THE NEXT GENERATION ALPHA MODEL:
INTRODUCING NATURAL LANGUAGE PROCESSING

VIRTUS SYSTEMATIC

NEXT FRONTIER OF SYSTEMATIC INVESTING

Natural language processing ("NLP") is a natural (no pun intended) extension to the behavioral finance ("BeFi")-based alpha generation model of the Virtus Systematic Team. As much as investors may think of themselves as rational actors, they are susceptible to "predictably irrational" behavior. This was the basis for our BeFi based alpha model, which primarily used *quantitative* data to capitalize on behavioral biases of investors, sell-side analysts, and company management.

In this research improvement, we demonstrate capabilities based on NLP that can effectively quantify textual inputs – including reports, writings, and speeches. This milestone adds a new dimension for our Befi models, opening an arena that has been largely untouched so far by investors. We leverage artificial intelligence ("AI") to accomplish the formidable task of creating a truly robust NLP model. Our starting point is using NLP to identify when the sell-side analyst sentiment contradicts their numerical forecasts on a stock.

KEY TAKEAWAYS

– The Virtus Systematic Team aims to capitalize on inefficiencies caused by human behavior. NLP can enhance our already robust alpha models.
– Not all NLP models are created equal and we carefully constructed our model in an effort to mimic the way humans read and interpret information.
– By incorporating artificial intelligence with NLP, the self-learning process is designed to generate continuous improvement as our training library expands.
– Results show that NLP can help an earnings revision alpha model avoid false positives and identify missed opportunities.

NAVIGATING TREACHEROUS BUZZWORDS

It is important to clarify some fuzzy technology jargon that has been casually—even haphazardly—discussed in popular media, across the dinner table, and around the water cooler. Following are the generally accepted definitions:

- **Artificial intelligence (AI)** is a broad term that encompasses a wide variety of concepts and practices within computer science that focus on systems that can interpret, adapt, and learn using external data to achieve specific goals.
- **Natural language processing (NLP)** is a field of study that focuses on programming that can analyze and process human language. Given the complexity of language, NLP heavily overlaps with artificial intelligence (AI) and is often considered a sub-field of AI.
- **Artificial neural network (deep learning)** refers to a framework of algorithms that are structured like a human brain – a biological neural network – and has machine learning ("ML") capabilities. This particular structure allows for a model that can continuously learn beyond explicit programming.

ROADMAP FOR INTEGRATING NLP

A common intuitive attempt to utilize NLP in asset management requires the following sequence: First is the objective, which is typically to forecast asset price movement. Next is ground truth, a gold standard that will guide the learning process of the machine-learning algorithm. Finally, and most importantly, it requires data, lots of data, in the form of "natural language." These can include regulatory filings, transcripts, news flows, sell-side equity research reports, and a wide range of other documents.

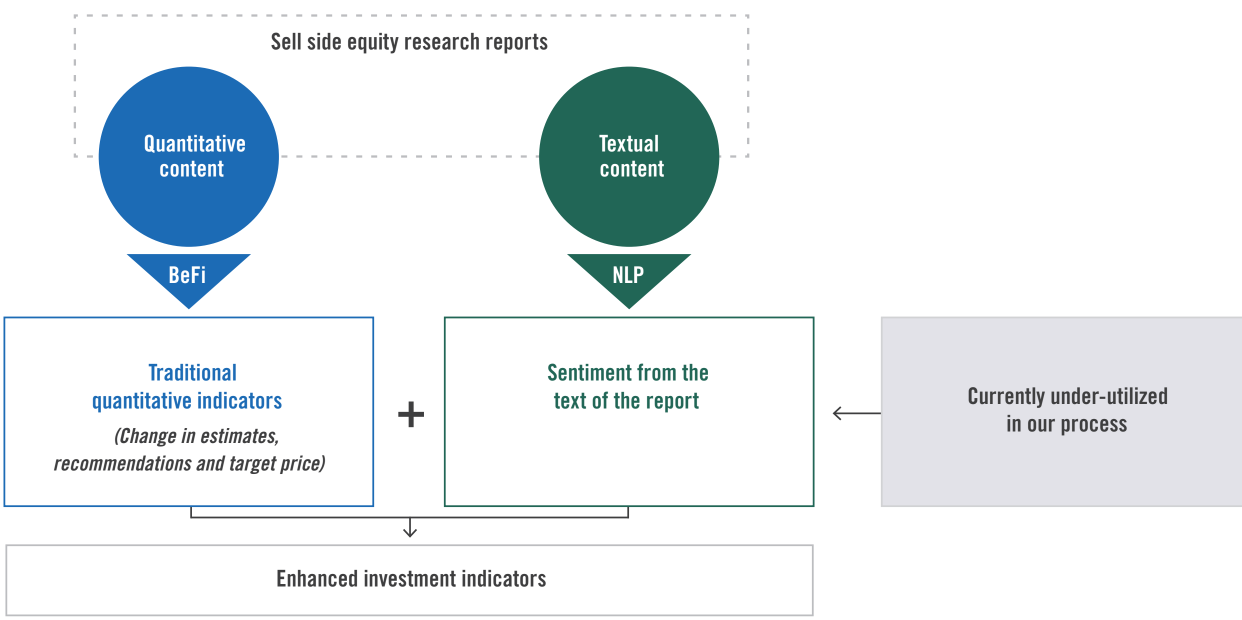
Note, however, that a direct asset price forecasting objective embodies an expansive scope and is susceptible to modeling weakness. Many quantitative investors can encounter "overfitting" issues, whereby the data has been mined to fit asset prices, generating in-sample results that are deceptively attractive. The live model, in contrast, can often fail to achieve anything similar. This effect is magnified for a machine learning NLP model that is trained using an asset return ground truth.

We took a more defined approach by building the NLP with sentiment detection as the objective and ground truth, effectively avoiding overfitting issues. Instead of directly forecasting stock returns, the goal is to have the NLP model enhance our already robust alpha models, helping to further capitalize on established behavioral finance relationships with equity prices. Currently, our models rely on quantitative data inputs from sell-side equity research estimates, which include earnings estimates, recommendations, and price targets. While our Befi-based models have generated strong results with these quantitative inputs from Wall Street research, they have underutilized the textual aspect of the sell-side research.

Sell-side analysts sometimes change their quantitative views for reasons that appear to be traced to behavioral psychology. An example is the propensity to "herd" and play it safe by following the behavior of other analysts. Other false quantitative indicators can include conflicts of interest from investment banking relationships; the desire to maintain a good relationship with company management; and financial one-offs in which earnings experience a temporary but unsustainable boost.

By incorporating analyst sentiment into the model, we add a new dimension and perspective for our alpha models as shown in exhibit 1. Using NLP, we believe that it is possible to materially reduce Type I Errors, or false positives, by identifying when earnings revisions do not fully represent the intent of the analyst, and therefore should be considered with circumspection. Further, this enhancement may reduce missed opportunities, or Type II Errors. As we show later, there are ample cases where an earnings estimate increase does not align with an analyst's view of the company outlook, and vice versa. By isolating companies with estimate changes that do not line up with analyst sentiment, we can expect to further improve our investment returns.

EXHIBIT 1: ENHANCING OUR BEHAVIORAL FINANCE MODEL



STRUCTURE OF AN NLP MODEL

Generally, an NLP model can be traced through the two following principal steps: data analysis and model (exhibit 2). In the data analysis step, the NLP model analyzes cleaned sell-side research report texts and generates sentiment prediction scores. Note that cleaned text is comprised of the core body of the reports and is devoid of any noisy text, such as disclosures, tables, or analyst contact information. In the model training phase, the model engages in a self-learning process, which is a positive feedback loop that creates a continually improved version of the model.

EXHIBIT 2: PATHWAY TO AN NLP MODEL

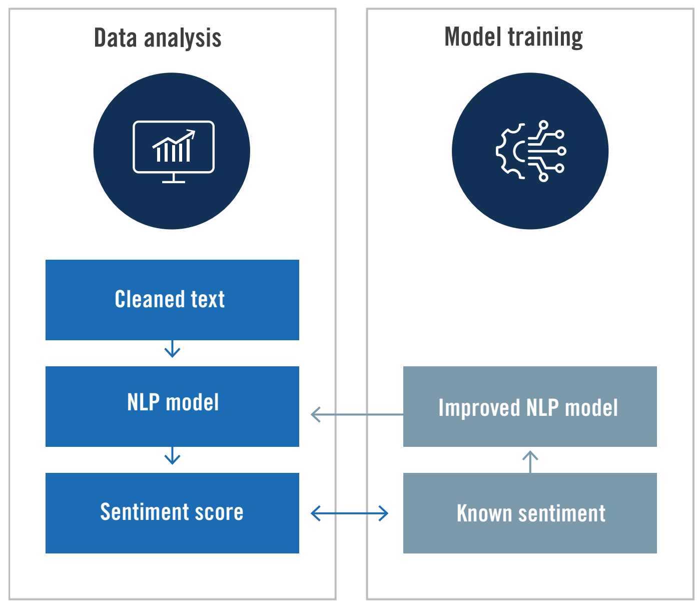
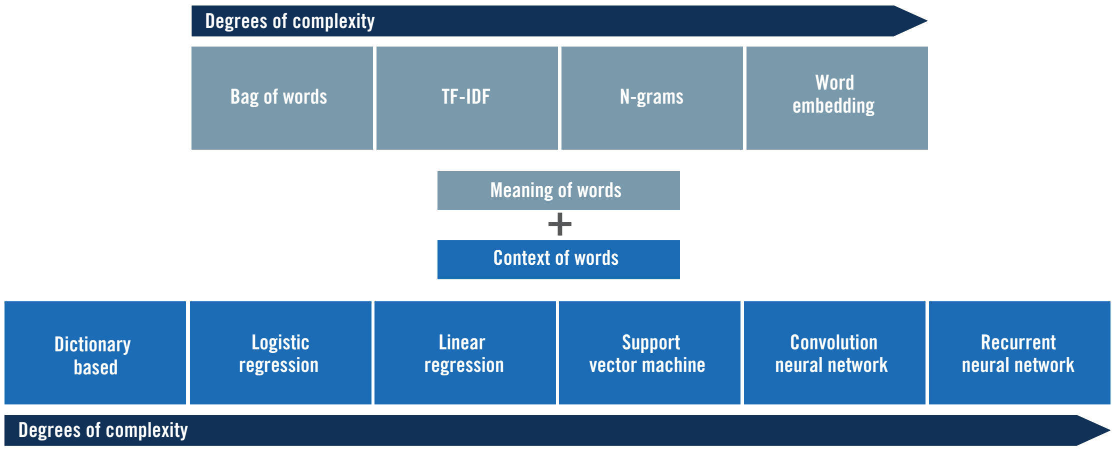


EXHIBIT 3: BUILDING BLOCKS FOR A NLP MODEL



DATA ANALYSIS: PROCESSING MEANING AND CONTEXT OF WORDS

The core of the NLP model is processing text into a sentiment score. This methodology is divided into two principal steps: understanding the meaning of words and discerning the contextual nuance of the words. We purposefully structured the methodology to mimic the way humans read and interpret information. Simply put, not all NLP models are created equal. As shown in exhibit 3, there is a range of options that we can utilize to construct the NLP model.

Consider the model in exhibit 4, which is comprised of the most elementary NLP techniques. For the model to process the meaning of words, it employs the widely used "bag-of-words" technique, which identifies the presence of known words. The model would assemble the greater context by using a dictionary-based word count, where every word has been pre-assigned a positive or negative implication.

We can plug in the following sentence from a restaurant review: "completely lacking of thoughtful service, good taste, and enjoyable food." It would be obvious to any human reader that this sentence has a negative sentiment. However, the "bag-of-words" based model would tally all the negative words (lacking) and positive words (thoughtful, good, enjoyable) to generate a positive—obviously incorrect—sentiment score. Clearly, this approach does not allow the model to distinguish the contextual nuance of the entire sentence.

Given the complexity of human language, we need to design an NLP model that has the flexibility to deal with these challenges. Going back to the different building blocks at our disposal, we opted for a combination of *word embedding and the recurrent neural network*. It is important to understand the interplay between these components.

Word embedding is a technique that identifies meaning by grouping words based on their relationship to each other.

EXHIBIT 4: BAG OF WORDS + DICTIONARY BASED NLP EXAMPLE

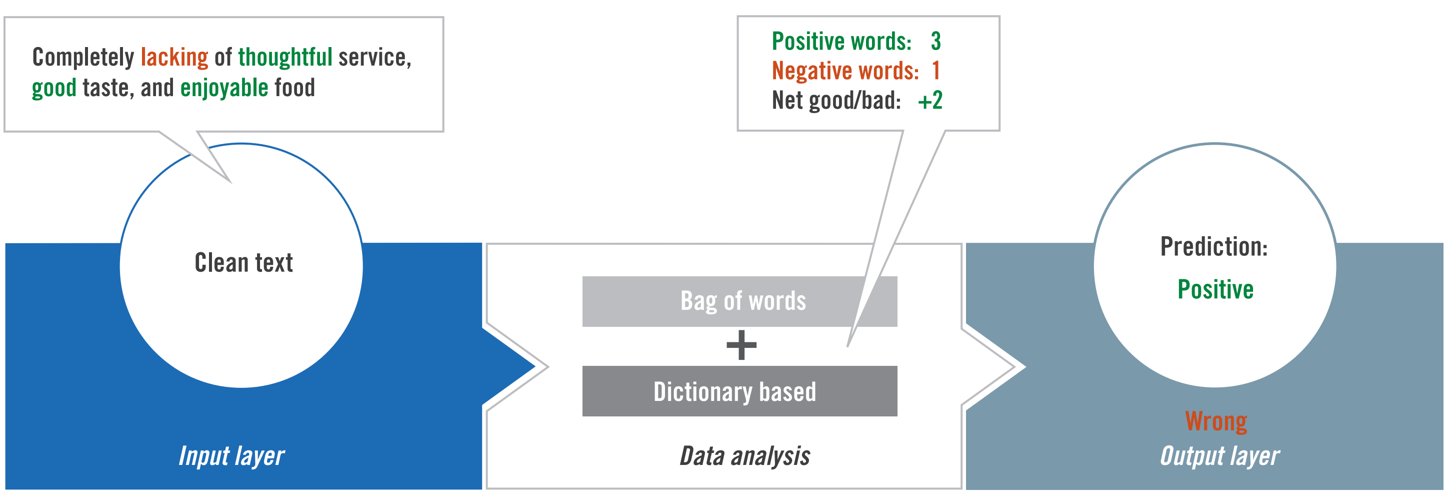
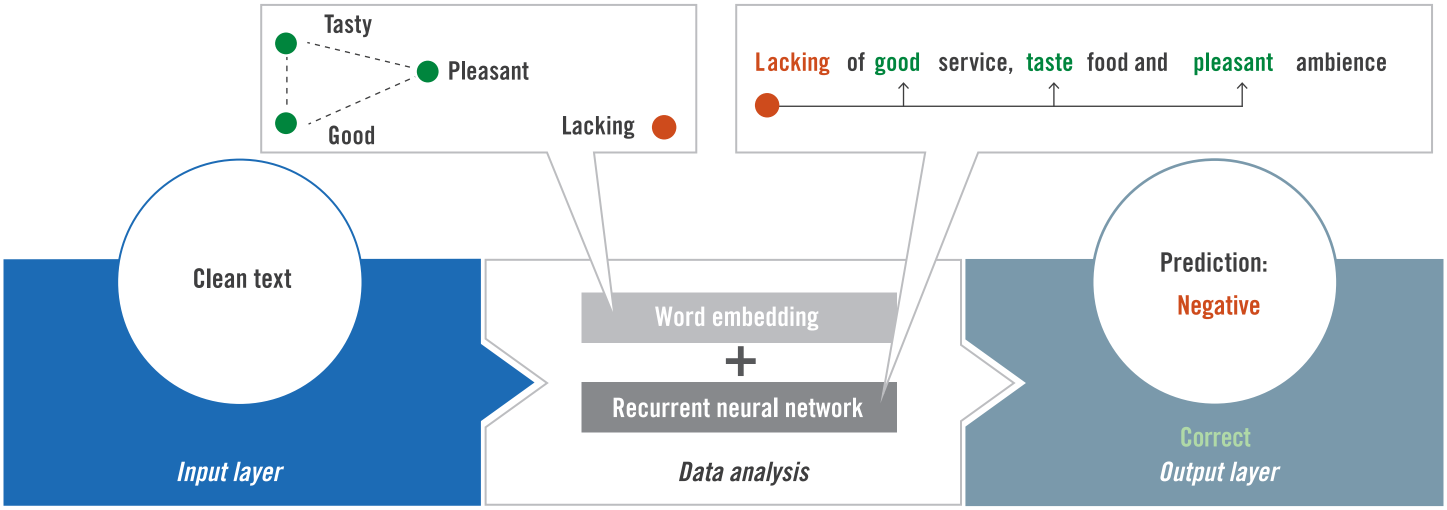


EXHIBIT 5: WORD EMBEDDING + RECURRENT NEURAL NETWORK NLP EXAMPLE



Going back to our restaurant review example, the words "thoughtful," "good," and "enjoyable" would have a close relationship because of their usage for a positive customer experience. This focus on relationship among terminologies allows word embedding to be calibrated for different genres of text. For example, "apple" would have a close relationship with "oranges" for culinary blogs, while equity research reports would have a closer association of "apple" with "smartphones."

While it is important for an NLP model to recognize the meaning of words, the model is still suboptimal if it cannot understand the greater context of a document. Returning to our restaurant review example, the overall statement had a negative sentiment because the word "lacking" negated all the subsequent positive words. But the interaction cannot be picked up if the NLP model uses a traditional method of aggregating positive and negative words. To overcome this obstacle, we use the recurrent neural network, which processes words in a sequential manner and retains them in memory as we progress. This allows the model to process each word differently based on the preceding text—which is exactly how humans read!

As shown in exhibit 5, bringing the overall data analysis process together, word embedding first identifies the relationship among words. This is then fed into the recurrent neural network, which processes the text in sequential order to generate a measurement for the sentiment.

MODEL TRAINING THROUGH SELF-LEARNING

Through the model training process, the NLP model has the ability to learn and evolve as it encounters new data. Also, this process takes place outside of the data analysis function described in the previous section. If data analysis is equated to competitive archery tournaments, model training would be a player spending endless hours at the range.

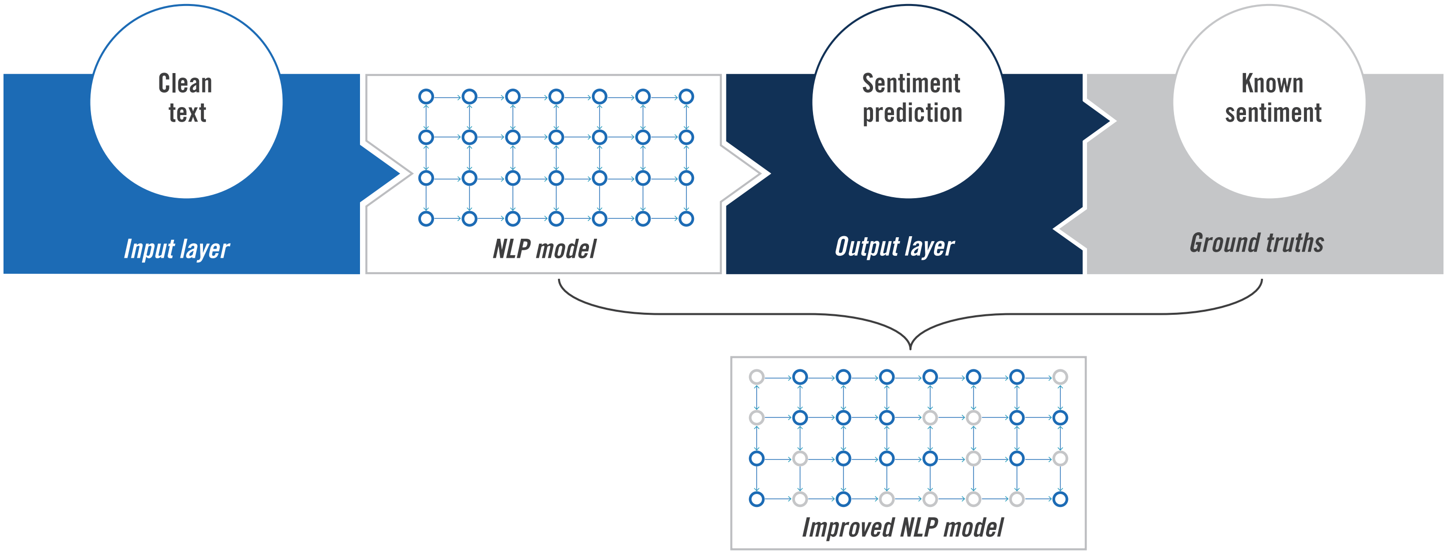
To fully realize the model's potential, we require extensive in- and out-of-sample datasets, which are used to train and validate the NLP model, respectively. At the time of this writing, our dataset is approximately 650,000 unique sell-side equity research reports on individual companies from emerging market countries. However, we are continually expanding the dataset to include new research reports from both recent and historical periods.

While intuitive, we want to highlight that data availability has an inverse relationship with a durable information advantage. Compiling a comprehensive and machine-readable dataset of sell-side research reports represents a significant challenge. The landscape of brokerage houses is highly fragmented and has been further complicated with the introduction of MiFID. Obtaining a decade of historical research and new reports requires significant access. However, this exercise is well worth the effort. Having a data source difficult to replicate protects the persistence of our NLP model's alpha enhancement potential and prevents it from getting easily arbitraged away.

To undergo the model training process, the NLP model's output is compared against the expected output. Going back to the archery analogy, suppose an archer's arrow continuously veers away from her intended trajectory. As the archer compares a videotape of her most recent shot with other videos with successful outcomes, she can course correct to eventually have mastery of the technique.

continued

EXHIBIT 6: MODEL TRAINING ALLOWS FOR CONSTANT MODEL IMPROVEMENT



In a similar manner, the model training process hones the NLP model's accuracy by comparing the sentiment prediction output against the known sentiment (as shown in exhibit 6). The NLP model will adjust its internal parameters to learn from past mistakes, thereby creating an improved iteration of the model. At the time of this writing, the most recent iteration of our NLP model has over 2.4 million parameters, which has been adjusted over 100 billion times. As we introduce new research reports, expand the learning library, and enhance the structure of the neural network, we expect the NLP model to improve.

We cannot emphasize enough the importance of establishing the ground truth as it is the gold standard for an AI-powered NLP model. Establishing a high-quality ground truth of a dataset requires a tremendous undertaking and demands significant resources. We did not limit ourselves to a rule-based ground truth (typically returns) as many of the other reported NLP research has demonstrated. Instead, we augmented a rule-based process by leveraging an extensive network of investment professionals and recruited finance students from a top university to authenticate the sentiment from sell-side research reports. By having a ground truth backed by the collective wisdom from a group with expertise, we optimally positioned the NLP model to achieve the goal of mimicking the mind of an experienced research analyst.

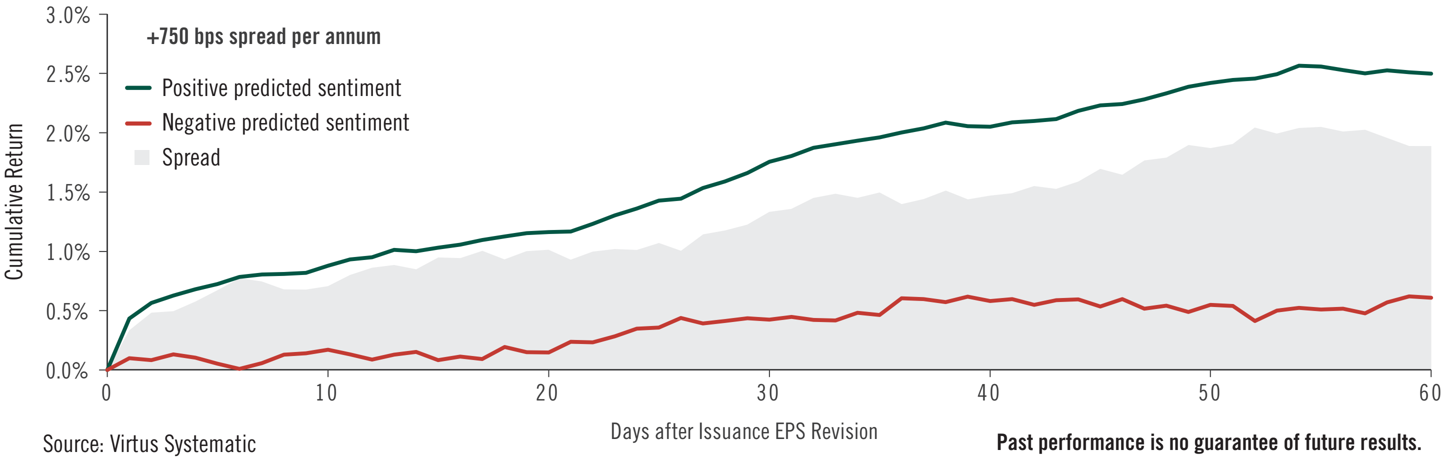
IMPACT OF NLP MODEL ON EARNINGS REVISION

To demonstrate the NLP model's potential, we put our new sentiment model to work alongside an earnings revision factor within the emerging markets equity universe. While this earnings revisions factor is already a robust indicator,

EXHIBIT 7: CASE EVENT STUDY

continued

EXHIBIT 8: EVENT STUDY OF MARKET REACTION WHEN ANALYST POSITIVE ESTIMATE CHANGE IS COMBINED WITH SENTIMENT



Source: Virtus Systematic

Past performance is no guarantee of future results.

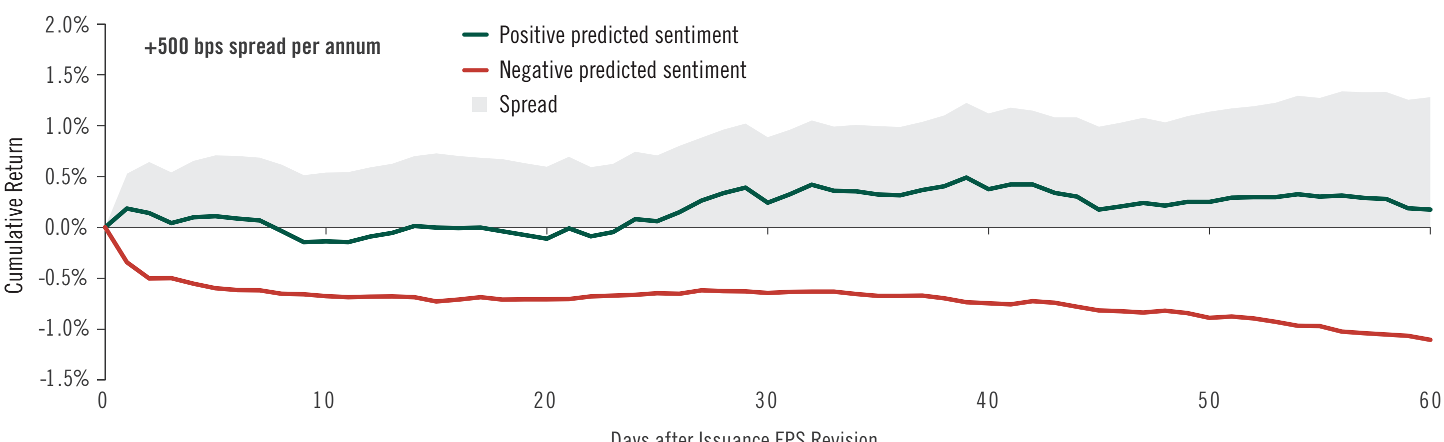
the results below show that the sentiment indicator can isolate analyst reports with conflicting earnings revisions and sentiment.

While positive earnings change is typically a buy signal, there are instances in which analysts have a negative outlook but do not want to deviate from the herd by having an outlier earnings estimate. Instead, analyst concerns could materialize in the report's text, which gets picked up by the NLP sentiment model. These are precisely the Type 1 Errors that we want to avoid, and are manifested in the exhibit 8 red line and exhibit 7 bottom left gray box. Further, by isolating opportunities with positive earnings revisions with positive analyst sentiment, we can generate significant performance improvement as shown in the green line in exhibit 8.

On the other side of the spectrum, our sentiment indicator also helps improve the factor's accuracy by identifying the hidden opportunities masked by negative earnings revisions. While negative earnings change typically is a sell signal, analysts may identify emerging trends or positive developments in the text of the report. These could be instances in which stock price movement has priced in a significant amount of bad news and the company may be approaching an inflection point. As shown in exhibit 9, the NLP sentiment indicator can help identify these hidden opportunities—and help further avoid investments with heavily deteriorating fundamentals.

Overall, we believe enhancing an earnings revision factor with a NLP model can generate durable excess returns, which has been shown to grow up to 60 days after an earnings revision is released. It is important to note that an NLP model's efficacy is highly dependent on the types of information it digests and the way it is integrated into an investment process. As we mentioned previously, the NLP model complements our BeFi alpha factors and should enhance the efficacy and durability of alpha potential.

EXHIBIT 9: EVENT STUDY OF MARKET REACTION WHEN ANALYST NEGATIVE ESTIMATE CHANGE IS COMBINED WITH SENTIMENT



Source: Virtus Systematic

Past performance is no guarantee of future results.

continued

ONLY THE BEGINNING

The NLP model described in this paper is only the starting point for the application of NLP for active managers. With an NLP model that is fully trained on an expansive library of sell-side research as a starting point, future iterations of the model could understand and form opinions on other text, including earnings call transcripts and management discussions in regulatory filings. The model could also generate views on the macroeconomic environment with central bank commentaries, economist speeches, and Wall Street strategist research notes. Further, the same model could be calibrated for different languages, opening inclusion of content generated from across the world. So far, the world of textual data appears to be a rich area for asset management innovation. We have only begun to explore and discover the next generation alpha models.

Kunal Ghosh
Senior Managing Director
Chief Investment Officer

Lu Yu, CFA, CIPM
Managing Director
Senior Portfolio Manager

Jie Wei, CFA
Director
Senior Portfolio Manager

Yang Zhang
Assistant Director
Data Scientist

The Virtus Systematic team of Virtus Investment Advisers, Inc. specializes in differentiated investment solutions, strategies, and outcomes across asset classes, regions, and securities. The Virtus Systematic team manages U.S. mutual funds for Virtus Investment Advisers, Inc. and other portfolios for Virtus Fund Advisers, LLC.

This report was originally published in July, 2019 while the Team was known as the Allianz Global Investors U.S. LLC (“AllianzGI”) Systematic Investment Team. Effective July 25, 2022, AllianzGI’s Systematic Investment Team joined Virtus Investment Partners, Inc. as the Virtus Systematic Team. The content has been edited from the original.

VIRTUS
SYSTEMATIC
virtus.com • 1-800-248-7971