# Sentimental quant

Sentiment Alpha Capital Management is an investment management firm that generates quantitative trading strategies from sentiment analysis techniques combined with large-scale media data – including sources such as Twitter, blogs, news and broadcast. CEO Jae Hong Kil tells *Automated Trader* about the firm's foundations and his hopes for the future.

Jae Hong Kil, CEO and Portfolio Manager, Sentiment Alpha Capital Management

Source: Sentiment Alpha Capital Management
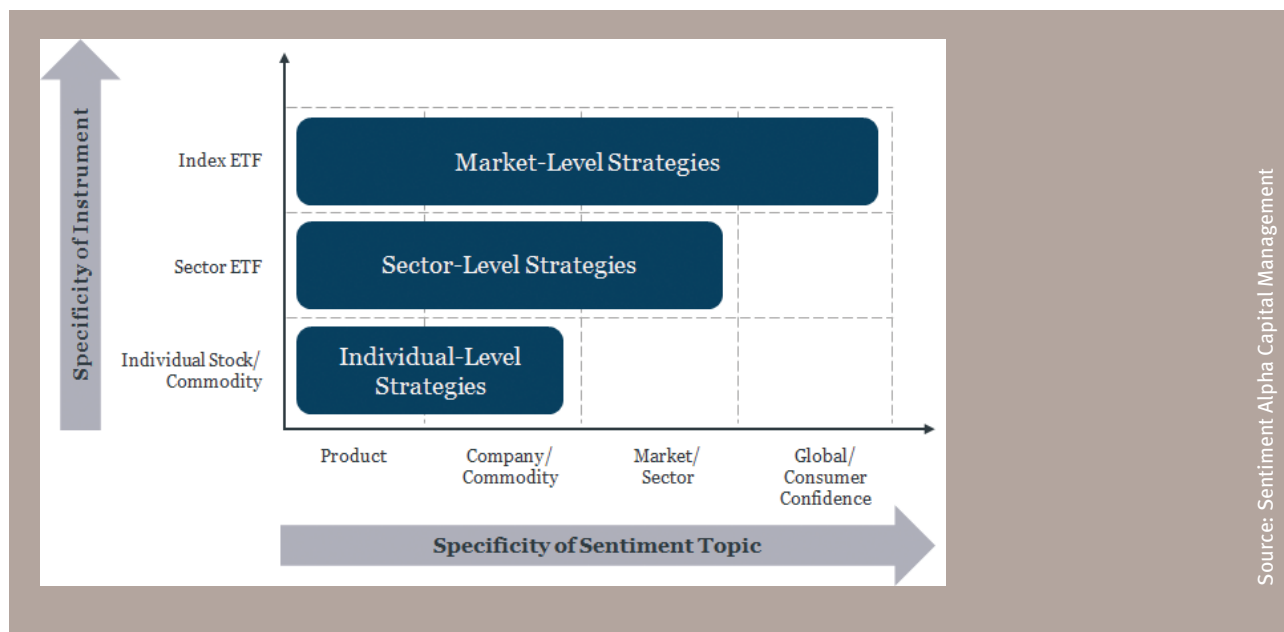
Automated Trader: Why did you choose a sentiment analysis-focused fund?

Jae Hong Kil: In the past, I was actually in the US military in intelligence. That's all about information and how it's connected. Then I got into computer science, and then finance. Right now all those different fields really go together.

I went to university at Stony Brook, I was a computer science graduate student from 2004 to 2006, and we developed a system to aggregate news and social media.

It was a big project and team. 11 years ago, that was pretty new. We tried to build a search engine a bit like Google, a named entity based engine, which means if you type a named entity, like a person's name, it builds results on the named entity – like how many times people mentioned the name over that day. We collected raw text and applied natural processing technology into the raw text and pull out the information and analysis. That's what's going on in the background.

One of the analytics is sentiment analysis. We looked at the polarity of the sentiment – positive or negative – and it shows the score of how it changed. That was a really big project, and now lots of the team works at Google. It was pretty successful.

AT: And then you worked as a quant for a bit?

JK: I joined Natixis and worked for about six years as a quant trader. But I was watching how the university project kept going. The technology became a company, and the company was spun out. They called me and were looking for a person who can run it. I knew the technology behind it and have a trading background, and that is how I ended up joining.

At the time there was no trading strategies, but I had a pretty good view, and was following the trend and growth of social media and saw that as a huge potential. I believe in the technology.

AT: Can you tell a bit more about the size and scale of the fund?

JK: AuM is still single digit millions and our strategy capacity is a billion dollars or more. Right now, the last couple of years we are really focused on building a track record.

Asset class is US equity and strategy type is long/short. We plan to launch long-only too. Average holding period is one month, so this is not short term strategies. We have a proprietary sentiment data engine, everything is built in-house and we get a market data feed from Bloomberg.

AT: What technology underpins the NLP engine?

JK: We use many algorithms including machine learning and there are many

Jae Hong Kil, CEO and Portfolio Manager, Sentiment Alpha Capital Management

different phases. The first one is marking up the raw text. Then we apply rule-based algorithms, and also we can apply a lexicon, for example, once we identify all the named entities, we classify those named entities into types – person or country, etc.

The big difference between us and data providers, like Thomson Reuters and Bloomberg, is we are not just getting sentiment analysis for a company, we are getting everything. A stock price is impacted by a lot of things. People think that we just monitor what people say about the company, but we also look at what the company's performance is directly and indirectly based on, for example, products and executives. For example, it matters what people think about iPhone 6 or iPads when it comes to Apple's reputation. But then we also differentiate from iPhone 4, the product is outdated.

The bottom line is we are capable of getting sentiment data for every single name. Then we can build any sort of analysis that can be sliced and diced on top of that.

AT: When you say you use social media, what does that mean exactly?

JK: Facebook, blogs, any other major names. But Twitter is the major one. We've archived and backtested since 2009. Twitter volume significantly increased only since 2011 though. Also the period was not natural because from

2009, the market was recovering. It would be much better if you could have 10 years.

AT: What kind of strategies are you developing?

JK: We are capable of developing micro and macro level strategies. A micro level means low level sentiment data. For example, get a sentiment score for each product and aggregate all products' sentiment scores. Then convert that into the company's sentiment score.

We can also develop strategies based on relevant entities – like executives. When Steve Jobs (former co-founder, chairman and CEO of Apple) passed away the stock went way down, so that is a factor.

But we can develop strategies on bigger levels, sector or market level strategies, which can be done by aggregating companies' sentiment scores

People talk about the US economy or China's economy, and we get the sentiment data for those terms. We look at kind of a macrostrategy in what people think about the market, or the US economy. It really depends on how you aggregate and what kind of data you are looking at.

AT: How is the fund raising environment?

JK: Some of the first funds launched in the past that have since shut down were influenced by what I consider shallow

research. There was one paper from Indiana University showing a significant correlation between the stock price and Twitter data.

But this space is really intuitive, it makes sense. If a million people say they want to buy iPhone 6, that is going to be translated into their sales. That will tell you something in the future.

So intuitively, people agree, however, there are so many questions about the quality of the data and people are sceptical. And also the lack of data. A little while ago Twitter didn't exist, and the volume was not significant, so backtesting period was not long enough.

A lot of institutions and quants didn't trust the data and the methodology. They thought it was intuitive and interesting, but wait a second, I am not ready to put my money into it because it's new and a lot of unknowns.

AT: What's been most successful for you?

JK: When I am really proud of our system is when we pick up on small news, news you have never heard about and not really anticipated but somehow people pick it up or create it.

This is especially true in consumer sectors. Let's say we believe there is a potential picking up in restaurant retail. You went to a restaurant, you loved it, you post the review on Yelp, your friends checked it out and they liked it, also ▶

posted about it and it spreads really fast. All of a sudden the restaurant's sales get bumped up, and that kind of thing is not based on their news, but based on the consumer's opinion and the expression of that opinion. Picking up that informational trend is really key.

When we pick things like that up I am really happy about it. We've also identified localised trends with our signals coming out from the local news. It's like identifying grass roots movements.

We believe that especially for consumer opinions and localised events spreading out takes some time. It's not going to happen within a day. That's why you should design a holding period into the model.

AT: Can you give me a specific example?

JK: We picked up buzz and a volume spike for negative mentions of this chain of restaurants in St. Louis, Missouri - Panera. People said they liked their food but one of their marketing campaigns was a big flop. That kind of spread out and the stock price fell.

AT: And what have been some of your worst moments?

JK: When we lose, when our signals don't work, that is not a good moment.

Other factors coming in based on expected or unexpected events, that can work against our favour. We cannot win 100%, so the timing and holding period really matters. A longer term like one month, or maybe three months too, a lot of other factors can come in and it can wash off the sentiment impact.

Right after we launched the fund, we had a strategy and we anticipated the market going bearish in September, 2013 based on sentiment analysis. Then in the middle of the month, Obama put on hold the military strike on Syria and Fed also decided to continue quantitative easing. After the events, the market moved up a lot. So basically, based on unexpected political events, which had nothing to do with our analysis based on sentiment, our strategy did not work well that month.

The unknown factor is kind of dangerous. It is what it is. ∎

DEUTSCHE BÖRSE GROUP

Intraday volatility forecasts for DAX, EURO STOXX 50 and Euro-Bund

Volatility is a key input for trading algorithms and risk methodologies. The capacity to forecast volatility is therefore enormously valuable for traders.

Providing insight for the next 10 seconds to 30 minutes

In November 2015, Deutsche Börse will launch the 'Intraday Volatility Forecast', a new analytic, that forecasts volatility for DAX, EURO STOXX 50 and Euro-Bund. The forecasts are based on trading data of the respective futures, the most important instruments traded on Eurex Exchange.

How it works

The Intraday Volatility Forecast is based on a proprietary mathematical model that uses historical futures data to forecast the direction (increasing or decreasing) and magnitude of volatility. The input going into the calculation can be grouped into three categories:

• Short-term memory of realized volatility

• Historical intraday seasonality

• Scheduled economic data releases

Utilizing the Intraday Volatility Forecast

Traders can use the forecast to support the assessment of the likelihood of price changes and therefore the risk involved in using certain automated strategies. Using an automated trading strategy set against a benchmark, such as a VWAP algorithm, is less risky if the likelihood of huge price changes in the underlying can be gauged. As such, the volatility forecast is a natural choice for

Schematic illustration Intraday Volatility Forecast for DAX

The following metrics will be provided:

• Average realised volatility: a reference for the last second is sent every second

• 10-second forecast: published every five seconds and covering the next three 10-second periods

• 1-minute forecast: published every 30 seconds and covering the next five 1-minute periods

• 10-minute forecast: published every five minutes, covering the three subsequent 10-minute periods

The short-term memory component of the volatility forecast model incorporates the fact that the future value of volatility is highly dependent on its current value. The second component of the model analyses historical data to identify the drivers of intraday seasonality in volatility. The model automatically accounts for relevant holidays, other calendar effects and daily events such as the opening and closing auctions of the different markets. The final model component dynamically evaluates the average impact of different scheduled releases. Releases with significant impact on volatility (e.g. US employment or GDP figures) are used in calculating the volatility forecast.

a state variable in the generation of a trading algo.

On the other hand, traders who are looking for high volatility opportunities and trade respective option strategies can use the forecast to help determine a good point in time for order submission.

The Volatility Forecast is not only interesting for algo traders, but also for screen traders who can use the forecast to support their live trading. Pre-trade transaction cost analysis (TCA) can likewise be enhanced considerably as the risk of price slippage can be assessed using the volatility forecast.

For more information, please go to www.mds.deutsche-boerse.com/volaforecast.