

FINANCIAL INNOVATION SERIES



FINANCIAL INNOVATION SERIES

Sentiment Analysis and Natural Language Processing in Finance: Applications, Implementations, Challenges

Author: Alexey Rubtsov
PhD, Senior Research Associate, Global Risk Institute

Recent advances in machine learning, increases in computational power, and reductions in data storage costs have revolutionized many computational techniques used by financial market players including Natural Language Processing (NLP) and Sentiment Analysis (SA), and we wish to keep members up to date on developments.

NLP is concerned with processing and analyzing human language data by computers and SA is a subfield of NLP that studies the mood, opinions, and attitudes expressed in written text (e.g., whether the author is positive or negative about a certain subject). According to a 2019 market study by

Table 1: Top 20 NLP solution providers

Top 20 NLP Solution Providers	Domiciled
3M	USA
Apple	USA
Automated Insights	USA
AWS	USA
Baidu	China
Bitext	USA
Conversica	USA
Dolbey	USA
Facebook	USA
Google	USA
Health Fidelity	USA
IBM	USA
Inbenta Technologies	USA
Intel	USA
Linguamatics	UK
Microsoft	USA
Narrative Science	USA
SAS Institute	USA
SparkCognition	USA
Veritone	USA

Figure 1: Expected Growth of NLP market



Research and Markets, the NLP market size is estimated to be USD10.2 billion and is expected to reach USD26.4 billion by 2024 with a Compound Annual Growth Rate (CAGR) of 21% (see Figure 1, NLP Report (2019)). This makes NLP the fastest growing technology in global AI for the financial asset management market. The major growth factors are: increase in smart device usage, growth of cloud-based solutions, and NLP-based applications to improve customer service.

Of the leading NLP solution providers, we note first that the Tech Giants (the so-called GAFA quartet: Google, Apple, Facebook, and Amazon) are among the leading participants and second, only 2 non-US companies are represented: Baidu (China) and Linguamatics (UK) (see Table 1, NLP Report (2019)).

The combination of NLP and Machine Learning (ML) provides an avenue to generate insights with a level of precision significantly higher than ever before. For instance, in 2013 a team of researchers at Google created an ML-based model Word2Vec that maps words to vectors in such a way that semantically close words have close vector representations. This, in turn, allows “arithmetic on words” which, loosely speaking, can be thought of as, for example,

Brother – Sister ≈ Man - Woman.

This example follows from the pattern “Brother is to Sister as Man is to Woman”.

The NLP/ML market is highly competitive and, for example, in December 2019 China’s Baidu announced that its ML-based software ERNIE achieved the highest score of 90.1 in the General Language Understanding Evaluation (moving Microsoft and Google to the 2nd and 3rd places, respectively).¹ In January 2020 the ranking changed again with Google’s T5 Team achieving the highest score of 90.3.

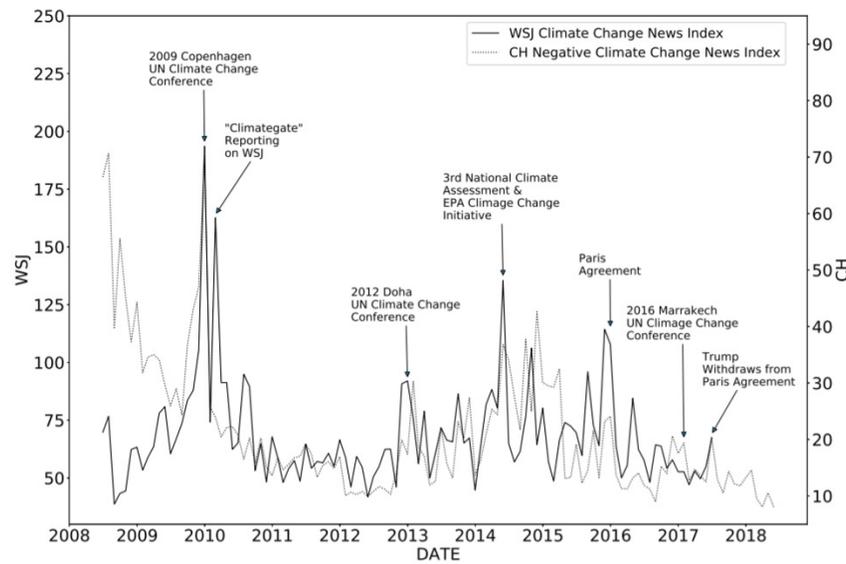
APPLICATIONS

NLP is being applied by financial institutions across a number of different areas of interest to our members.

Climate change. In GRI-funded research, "[Hedging Climate Change News](#)" Engle et al. (2019) construct a measure of climate change risk using sentiment analysis of over one trillion news articles from more than 1,000 outlets including The Wall Street Journal, The New York Times, BBC, CNN, etc. The constructed measure is used to construct portfolios with payoffs that compensate investors for losses that arise from the realization of climate risk, thereby providing a “hedge” against climate risk.

The researchers use NPL to construct two indices: one that measures negative sentiment of the climate news and the other one that simply reflects the intensity of climate-related news, both positive and negative (see Figure 2).

Figure 2: The CH Negative Climate Change News Index from 2008 to 2017, overlaid against the WSJ Climate Change News Index, and annotated with climate-relevant news announcements (source: Engle et al. (2019))



¹ General Language Understanding Evaluation (GLUE) is a collection of datasets used to train, evaluate, and analyse NLP models relative to one another.

Asset Management. Using news articles about S&P500 companies published between 2000-2018 by Reuters News, Daetz et al. (2019) study the effect of articles' sentiment on stock market performance of corresponding companies. The researchers found that news sentiment does predict the short-term performance of stocks. The authors also reported that the relationship has become weaker over time, and a trading strategy based solely on news sentiment would generate much smaller returns than it would have generated fifteen years ago.

Geopolitics. By applying an automated keyword search, Caldara and Iacoviello (2018) constructed a monthly Geopolitics Risk Index based on 11 leading English-language newspapers since 1985. They found that high values of the index lead to a decline in real activity, lower stock returns, and movements in capital flows away from emerging economies towards advanced economies.

Economic policy. Baker et al. (2016) develop an index of Economic Policy Uncertainty that is based on the keyword search in the articles published in 10 leading US newspapers. They also develop category-specific policy uncertainty indices for the US such as healthcare policy uncertainty and national security policy uncertainty. It is found that spikes in the indices foreshadow declines in investment, output, and employment.

Monetary policy. Do the monetary policy-related tweets of the US President Donald Trump influence market expectations about monetary policy? Using fed funds futures data, Bianchi et al. (2019) found strong evidence that tweets that advocate looser monetary policy result in the market expectations of a lower target rate. In their analysis the researchers consider tweets that unequivocally advocate looser monetary policy.

Cyber risk. NLP can be used in highly advanced phishing attacks (see Baki et al. (2017), Shropshire (2018)).² An NLP algorithm can emulate a specific person's

² Phishing is a cyber attack used to steal user's data (e.g., login credentials, credit card numbers). It occurs when an attacker, masquerading as a trusted entity, dupes a victim into opening an email, instant message, or text message.

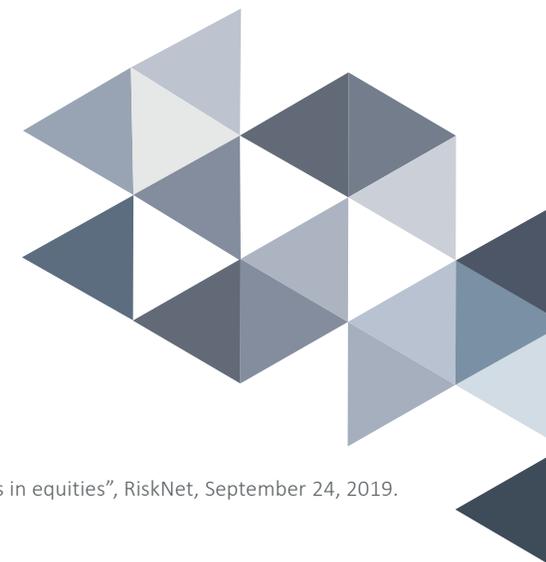
communication patterns and style making it difficult for the other party to determine whether they are communicating with someone they know or with an algorithm masquerading as that individual. An unsuspecting person could assume they are communicating with a spouse, family member, friend, or co-worker and end up sharing sensitive information. On the other hand, NLP can also be used in recognizing such sophisticated phishing attacks.

IMPLEMENTATION AND CHALLENGES FOR FINANCIAL INSTITUTIONS

Financial institutions have been taking advantage of the opportunities offered by NLP technologies either by building in-house teams or by employing external resources. For instance, Allianz Global Investors is internally developing NLP models with signals generated by analyzing

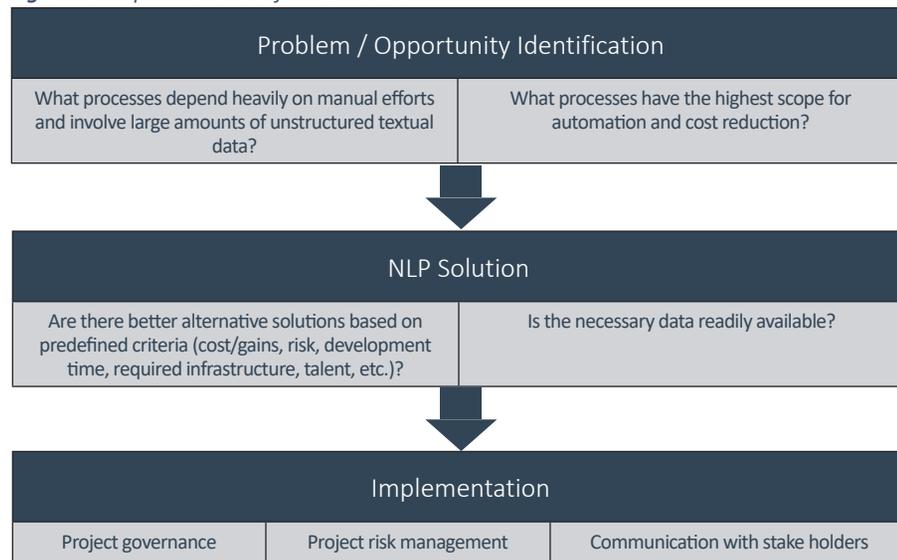
- *sell-side analyst reports;*
- *earning calls*³.

In general, to successfully deploy NLP solutions and capitalize on the business impact, financial institutions may consider the steps shown in Figure 3.



³ See "Allianz Global Investors adopts NLP signals in equities", RiskNet, September 24, 2019.

Figure 3: Steps to consider for NLP solutions



Financial institutions should also be aware of the challenges underlying NLP solutions (see Figure 4). In this respect, the data science expertise and finance domain experience of development team are essential for successful applications.

First, it is important to determine what NLP approach is appropriate to use for a specific application. Indeed, there are applications where a simple keyword search is sufficient, whereas for other applications more sophisticated techniques (e.g., SA) may be warranted. For example, when constructing their climate risk measure, Engle et al. (2019) show that the results of the keyword search approach could be quite different from the SA technique (see also Figure 2). In particular, the researchers point out that the keyword search model implicitly assumes that “...when it comes to climate change, no news is good news”, which is clearly a highly questionable assumption.

At its current stage of development, SA can quite frequently produce inaccurate results which is especially true for texts written in languages other than English.⁴

⁴ For challenges of NLP applications to languages other than English see, for instance, “PanAgora uses NLP to cut through Chinese cyber speak”, RiskNet, July 18, 2019

Figure 4: Major challenges in the use of NLP



In this respect, it is important to realize that while SA can be used to provide insights, its results have to be validated with other evidence. For example, to construct a measure of Economic Policy Uncertainty Baker et al. (2016) conducted an extensive audit study of 12,000 randomly selected articles drawn from major US newspapers. According to such an audit, the authors confirmed that there is a very high correlation between human- and computer-generated results.

Many SA approaches start with building dictionaries of “negative” and “positive” words and then use positive and negative word counts to construct a sentiment measure. In this respect, one should be aware of the fact that the meaning of words could change depending on the context. For instance, according to a widely used financial dictionary developed by Loughran and McDonald (2011), the words “antitrust” and “concede” have negative meaning, whereas they could have neutral or even positive meaning in non-financial contexts. Thus, teams working on NLP applications for financial institutions should be very careful when developing lexicons for their SA algorithms.

To avoid biased results, developers of NLP applications should also be careful with the choice of data sources used to feed into an NLP algorithm. There are several ways in which data source concerns can be addressed. Baker et al. (2016), for example, find similar movements in their Economic Policy Uncertainty index based on right-leaning and left-leaning newspapers, suggesting that political slant does not seriously distort their measure.

NLP solutions could also pose data security issues. This is especially true when services are provided by a third party and the data necessary for NLP algorithms is stored and processed outside of the organization. Thus, an organization may consider establishing internal teams that develop and implement NLP algorithms.

Last but not least, NLP algorithms should have built-in mechanisms for “fake news” detection that would remove false information from the data fed into the algorithms. In simple terms, once it is realized that NLP is used in decision-making, it might become tempting to manipulate the process by creating disinformation. The importance of such a mechanism is recognized by the Leaders Prize competition that will award \$1,000,000 to the Canadian team that most effectively automates the fact-checking process using artificial intelligence (winners to be announced in June 2020).⁵

© 2020 Global Risk Institute in Financial Services (GRI). This “Sentiment Analysis and Natural Language Processing in Finance” is a publication of GRI. This “Sentiment Analysis and Natural Language Processing in Finance” is available at www.globalriskinstitute.org. Permission is hereby granted to reprint the “Sentiment Analysis and Natural Language Processing in Finance” on the following conditions: the content is not altered or edited in any way and proper attribution of the author and GRI is displayed in any reproduction. All other rights reserved.

⁵ <https://leadersprize.truenorthwaterloo.com/en/>

REFERENCES

1. Baki, S., Verma, R., Mukherjee, A., Gnawali, O.: Scaling and Effectiveness of Email Masquerade Attacks: Exploiting Natural Language Generation. ACM on Asia Conference on Computer and Communications Security, Abu Dhabi, UAE (2017)
2. Bianchini, F., King, H., Kind, T.: Threats to Central Bank Independence: High-Frequency Identification with Twitter. NBER Working Paper Series (2019)
3. Caldara, D., Iacoviello, M.: Measuring Geopolitical Risk. Board of Governors of the Federal Reserve System. Discussion paper (2018)
4. Daetz, S., Hvid, A., Martinello, A., Matin, R.: Seeing through the spin: the effect of news sentiment on firms’ stock market performance. National Bank of Denmark. Discussion paper (2019)
5. Engle, R., Giglio, S., Lee, H., Kelly, B., Stroebel, J.: Hedging Climate Change News. Working paper (2019)
6. Global AI report: Global AI in Financial Asset Management Market – Drivers, Restraints, Opportunities, Trends, and Forecast up to 2025. Research and Markets (2019)
7. Loughran, T., McDonald, B.: When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *Journal of Finance*, 66(1), 35-65 (2011)
8. NLP report: Natural Language Processing Market by Component, Deployment Mode, Organization Size, Type, Application (Sentiment Analysis and Text Classification), Vertical (Healthcare and Life Sciences, and BFSI), and Region - Global Forecast to 2024. Report by Markets and Markets (2019)
9. Shropshire, J.: Natural Language Processing as a Weapon. Proceedings of the 13th Pre-ICIS Workshop on Information Security and Privacy, San Francisco (2018)