

Generating Brand Insights from Millions of Social Media Posts with AI

QUICK FACTS

Situation: A pharmaceutical company needed to extract insights from noisy social media data across multiple languages and regions to inform their marketing strategies. The data informed decisions of various internal teams and was updated regularly. They proposed using NLP for analyzing this large, unstructured data as part of a digital intelligence initiative.

Solution: IQVIA developed an AI-driven data feed to analyze social media data, providing key insights for brand teams. The system used a blend of methods, including BERT for persona identification, and was designed for flexibility to accommodate future AI advancements.

Success: IQVIA's work significantly impacted the program, providing valuable insights for brand teams' marketing and engagement strategies. Their use of AI tools enabled accurate post tagging and categorization, facilitating efficient identification of key topics. This led to improved stakeholder engagement, operational efficiency, and cost reductions.

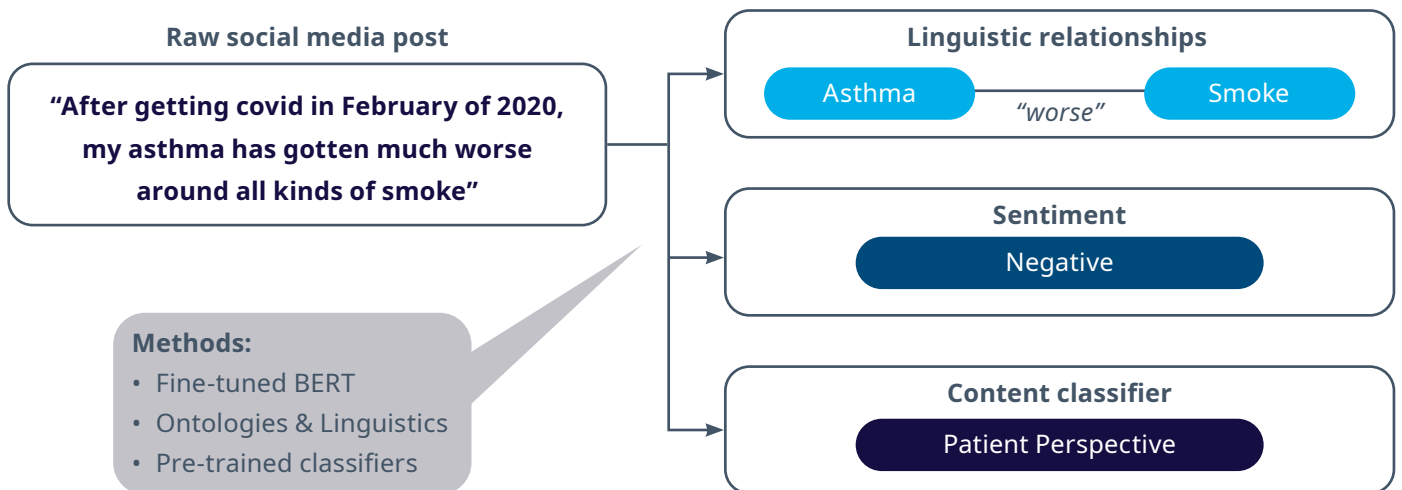
Situation

A pharmaceutical company's global, regional, and country brand teams needed to gather insights on multiple products across key therapy areas due to the high volume of noisy and irrelevant posts on social media and other sources. They needed to understand the key needs, concerns, and experiences of patients and other stakeholders to inform their marketing and stakeholder engagement strategies. The data extraction was complex because there were millions of social conversations across multiple languages and geographies covering several therapy areas. As the insights from social data informed the routine decisions of various internal stakeholders, such as brand marketing teams, medical affairs, patient insights, competitive intelligence, and others, they were needed on an ongoing basis with a process for regular refreshes. The proposed NLP-driven analysis of large unstructured social and online information was a part of the broader digital intelligence initiative.

Solution

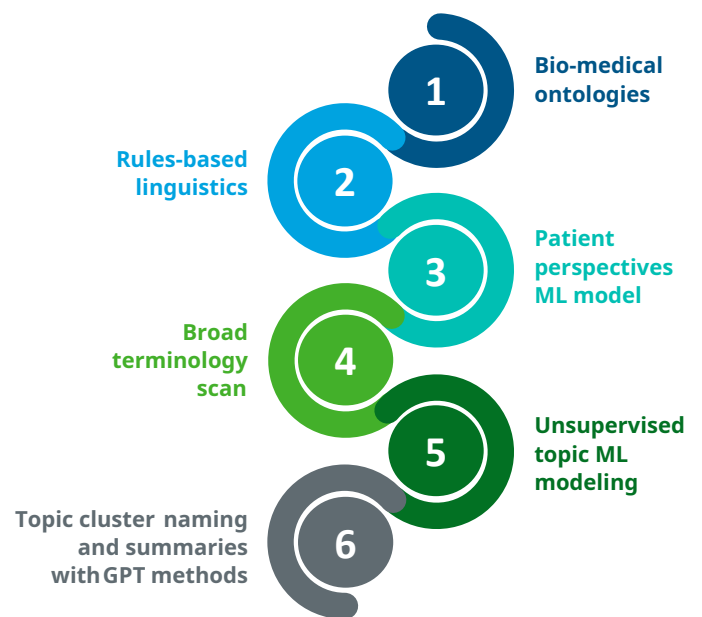
To address these challenges, IQVIA developed an enhanced data feed using a combination of social listening tools, machine translation, and AI including natural language processing and machine learning. The resulting enhanced data feed was ingested into a data warehouse and fed into interactive visuals that surfaced trends and highlighted key topics. To ensure the accuracy and relevance of the insights, IQVIA's algorithms utilized comprehensive biomedical ontologies, linguistics, pre-trained models, and supervised machine learning to create powerful tags and remove noise. This approach allowed the brand teams to gain a better understanding of patient and stakeholder needs, concerns, and experiences across multiple therapy areas. Thanks to the noise reduction and enhanced tagging, the teams could quickly identify core topics and adjacent topics for targeted marketing and communication activities. Given the sensitivity of working with data generated by the public, we also put in place a process to manage the data storage to ensure we were within the policies of the underlying data providers.

Figure 1: Example post with classifications



At the core of the data pipeline developed were the models themselves, and we used an approach we call a 'blend of methods' which allowed our data scientists to select the most appropriate method to achieve a specific result. For example, we selected low-cost, highly scalable ML models for unsupervised topic analysis to uncover novel or edge case topics from within the data. To give users a view of first person experiences we used BERT, a large language model that can be trained to score each post against the likelihood the content is written by a specific persona such as a physician, patient, caregiver or other. Using our vast, multi-million strong terminology databases we were able to tag all posts very quickly and efficiently with biomedical terms such as drug names, diseases, and symptoms. In a fast-moving AI environment, it's important to have flexible pipelines that allow new techniques to be deployed in a modular fashion which balances the need for results today, without losing out on the potential of tomorrow's technology.

Figure 2: Blend of AI and NLP methods



The insights generated allowed global and regional teams to improve stakeholder engagement, collaboration across markets and efficiency in marketing operations, all powered by a truly data-driven approach.

Success

The work done by IQVIA had a significant impact on the program providing brand teams with valuable insights that helped inform their marketing, strategy and planning and stakeholder engagement strategies. By leveraging social listening tools, machine translation, and natural language processing, IQVIA was able to create a multi-country, multi-language solution that met the needs of the broad user group. The algorithms used by IQVIA were able to accurately tag and categorize posts, removing irrelevant and noisy content. This allowed the brand teams to quickly identify core topics and adjacent topics for targeted marketing and communication activities.

The work also allowed the brand teams to move from insight to action more efficiently and effectively, resulting in improved stakeholder engagement, collaboration across markets, efficiency in marketing operations and a data driven approach being used widely by different teams.

Overall, IQVIA's contribution to this initiative led to significant reduction in tool subscription and data processing costs for the client. More importantly, IQVIA's deliverables helped democratize the social listening process across various stakeholders across the client's organization.

