Sentimental education

Alternative investment managers may have been first to seize on the opportunities presented by news-driven algorithms, but sentiment engines and related tools are beginning to support more traditional strategies.

Bob Giffords

Automated news feeds driving alpha have long been the dream of traders. For some it has become a reality. "Machine capture of fundamental economic data and analyst forecasts is fairly commonplace now for quantitative traders," says Adam Toms, head of the market access group in EMEA at Nomura. "Sentiment and other news metrics, however, are less widely used in a systematic manner and demand still appears fairly limited. It's more a 'nice to have'."

"The key idea is to combine news with traditional data sources as an added component of alpha signaling," explains David Leinweber, Haas fellow in finance at the University of California, Berkeley and founder and CEO of

Leinweber & Co. "It's especially good at predicting volatility spikes together with GARCH[1] and implied volatility models. It can also affect beta and covariance with marked differences across sectors."

"Traders need to be wary of 'quantcentration' effects, where most fund managers have access to the same sources of data," warns Professor Gautam Mitra, director of OptiRisk Systems, an optimisation and risk solutions provider, in his latest review of emerging news technologies[2]. "This makes it difficult to distinguish

<hr>

1 Generalised AutoRegressive Conditional Heteroskedasticity

2 'News Analytics in Finance' by Professor Gautam Mitra, Leela Mitra, (OptiRisk Systems), based on research undertaken in: CARISMA, Brunel University, John Wiley, 2010

© Digital Vision

between their multi-factor models and performance. News offers real diversification benefits in this scenario."

However, it is still early days. "We've rarely seen machine-readable news analytics used in anger," says Steve Wilcockson, industry manager for financial services at The MathWorks, the global software provider for model-based design, including news integration products. "Where it is used, or at least tried, the conclusion seems to be that it's very challenging. You need a lot of skill and a big risk infrastructure around news to make sense of it since the significance of any event depends so much on both the context and the strategy."

Nevertheless, algorithmic news feeds are becoming big business, as Armando Gonzalez, president and CEO of RavenPack, describes. "We built a fully automatic system with Dow Jones covering the Wall Street Journal, Barron's and third-party content, annotating over 15,000 stories per day," says Gonzalez. Ravenpack identifies all companies mentioned in a story with separate metadata for each, including sentiment, topics and novelty. Ravenpack also provides independent services

■ "News data can give you advantages of speed, scope and better risk control."

Gautam Mitra, director, OptiRisk Systems

scanning over 100,000 news items a day, adding metadata and filtering out duplicates unless there is new data. "The results are fed in real time to our clients," says Gonzalez, "over the internet, leased lines or direct cross-connects to co-located servers in our various data centres, and they then plumb them into their pre-tested models." Ravenpack has built a massive companies database that allows clients to track firms

and their managers through various mergers and spin-offs, handling all the synonyms, multiple spellings and brand names.

To increase the diversification benefits, stories can be analysed in different ways, so clients get different signals depending on the factors they use. Lower-frequency traders might focus on the longer-term factors while high-frequency traders leverage short-term impacts. "Scheduled events, like earnings statements, tend to have a very short lifecycle in customer algorithms with typically small impact or a sudden spike," says Gonzalez, "while unexpected events tend to show a longer lifecycle that might build for days or even months. Sentiment data also tends to accumulate over time."

Down to the wire
For most news algorithms, speed is essential. "Economic data releases illustrate the challenge," says Peter van Kleef, CEO of Lakeview Capital Market Services. "You have expectations and then the real numbers that might of course be revised. To what extent have expectations been discounted into market prices before the release? To what extent will

unexpected deviations move sentiment and therefore the price? It may or may not increase uncertainty and thus volatility. There are no simple rules.

"Your model always has to decide first what the numbers mean relative to expectations," continues van Kleef, "and then relate that to the market reaction, which is actually more important. Markets of course may move without specific news as investors digest the total context or simply adjust their positions on non-public information."

For economic data releases, being 'first to market' matters. Deutsche Börse has moved boldly into this sector, as Georg Gross, head of market data and analytics explains. "To strengthen our ties with the algorithmic trading community we made two strategic acquisitions in the automated news space last year," says Gross. "First we acquired Market News International, one of the four international newswire services with fully accredited journalists in all the global market centres. Then we bought Need to Know News, the market leader for low-latency macro-economic data feeds, whose Lightning Bolt product had acquired a substantial client base in North America and Europe. We've now merged their offerings into a best-of-breed AlphaFlash service with a new low-latency infrastructure co-located with our customers in the key trading hubs."

"The AlphaFlash feed currently covers over 150 global events each with up to 20 data points," says Clint Rhea, COO of Need to Know News in Chicago. "We're already beta testing our first Asian events, which will provide similar data to the other regions, but with local specifics. We try to offer everything at a macro level that moves the markets including inflation, growth, spending, labour markets, policy interest rates and much more.

"Our aim is to be the fastest feed in the race to zero latency, with competitive pricing and a tight development cycle to continuously adapt our feed to market needs," continues Rhea. "Every couple of weeks we add data points in response to trader requests and the publishers' release schedule. Since it's all tags and numbers, no text, it's highly compact and efficient, but easy to change. Our clients just ignore what they don't want."

> ■ "Scheduled events, like earnings statements, tend to have a very short lifecycle in customer algorithms."
>
> *Armando Gonzalez, president and CEO, RavenPack*

Like all low-latency strategies it is a never-ending battle. "We're constantly optimising and tuning the whole supply chain as faster routes or technologies become available," says Gross. "From the lock-up rooms in Washington or Europe or secure access points for embargoed data releases to the global trading hubs where our customers are waiting, it's a race every day for every release to save microseconds, and everything

keeps expanding: inputs; outputs and distribution networks to all the major financial centres."

Diverse applications

There may, however, be more to it than just high-frequency trading. "Getting precise economic data fast can also give us a context for evaluating the newswire services," says Victor Lebreton, COO at Quant Hedge, a Paris-based hedge fund. "That looks potentially interesting." At the other extreme, Quant Hedge may use pure statistical data about the news flow itself, and correlate that to market behaviours without really looking at individual stories. "There are clearly many different strategies to explore," says Lebreton.

"Along with other investment managers, today we trade using news analytics, focusing mainly on the more statistical metrics of newsflow," says Andrew Lo, founder and chief investment strategist of AlphaSimplex Group, a Massachusetts-based absolute-return specialist, owned by Natixis Global Asset Management. "We also license some of these analytics to Thomson Reuters."

Lo argues that since the best way to make money is

■ "News appears to be good at predicting both volumes and volatility, but not necessarily price levels."

Dan diBartolomeo, president, NorthField Information Services

not to lose it in the first place, AlphaSimplex focused first on applications for event risk management rather than alpha generation. He continues, "The question then becomes, how do we recognise abnormality? We start by defining what it means to be an outlier in terms of normal behaviour measured by word counts or the frequency of sentence patterns in the news flow. The outliers then

simply become patterns more than two standard deviations away from benchmarks established over the last three to six months."

AlphaSimplex has established 40 categories of statistical benchmarks covering political, military, macro-economic, commodity, weather and company-specific measures of normality to address the needs of different investors.

"Now we really could not do without it," says Lo. "Sudden spikes in these indices indicate that something real is happening. Global macro or distressed asset investors for example find them very useful."

Risk and reward

"Sentiment is far more interesting than economic figures, because it tells you something about trends as interest builds," says van Kleef at Lakeview CMS. "However, it's much more complex to dig up and interpret. So the more feeds you have the better, but everyone uses different metrics, which is a challenge. We're learning all the time what's useful and what isn't."

Thomson Reuters has been providing sentiment data with their News Analytics service for several years now. "News can be a

powerful indicator in predicting future trading volume and price volatility," says Richard Brown, global business manager for machine-readable news. "Incorporating the News Analytics data can increase that predictive power so that the algorithms respond more effectively to the most relevant, important, market moving news. For example, items with higher relevance scores or a more emotive tone, and those that are more unique, can better predict volume and volatility. And when used in a circuit breaker capacity, they can prevent the system from crying wolf too often, especially for those firms with higher amounts of news."

Mitra at OptiRisk agrees that overreacting to news items or alternatively underreacting if the algorithm shouts wolf too often can be a challenge. "So we need to include so-called wolf detectors or circuit breakers in the model both to recognise when trading a stock may be dangerous but also to avoid too many false positives," he argues. "Determining what is substantive however is never easy."

"When the system scores a news item, it can track how unique it is," says Brown at Thomson Reuters.

"In addition to noting whether it is an update to a previous story, it also takes a 'vocabulary fingerprint' of the article and compares it against other stories from the same and different news sources over various time periods. This information can be used to determine if other stories were similar and perhaps should be treated differently, depending on the strategy."

"In my experience, the Thomson Reuters sentiment indices are usually right, but the issue is timing since news typically follows most of the increase," says Leinweber. "So the effort now is to move closer to the primary sources, government filings, local papers or specialist journals. Instead of reading about a CEO's statement, the aim is to get the quote directly. Only first-hand news is really actionable."

"Some studies suggest that price changes accompanied by supporting news would better signal momentum," notes Brown, "while price movements without news are more likely to revert. With the News Analytics system, you can gain further confidence in the continuation of those movements. We're discovering many interesting correlations as we examine

■ "Only first-hand news is really actionable."

David Leinweber, Haas fellow in finance, University of California, Berkeley

the data in different ways for various strategies."

News sentiment metrics seem particularly good at analysing risk. "Ninety per cent of asset managers use third-party risk models to avoid internal bias," says Dan diBartolomeo, founder and president of NorthField Information Services, a supplier of risk management analytics. "Internal models are all about portfolio returns, not risk. The financial crisis is pushing us all to come up with new

solutions to deal with the new uncertainties."

diBartolomeo explains how investors measure portfolio risks by monitoring variance in returns. "When conditions change, we need many new observations to establish a new risk assessment, which is a slow process," he notes. "Implied volatility can also be taken from option prices, where they are stable, but most financial assets are not optionable, and options trade with frictions anyway, due to limited liquidity and trading costs. So in a period of instability, they offer an imperfect reflection of expectations, and we really need an alternative measure of risk."

Alternative metrics

NorthField's published studies suggest that news sentiment might give that alternative metric and even improve the predictability of implied volatilities where stock option data is also available.

"News appears to be good at predicting both volumes and volatility," says diBartolomeo, "but not necessarily price levels, because I never really know how others will react, how much is already baked into the price, or how future developments will unfold."

The sell-side is beginning to take these risk strategies on board. "We have also looked to use news-driven alerts to control the tolerance factor on the risks we take in VWAP algorithms; for example, how far ahead or behind the target participation rate we might allow," says Toms at Nomura. But cost remains a factor. "Latency mitigation is still delivering alpha so we continue to invest in the fastest platform, while for news cost is currently inhibiting wider usage," says Toms. "News redistribution is expensive. As costs come down, usage will no doubt spread."

OptiRisk's Mitra confirms that news usage in trading is likely to grow for both low- and high-latency strategies.

"Overall I would say that news data can give you advantages of speed, scope and better risk control than models driven purely by market data," he says. "Portfolio models essentially can take more factors into account and respond more reliably with this additional perspective to drive correlations. This enables a more diversified set of holdings and less crowded trades than might otherwise be possible." ■

Bob Giffords is an independent banking and technology analyst Bob.Giffords@btinternet.com