

DOI: https://doi.org/10.48009/4_iis_2024_110

Information extraction from the reviews of AI applications using SAS text Mining Process

Imranul Kabir Chowdhury, *Missouri University of Science and Technology*, ichowdhury@mst.edu

Wen-Bin Yu, *Missouri University of Science and Technology*, yuwen@umsystem.edu

Abstract

The recent surge in AI advancements, particularly generative AI, has grabbed the attention of major tech companies who are all striving to integrate it seamlessly into their products. This highlights the need to evaluate the various AI products on the market and identify the features that resonate most with users. SAS Enterprise Miner, a powerful text analysis tool, can be used to uncover hidden insights from user reviews. In this study, we will leverage various SAS functions to analyze reviews of different AI applications. By analyzing these reviews, we aim to pinpoint the unique strengths of these AI applications and propose improvements for existing ones. Furthermore, we hope to shed light on the potential social and economic impacts of these advancements based on our findings.

Keywords: feature extraction, text mining, SAS Enterprise Miner, AI products, customer reviews, information retrieval.

Introduction

This paper aims to identify the most appealing features of AI products by analyzing user reviews with SAS Enterprise Miner, a powerful tool for handling large amounts of text data. By pinpointing frequently mentioned features and functionalities, this analysis can empower users to make informed choices and guide developers in prioritizing future features. Additionally, the project will compare the capabilities of different AI products, explore their unique strengths and weaknesses based on user sentiment, and identify potential social and economic impacts. This comprehensive analysis will provide valuable insights to improve existing AI applications, inform marketing strategies, and ultimately guide logical decision-making within the AI product landscape.

Literature Review

Generative AI, a type of artificial intelligence that can create new content, has become significantly more powerful and efficient merit to recent advancements in deep learning. Examples like Bard (Gemini) and ChatGPT have garnered significant attention. This technology has the potential to revolutionize various industries like healthcare, finance, and entertainment. While still in its early stages, businesses are actively looking to integrate generative AI engines into their applications to better serve their customers reported by IBM (IBM, 2024). Developers are also busy creating specialized generative AI solutions for the market as per marketplace named Product Hunt (Wright, 2024). However, to ensure effective integration, it's crucial to review and analyze the requirements for these AI tools/ plugins developed based on generative AI engines as Parti mentioned the benefit of using no-code tools for automation (Parti, 2024). Text mining techniques, which involve analyzing large amounts of textual data, can be used to gain valuable insights from user

reviews. This analysis can help us understand how to best utilize text mining on user comments and reviews found on webpages and blogs.

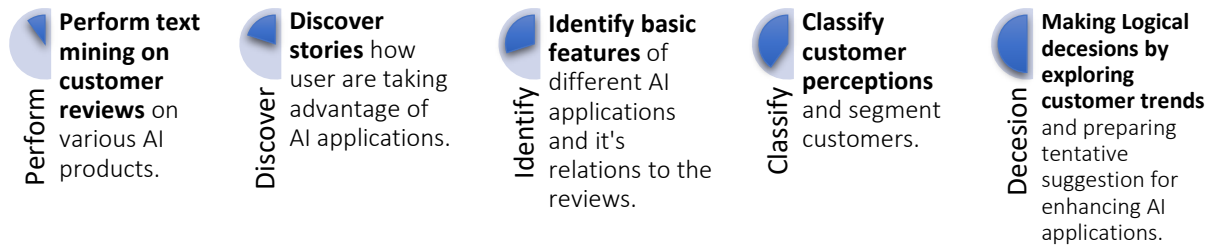


Figure 1: Info Retrieval Process.

Chowdhury with his associates (Chowdhury, Latif, & Saddam, 2016) explores the use of microblogging platforms (like Twitter) to analyze informal text for opinion mining on various topics. Businesses can leverage this extracted information, such as public sentiment on a particular product or service, to make informed decisions. This research lays the groundwork for an information extraction framework that can be used to verify sentiment intensity of public opinion on a larger scale using unstructured data. In one study Segall and his team (Segall, Zhang, & Cao, 2009) compared SAS Enterprise Miner to another software program, Megaputer PolyAnalyst, using hotel survey data. Their research revealed that SAS Enterprise Miner offered a more in-depth analysis of the customer reviews, providing businesses with richer insights into guest experiences. The ability to integrate with social media platforms is another advantage of SAS Enterprise Miner. Yap with his colleagues (Yap, Abdullah, Abdul-Rahman, & Tan, 2018) conducted a study where they analyzed Malaysian customer reviews of Proton cars retrieved directly from the company's Facebook page. By employing sentiment analysis techniques within SAS, they were able to gauge overall customer satisfaction, revealing a positive sentiment towards the car brand. Mobile app reviews also hold valuable insights. Liu and team (Liu, Sarkar, & Chakraborty, 2013) investigated reviews of Android apps using SAS Enterprise Miner. Their study compared different sentiment analysis models and found that NLP-rule based models provided a deeper understanding of customer sentiment compared to standard statistical models. This highlights the importance of using the right tools to extract the most meaningful information from customer reviews.

Social media listening in real-time allows businesses to stay on top of customer sentiment. Jha and Guggila (Jha, Guggila, & Chakraborty 2016) explored tweets about Tesla cars using SAS Enterprise Miner. Their research uncovered a positive sentiment towards the brand and a strong customer preference for Tesla vehicles. This real-time feedback allows businesses to address any emerging concerns and capitalize on positive trends. Understanding the relationship between customer experiences and expectations is crucial for improvement. Tolety and Choudhary (Tolety & Choudhary, 2018) analyzed American Airlines customer reviews with SAS to identify these connections. This valuable information can be used by airlines to tailor their services to better meet customer needs and expectations, ultimately leading to increased customer satisfaction.

Finally, Mellachervu (Mellachervu, n.d.) showcased how SAS Enterprise Miner can be used to inform product launch strategies. Their study examined Amazon customer reviews to identify product strengths and weaknesses. By leveraging these insights, businesses can develop products that resonate with their target audience and achieve greater success upon launch.

Methodology

In this project we have followed the below structure as our methodology for feature extraction of AI products. The methodology includes exploring data by collecting data manually in excel format to gather user comments. This data then undergoes text mining, where SAS Enterprise Miner, a data mining tool, extracts and analyses the comments. The tool identifies relevant pieces of information, or features, from the comments. By analyzing these features, the process aims to find valuable insights, such as user sentiment or common topics. Finally, these insights are used to draw conclusions that can improve the deeper understanding of user opinion about the AI applications and market trend.



Figure 2: Research Methodology.

Data Exploration

The first challenge is to find the dataset and identify the tools for extracting & preparing the dataset to fit for handling and running our desired functions in SAS text miner client software. The data collection method that we followed is provided below in figure.

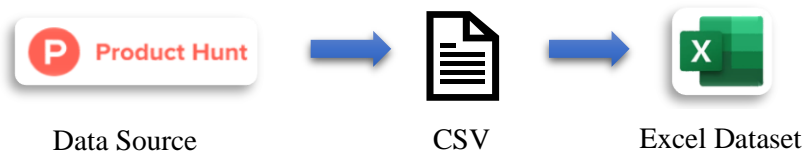


Figure 3: Data Collection Method.

Data have been collected manually in a CSV format. For this project it is decided to collect the first ten products to see how people are reacting to the AI plugins or tools. After running the web tool, got to export the dataset as excel format shown below:

Product	Comments	Length
Emma AI	The navigation on the Emma AI website is intuiti	756
Pitch N Hire Applicant Tracking	Promptly, cheerfully, and professionally. Folkw	725
ChainGPT	ChainGPT and ChatGPT are the same entity.Afte	708
Movie Deep Search by AI Keyta	I'm a participant of the Movie Deep Search Beta	691
GPTBots.AI	This is an amazing product! It seamlessly conne	677
Venturefy.ai	I think Venturefy is going to be a deal breaker	660
ChatGPT	I dont think so anyone currently doesnt know w	659
GPTBots.AI	It offers a user-friendly interface and seamless	647
Emma AI	I can't believe how amazing my experience has	643
Canva	As a non-designer, it's made my life very easy v	643
Zipy	Zipy has truly revolutionized my debugging exp	642
GPTBots.AI	GPTBots.AI refers to artificial intelligence techn	637

Figure 4: Exported dataset in excel.

Text Mining Process

SAS Enterprise Miner unlocks valuable insights from user reviews through a text mining process. It starts with the text parsing node, which acts like a word separator, breaking down each comment into individual terms. These terms are then organized by frequency to identify the most commonly used words. Next, the text filtering node refines the analysis by eliminating irrelevant terms, similar to decluttering your closet to focus on the most important items. Finally, the text clustering node takes center stage. Here, a technique called Singular Value Decomposition (SVD) helps manage the data by transforming it into a more manageable format. Different clustering methods are then used to group similar terms together, revealing hidden patterns within the reviews. This process allows us to extract valuable knowledge from seemingly unstructured text data. Below is the setup for the SAS text miner application:

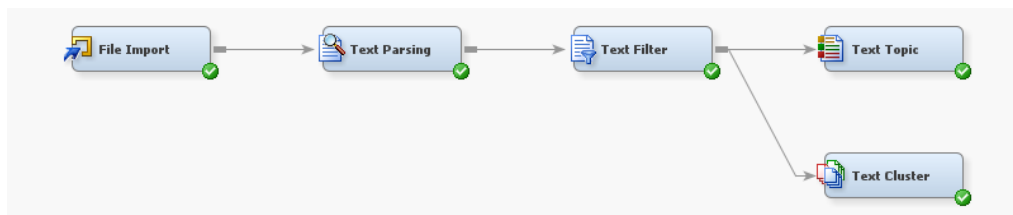


Figure 5: Text mining process using SAS Enterprise Miner.

Result

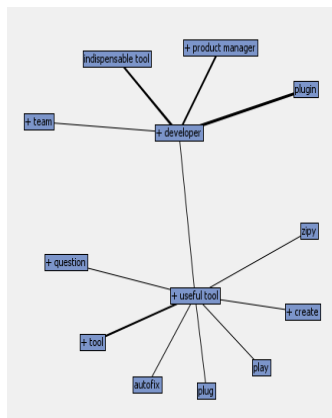
This project delves into user reviews on Product Hunt, a platform for discovering new products, to understand what features resonate most with users in the realm of AI. By analyzing comments about recent AI developments and how they optimize product features, the researchers aim to identify the most valuable and interesting aspects of AI products. Following a step-by-step approach, they first use an "interactive filter viewer" to pinpoint frequently mentioned terms, likely reflecting key AI features based on user discussions (details likely shown in Table 1). Next, they plan to delve deeper into each term, analyzing comments to tell a "story" that explains how AI is applied in that context (e.g., automating hiring) and how users perceive this application (e.g., positive comments suggesting user appreciation). Ultimately, this project leverages user voices to gain insights into successful and user-friendly AI features.

Table 1: High Weighted Terms.

Term	Frequency	Doc	Weight
Hire	27	20	0.613, 0.66
Integrate	22	21	0.599
Struggle	5	5	0.787
issue	14	14	0.65
Problem	20	20	0.603
developer	31	27	0.569
Useful tool	10	10	0.695
zeda	13	11	0.688

Concept Link Results

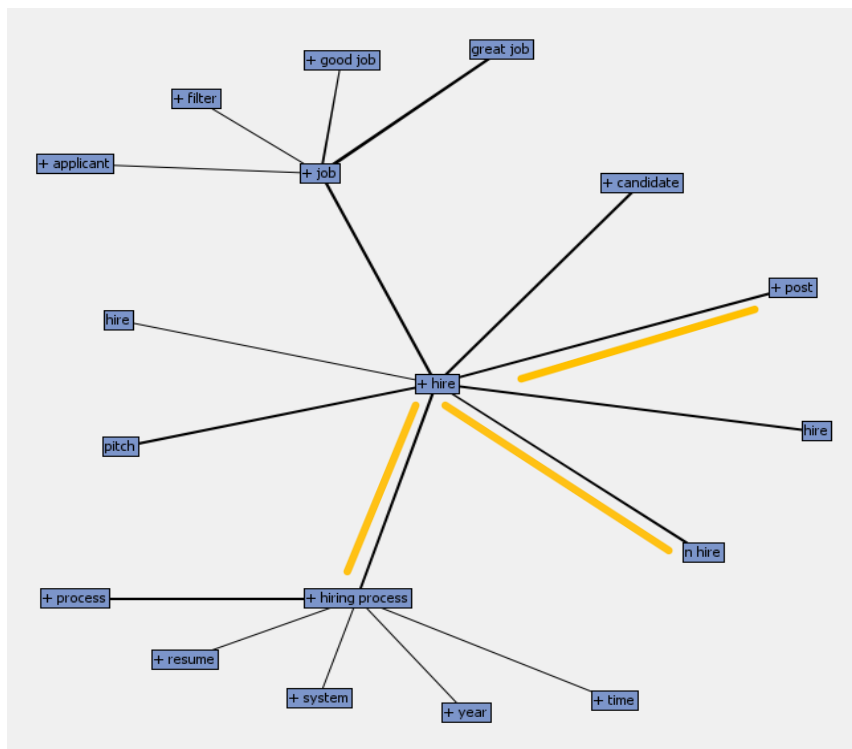
As per the findings, the result had a view of 2-3 order concept link. It can easily draw the conclusion from the below concept link scenario of platform integration, for example ai platform is capable of integrating enterprise data and user-friendly interface and help prepare chatbot. The concept link “developer” says that zipy is a useful tool/ plugin for developer that can auto fix customers bug quickly and text relevancy is 0.4.



COMMENTS	TEXTFILTER_SNIPPET ▲	TEXTFILTER_RELEVANCE	LENGTH	PRODUCT
This is truly remarkable! We use Zipy for a variety of purposes, such as pinpointing and reproducing errors, observing user behaviors, and obtaining feedback on user experience. Despite identifying the errors, our team of developers used to spend hours, if not days, searching online for solutions or consulting with experts to address them. However, with the integration of the ChatGPT plugins, our developers can now find solutions in the same context as their work.	... I We use Zipy for a variety of purposes ... Kudos to the Zipy team for their excellent work ...	0.4	566.0	Zipy
Kudos to the Zipy team for their excellent work! I would gladly recommend this product to others.				
Zipy Plug&Play with AutoFix is a must-have tool for anyone who wants to solve customer bugs quickly. I'm impressed with how it runs on any website or JS-based app, and how it provides AutoFix suggestions in your local Zipy sessions.	Zipy Plug&Play with AutoFix ... in your local Zipy session. ...	0.4	231.0	Zipy
Zipy has completely transformed my debugging game.	Zipy has completely ...	0.2	50.0	Zipy
Zipy has truly revolutionized my debugging experience. As a unified platform for user session replay, frontend, and network monitoring, it offers an efficient and integrated solution for all my debugging needs. The easy and quick installation process is commendable, and being able to replay error sessions in real-time is a game-changer. Its vast array of features, including developer tools, Stack Traces, Console, and Network Logs, make it a powerhouse for identifying and fixing issues. Plus, the Maker's readiness to assist with queries further enhances the overall user experience. Truly, Zipy is an indispensable tool for any developer.	Zipy has truly revolutionized my debugging ... Truly, Zipy is an indispensable tool for ...	0.4	642.0	Zipy
Zipy's Session Playback and Monitoring is a game-changer for Chrome Extensions. Debugging made easy	Zipy 's Session Playback and Monitoring ...	0.2	98.0	Zipy
Amazing work team Zipy. I have been using Zipy from it's early days for replaying user sessions and understanding the behaviour or steps by which user encountered errors. Zipy's ChatGPT	... Amazing work team Zipy . I have been using ... have been	0.6	254.0	Zipy

Figure 6: Concept link "Zipy".

The term “Hire” influenced by the reviewer as per screenshot below. To check the story behind the concept link it was required to search further the concept link and get the scenario, which verifies the information as desired and for this search. Finally, SAS responded with a relevancy score 1.0 and discovered the product is Pitch N hire.



COMMENTS	TEXTFILTER_SNIPPET	TEXTFILTER_RELEVANCE	LENGTH	PRODUCT
Pitch N Hire is actually definitely beneficial at pushing candidates to perform well since the total complexity of the evaluations makes the test difficult to complete because you have to be extremely focused. This is really beneficial to us because it makes the filtering process easier to go forward in successive stages of the recruitment process.	... Pitch N Hire is actually definitely beneficial at ...	0.182	350.0	Pitch N Hire Applicant Tr...
Companies may employ the top people by using Pitch N Hire. Pitch N Hire features a highly user-friendly UI and is very well-organized. Candidates can search through a wide variety of jobs on the website to discover one that best suits them. I heartily urge everyone to utilize Pitch N Hire because it's wonderful for both employers and job seekers.	... using Pitch N Hire. Pitch N Hire features ... wide variety of jobs on the website to discover ... utilize Pitch	1.0	348.0	Pitch N Hire Applicant Tr...
The site is simple to use, has a nice front page dashboard with solid metrics, can post jobs and track applications rapidly, and interfaces with the HRIS system. The hiring process is efficient and simple to move candidates through. It is also simple to search for candidates and maintain records. excellent customer and technical support.	... can post jobs and track applications rapidly, ... system. The hiring process is efficient and simple	0.566	339.0	Pitch N Hire Applicant Tr...
Pitch N Hire's versatile flexibility offers a distinctive, all-encompassing solution for a range of project needs. It has been used to create forms for task assignment, progress monitoring, MVP creation for field testing, and other purposes. There isn't much to criticize, either, thanks to the excellent user interface.	... Pitch N Hire's versatile flexibility offers a ...	0.182	320.0	Pitch N Hire Applicant Tr...
I use Canva for basically everything! From social media posts to presentation to even writing a resume.	... Canva made my job so easy and enriched the ...	0.159	310.0	Canva
I'm a visually sensitive person and before Canva, I sept hours finding the right icon or picture for my presentations. Canva made my job so easy and enriched the quality of my work.				
Thank you Canva team.				
I recently needed to employ an associate to assist me with my work and chose to do so via Pitch N Hire. I received many applications, and Pitch N Hire was able to help me keep track of the applicants and create my evaluations. Workflow management is another feature of this tool that is quite beneficial.	... via Pitch N Hire. I received many applications ... and Pitch N Hire was able to help me ...	0.364	304.0	Pitch N Hire Applicant Tr...
My onboarding procedure was handled by my organization using Pitch N Hire, and it was really simple. Through the portal, they also let us know about payroll/PTO and perks, so I check it frequently. I've never had a problem with it. Compared to other platforms, operates a lot better in my experience.	... using Pitch N Hire, and it was really ...	0.182	300.0	Pitch N Hire Applicant Tr...
Pitch N Hire offers users an intuitive platform for monitoring time and receiving pay information. I continue to use the service	... Pitch N Hire offers users an			

Figure 7: Concept link "hire".

Cluster Result

Different cluster settings were tried to reach the best mutual exclusive combination of the cluster. The optimal one (Figure 8) we got with train settings: for variable SVD resolution low, dimensions 200, and for cluster option maximum and cluster number 50. SAS Enterprise Miner responded with 12 clusters, which is a better combination as it can distinguish the mutually exclusive information to the limit. During the process the following screenshot has been taken for demonstration.

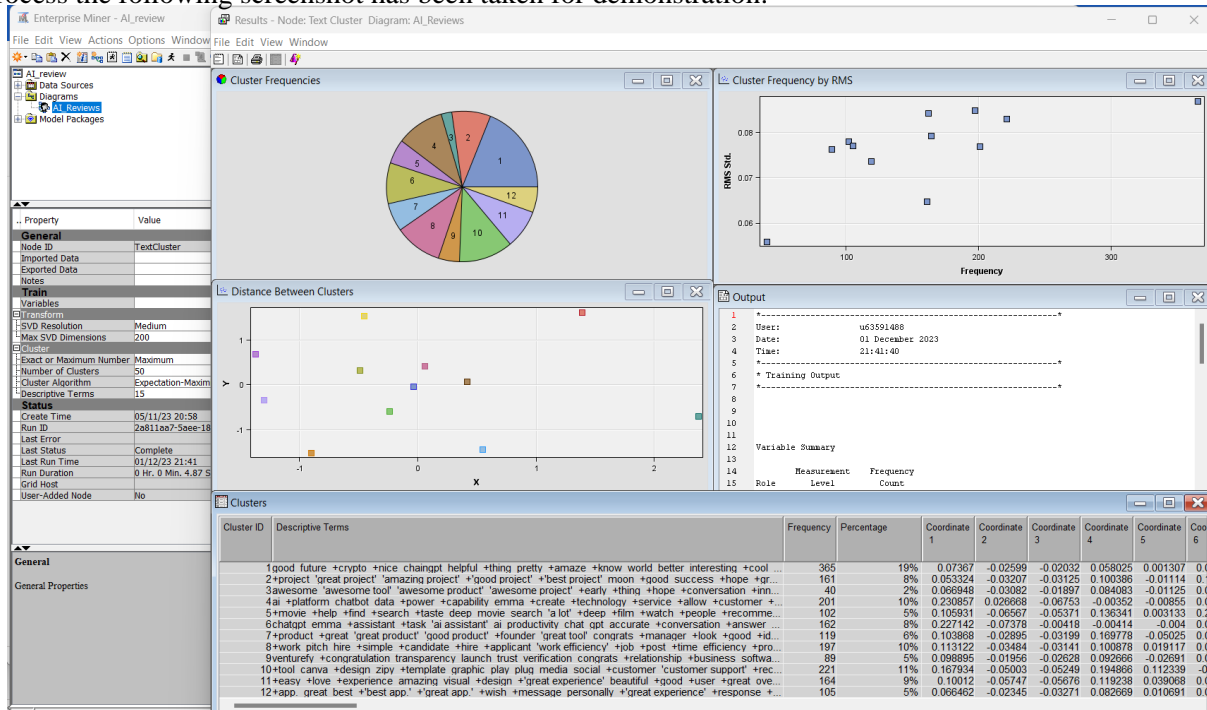


Figure 8: Cluster Result (12 clusters).

Below are our cluster descriptions from the better result (Figure 26):

Table 2: Cluster Description.

Cluster	Description	Cluster	Description
Group 1	reflecting future of crypto and product chainGPT.	Group 7	is about customer good product.
Group 2	reflecting about good project.	Group 8	is reflecting hiring process application pith hire.
Group 3	reflects review of awesome tool.	Group 9	is about business relation product venturify.
Group 4	reflects customer service tool emma.	Group 10	talks about graphic design tool Canva.
Group 5	talks about deep movie search.	Group 11	represents about user experience.
Group 6	is about chainGPT (blockchain based)	Group 12	discuss more about great app.

Analysis from the Result

A significant portion of users discuss specific products and services, including Emma, Venturify, Canva, and a category of "Great Apps." This indicates a strong trend of user feedback within the analyzed data. Businesses can leverage this information by conducting sentiment analysis within these clusters. By understanding how users feel about these products and services, businesses can identify areas for improvement, address user pain points, and ultimately enhance customer satisfaction.

Another interesting insight comes from user discussions on future trends and innovation. Groups focus on topics like cryptocurrency (future of crypto) and innovative projects, suggesting users are engaged with advancements and have expectations for progress. Analyzing these clusters can provide valuable insights for businesses. They can learn about user expectations for future developments and potentially identify new product development directions that align with user interests. The analysis also reveals user focus on process improvement. Discussions around deep movie search and the hiring application Pith Hire highlight a desire for efficiency and effectiveness in specific tasks. Businesses can use insights from these clusters to streamline user workflows. By understanding the aspects of processes that users find cumbersome or inefficient, businesses can implement changes that make it easier for users to achieve their goals.

Finally, a dedicated cluster focused on user experience (UX) suggests that users value intuitive and user-friendly interfaces. Analyzing this cluster can provide valuable feedback for businesses. By understanding how users interact with their products and services, businesses can identify areas where the user interface can be improved. This can lead to a more positive user experience overall, potentially increasing user satisfaction and engagement.

Conclusion

In conclusion, by understanding user opinions through these clusters, businesses can make data-driven decisions to improve user experience, product development, and overall customer satisfaction. Analyzing user feedback allows businesses to stay attuned to user needs and expectations, ultimately leading to a more successful and user-centric approach.

References

- Chakraborty, G. (2015, April 26-29). *Alien Nation: Text Analysis of UFO Sightings in the US Using SAS® Enterprise Miner™ 13.1*. SAS Global Forum, Dallas, TX, USA.
- Chowdhury, I. K., Latif, S., & Saddam, Md. S. H. (2016). Sentiment intensity analysis of informal texts. *International Journal of Computer Applications*, 147(10), 24-31.
<https://doi.org/10.5120/ijca2016911195>
- IBM. (2024, June). *GenAI Platforms Buyers Guide [2024 Software Provider and Product Assessment]*.
<https://www.ibm.com/downloads/cas/8Q705AM3>
- Jadhavar, A., Sirolkar, P., & Chakraborty, G. (2016, April 18-21). *Analysis of IMDB reviews for movies and television series using SAS® Enterprise Miner™ and SAS® Sentiment Analysis Studio*. SAS Global Forum, Las Vegas, NV, USA.

- Jadhavar, R., & Komarraji, A. K. (2018, April 8-10). *Sentiment analysis of Netflix and competitor tweets to classify customer opinions*. SAS Global Forum, Denver, CO, USA.
- Jha, T., & Guggila, P. *Analyzing sentiments in tweets for Tesla Model 3 using SAS Enterprise Miner and SAS Sentiment Analysis Studio*. The South-Central SAS Users Group (SCSUG) G Student Symposium 2016, San Antonio, TX, USA.
- Joung, J., Jung, K., Ko, S., & Kim, K. (2018). Customer complaints analysis using text mining and outcome-driven innovation method for market-oriented product development. *Sustainability*, 11(1), 40. <https://doi.org/10.3390/su11010040>
- Liu, J., Sarkar, M. K., & Chakraborty, G. (2013, April 28-May 1). *Feature-based sentiment analysis on Android app reviews using SAS® Text Miner and SAS® Sentiment Analysis Studio*. SAS Global Forum, San Francisco, CA, USA.
- Mellachervu, M., & Minukur A. R. (2018, September 30 - October 2). *Analyzing Amazon's customer reviews using SAS® Text Miner for devising successful product launch strategies*. MidWest SAS Users Group (MWSUG) 2018 Conference, Indianapolis, IN, USA.
- Kakarla, R., Chakraborty G. (2015, April 26-29). *Donor sentiment and characteristic analysis using SAS® Enterprise Miner™ and SAS® Sentiment Analysis Studio*. SAS Global Forum, Dallas, TX, USA.
- Parti, A. (2023, August). What is no-code automation? Types, benefits and tools. Product Hunt. <https://www.producthunt.com/stories/the-top-15-ai-product-launches-from-2023>
- Segall, R. S., Zhang, Q., & Cao, M. (2009). *Web-based text mining of hotel customer comments using SAS® Text Miner and Megaputer PolyAnalyst®*. 2009 Southwest Decision Sciences Institute Annual Conference SWDSI, Oklahoma City, OK, USA.
- Singh, M. K., Batti, N., & McGaugh, M. (2016, October 16–18). *Text mining and sentiment analysis on video game user reviews using SAS® Enterprise Miner™*. SouthEast SAS Users Group (SESUG) Conference 2016, Bethesda, MD, USA.
- Tolety, R., & Choudhary, S. K. (2018, September 7-9). *Text analysis of American Airlines customer reviews*. Western Users of SAS Software 2016 Conference, September 7-9, 2016, San Francisco, CA, USA.
- Woodfield, T. J., & Irvine, C. A. (2005, April 10-13). *Predicting workers' compensation insurance fraud using SAS Enterprise Miner 5.1 and SAS Text Miner*. SAS Users Group International Conference (SUGI 30), Philadelphia, PA, USA.
- Wright, S. (2024, January). *The top 15 AI products from 2023*. Product Hunt. <https://www.producthunt.com/stories/the-top-15-ai-product-launches-from-2023>
- Yap, B. W., Abdullah, N., Abdul-Rahman, S., & Tan, M. L. P. (2018). Text mining and sentiment analysis on reviews of Proton cars in Malaysia. *Malaysian Journal of Science*, 37(2), 137-153. <https://doi.org/10.22452/mjs.vol37no2.5>