How a biopharmaceutical company evaluated social media to monitor off-label drug use

SemanticPro Classify & Automate Case Study

The Challenge

The prescribing of medications in a manner not specified by the FDA – known as off-label use – is widespread: according to a 2014 US study, it accounts for 10 to 20 percent of all prescriptions written. However, its exact extent is difficult to measure because pharmaceutical companies have no direct feedback mechanism to track what medical conditions their products are prescribed for. However, being aware of how the medications they produce are used is of high importance for regulatory purposes as well as general pharmacovigilance activities. Since pharmaceutical companies do not sell directly to consumers, they have no means to monitor how their medications are used. Patient records are subject to privacy restrictions and when available they do not allow the company to reconstruct which medication was prescribed for which condition.

Our customer had the idea to listen to social media, in particular Reddit, as as a source of knowledge about medication usage, and as a means to take the voice of patients into account. The company turned to Cortical.io for their expertise in developing Natural Language Understanding (NLU) based-solutions to develop a tool able to overcome the hurdle of ambiguity and vague wording inherent to social media posts and to correctly interpret them despite the limited number of posts available for training.

Company Profile

Global biopharmaceutical company

The Goal

Gather intelligence about off-label drug usage by screening patient comments in social media

The Cortical.io Solution

During the initial phase of the project, Cortical.io created a prototype based on SemanticPro Classify & Automate to identify mentions of on- and off-label medication usage in a static set of Reddit posts. Leveraging Cortical.io meaning-based algorithms, the application automatically and accurately filters and classifies Reddit messages and summarizes the results. Cortical.io trained classifiers for each of the example drugs specified by the company using publicly available information.

The Cortical.io Impact

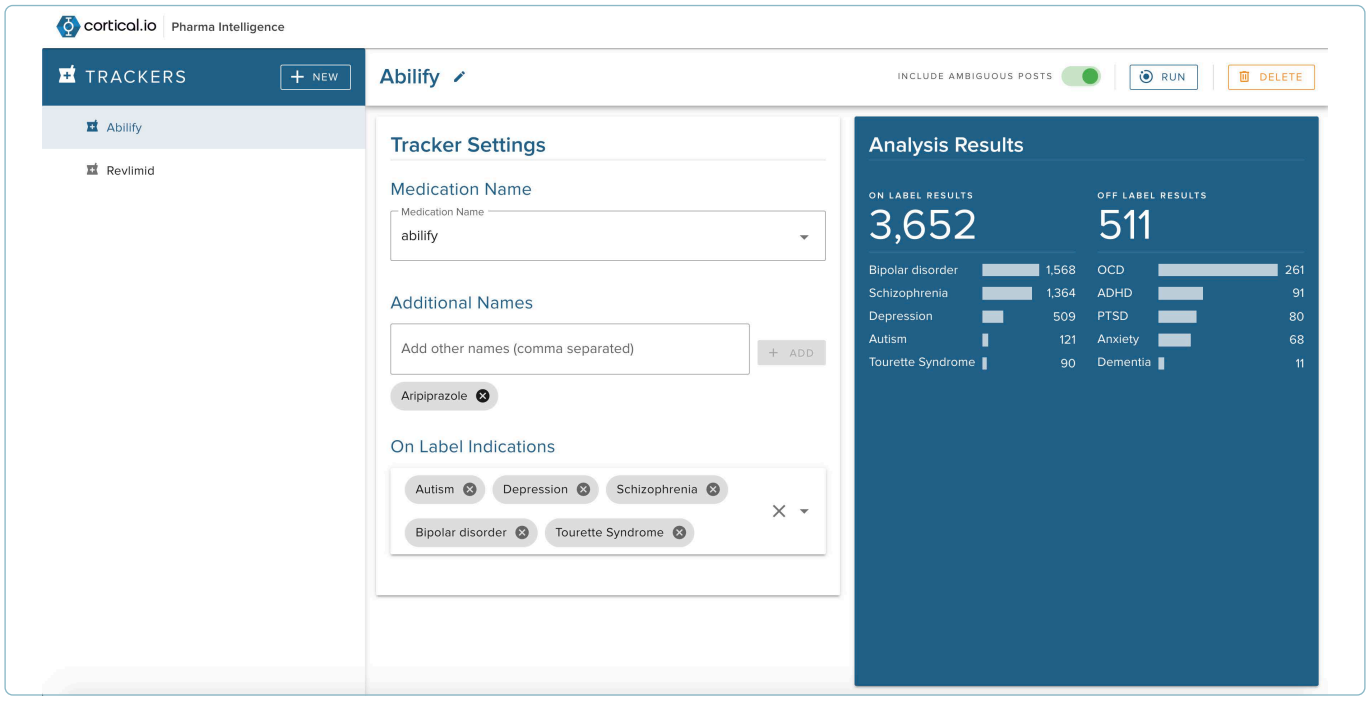
The Cortical.io application achieved a very high level of accuracy based on standard NLU metrics (92-100%) in classifying on-label vs. off-label usage. Both the quality and interpretability of results convinced the company to pilot the prototype with additional drugs and live Reddit streaming and, depending on results, move to production in the course of 2021.

SemanticPro Classify & Automate Study

By using SemanticPro Classify & Automate, this biopharmaceutical company was able to automatically:

- **Review** over 2.2 million Reddit posts mentioning medications
- **Filter for trade or generic names** of medication
- **Filter out off-topic posts** based on the semantic meaning of the post
- **Filter out ambiguous posts** (e.g. mentioning multiple medications)
- **Classify posts by medical conditions** based on semantic meaning
- **Identify on- vs off-label usage**

A dedicated UI provides a summary of analysis results for each drug, with the possibility to inspect on-label and off-label results.



For more information about this case study and Cortical.io solutions, visit
www.cortical.io or email **info@cortical.io**

cortical.io