The Wisdom of Collective Knowledge: a new approach to model markets.

The Causality Link Team

In April 2017, Bill Peters wrote: “Last month, Ford cited **higher commodity costs** as among the reasons for its **weak Q1 profit** guidance.” (https://www.investors.com/news/ford-reports-first-quarter-earnings/) It would be easy to discard this now, in August 2019, since the Q1’17 profit guidance is in the distant past. However, commodity costs impacting Ford may be relevant in the light of new import tariffs rocking global markets. Moreover, realizing that commodity costs also impact Ford’s automotive peers, without it being explicitly written, is another long-lasting benefit of analyzing this single sentence.

Financial news is a key resource for investors to base decisions on. While quarterly results are the measure of the health of a company, the experts’ interpretations of the reported numbers, as written down in financial analyses, create a lot of additional nuance. Humans are very good at collecting information from written statements, however, to do this consistently and at scale is a challenge. In recent years, Natural Language Processing (NLP) has become a hot topic and various commercial offerings are now available that interpret financial news in a more automated and structured fashion. The output from these approaches is mostly centered on key word counts and sentiment scores.

At Causality Link we believe there’s far more information in text than is currently extracted by the current class of NLP solutions. This article discusses which elements we extract from text, how we convert those into actionable statistics and how this enables financial institutions to get more value out of financial news.

Extracting long-lasting knowledge from financial documents

The human capacity to read a text focuses on answering the ‘who/what/when/where and why’ questions. For AI, this is not different. To understand a (financial) text, three elements need to be captured:

1. What is the **topic** being talked about? Note that this is not a keyword search. Picking up the word ‘demand’ is meaningless, unless it’s connected to its proper context. An example of a proper context for demand is: ‘the demand for the Ford F150 in the USA in 2019 Q2’. In the Causality Link platform, a topic falls in either of two categories: events and indicators. In the later section *‘It’s all about context’*, the details of collecting the context for a topic are discussed.
2. What is the **past change** or **predicted future change** of the topic? Outside of events that may have the capacity to shock the system, only indicators that change, i.e. are trending over time, can be causes (their change impacts other indicators) or consequences (the change of other indicators is causing a trend for the topic at hand). More details are discussed in the section *‘Causality needs trends’*,
3. Is the topic at hand mentioned as a **‘cause’ or as a ’consequence’**? Our unique angle is the extraction of human stated causal relationships, which we re-use in an automated reasoning process to deduce future consequences of current trends. In the section *‘Learning from causal statements’*, the details are discussed.

Thus, all financial texts are reduced to their topics (including context), their associated trends and their role in causal statements. Subsequently, the data can be aggregated into summary statistics without losing the ability to drill down from that summary statistic to the original underlying quote that gave rise to the extraction. This is discussed in the section *'how to numerically represent an expert opinion'.* Finally, the extraction process is constrained by a very flexible ontology developed by a combination of machine learning and human experts. Details are discussed in the section '*Ontology'.*

It's all about context

The difference between a fact and an (expert) opinion is more subtle than meets the eye. If a financial analyst with a background in automotive discusses the revenue of Ford, while the numbers may be factual, it's the expert's decision to draw attention to, say, the positive sales trend in China or to the increased margin on Ford's F150. Thus, the choice of context represents the expert opinion as to which facts deserve the most attention.

And, while one may be pleased with Ford's worldwide profit, it takes a much deeper look inside the drivers of Ford profit to understand the health of the reported number. Again, the many experts writing research articles on Ford have done us the favor of doing that deeper analysis by enumerating relevant contexts, if only reading didn't require so much mental effort to organize and remember the details!

Given any financial KPI (e.g., 'demand'), the key is to have an NLP engine recover the complete context from the text. The context for a financial KPI can be one or more of - but not limited to - the following: an industry, a company, a location, a product or a segment but also the source, the language and the author as well as periods referred to can be considered as part of the context. Clearly, for a human reader, or for an NLP engine, the full context does not have to be limited to one sentence. For example, consider the fragment: "Ford did well this quarter, especially in the SUV segment. Their sales in Europe was up by 3%". The KPI is 'sales' and the full context is: Ford SUV sales in Europe in current quarter.

Suppose, using the extracts of the NLP engine, you were able to list all contexts found in financial writings for a specific topic (i.e. 'Ford demand Europe'), this would mimic the capability of having read every article and captured the agreement and disagreement between all relevant experts.

One can argue that all (numerical) data that is discussed in financial news articles starts from structured data or is available as structured data. While that may be true, lacking from the structured data is the expert pointing out which numbers are most relevant. *That* expert word choice selection is what is captured as context for the indicator.

Ontology

An ontology is a set of facts and constraints for the NLP engine, specifying what and how to extract information. Our ontology consists of a collection of tree structures that specify single context extraction aspects. The full indicator as 'Ford profit in the USA' cannot be found as a single item in the ontology. However, in the ontology's company structure, 'Ford' is a company within the automobile manufacturer industry with descendants such as 'Lincoln'. Extraction rules associated with 'Ford' guide the extraction to accept or reject specific contexts, such as industries and products. The KPI 'profit' is situated below 'microeconomic indicators' and contains descendants like 'adjusted EBIT', 'EPS', and 'net income'. The location is extracted using the ontology location tree, starting at city/state/region level, rolling up to country and continent.

Thus, while ‘Ford’, ‘profit’ and ‘USA’ are all single entries in the ontology trees, it’s the NLP engine that can connect all appropriate elements to complete the context. This method is very flexible and avoids an in-advance enumeration of all full contexts (imagine the list of all companies multiplied by all KPIs multiplied by all locations, and so on).

The tree structure of the ontology facilitates generalizations, for example, if the text contains statement of profits for separate countries of the EU, the ontology allows one to generalize this and aggregate statements on the continent level. This type of ontology is also much more aligned with the human understanding of words rather than a single keyword extraction approach. For example, in discussing the various aspects of profit, an analyst may point at the EBITDAR or the Net Interest Margin. While it’s important to realize that those KPIs are not synonyms of profit, it is very helpful to collect and display these trends using profit as a category label. If a more detailed understanding is required, one queries to see the detail KPIs within profit and gets the requested details right away.

The creation of the ontology is facilitated by machine learning and manually curated. Currently, we recognize 20K companies, 1,300+ KPIs and virtually every location and industry.

How to numerically represent expert opinions?

The Causality Link NLP engine ingests financial articles of any kind, in up to 17 languages, and for every concept encountered in text:

- Collects the context
- Determines if the concept is associated with a trend,
  - If yes, determines the direction of the trend (up or down).
- Determines if the concept is mentioned as cause or consequence of a causal statement,
  - If yes, determines the sign of the relationship (positive or negative).

Subsequently, we aggregate both the trends and the relationships into a number representing the percentage of up-trends and the percentage of positive relationships. We call those numbers Positive Trend Percentage (PTP) and Positive Relationship Percentage (PRP), respectively. The PTP and PRP are percentages between 0 and 100% and are replaced with 50% (neutral) when not statistically significant differing from 50% (at a user-determined confidence level). The significance filter ensures only looking at outcomes that are salient enough.

The Positive Trend Percentage (PTP)

Imagine that both Ford’s profit as well as the Benchmark Interest Rate increase with 1%. Clearly, the impact of the 1% increase in Interest Rate will be discussed far more frequently than the 1% associated with Ford profit. Thus, expert opinions introduce a natural scaling for the perceived importance of a change.

Suppose the PTP for Ford’s revenue is 80%. This does not mean that 20% of experts say that Ford’s worldwide revenue is down and 80% reports it is going up. The revenue of Ford is a factual number available as structured data. It’s the full contexts that are mentioned that make up the 80%. For example, one group of experts calls out the positive sales (a financial KPI that in our system rolls up to revenue) of the F150s while other experts echo the lower revenue guidance mentioned by Ford in the quarterly earnings call transcript. As such, **the PTP associated with Ford revenue is the expert vote on the trend**

for all relevant contexts within Ford revenue. Thus, the PTP gives the ability to get a single summary statistic of all relevant topics that mention Ford revenue or more specific aspects of Ford revenue.

Note that the topic for the context is very flexible. For example, for the PTP of ‘Ford demand’, it is very natural to specify the topic to the exact domain of interest: ‘Ford demand for the SUV segment in the USA by this publisher’. That context can even be suggested by our analytic engine as it may receive the most attention (frequency-wise) or the most controversy by experts (PTP-wise).

The PTP also allows seamless generalization. A topic could specify: ‘profit for automobile manufacturers in the USA’ and the PTP represents the automatic summation of the uptrends regarding all profit statements of USA automobile manufacturers. This allows us to understand industry or sector level performance indicators that are computed bottom-up and without the need to have an explicit industry or sector statement.

Finally, how would one compare two quantities that are not on the same scale? To stay in the Ford’s domain: maybe both Ford expenses and Ford demand increased. Naturally, one would ask: will the increase in demand outweigh the increase in expenses? The percentage uptrend for either KPI simply states how many experts acknowledge the respective trends going up, making a direct comparison possible.

Note that an increase of the number of mentions does not have to change the value of the PTP; it will, however, increase the accuracy of the estimate. For example, if a concept (‘Ford profit’) is associated with three trends, of which two are uptrends, the PTP is 2/3. If the same concept is associated with 30 trends of which 20 are uptrends, the PTP is still 2/3, however, a confidence interval around the latter is tighter than around the former. (Note: this computational example is simplified by leaving out the smoothing that we apply to deal with low counts.)

The predictive nature of the PTP

We distinguish two types of PTPs: forward looking and backward looking. The forward-looking PTPs are based on statements that discuss a future trend, either implicit (“Once Tesla’s demand will take off”) or explicit (“The oil price is set to fall toward the end of 2019”). The backward looking PTPs typically interpret the current or past (“It is not surprising to see investors drop Heinz stock”).

A natural question to pose when contemplating the PTP is: ‘Does the (forward-looking) PTP predict the future?’ If the answer is ‘yes’, it means that the collective expert opinion is mostly right about how things will pan out. For some topics, the PTP may be a predictor of the future. For every topic one would need to have a viewpoint on common belief systems and biases of the actors involved. Thus, the question about direct predictability turns into the questions: does the PTP represent the market opinion and does the market opinion drive the market? **Causality Link enables market participants to improve their financial models by including the market opinions on any topic imaginable.**

Understanding what other people believe can be a great advantage, and even more so if the degree of those beliefs is calibrated, i.e., based on objectively observed data. It may well be that a large group of authors is not foreseeing a recession. If your viewpoint on a looming recession is different from the crowd, then 1) it’s good to take note of their arguments (mentions of contexts) and 2) if more people take the opposed view while you still believe in your view, it may be good time for a contrarian bet.

The father of value investing, Benjamin Graham, said that in the short run, the market is like a voting machine--tallying up which firms are popular and unpopular. But in the long run, the market is like a weighing machine--assessing the substance of a company. We see many more authors discussing Apple than any mid-cap company. If our main statistics would report on counts, indeed we would monitor topic popularity. However, since our summary statistics are proportions of up-trends, an increase in popularity would result in a smaller confidence interval around the PTP statistic. The trending contexts an author chooses to point out are factual observations and the authors try to shine with their gathered insights. As such, we believe the PTP assesses the substance of the company.

Furthermore, the evidence for the relevance of the PTP is seen from the following procedure: our causal inference engine can list all intra-company KPIs that have been mentioned to impact the stock price. We use that list to compile a single backward looking PTP. For example, the revenue, the demand and the expenses are KPIs that authors indicate have an impact on the stock price. We can sum the uptrends of revenue and demand to create a single 'performance' KPI. Since lower expenses drive a higher stock price, we sum the uptrends of revenue and demand and the downtrends of expenses into the single performance KPI. The list of stock impacting KPIs contains about 150 items. The combined performance PTP serves as a health indicator for a company. If we plot this daily health indicator against the stock price, we observe large co-movements. On the one hand, this is not surprising, because obviously, what is discussed is what is impacting the stock price. However, what is surprising is that we do not use a single numeric detail (e.g. the revenue in dollars), but rather determine the proportion of uptrends of all contexts for all KPIs that (according to experts) link to the stock price.

Causality needs trends

So far, mentions and trends have been discussed. We chose the name 'Causality Link' based on the unique capability of extracting written causal statements. Before discussing the causal statements, we outline the arguments that are the basis of our causal inference engine.

In the language of the financial world, causes come in two extremes and many phenomena are a mix of both: either an event sparks a consequence (import tariffs cause the GDP growth to slow down) or a slow-moving market shift results in an effect (increasing cost of commodities results in margin pressure). Many events are very unpredictable by nature; slower changing forces, in contrast, are not unpredictable, however, they may be more hidden from direct observation, both from the perspective of their value change as well as how their influence is carried over into consequences.

While there may be elaborate ways to establish the presence of causality, there's one simple rule on the absence of causality: if A (say, economic uncertainty) impacts B (say, the gold price), and B changes, while A is not changing, then, in that instance and time, A is not the cause of B (yet, it may be that the change of B is caused by a change of A in an earlier time frame due to delay effects). This says: for a cause to be impacting, it needs to be changing. In the financial world, these changes are written down as trends. The sentence: *"In 2009:Q4, with only a 0.5 percent decrease in GDP, the unemployment rate rose by 3 percentage points relative to 2008:Q4."* contains two trends and therefore, two inputs as possible causes or consequences for inference.

In the Causality Link platform, we establish trends and events as drivers or impulses which propagate their influence via causal links to other events and trends. For example, if, from another article, we learn that the rising unemployment rate is a driver for house prices, we can infer using the above trend statement on the unemployment rate that in 2010 the housing prices would have been impacted by the 2009 Q4 increasing unemployment numbers.

Note that many NLP approaches in the financial world only count the mention of financial concepts rather than their (potentially) associated trend. While understanding the evolution of the frequency of mentions may be interesting, it's only if one understands the trend direction for a financial concept that one can 'compute' the influences over causal pathways, hence our preoccupation with trends.

Learning from causal statements

Collecting all statements around Ford profit does give the 'what', but not the 'why'. Financial analysts explain how they come to their conclusions: "with increasing cost of steel, Ford's margin is getting thinner…" An NLP engine that performs a keyword search around 'Ford' never provides the insight of the impact of commodity prices. It's exactly the link between commodity prices and Ford profit that is the long-lasting knowledge: while the steel price may be stable for some time, once it's changing again, one better remembers that it is impacting Ford.

A causal statement has a cause (we call this a 'driver') and an effect ('consequence'). Given our platform's automated extraction of millions of causal statements, our solution lists the drivers for a set of consequences or lists the consequences for a given set of drivers. This automation by itself has is a useful checklist for every company, industry or macro concept.

In addition to the causal statement, the direction of the relationship is extracted. The sign of the relationship for the above example of the cost of steel on profit is negative. The summary statistic Positive Relationship Percentage (PRP) is the percentage of authors claiming a positive relationship. One can argue that the cost of steel is always negatively related to Ford's profit, however, authors claim contrarian views, with statements as 'despite the increasing cost of steel, Ford's margin is growing', indicating a declining relevance of the cost of steel due to other factors. The PRP displayed over time automatically captures these nuances and makes them structurally available for discussion. To the best of our understanding, there is no simple way to access changes in relationships between financial concepts that do not require heavy statistical modeling.

It must be a matter of trust to rely on a single expert's opinion on a changing relationship and act on its logical conclusion. Yet, if over time it shows that most experts believe a relationship is changing, it most certainly pays off to validate existing financial models against that change.

Use cases

Reading financial articles has been the traditional way of collecting enough expert opinions, and, categorizing and weighting the perceived information has been the thinking process that happens in the reader's mind. One only hopes that the few viewpoints one considered are enough to make a good decision about an expected outcome.

For many concepts, the exhaustive list of contexts for a given query is larger than a single mind could ingest. Various filter mechanisms shorten the list: increase the significance level, take the most commonly mentioned concepts, or take the concepts most agreed upon.

The most straightforward use case is enabling an expert to 'read everything'. We find it very helpful to be able to consult a list of all possible contexts and causes on any financial topic as discussed by experts.

Furthermore, comparing the contributions of the various authors and sources shows different points of view and relating one's own view to the collective of experts helps one to be more informed and more objective in one's perspective.

If the PTP represents the trend, and the PRP the relationship, we can take the PTP of a driver and multiply this with the PRP to compute an impact statistic. (Note that this is not a straightforward multiplication, but it's an operation that yields the desired properties for the impact statistic, for example, if both the PTP and PRP are lower than 0.5, the resulting impact should be larger than 0.5.) The impact is the implied logical consequence of the expert opinion carried forward over the causal chain of events. The discrepancy of an impact and the current target PTP may serve as an early warning system for portfolio risk, showing consequences that are about to impact, yet have not manifested in the market yet. Computing consequences on large scale is outside the realm of human capabilities and thus forms a welcome addition to existing early warning signals.

In this context, very specific use cases may exist: using the inference system, we may ask for a cause that has an extreme PTP, while the PRP indicates that in recent times no author has discussed that link. The opportunity is to drive the cause to its logical consequence before others do so. Again, this process is fully automated and yields candidate hypotheses in scale.

In many use cases, we are learning to reason on the reasoning of experts. While the arguments of a single author may be countered by another argument, the collection of arguments of all authors requires a new type of reasoning which, we believe, is only in its infancy and has great potential to understand how collective opinions move the market.

Most of these use cases center around a single person trying to make smarter decisions by having better (or more complete) information, however, we believe that machine learning approaches could also benefit from our approach. Note, for machine learning, having any KPI for any context represented uniformly with PTPs and PRPs as proportions between 0 and 1 greatly reduces the complexity to make the inputs suitable for machine learning. Furthermore, as 'traditional' financial models take more structured data into account, our structured output may give rise to several benefits:

- Rather than only incorporating 'hard facts', incorporating how experts view the hard facts may enhance the predictive capacity. For example, if the profit of a company is good, however all experts claim that the underlying demand is deteriorating, it's more likely that the stock will go down. Without taking the expert opinion on faltering demand into account, the model will over incorporate positive signals on profit and thus incorrectly predict a stock price increase.
- Every company has its own set of drivers and this can be used to (automatically) inform which factors needs to be considered in the model. Here, our NLP extracted relationships can be used as Bayesian priors. This gives rise to the automatic creation of company specific models. A similar approach is Judea's Pearl view of causality where one starts (scientific) research with "the hard step of constructing or acquiring a causal model"[1], which is what we deliver, constructed from collective wisdom.
- Every company has peers in the industry, and while not every driver may have been mentioned for every company**,** we may conduct graph inferences to 'learn' drivers from peers. This is like the previous example, with the additional step to 'generate' new knowledge from the graph that was never mentioned by an expert.

Statistics on the size of the system

The value of the system increases with the number of ingested articles. Currently, we have processed about 70M articles, covering financial analysis and world news, in 17 languages that span the last 5 years. On a daily basis, we ingest and process about 30K articles in an agile cloud infrastructure that is fully customer provisioned. The process is easy to scale up or down as many processes are parallelizable by nature. As mentioned earlier, the articles remain linked to the presented statistics and are directly accessible as reference for any query.

Moreover, we can create a dedicated AWS instance of our system per customer, and safely ingest some of the customer's content for his own consumption, which can be for example to compare his experts' point of view with the rest of the publicly available information.

Concluding remarks

This document presented insight in how we use automatically extracted expert knowledge and transform it into structured data. The process of learning to draw inferences on the collective reasoning on financial concepts has given us the ability to make expert opinions computable. Tapping into that collective knowledge wisdom can make any of us smarter.

In recent ML/AI discussions, the call for explainable AI has received increasing attention. In the context of our PTP and PRP statistics: every sentence, and thus every article, that aggregates into those numbers is directly linked to it, enabling direct inspection and validation of the statistic. We also found it enables hyper-focused reading: after specifying the query (i.e. 'Ford revenue China'), the PTP shows the expert agreement and all relevant passages of text for all articles that added up to that PTP are available as underlying evidence, enabling very quick navigation to divergent opinions, for example.

Our AI is a new form of collective intelligence rather than a replacement of human experts. It also shows that despite what people may be saying about the intelligence of AI, the subtlety of an expert opinion mentioning a specific context from the perspective of his/her expertise, is not likely to be replaced by AI any time soon. The ability to use AI to create a collective point of view is a great start to act on the 'big picture'.

Reference

1 Judea Pearl and Dana Mackenzie. 2018. The Book of Why: The New Science of Cause and Effect (1st ed.). Basic Books, Inc., New York, NY, USA. P16.