



Natural Language Processing (NLP) in Fraud Analytics

Jayeeta Putatunda
Senior Data Scientist

June 10, 2020



Natural Language Processing with the help of Machine Learning is the current win-win combination used to detect fraud and misinterpreted information. One of the biggest challenges of the free and anonymous internet that we have constant access to and basically drives our life is “Fraud”.

Fraud has many forms – ranging from fake news spread on social media and doctored images/videos to manipulate hateful ideas. Even instilling fear with numerous hoax calls about SSN blockades; and how you can be deported if you do not pay the fine immediately over the call! Eh, really?

A long-standing open problem of fraud has always existed in major industries like banking, medical, and insurance. Unfortunately, it is becoming increasingly common. Not to mention the magnitude and novelty of fraud attacks. With the surge in online transactions all over the world, systems are more vulnerable than ever.

But it's not all bad news! An exponential increase in computing powers and advancement in statistical modeling has ensured that we are a step ahead from the attackers. New design preventive measures have been made to counter fraud in real-time. [Machine Learning](#) has become the go-to strategy for creating and updating various supervised/unsupervised fraud detection algorithms. This has been a step up from traditional rule-based approaches which were more time-effort consuming and led to higher rates of false positives.

The [data science team](#) at [Indellient](#) has applied many different analytical approaches in fraud detection applications. In this blog, we will see how the industry is starting to leverage Natural Language Processing along with [ML algorithms](#) to counter fraud in various use-cases:

What is Natural Language Processing?

Natural Language Processing (NLP) is a field of Artificial Intelligence that combines computational linguistics with statistics models. This gives the machines the ability to automatically read, understand and derive meaning from human languages. The

development of underlying computational hardware like [Tensor Processing Unit \(TPU\)](#) has led to an enormous leap in the area of research resulting in some state of the art language models.

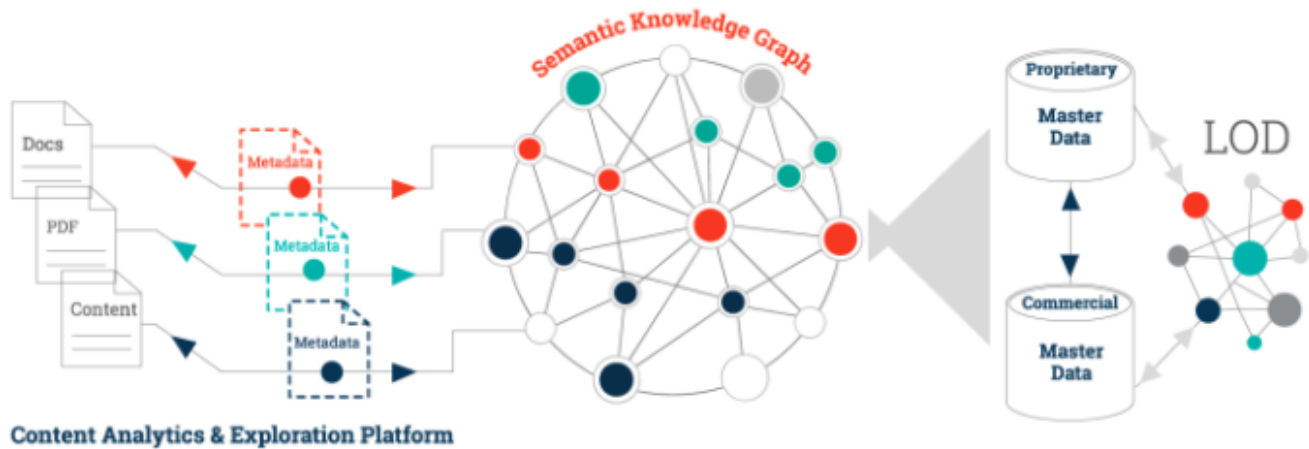
Natural Language Processing for the Insurance Industry

For the insurance sector, identification of fraudulent claims are one of the key factors of success. Insurance firms, working with brokers, agents and investigators have also deployed various IoT technologies and have vast records of social media, geographical, and user-sentiment data. It's not feasible for human agents to review all claims manually with so much background information verification. This is where natural language processing can step up the scale as well response time through Text Mining and Sentiment Analysis.

Natural Language Processing with Text Mining

Text mining is the way to process patterns and assertions from a huge block of textual big data. Analyzing insurance applications derive insights from similar claims by creating a knowledge database. This also helps insurance investigators detect fraudulent cases like common keywords or descriptions of an accident across different geo-locations by multiple claimants, which is possibly a red alert of organized fraud. Processing genuine claims faster also lead to better customer experience.





Source: <https://medium.com/@ontotext/it-takes-two-to-tango-knowledge-graphs-and-text-analysis-66c0e97a9641>

Natural Language Processing with Sentiment Analysis

Sentiment Analysis – This is a method of identifying and categorizing opinions present in a body of text to determine whether attitude towards a particular context is positive, negative, or neutral. It is crucial for companies to keep a sense of the market and their customers (both current and potential) to stay a step ahead in planning new policies and outreach efforts. Knowing what the industry is talking about would also help in identifying any potential scam issues and warn your customers in advance.

Source: <https://www.knime.com/blog/sentiment-analysis>

Natural Language Processing for the Banking Industry

Similarly in the banking industry, the use-cases of NLP are implemented at scale. [Emerj](#), an artificial intelligence market research firm stated that NLP-based products make up 28.1% of the total AI Approaches across various product offerings. The biggest share of these NLP products is for Information Retrieval or document search based products.



Source: <https://emerj.com/ai-sector-overviews/natural-language-processing-banking-current-applications/>

Search capabilities to quickly find crucial highlights amongst stacks of digital documents can be crucial for external compliance tests. This is also important for restricting fraudulent behavior against customers by wealth managers and financial advisors.

For example, [Long-Term Capital Management](#), the most famous hedge fund, collapsed with a bailout of \$4billion dollars from the government and thousands of penniless customers. It was founded by [John Meriwether](#) (of Salomon Brothers fame) and its principal players included two Nobel Memorial Prize-winning economists.

General rules to apply NLP in Fraud Detection

While NLP technology is still under continuous research to solve the challenges of ambiguity, co-reference, synonymity as well as syntactic language-based rules to ensure large scale accuracy; we do have a lot of available NLP technology to deploy in various fraud analytics use-cases and derive value from terabytes of unstructured text.

Here's a quick visual summary of the key points we discussed on how NLP can assist in fraud analytics:

Refer to this nice descriptive paper if you want to dig deeper on the NLP applications

—

<https://pdfs.semanticscholar.org/343c/2a0499cb204f3de6466d908a92ec0fc24ca9.pdf>

Collect and Analyze your Data with Indellient

Indellient is a Software Development Company that specializes in [Data Analytics](#), [Cloud Development Application](#), [DevOps Services](#), and [Business Process](#)

Management.

How to Build a Fraud Detection System using Machine Learning Models

How to Use Machine Learning to Detect Fraud

Analytics Techniques Data Science fraud detection machine learning

About The Author



Jayeeta Putatunda

Hi, I am Jayeeta Putatunda, Data Scientist at Indellient. I work mostly on NLP projects where I get to explore a lot of state-of-the-art models and build cool products. I am passionate about exploring new concepts in the data science domain and firmly believe that data is the best story teller. An active speaker in various meetups and conferences and also hosted workshops on various Machine Learning topics. I am also engaged with some amazing organizations to promote and inspire more women to take up STEM. I have received my Master of Science in Quantitative Methods and Modeling from City University of New York, NY and my Bachelor of Science in Economics and Statistics from West Bengal State University, India.



Liked this post? Share It:

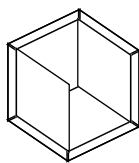


2630 Bristol Cir #100 Oakville, On L6H 6Z7

+1-855-544-9611

Get in touch

Learn More



Indellient is a Cloud, DevOps, and Data consulting company. We partner with our clients to provide end-to-end solutions on cloud migration, cloud cost management, continuous integration, continuous delivery, data consolidation, data management, advanced analytics and more. As a complete lifecycle partner, Indellient works across all industries and company sizes ranging from startups to large enterprise.

Copyright © 2023 Indellient Inc.

[Policies](#) | [Privacy Policy](#) | [Terms of Use](#)

