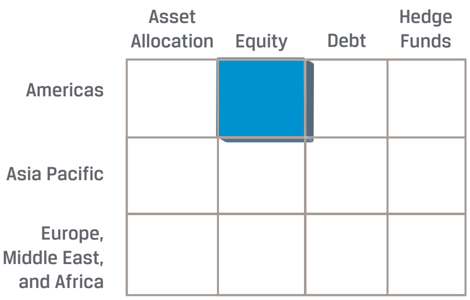
8. SENTIMENT ANALYSIS: BLOOMBERG

Contributor: Gary Kazantsev[31]



Background

Founded in 1981, Bloomberg provides financial software tools, such as securities trading platforms, data services, news, and analytics to financial companies and organizations through the Bloomberg Terminal and its Enterprise products and data feeds. The company started investing in ML and AI applications over a decade ago, ranging from improving the customer support experience to improving the Bloomberg Terminal user experience through natural language understanding and question answering, as well as information extraction and innovative news and analytics applications.

One of the earliest such applications was the sentiment analysis product, which has been available to customers since 2009. Since then, the product has grown and expanded from English to other languages (e.g., Japanese), from equities to commodities and currencies markets, and from news to other content types, such as social media (e.g., Twitter). The product can now also tackle other issues—named entity recognition and disambiguation, topic clustering, theme detection, market impact analysis, and others. While the application of machine learning in this context was novel back in 2009, it is now used by key players across the global capital markets for a variety of applications, from risk analysis and portfolio construction to alpha generation.

Today, this product is just one of a spectrum of advanced analytical tools powered by machine learning that Bloomberg offers.

Methodology

A variety of approaches to sentiment analysis have been explored in academic literature and in the industry. When we started development on this product in 2009, Bloomberg chose to approach it using cutting-edge technology. To this end, we used supervised machine learning and developed novel ensemble construction methods.

The key issue in text analysis problems is the selection of the target variable—that is, precisely formulating the question we wish the machine to answer. Many possible perspectives on this problem exist. For example, some seek to assess the internal state, or opinion, of the author of the document. There are applications where that perspective is of interest. However, in finance, this introduces additional ambiguities and complications to an already difficult problem. In contrast, we chose to predict the opinion of the reader—that is, a consumer of the news story who is a participant in the market. The question we chose to ask is, roughly, If you are a long-position investor in the underlying company and you read the story in absence of other information, would you take it as a positive, negative, or neutral event? This change in perspective, and thereby the target variable for the ML algorithm, makes a big difference in practice.

Training data for these models were collected in-house, and we trained annotators who had an appropriate background in finance. Special care had to be taken with respect to sampling to ensure that a sufficient diversity of sources, document lengths, topics, market caps, industries, financial periods, and so forth, were represented in the data. The key measure to inform model development is the inter-annotator agreement—that is, How often do human annotators agree on the answer to the posed question? If the agreement rate is low (i.e., no different from chance), the problem is ill posed. If the agreement rate is high, then there is a high likelihood we can build a machine to reliably reproduce this aspect of human judgment.

Using this carefully chosen question, we are able to obtain agreement rates between 80% and 92%—for news. For social media like Twitter, agreement rates for every type of language understanding task are generally lower. This number is naturally an intrinsic boundary to the accuracy of any classifier we build. It is important to note that simple methods, such as lexicons (dictionaries), are unable to solve this complex and ambiguous problem with the required degree of accuracy. We end up building much more complex models—nonlinear support vector machines, models incorporating complex linguistic features, and, most recently, recurrent neural networks—and integrating them into ensembles using data-driven methods. The resulting ML ensembles are real-time systems, processing a stream of more than 2 million

31 Head of Quant Technology Strategy, Bloomberg. This case is prepared based on our interview with him and materials submitted by Bloomberg.

documents per day—in just tens of milliseconds per document. The per story and entity sentiment data (e.g., for one document mentioning both Google and Microsoft, you may see positive sentiment for Google with 75% probability and negative for Microsoft with 50% probability) can then be aggregated into a confidence-weighted per-security time series, or index, to be used in portfolio construction, strategy development, or risk analysis.

Sentiment analysis data have been used by a wide variety of clients, both on the Bloomberg Terminal and in context of our Enterprise Event Driven Feed (EDF) product. On the Bloomberg Terminal, there are functions that allow clients to visualize the data, correlate them with other variables, set alerts on them, or use them to monitor a portfolio. The sentiment data are also used for automated story generation. In the Enterprise space, sentiment data have been used as a factor in portfolio construction and optimization, risk analysis, trading strategy development at various periodicities and for various holding periods, execution, and statistical arbitrage. Over the years, the composition of the consumers of sentiment data has changed. Initially, quant hedge funds and market makers were the primary users of this "alternative" data, whereas today, most institutional clients either use it or are looking to incorporate it into their workflows.

AI/Big Data Technology

Named entity recognition and disambiguation (NER/NED), topic classification, sentiment analysis, and clustering are all examples of the application of modern machine learning in the NLP domain. These tools can be used much more broadly across the industry to understand unstructured data, be that instant messaging chats, emails, analyst reports, quarterly filings and annual reports, compliance notes, or other trade-related information. For example, publicly available APIs from a variety of cloud vendors now make it possible to effectively process the audio stream of an earnings call into a text transcript. Once the data are in textual form, these same tools can be applied to build integrated market impact models that update as the call progresses.

Team Structure and Development Process

The overall system that supports and enables these advanced analytics is a large and complex technological solution that enables clients to perform many other workflows—search, alerts, research, and so on. Hundreds of people were involved in its development over the span of more than 10 years. At this point, the AI and Data Science Engineering organization at Bloomberg numbers over 200 people with advanced backgrounds in software engineering and high-performance distributed systems, mathematics, computational linguistics, machine learning, and natural language processing.

The development of these products is a very client-driven activity, and we work closely with customers to design and build solutions to these problems. For example, a small engineering team consisting of just two people started to build sentiment analysis in response to a client request. Having successfully demonstrated an initial prototype in the first six months of development, the team grew and expanded to deliver the product to clients and extend the technology to other aspects of the problem.

A decade ago, machine learning was exotic. The state of tools and libraries for computational linguistics, natural language understanding, and machine learning was much more basic. As a result, the majority of the development had to be undertaken in-house. The development process emphasized the need for a distributed computation environment and natural language understanding tools tuned to the finance domain, which spurred a major, multi-year investment in these areas. Given these investments in infrastructure and data processes (i.e., distributed computing, GPU hardware), the team can now execute these and similar projects much more rapidly, drastically minimizing our time to market with new solutions.

Key Takeaways

There were several key takeaways from the sentiment analysis product's development process and the broader text analysis initiative that followed. As previously alluded to, end-to-end evaluation is important. Careful selection of the target variable is critical. It is important to start with simple baselines and perform error analysis at every iteration of the model development; otherwise, we can easily find that a particular date or quarter is positive or negative by itself—because of inherent bias in the underlying data, for example. For this reason and others, interpretable models and human-in-the-loop are very important to us.