را نشان می‌دهد.

^۹ Big Data



3

شکل ۱: مدل پیشنهادی برای تحلیل کلان داده ها به منظور نوآوری در مدل کسب و کار با استفاده از الگوریتم LDA

در این مدل پس از انتخاب منبع داده کلان مانند شبکه های اجتماعی، وبسایت ها و فروم ها، داده از منبع مورد نظر استخراج می شود. سپس بعد از انجام مراحل اولیه پیش پردازش داده، روش تحلیل احساس و روش مدل سازی موضوعی انتخاب می شوند. یکی از الگوریتم های مدل سازی موضوعی الگوریتم LDA است که برای مثال در این تحقیق به آن پرداخته می شود. پس از انتخاب الگوریتم،



متن‌ها توکنیزه می‌شوند، سپس از آن واژه‌نامه و مجموعه زبانی ساخته می‌شود و مدل موضوعی با استفاده از الگوریتم LDA بر روی داده‌ها اعمال می‌شود. در نهایت ستون توزیع موضوع به مجموعه داده اضافه می‌شود. برای تحلیل احساس پس از انتخاب روش تحلیل احساس و محاسبه حس چیره بر هر سند، ستون برچسب احساس به مجموعه داده اضافه می‌شود. در نهایت خروجی این دو روش در ستون احساس موضوع با یکدیگر ادغام می‌شوند تا در ادامه به تحلیل آن‌ها پرداخته شود. در ادامه هر مرحله به صورت گسترده‌تر بررسی خواهد شد.

انتخاب مجموعه داده

ظهور شبکه‌های اجتماعی تأثیر زیادی روی تعامل افراد با برندها داشته است و افراد اطلاعات خصوصی و نظراتشان را درباره برندهایی که با آن‌ها تعامل دارند در آن‌ها به اشتراک می‌گذارند، (Pitt et al, 2019) نظر دیگر مشتریان را درباره برند می‌خوانند و محتوایی که برندها برای تعامل با آنان ایجاد کرده‌اند را بررسی می‌کنند. (Robertson et al, 2019) مهم‌ترین شبکه‌های اجتماعی در این زمینه عبارت‌اند از:

۱. شبکه‌های اجتماعی عمومی مانند فیس‌بوک^{۱۰}، توئیتر^{۱۱}، یوتیوب^{۱۲}، اینستاگرام^{۱۳} و اسنپ‌چت^{۱۴}
 ۲. شبکه‌های اجتماعی تخصصی مانند لینکدین^{۱۵}، گلس‌دور^{۱۶} و تریپ‌ادوایزر^{۱۷} (Pitt et al, 2019; Robertson et al, 2019)
- علاوه بر شبکه‌های اجتماعی فروم‌ها و وبسایت‌ها نیز منابع داده متنی مهم برای شرکت‌ها شمرده می‌شوند. (Tang and Guo, 2013)

پیش‌پردازش داده

در این مرحله از تحلیل، داده‌های خام استخراج شده به کمک پاک‌سازی، اعمال ویژگی‌ها و یکپارچه‌سازی به داده‌های قابل استفاده در تحلیل تبدیل می‌شوند. (He et al, 2013) در روش پیشنهادی این مدل برای آماده‌سازی داده خام برای تحلیل، در ابتدا متن داده‌ها به حروف کوچک تبدیل شده، آدرس‌های وب^{۱۸}، اشاره‌ها^{۱۹}، کلمه‌های تک‌حرفی، حروف تکراری و علائم نگارشی حذف شده، کلمات ایستای عمومی و تخصصی که ارزشی را در تحلیل ایجاد نمی‌کنند پاک‌سازی شده، پرتکرارترین و نادرترین کلمات از متون حذف شده و در نهایت کلمات توکنیزه^{۲۰} می‌شوند.

تحلیل داده

با کمک استفاده از روش‌های تحلیل خودکار متن می‌توان هزینه و زمان به دست آوردن اطلاعات و بینش از داده‌ها را کاهش داد و بینش مورد نیاز را از داده کسب کرد. (Berger et al, 2020) در این مرحله خروجی به دست آمده از مرحله پیش‌پردازش داده به الگوریتم‌های تحلیل موردنظر داده می‌شود تا از آن‌ها به اطلاعات مفید برای کسب‌وکار دست‌یافت. تحلیل احساس و مدل‌سازی موضوعی با استفاده از الگوریتم LDA دو روش برای تحلیل داده‌های خام‌اند که در این مدل استفاده شده‌اند.

^{۱۰} Facebook

^{۱۱} Twitter

^{۱۲} Youtube

^{۱۳} Instagram

^{۱۴} Snapchat

^{۱۵} LinkedIn

^{۱۶} GlassDoor

^{۱۷} Trip Advisor

^{۱۸} URL

^{۱۹} Mention

^{۲۰} Tokenization

یکی از اطلاعاتی که متون می‌توانند به سازمان‌ها ارائه کنند احساس و عواطف نویسنده متن است که نشانگر میزان لذت یا عدم لذت نویسنده از یک موضوع است. تحلیل احساس امکان بررسی احساس موجود در متن، قطبیت آن و میزان ارتباط عاطفه به احساس نویسنده و رسیدن به بینش کمی از داده‌های کیفی را می‌دهد. (Ray et al, 2021)

الگوریتم LDA نوعی روش مدل‌سازی موضوعی است که به کمک آن می‌توان موضوعات نهفته یا تم‌های مورد استفاده توسط کاربران در توییت‌هایشان را شناسایی کرد. در این روش مولد احتمالی، هر موضوع یک برچسب دارد که توسط آن ساختار معنایی متن نشانه‌گذاری می‌شود. (Kushwaha et al, 2021; Bastani et al, 2018; Tresnasari et al, 2020) هدف الگوریتم LDA آشکارسازی ساختارهای معنایی در مجموعه اسناد متنی است.

ادغام نتایج و بررسی احساس موضوع

در این مرحله خروجی‌های این دو مدل مجزا به منظور افزایش بینش و اطلاعات به دست آمده از داده ترکیب می‌شوند. در این مرحله پس از محاسبه احساس چیره بر هر سند، نتایج آن در کنار توزیع موضوعی آن سند قرار داده می‌شود تا با به دست آمدن احساس موضوع در تحلیل‌های آتی و مصورسازی داده‌ها از آن بهره برده شود.

مصورسازی و تحلیل نتایج

در گام آخر متن کاوی، نتایج به دست آمده از تحلیل داده‌ها به اطلاعات قابل ارائه تبدیل می‌شود. این ارائه می‌تواند شامل جداول، نمودارها، تصاویر و متون قابل درک توسط مدیران ارشد سازمان‌ها و دیگر ذینفعان باشد. با استفاده از نتایج به دست آمده می‌توان تدابیری اندیشیده تا با استفاده از آن به پیش‌بینی تقاضا، شناسایی رفتار، تجربه و رضایت مشتریان پرداخت؛ جریان درآمدی و تعامل با مشتریان را بهبود داد؛ مصرف مشتریان و عملکرد سازمان را پیش‌بینی کرد؛ عملیات تولید و قانون‌گذاری‌های سازمان را بهتر کرد؛ لجستیک، بازاریابی و قیمت‌گذاری را بهبود داد و توسعه محصول و فروش را بهینه کرد و در نهایت ارزش ارائه شده به مشتری و ارزش اکتسابی توسط کسب‌وکار را افزایش داد. موارد بالا به سازمان‌ها اجازه می‌دهند تا با نوآوری در مدل کسب‌وکار خود مزیت رقابتی خود را حفظ کنند و کسب‌وکار خود را به سمت پیشرفت ببرند.

نتیجه‌گیری

در تحقیقات بررسی شده (Wang and Goh (2019), Liu (2020) و Kushwaha et al (2021) از الگوریتم LDA در کنار تحلیل احساس به صورت مجزا استفاده کرده بودند اما ترکیب این دو روش برای کسب بینش بیشتر از داده‌ها در تحقیقات پیشین استفاده نشده بود. در این تحقیق با ارائه مدلی برای ترکیب نتایج به دست آمده از تحلیل احساس و مدل‌سازی موضوعی با استفاده از الگوریتم LDA احساس موضوعات به دست آمده از داده‌ها بررسی شد. تحلیل احساس به تنهایی راهی برای شناخت احساسات مثبت، منفی یا خنثی افراد درباره محصولات، خدمات و ارتباط با شرکت‌ها را فراهم می‌کند. (Kumar et al, 2021) مدل‌سازی موضوعی نیز امکان جستجو و کشف تم‌های موجود در داده‌ها را فراهم می‌کند. (Kwayu et al, 2022; Kushwaha et al, 2021; Wang and Goh, 2020)

هر دو روش به محققان و مدیران اطلاعات مفیدی درباره داده‌ها ارائه می‌دهند اما مدل‌سازی موضوعی به تنهایی فقط موضوعات اصلی موجود در متن را نشان می‌دهد و تحلیل مثبت و منفی بودن هر موضوع بر عهده تحلیل‌گر است که باعث ایجاد سوگیری‌ها می‌شود و همچنین نیاز به شناخت حوزه محتوای بررسی شونده توسط تحلیل‌گر را ایجاد می‌کند. به همین شکل تحلیل احساس به تنهایی فقط احساس کلی متن را نشان می‌دهد. در مدل پیشنهادی این تحقیق با ادغام نتایج این دو روش به صورت خودکار و بررسی احساس موضوعات می‌توان حس کلی مشتریان نسبت به موضوع مورد بررسی را شناسایی کرد و با توجه به آن تصمیماتی با بینش بیشتر، با سوگیری کمتر و با نیاز کمتر به شناخت حوزه مورد بررسی اخذ کرد. با مصورسازی این خروجی امکان شناسایی موضوعات



مهم موجود در متون مانند تقاضای مشتریان، نیازها و خواسته‌های آن‌ها و احساساتشان به خدمات و محصولات شرکت، می‌توان پیشنهادهایی جهت بهبود و نوآوری در مدل کسب‌وکار به‌منظور افزایش ارزش و مزیت رقابتی بهره‌ارائه‌کرد. در این تحقیق از ترکیب تحلیل احساس و مدل‌سازی موضوعی و بررسی احساس موضوعات برای ارائه پیشنهادها برای نوآوری در مدل کسب‌وکار استفاده شد. تحقیقات آتی می‌توانند الگوریتم‌های دیگر متن‌کاوی و یادگیری ماشین و ترکیب آن‌ها را در این حوزه بررسی کنند و نتایج به‌دست‌آمده را با نتایج مدل پیشنهادی این تحقیق مقایسه کنند.

منابع

- Acciarini, Chiara. Cappa, Francesco. Boccardelli, Paolo. And Oriani, Raffaele. (2023). **How can organizations leverage big data to innovate their business models? A systematic literature review.** Technovation. Vol. 123
- Bastani, Kaveh. Namavari, Hamed. And Shaffer, Jeffry. (2018). **Latent Dirichlet Allocation (LDA) for Topic Modeling of the CFPB Consumer Complaints.** Version 1
- Berger, Jonah. Humphreys, Ashlee. Ludwig, Stephan. Moe, Wendy. Netzer, Oded. And Schweidel, David. (2019). **Uniting the Tribes: Using Text for Marketing Insight.** Journal of Marketing. Vol. 84, Issue 1, pp. 1–25
- Chapman, Chris. (2019). **Commentary: Mind Your Text in Marketing Practice.** Journal of Marketing. Vol. 84, Issue 1, pp. 26–31
- Chatterjee, Swagato. Goyal, Divesh. Prakash, Atul. And Sharma, Jiwan. (2021). **Exploring healthcare/health-product ecommerce satisfaction: A text mining and machine learning application.** Journal of Business Research. Vol. 131, pp. 815–825
- Chaudhary, Rashi. Pandey, Jalaj Ranjan. And Pandey, Prakhar. (2015). **Business model innovation through big data.** 2015 International Conference on Green Computing and Internet of Things (ICGCIoT). 2015 International Conference on Green Computing and Internet of Things (ICGCIoT)
- Dai, Bingqin. And Liang, Wenquan. (2022). **The Impact of Big Data Technical Skills on Novel Business Model Innovation Based on the Role of Resource Integration and Environmental Uncertainty.** Sustainability. Vol. 14, Issue 5
- Hallikainen, Heli. Savimäki, Emma. And Laukkanen, Tommi. (2020). **Fostering B2B sales with customer big data analytics.** Industrial Marketing Management. Vol. 86, pp. 90–98
- He, Wu. Zha, Shenghua. And Li, Ling. (2013). **Social media competitive analysis and text mining: A case study in the pizza industry.** International Journal of Information Management. Vol. 33, Issue 3, pp. 464–472
- Kumar, Sunil. Kar, Arpan Kumar. And Ilavarasan, P. Vigneswara. (2021). **Applications of text mining in services management: A systematic literature review.** International Journal of Information Management Data Insights. Vol. 1, Issue 1
- Kushwaha, Amit Kumar. Kumar, Prashant. And Kar, Arpan Kumar. (2021). **What impacts customer experience for B2B enterprises on using AI-enabled chatbots? Insights from Big data analytics.** Industrial Marketing Management. Vol. 98, pp. 207–221



- Kwayu, Keneth Morgan. Kwigizile, Valerian. Lee, Kevin. Oh, Jun-Seok. And Nelson, Trisalyn. (2022). **Automatic topics extraction from crowdsourced cyclists near-miss and collision reports using text mining and Artificial Neural Networks**. International Journal of Transportation Science and Technology. Vol. 11, Issue 4, pp. 767–779
- Liu, Xia. (2020). **Analyzing the impact of user-generated content on B2B Firms' stock performance: Big data analysis with machine learning methods**. Industrial Marketing Management. Vol. 86, pp. 30–39
- Pitt, Christine. Plangger, Kirk. Botha, Elsamari. Kietzmann, Jan. And Pitt, Leyland. (2019). **How employees engage with B2B brands on social media: Word choice and verbal tone**. Industrial Marketing Management. Vol. 81, pp. 130–137
- Ray, Arghya. Bala, Pradip Kumar. And Rana, Nripendra P. (2021). **Exploring the drivers of customers' brand attitudes of online travel agency services: A text-mining based approach**. Journal of Business Research. Vol. 128, pp. 391–404
- Robertson, Jeandri. Lord Ferguson, Sarah. Eriksson, Theresa. And Näppä, Anna. (2019). **The brand personality dimensions of business-to-business firms: a content analysis of employer reviews on social media**. Journal of Business-to-Business Marketing. Vol. 26, Issue 2, pp. 109–124
- Sorescu, Alina. (2017). **Data-Driven Business Model Innovation**. Journal of Product Innovation Management. Vol. 34, Issue 5, pp. 691–696
- Tang, Chuanyi. And Guo, Lin. (2013). **Digging for gold with a simple tool: Validating text mining in studying electronic word-of-mouth (eWOM) communication**. Marketing Letters. Vol. 26, Issue 1, pp. 67–80
- Tresnasari, Nur Annisa. Adji, Teguh Bharata. And Permanasari, Adhistya Erna. (2020). **Social-Child-Cas e Document Clustering based on Topic Modeling using Latent Dirichlet Allocation**. IJCCS (Indonesian Journal of Computing and Cybernetics Systems). Vol. 14, Issue 2, p. 179
- Wang, Jiaqi. (2022). **Research on Business Model Innovation Based on Big Data Analysis**. Frontiers in Business, Economics and Management. Vol. 6, Issue 2, pp. 114–116
- Wang, Xiaohui. And Goh, Dion Hoe-Lian. (2020). **Components of game experience: An automatic text analysis of online reviews**. Entertainment Computing. Vol. 33