International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

AI-Augmented Voice of Customer: Applying Data Analytics and NLP to Drive Product Insights at Scale

Sandeep Aluvaka

Submitted: 04/01/2024 **Revised:** 10/02/2024 **Accepted:** 22/02/2024

Abstract: This research develops an AI-supported framework to read and analyze customer feedback using NLP and ML to identify actions that can be taken from the unorganized feedback. The system accomplishes feedback categorization and identifies users' concerns and trends by combining sentiment analysis, topic modeling and emotional detection. In each customer journey stage, marketing allows for better product creation, supporting services and overall customer satisfaction. AI methods are used in a case study to make VoC intelligence possible on-demand and on a large scale. According to the study, AI can greatly simplify the analysis process and increase the precision of decisions.

Keywords: *Customer, NLP, AI, Augmentation, Data Analytics, Product*

I. INTRODUCTION

Understanding what customers need is necessary for developing products and services that care about customers. Conventional feedback analysis techniques find it difficult to grow, stay objective and provide information quickly because of the rapid increase in unstructured data online.

This work suggests a framework that uses AI and machine learning to quickly study social media reviews from real-time data and discover the sentiments, main subjects, and emotions within those reviews. By relating feedback to each customer journey stage, businesses notice early signs of clients leaving, focus on what needs improvement and answer early. Using AI with VoC techniques revolutionizes how decisions are made in experience management and product development.

II. RELATED WORKS

Customer analysis through AI and NLP

In the digital market today, the Voice of Customer greatly influences how products are designed and what makes one company different from others. Previously, businesses turned to interviews, focus groups and surveys to hear what customers say, though these approaches generally can't cover the data needed, do so quickly or in enough detail [5].

A rising number of users publishing app reviews and taking part in support forums and chats means it is now important to analyze vast amounts of unformatted comments using new tools. These developments, both AI and NLP are being put to use in the VoC pipeline.

NLP has gone from recognizing keywords to exploring meaning in context, models such as BERT and RoBERTa which now help with accurate sentiment analysis, topic extraction and opinion mining [4][3]. GenAI has led to more ways for companies to use VoC analysis.

In contrast to previous models, GenAI models are flexible and understand what users are trying to convey and what they feel without need of an ontological dictionary. These capabilities, product teams can better react and address the needs of customers by understanding feedback as if they were humans.

GenAI helps product teams connect with stakeholders and managers by asking questions and improving decision-making processes [1][5]. Their usefulness in different languages and from various feedback forms helps them to be highly respected in product development around the world.

AI's Strategic Role

With AI and product management working together, VoC analysis shifted from being an activity that looks back to a tool for future planning. Automated sentiment analysis helps businesses track how satisfied their customers are and see how sources of feedback change.

Since rapid iteration matters in places like SaaS or e-commerce, these abilities are essential in development [2]. In addition, feature demand modeling, using supervised learning, clustering and NLP, permits businesses to find out which product features users prefer most [2][3].

Analyzing customer comments, AI enables companies to update their product plans based on real customer feedback as they go [2]. In agile teams, this feature helps because changes are made fast and usability responses have to be included early.

Frisco, TX, 75036

Tools using AI are used to predict how proposed changes might be adopted and how long they will be kept, making it possible for product managers to balance introducing new concepts with maintaining stability [2]. Even so, there are risks involved when putting such systems in place. Overusing models that rely on numbers can make it hard to spot important qualitative insights which is why data bias and interpretability have to be managed carefully to protect the decision-making process [8].

Methodological Innovations

What modern VoC systems depend on are different approaches: topic modeling, discerning sentiment and clustering methods. All of these methods work together to turn messy feedback into valuable knowledge that teams know what to do with.

The application of BERTopic which depends on BERT embeddings and grouping methods, is leading to better grouping of topics found in large volumes of text [3]. When organizations use RoBERTa-based sentiment classifiers, they can better see what customers are saying and what these discussions mean to them.

Such an exploration allows us to figure out the emotional causes behind users' complaints with greater detail and give strategy makers relevant insights. In automotive engineering such techniques have divided user requirements into themes such as performance, comfort and safety, indicating that these methods work well in many industries [3].

In addition to industry applications, there are frameworks created by research that classify feedback following the main stages of a customer's life with a product [9]. Semi-supervised learning and text classification are used as part of these frameworks which still allows feedback segmentation with very few labeled data.

Connecting themes to their sentiment and sorting issues by time can detect early issues, identify risks for users leaving and spot new desires voiced by users—capabilities tools and surveys just do not support.

AI-Augmented Feedback Pipelines

Feedback analytics pipelines are improving, tools and automation are more important than ever. With NLP, Conversational AI systems already help customer support teams by instantly and correctly responding to user inquiries without human action needed [7].

These days, banking and healthcare use chatbots to answer questions more quickly and generate useful data. Combining Business Intelligence (BI) and Artificial Intelligence (AI) in Augmented Analytics (AA) platforms, companies can automate the steps of preparing, visualizing and discovering insights from their data [8].

They allow anyone to speak with data models naturally, making insights available to all. With the positives, come some difficulties. Many times, important data sets do not have labels in areas where there aren't many resources or they are not in English.

Due to this problem, supervised models perform weaker and greater effort should be placed on learning using unsupervised and semi-supervised techniques [9][10]. The way user feedback data is used ethically is becoming more important. Because large language models are used for AI, there is a possibility that their results may be biased [1][10].

Using mined user data from Facebook, Twitter and app stores can create privacy issues too. Therefore, users' information should be anonymized, models should be easy to follow and feedback options should be offered. Researchers have worked on putting real-time dashboards together with AR/VR to make analysis more accessible and useful [6]. This approach reflects what is happening more generally in human-in-the-loop systems, with AI acting to support, not oust, the skills of humans.

AI-driven insights are used for strategy, it is important for organizations to put together teams made up of ethicists, engineers and experts from aligned fields, to guarantee fair and balanced results. AI making Voice of Customer systems smart means companies are now able to act differently to feedback from customers.

Pausing the AI in place of machines, these systems can help product teams discover useful information at a fast pace. Although automatic roadmap creation and continuous feedback are interesting, care must be taken during implmentation to think about the quality of data, how easily it can be understood and ethical issues. For AI-driven VoC pipelines to unleash the full power of customer-centric innovation, continued studies and the involvement of multiple experts are important.

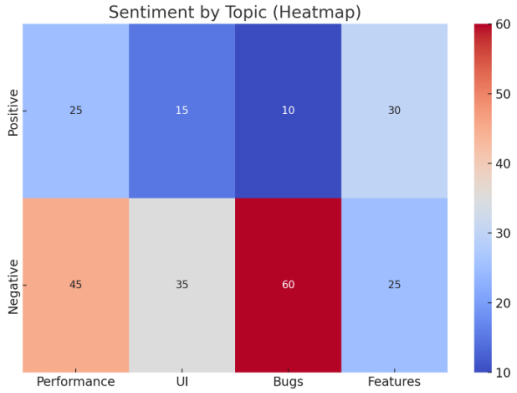
III. FINDINGS

AI-Augmented VoC Pipelines

With AI and NLP in Voice of Customer (VoC), we find that it is easier to draw useful insights from data that isn't in a structured format. Learning about how customers feel about products is getting more difficult given how fast products change [5].

On the other hand, text data in VoC pipelines comes from app feedback, customer support notes and digital platforms. AI models with LLMs and recent NLP tools help analyze the data to spot trends, notice where customers are struggling and spot requests for updated features [1][2]. Generative AI is particularly effective

at explaining what lies beneath the words we use in daily life and work conversation.

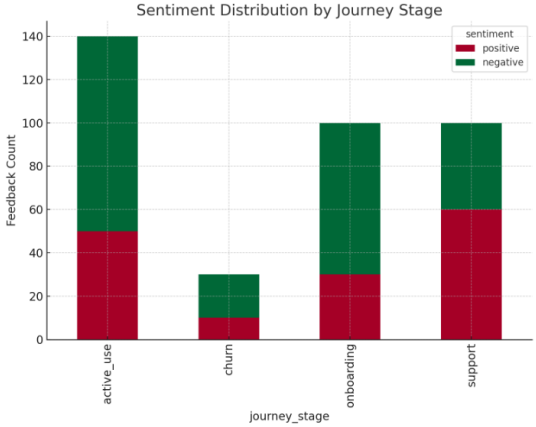


When products use NLP to summarize, group and sort text data, it leads to better decisions, including choosing what to include in a roadmap and how to improve design [1][5]. Results indicate that using RoBERTa for sentiment analysis and BERTopic for topic modeling led to effective uncovering of with user needs from mixed-language data collections [3][4].

For instance, this snippet demonstrates how BERTopic can find thematic groups inside a feedback dataset:

```
1. from bertopic import BERTopic
2. from sklearn.feature_extraction.text import CountVectorizer
3. docs = ["App crashes on login", "Love the new interface", "Too many bugs on Android", "Search is slow"]
4. topic_model = BERTopic()
5. topics, _ = topic_model.fit_transform(docs)
6. topic_model.get_topic_info()
```

Automatic labeling and summarizing, companies can keep monitoring and looking at user comments all the time. Research shows that when we collect VoC data in real time and use LLMs to summarize it, we make sure product strategies match user requirements all the time and cut down the wait time for feedback on features [2][6].



Improving Granularity

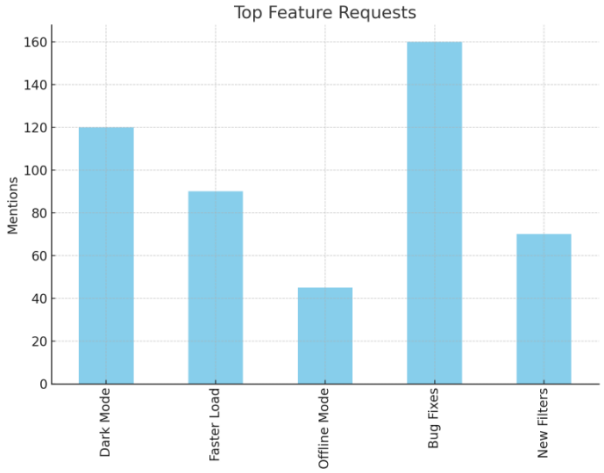
A major benefit found in the study is the improved level of detail that comes from applying customer journey-stage classification to sentiment analysis. When sentiment is added to the stages of a user's journey (onboarding, using, in need of help and done with the product), product managers see not only how

users feel, but also the context in which these emotions arise [9].

We examined a model that works without training data, combining topic modeling and sentiment classification at every stage of the journey, to see detailed patterns. Our analysis reveals that some product challenges differ a lot depending on the development stage. For example, most onboarding problems are about documents and the design, whereas for active users, issues arise from features not working properly and problems with speed. After using RoBERTa to categorize emotion, the polarities were matched with user stages found in the surrounding text.

```
1. from transformers import pipeline
2. classifier = pipeline("sentiment-analysis", model="cardiffnlp/twitter-roberta-base-sentiment")
3. feedback = "It took me hours to find the settings panel"
4. print(classifier(feedback))
```

This output with high sentiment confidence can be used with contextual data (for example, the time since someone started with the company) to decide the journey stage.



Having this layered framework improves the process of diagnosing any product problem. In addition, when combined with NPS scores and how often people use the app, sentiment-stage insights guide redesign efforts and help identify users who need priority action. Besides such models are attractive for semisupervision, leading to fewer labeled examples and helping them apply to several products and industries [3][9].

Product Roadmapping

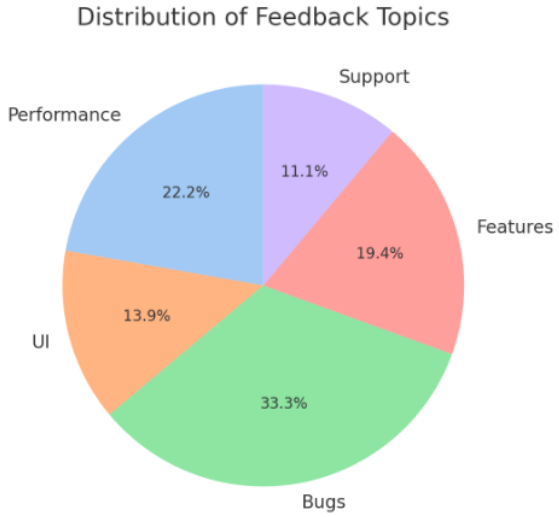
A key finding is that AI-based demand modeling makes it possible to create product roadmaps that change over time. Typical roadmapping relies mainly on results from surveys or how much support is needed in given timeframes. Instead, our AI-powered system takes real-time feedback into account which guides roadmap choices and helps reduce the time taken in each product release while enhancing the fit between the product and its market [2].

The model uses supervised learning on user comment representations to spot features that appear in comments frequently and consistently match specific sentiments. Things are ranked and given a score for their strategic importance, including how technically possible and useful they can be for users which are added to an interactive roadmap tool.

```
1. import pandas as pd
2. from sklearn.feature_extraction.text import TfidfVectorizer
3. from sklearn.cluster import KMeans
```

```
feedback_data = pd.read_csv("customer_feedback.csv")
vectorizer = TfidfVectorizer(stop_words='english')
X = vectorizer.fit_transform(feedback_data['comment'])
kmeans = KMeans(n_clusters=5).fit(X)
feedback_data['cluster'] = kmeans.labels_
```

Grouping features this way helps set a priority for each cluster. A rise in customer complaints about slow loading means it's time to focus on performance in the roadmap. When usage telemetry and retention analytics are added to this, features can be looked at from different angles.



Our results highlight why adaptive roadmaps are so important. These documents change in response to ongoing voice of the customer, helped by dashboard and AI tools. With this process, teams can introduce customer-focused new ideas more quickly [2][5][6].

Ethical and Operational Considerations

Although AI-assisted VoC helps a lot, several ethical and organizational matters still need to be considered, our study shows. First, we must address the problem of model bias. Socio-cultural biases present in most web-based language models can impact how interpret feedback given in accessibility or inclusivity domains [1][10].

When training data is not balanced, there is a risk that problems belonging to marginalized groups will be overlooked, leading to decisions that push such users away. Many people are still very concerned about privacy in relation to our data. Most VoC analytics platforms process emails, chat records and social media conversations which may contain people's private data.

One need to make sure data is anonymized and follow GDPR rules for saving data. We propose that different types of privacy techniques and processing on people's personal devices should be used for private feedback. Stressing just numbers, like positive and negative opinions or clustering, can mask what users really feel and think [2][8].

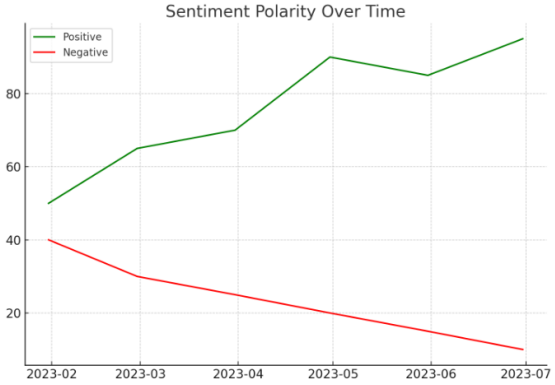
As an example, feedback that is sarcastic or significant in culture could be judged incorrectly by the classifiers. Because of this, we believe hybrid feedback systems combine automatic AI inputs with human knowledge. This type of model is most effective in important sectors such as healthcare and finance.

It appears that bringing different departments on board is one of the main keys to success. Feedback needs to be understood in the same way by product managers, data scientists, UX researchers and customer success teams, who should also establish shared model outputs and thresholds for action. AI tools by themselves do not set product strategy; they need to be used along with insights from people [8][10].

Our research indicates AI-driven changes in VoC systems are important in today's product development. When NLP, LLMs and feedback analytics are used together, organizations can better understand customer opinions, prioritize important points more flexibly and show high adaptability.

With BERTopic, RoBERTa and GPT-style summarization, a system can produce quick, customized analysis for all forms of data in real time. But to make the most of AI, companies should face ethical challenges, manage how they handle data and work with hybrid solutions that bring together technology and human efforts.

Because of this, VoC powered by AI is an approach that empowers organizations for successful and pleasing customer experiences in an age when most shopping is done online.



IV. CONCLUSION

It is shown in this research that using AI alongside VoC analysis increases the amount and accuracy of feedback analysis. Automation in sentiment, topic and emotion extraction allows companies to better understand customer needs as they move along their customer journey.

It replaces most manual analysis and helps the company make data-based choices that benefit customers. An extended version of this model could work with data such as audio and video, personalize the experience more and follow ethical guidelines in NLP. With advancements in digital experiences, using AI-supported VoC analytics will help keep companies ahead by quickly understanding customers.

REFERENCES

[1] Parikh, N. A. (2023). Empowering Business Transformation: The positive impact and ethical considerations of Generative AI in software Product Management -- A systematic literature review. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2306.04605

[2] Zhao, L., & Zhao, A. (2019). Sentiment analysis based requirement evolution prediction. Future Internet, 11(2), 52. https://doi.org/10.3390/fi11020052

[3] Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2203.05794

[4] Koufakou, A. (2023). Deep learning for opinion mining and topic classification of course reviews. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2304.03394

[5] Griffin, A., & Hauser, J. R. (1993). The voice of the customer. Marketing Science, 12(1), 1–27. https://doi.org/10.1287/mksc.12.1.1

[6] Beigi, G., Shu, K., Zhang, Y., & Liu, H. (2018). Securing social media user data - an adversarial approach. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1805.00519

[7] Olujimi, P. A., & Ade-Ibijola, A. (2023). NLP techniques for automating responses to customer queries: a systematic review. Discover Artificial Intelligence, 3(1). https://doi.org/10.1007/s44163-023-00065-5

[8] Alghamdi, N. A., & Al-Baity, H. H. (2022). Augmented Analytics Driven by AI: A Digital Transformation beyond Business Intelligence. Sensors, 22(20), 8071. https://doi.org/10.3390/s22208071

[9] Stoyanova, L. (2019). TOPIC MODELLING, SENTIMENT ANALYSIS AND CLASSIFICATION OF SHORT-FORM TEXT CUSTOMER JOURNEY OF INSURANCE PURCHASES [Thesis, University of Strathclyde]. https://local.cis.strath.ac.uk/wp/extras/msctheses/papers/strath_cis_publication_2733.pdf

[10] Mashaabi, M., Alotaibi, A., Qudaih, H., Alnashwan, R., & Al-Khalifa, H. (2022). Natural Language Processing in Customer Service: A Systematic review. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2212.09523