

Beating alpha decay in the sentiment arms race

News trading technology & trading returns

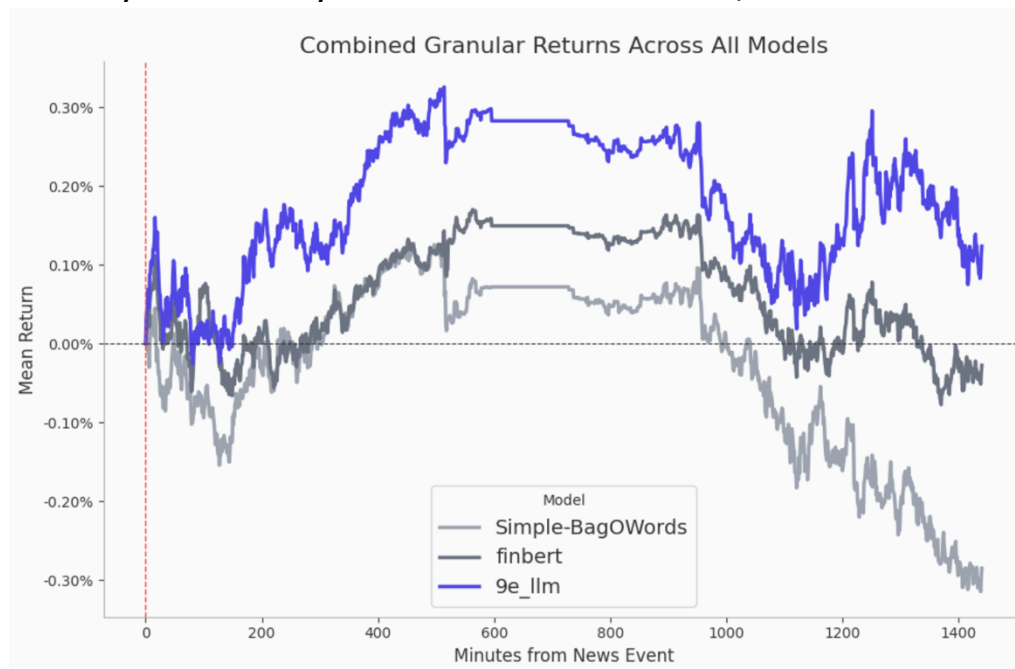
Working as a quant consultant on various news-based strategies I've always tried to stay ahead of the game when it comes to sentiment analysis models, I mean these are the beating heart and most critical component of a systematic news trading strategy. If the language model does a poor job identifying the news you want to trade (and the right direction) you may as well pack it up and call it a day.

Recently admiring the performance of our latest sentiment model I was inspired to do some background research into the relationship between technology and alpha decay in the space

How valuable is staying on the cutting edge of technology in trading? (spoiler - very)

How long do you have to upgrade when the paradigm shifts? (not long)

Event study – trade returns by sentiment model – the better the tech, the better the trade



* 2015-2025, minute by minute trade returns (long/short) based on sentiment signal sign. Methodology as detailed on page 2.

9epoch Quantitative Solutions

Nathan Szeitli

Monday, 16 June 2025



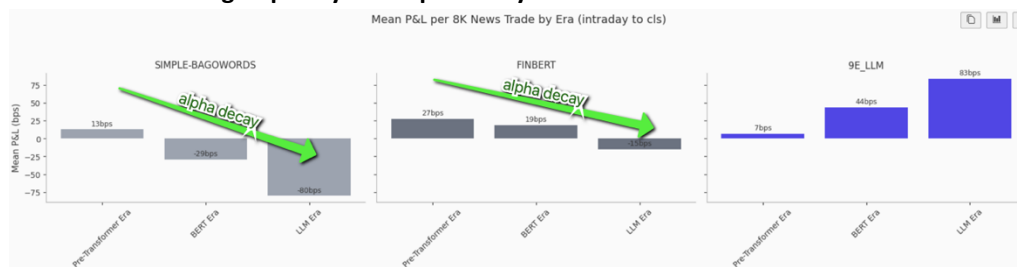
Hypothesis

- More advanced sentiment technology should drive better trading returns
- Older technology should show alpha decay after being superseded.

Experiment design

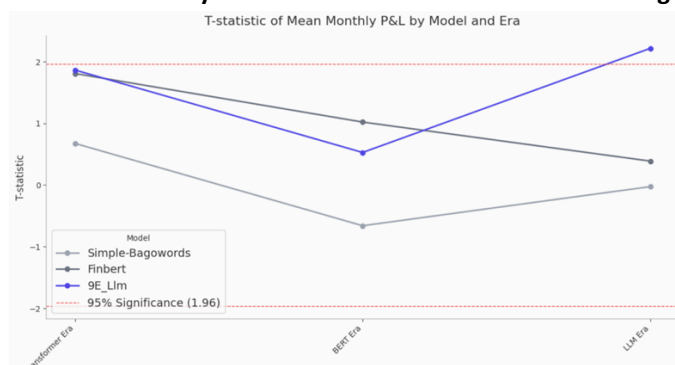
- 8k filings from EDGAR released intraday 2015-2025
- We analysed the company commentary text within the 8k filing for sentiment.
- Top 1000 US stock universe
- Three sentiment models from different tech paradigms
 1. Simple Bag-of-Words sentiment model from 2015 (VADER library)
 2. Financial document trained FinBERT model from 2018
 3. 9epoch LLM based sentiment engine
- Returns taken from 1 minute after 8k release to market close.
- Portfolio weighted by sentiment strength decile.
- Performance analysed with respect to the years the technology was dominant.

Mean trade returns grouped by era - alpha decay evident



* 2015-2025, mean trade returns (long/short) based on sentiment signal sign. Methodology as detailed on page 2.

T-stat of returns by model and era - newer tech showed stronger significance



* 2015-2025, monthly trade returns (long/short) based on sentiment signal sign. Methodology as detailed on page 2.

Mean returns summary table

Mean Returns by Model and Era						
Era	Model	Mean Long	Long Trades	Mean Short	Short Trades	Mean All
Pre-Transformer Era	Simple-Bagofwords	0.23%	202	-0.17%	79	0.12%
	Finbert	0.47%	210	-0.03%	196	0.23%
	9E_Llm	0.28%	148	0.19%	80	0.25%
BERT Era	Simple-Bagofwords	-0.25%	168	-0.35%	56	-0.27%
	Finbert	0.20%	187	-0.29%	164	-0.03%
	9E_Llm	0.03%	141	-0.13%	93	-0.03%
LLM Era	Simple-Bagofwords	0.03%	61	-1.08%	12	-0.15%
	Finbert	0.41%	53	-0.19%	40	0.16%
	9E_Llm	1.29%	42	1.56%	17	1.37%

* 2015-2025, mean trade returns (long/short) based on sentiment signal sign. Methodology as detailed on page 2.

THE PRE-BERT ERA (bad words make the number go down)

I remember back in the ZIRP years of 2015 and 2016 manually coding up scored dictionary based methods analysing the positive or negative sentiment from company news and this actually worked surprisingly well for the time however there were two big weaknesses in these simple models:

1. The inability to understand the relationship between words and sentences (ie interest rates 'DOWN' and 'FALLING' is positive for equities but profits 'DOWN' and 'FALLING' is negative)
2. The propensity of authors from all walks of life to linguistically spin the news in a way that suits their agenda but runs opposite to the actual fundamental impact on the stock.

THE BERT ERA

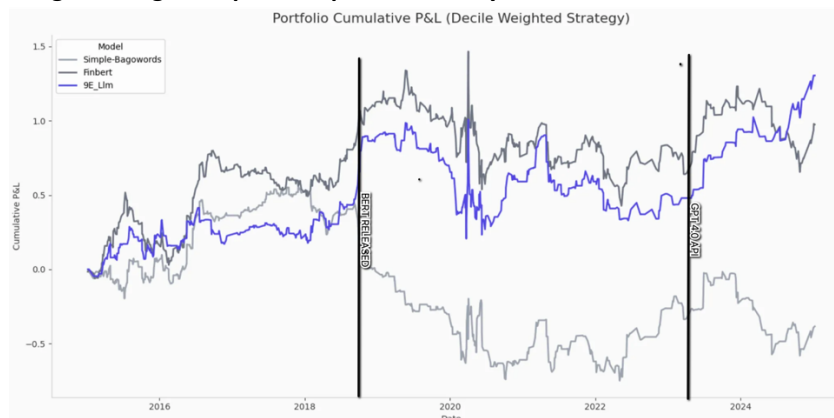
A big breakthrough came in late 2018 when Google released their BERT model which understood the relationship between words and could deliver a much more accurate and sophisticated sentiment analysis compared to scored dictionary models. This was a big technological upgrade and although it took some above average infrastructure to run properly (GPU's before they were big) BERT and the financially tuned FINBERT were significantly better at very quickly producing quantitative sentiment scores that were fundamentally correct about 90-95% of the time. As you can imagine the 5% it got wrong were some truly catastrophic shorts but overall the alpha was there especially in small caps where expert human analysis was sparser and slower.

THE LLM ERA (praise our robot overlords)

LLM's which became widely available in late 2022 are as you can imagine, significantly smarter than BERT models. BERT has x million neurons whereas an LLM has x billion neurons, it's like comparing the brain of an insect to a mammal. LLM's have a pretty in-depth understanding of many expert domains such as biology, geology and business which makes its comprehension of news and company filings extremely good (approaching human level).

The quantification of LLM output can be solved with quality prompt design and agentic logic. We have spent a several years refining our framework which is now battle tested and very good at scoring the fundamental impact of the news in multiple categories. Earnings change, investor sentiment and sector risk all gets boiled down into a single score for decision making or flow through into other models.

Weighted long-short portfolio performance by model



* 2015-2025, weighted mean trade returns (long/short) based on sentiment signal. Methodology as detailed on page 2.

Its not just the direction accuracy, it's the quality of the 'neutral' prediction.

A huge improvement of LLM's over BERT models is the ability to truly discard the bias and hype of the author and identify a genuine nothing-burger event. A glowing article about something relatively minor will generate a strong sentiment score from BERT but the LLM is unfazed. We can see evidence of this in the consistently high mean returns, lower number of trades and outstanding short returns from our news strategy dataset:

Complex domains are much better understood now, although the market can be slow to catch on

Going through individual examples where the LLM disagreed with the simpler models really illustrated the in-depth understanding of the new technology, for example this biotech release:

"On September 4, 2018, BioCryst Pharmaceuticals, Inc. (the "Company") announced initial results from the ZENITH-1 trial showing that a single 750 mg oral dose of BCX7353 was well tolerated and superior to placebo ($p < 0.05$) against the majority of efficacy endpoints evaluated in HAE patients suffering an acute attack. BCX7353 is a novel oral plasma kallikrein inhibitor being developed for both prophylactic and acute treatment of HAE attacks."

Sentiment Model Scores:

Simple: -92%

Finbert: -11.5%

LLM: +10.5% (correct)

This is a pretty indicative example, I read through dozens of these and they often involved litigation, geology, biotech or important releases without a clear business language explanation. Company-sensitive news where deep domain understanding is critical to sentiment evaluation is handled well by LLMs (and poorly by older models).

The market may react quickly to a "simple" evaluation of the news, but only a fundamentally correct interpretation will sustain the price move

It seems we are early - but the rate of change is accelerating

Technology	Time to deprecation
Bag-of-Words era	10+ years
BERT transformer era	4.5 years
Text LLM (3.5) → Multi-modal (4o) →	<3 years
Thinking LLM (o3) → Agentic swarms	

LLM technology is highly accessible and already making markets more efficient everywhere, but to use them systematically requires infrastructure and expert prompting to suit the use case and avoid pitfalls like the dreaded hallucination problem. This requires investment and expertise.

With the right approach we can significantly improve returns in any strategy that uses qualitative data as an input either by improving the information density of the features or good old fashioned operational efficiency. Anecdotally, it feels like we're still a few years away from this being commonplace across most buy-side firms



Disclaimer

9epoch Pty Ltd ("9epoch") (ABN: 78 637 136 844) provides quantitative solutions for financial markets and does not hold an Australian Financial Services Licence ("AFSL"). The information provided to you by 9epoch is factual information only. It does not contain any general advice or opinion. The information has been prepared without taking into account your objectives, financial situation or needs. It is made available "as is", without any representation or warranty, either express or implied, being given as to the accuracy, completeness, timeliness, reliability or otherwise of its content. No responsibility is accepted by us for any errors, omissions or any inaccurate information provided to you in this document. The material provided to you is not intended to amount to or provide, constitute or comprise, investment, financial, legal or other advice on any particular matter and no action should be taken or omitted to be taken in reliance on the information contained in this document. You should, before acting on any advice, seek financial advice from a suitably qualified professional in relation to any particular matter or circumstance.

To the fullest extent permitted by law, we expressly exclude, and accept no liability for any loss or damage which may be suffered by you, or any third party, and/or which may arise from the use, or the inability to use, this document or the use of, or any reliance placed on, any information contained in this document, including, but not limited to direct, indirect, consequential or other loss or damage, loss of data, loss of income, profits, contracts or opportunity, loss of goodwill or claims by third parties and whether caused by tort (including negligence), breach of contract or otherwise, even if such loss or damage was reasonably foreseeable by us.

The information in this document is not intended for distribution in or into, or for use by any person or entity in, any jurisdiction where such distribution or use would be contrary to any local law, order, directive or regulation. You should acquaint yourself with any local laws, order, directives or regulations that might apply to the use of this document.