# Expected Returns and Large Language Models

Yifei Chen,Bryan Kelly,Dacheng Xiu

解读：王梦涵

2025 年 4 月 18 日

# Motivation

- Finance literature has begun to extract info from economic text but limited
  - Only a limited portion of market-relevant textual data was used
  - Focus on a single specialized data source at a time
  - Text info is often represented in rudimentary ways
- There are valid reasons why text data has been underexploited
  - Language is an extremely nuanced info encoding scheme
  - $\rightarrow$ Highly complex models are necessary, yet prohibitive for many researchers

    *Technological barriers;High computational cost*
- Construct refined news text representations derived from LLMs and then use these to improve models of expected stock returns

# Why LLMs?

- Financial text mining is comprised of two general steps
  - **Step1:** Get numerical representation of the text data
  - **Step2:** Construct econometric model from numerical representation
- LLMs delegate **Step 1** to a handful of individuals globally who are best equipped to perform it
  - Trained on large-scale text datasets
  - Contain billions of parameters and support nonlinearity
  - Remarkable capacity for transfer learning
  - Accessible to non-specialist researchers

# Question

- Q1: How do LLMs & word_based models representations model stock returns?
    - They can draw useful info from text data
    - LLMs outperform word_based models
- Q2: What causes different performance between LLMs and word_based models?
    - LLMs can understand the context of article better
- Q3: How does the complexity of LLMs affect their performance?
    - Larger models perform better, but with diminishing returns
- Q4:Do above results hold on Globally?
    - Yes

# Contribution

- Literature on financial text-data analysis and return prediction
  - Systematically leveraging various LLMs to analyze text data
  - Showcase the advantages of LLM representations for modeling stock returns
  - Extend the financial text-data analysis to international text
- Literature on efficient market hypothesis
  - Provide novel insights and empirical evidence on deviations from the EMH

# Data

- News Text: From Refinitiv, 1996.1-2019.6
  - Types: Articles-a headline and a body of text; Alerts-only a headline
  - Filters
    - Retained news associated with a single stock for which three-day close-to close returns are available($t-1, t, t+1$)
    - Removed excessively short/detailed news($N_{characters} < 100\& > 100000$)
    - Remove redundant articles
- Equity data:
  - US:CRSP
  - International: EIKON

# Text Mining Methods:LLMs

- LLMs: BERT (developed by Google), RoBERTa (by Meta), LLaMA(LLaMA2) (by Meta),ChatGPT
- Model training steps:
  - Tokenization:word ->token: characters,words or **subwords**
    eg:surreptitious -> 'sur', '##re', '##pt', '##iti', '##ous'
  - Transformer:token->contextualized embedding
- The LLMs used in this paper are all **pre-trained**
  - Minimizes computational efforts
  - Easier to replicate
- Article $\rightarrow X_{N_{max-tokens}, N_{vector-dimensions}} \rightarrow$ Article-level representations $x_{i,t}$

# Text Mining Methods:word_based methods

- Word Embeddings:Word2Vec
  - Train: Parse→ Exclued nums、stopwords,...→Learn word from fixed-size window→ word embeddings(300d)
  - Use **pre-trained** model from fastText
- Bags of words
  - SESTM: Only for sentiment analysis, **need to be trained**
    - Identify a list of terms most closely correlated with sentiment
    - Assign weights to these words by estimating a topic model
    - Aggregate these terms into an article-level sentiment score
- LLMD: Only for sentiment analysis

# Design: Econometric models

- **Sentiment Analysis**:Classification Problem

$$E(y_{i,t}|x_{i,t}) = \sigma(x'_{i,t}\beta)$$

  - $x_{i,t}$:text_based features;$y_{i,t}$:binary var,pos or neg;$\sigma(\cdot) = \frac{e^x}{1+e^x}$
  - Sentiment label: the sign of three-day stock's return

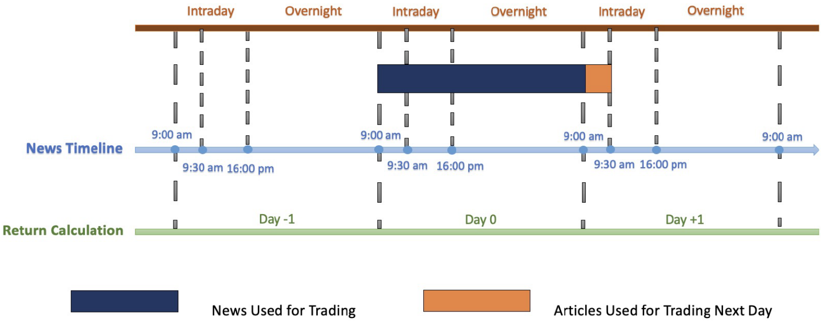
- **Return Prediction**: Regression Problem

$$E(r_{i,t+1}|x_{i,t}) = x'_{i,t}\theta$$

  - $x_{i,t}$:text_based features;$r_{i,t+1}$:stock return
  - Use simple panel regression to emphasize the significance of text-based representations

# Design: Construct Portfolio

- Exclude articles published between 9:00am and 9:30am
- For news occur on day 0, build positions on day 1



# Q1: Model stock returns–Sentiment Analysis

| | ChatGPT | | | | | | LLaMA2 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.34 | -0.14 | 0.48 | 0.19 | 0.04 | 0.15 | 0.35 | -0.10 | 0.45 | 0.18 | 0.07 | 0.11 |
| Std | 0.20 | 0.22 | 0.10 | 0.19 | 0.22 | 0.11 | 0.20 | 0.23 | 0.11 | 0.19 | 0.22 | 0.11 |
| SR | 1.71 | -0.62 | 4.62 | 1.03 | 0.18 | 1.41 | 1.75 | -0.43 | 4.16 | 0.97 | 0.33 | 0.98 |
| | LLaMA | | | | | | RoBERTa | | | | | |
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.34 | -0.07 | 0.41 | 0.19 | 0.08 | 0.11 | 0.33 | -0.06 | 0.39 | 0.20 | 0.09 | 0.11 |
| Std | 0.20 | 0.23 | 0.11 | 0.19 | 0.22 | 0.11 | 0.20 | 0.22 | 0.10 | 0.19 | 0.22 | 0.11 |
| SR | 1.67 | -0.33 | 3.89 | 1.02 | 0.36 | 1.04 | 1.62 | -0.29 | 3.75 | 1.08 | 0.43 | 0.94 |
| | BERT | | | | | | Word2vec | | | | | |
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.32 | -0.04 | 0.36 | 0.16 | 0.07 | 0.10 | 0.29 | -0.01 | 0.30 | 0.18 | 0.08 | 0.09 |
| Std | 0.20 | 0.22 | 0.10 | 0.18 | 0.21 | 0.10 | 0.21 | 0.22 | 0.10 | 0.19 | 0.21 | 0.10 |
| SR | 1.59 | -0.19 | 3.60 | 0.89 | 0.31 | 0.92 | 1.41 | -0.05 | 3.06 | 0.93 | 0.40 | 0.92 |
| | SESTM | | | | | | LMMD | | | | | |
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.31 | -0.03 | 0.34 | 0.18 | 0.09 | 0.09 | 0.24 | 0.01 | 0.22 | 0.14 | 0.10 | 0.04 |
| Std | 0.20 | 0.22 | 0.10 | 0.19 | 0.21 | 0.11 | 0.20 | 0.23 | 0.10 | 0.18 | 0.21 | 0.10 |
| SR | 1.53 | -0.14 | 3.43 | 0.97 | 0.42 | 0.86 | 1.18 | 0.06 | 2.29 | 0.77 | 0.47 | 0.39 |

- LLMs outperforms word_based models; EW>VW; Long > Short

# Q1: Model stock returns–Return Prediction

| | ChatGPT | | | | | | LLaMA2 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.39 | -0.04 | 0.43 | 0.22 | 0.11 | 0.10 | 0.46 | -0.11 | 0.57 | 0.23 | 0.09 | 0.14 |
| Std | 0.21 | 0.22 | 0.10 | 0.20 | 0.20 | 0.11 | 0.21 | 0.22 | 0.11 | 0.20 | 0.20 | 0.11 |
| SR | 1.87 | -0.21 | 4.23 | 1.07 | 0.58 | 0.91 | 2.22 | -0.50 | 5.31 | 1.14 | 0.44 | 1.32 |
| | | | 4.62 | | | 1.41 | | | 4.16 | | | 0.98 |
| | LLaMA | | | | | | RoBERTa | | | | | |
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.45 | -0.11 | 0.56 | 0.23 | 0.08 | 0.15 | 0.37 | -0.05 | 0.42 | 0.21 | 0.08 | 0.13 |
| Std | 0.21 | 0.22 | 0.11 | 0.20 | 0.20 | 0.11 | 0.20 | 0.21 | 0.10 | 0.20 | 0.20 | 0.11 |
| SR | 2.17 | -0.51 | 5.17 | 1.12 | 0.41 | 1.35 | 1.81 | -0.23 | 4.36 | 1.06 | 0.42 | 1.17 |
| | | | 3.89 | | | 1.04 | | | 3.75 | | | 0.94 |
| | BERT | | | | | | Word2vec | | | | | |
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.33 | -0.03 | 0.37 | 0.16 | 0.09 | 0.07 | 0.33 | -0.00 | 0.33 | 0.20 | 0.13 | 0.08 |
| Std | 0.21 | 0.22 | 0.09 | 0.19 | 0.20 | 0.10 | 0.21 | 0.22 | 0.10 | 0.19 | 0.20 | 0.10 |
| SR | 1.62 | -0.15 | 3.94 | 0.85 | 0.45 | 0.68 | 1.61 | -0.00 | 3.20 | 1.06 | 0.63 | 0.74 |
| | | | 3.60 | | | 0.92 | | | 3.06 | | | 0.92 |

- LLaMA2, LLaMA and RoBERTa outperform their sentiment portfolio.

# Q1: Model stock returns–News Momentum

| | Day +1 Portfolios | | | | | | Day +2 Portfolios | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.35 | -0.10 | 0.45 | 0.18 | 0.07 | 0.11 | 0.17 | 0.05 | 0.12 | 0.06 | 0.06 | 0.00 |
| Std | 0.20 | 0.23 | 0.11 | 0.19 | 0.22 | 0.11 | 0.20 | 0.22 | 0.10 | 0.19 | 0.21 | 0.11 |
| SR | 1.75 | -0.43 | 4.16 | 0.97 | 0.33 | 0.98 | 0.86 | 0.22 | 1.26 | 0.30 | 0.26 | 0.02 |

| | Day +3 Portfolios | | | | | | Day +4 Portfolios | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.16 | 0.09 | 0.07 | 0.06 | 0.08 | -0.01 | 0.12 | 0.10 | 0.02 | 0.09 | 0.08 | 0.01 |
| Std | 0.20 | 0.22 | 0.09 | 0.19 | 0.22 | 0.11 | 0.20 | 0.22 | 0.09 | 0.20 | 0.21 | 0.11 |
| SR | 0.80 | 0.39 | 0.81 | 0.34 | 0.36 | -0.12 | 0.62 | 0.48 | 0.21 | 0.46 | 0.40 | 0.06 |

| | Day +5 Portfolios | | | | | | Day +6 Portfolios | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.15 | 0.16 | -0.01 | 0.11 | 0.05 | 0.06 | 0.11 | 0.13 | -0.02 | 0.08 | 0.05 | 0.02 |
| Std | 0.20 | 0.22 | 0.09 | 0.19 | 0.21 | 0.10 | 0.20 | 0.22 | 0.09 | 0.19 | 0.20 | 0.10 |
| SR | 0.75 | 0.72 | -0.12 | 0.56 | 0.25 | 0.54 | 0.54 | 0.58 | -0.22 | 0.39 | 0.25 | 0.24 |

- A significant momentum effect–> market inefficiency

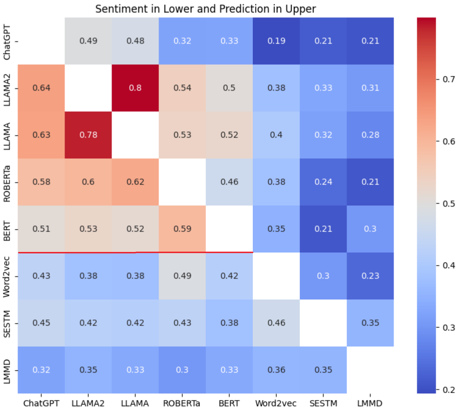
# Q1: Model stock returns–Text-info VS Past-Returns

| | Stocks with news | | | | | | ChatGPT | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.35 | 0.06 | 0.29 | 0.29 | 0.08 | 0.20 | 0.38 | -0.16 | 0.54 | 0.22 | 0.04 | 0.18 |
| Std | 0.27 | 0.23 | 0.18 | 0.30 | 0.23 | 0.24 | 0.21 | 0.22 | 0.11 | 0.19 | 0.22 | 0.11 |
| SR | 1.29 | 0.25 | $\boxed{1.58}$ | 0.95 | 0.37 | 0.83 | 1.86 | -0.71 | $\boxed{5.03}$ | 1.13 | 0.20 | 1.58 |
| | LLaMA2 | | | | | | LLaMA | | | | | |
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.40 | -0.12 | 0.52 | 0.21 | 0.07 | 0.14 | 0.37 | -0.08 | 0.45 | 0.21 | 0.09 | 0.12 |
| Std | 0.21 | 0.23 | 0.12 | 0.20 | 0.21 | 0.12 | 0.21 | 0.22 | 0.11 | 0.20 | 0.22 | 0.12 |
| SR | 1.88 | -0.53 | $\boxed{4.43}$ | 1.06 | 0.34 | 1.19 | 1.79 | -0.34 | $\boxed{4.00}$ | 1.06 | 0.42 | 1.00 |
| | RoBERTa | | | | | | BERT | | | | | |
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.39 | -0.11 | 0.49 | 0.24 | 0.08 | 0.16 | 0.38 | -0.07 | 0.44 | 0.20 | 0.06 | 0.14 |
| Std | 0.21 | 0.22 | 0.11 | 0.20 | 0.22 | 0.12 | 0.21 | 0.22 | 0.11 | 0.19 | 0.21 | 0.11 |
| SR | 1.84 | -0.48 | $\boxed{4.46}$ | 1.22 | 0.36 | 1.38 | 1.80 | -0.31 | $\boxed{4.13}$ | 1.01 | 0.28 | 1.22 |
| | SESTM | | | | | | LMMD | | | | | |
| | EW | | | VW | | | EW | | | VW | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.42 | -0.03 | 0.45 | 0.30 | 0.05 | 0.26 | 0.26 | -0.01 | 0.28 | 0.17 | 0.09 | 0.07 |
| Std | 0.26 | 0.22 | 0.17 | 0.27 | 0.23 | 0.21 | 0.20 | 0.22 | 0.10 | 0.18 | 0.21 | 0.10 |
| SR | 1.63 | -0.12 | $\boxed{2.61}$ | 1.10 | 0.21 | 1.21 | 1.30 | -0.06 | $\boxed{2.79}$ | 0.90 | 0.44 | 0.71 |

- Text data can provide more information than past returns

# Q2: LLMs VS Word_based models

- Analyze the corr of daily portfolio returns between different models



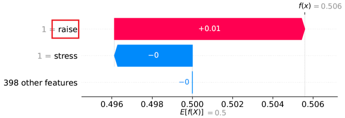

- Correlation between LLMs and word-based models is notably lower

# Q2: LLMs VS Word_based models–An Example

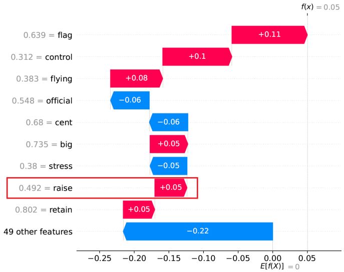
- SHAP: expected change in prediction based on specific features

Brussels has warned British Airways owner IAG ICAG.L that its favoured strategy to allow it to continue flying freely in and around Europe in the event of a nodeal Brexit will not work, the Financial Times reported on Tuesday.After Brexit, European carriers will have to show they are more than 50 per cent EUowned and controlled to retain flying rights in the bloc, the FT said.IAG, which also owns the Spanish flag carrier Iberia, is registered in Spain but headquartered in the United Kingdom and has diverse global shareholders.The FT said part of IAG's strategy to retain both EU and UK operating rights is to stress that its important individual airlines are domestically owned through a series of trusts rather than being part of the bigger a high proportion of nonEU investors.The FT quoted an unnamed senior EU official as saying, "For IAG, I can't see how it can be a solution."Concerns have been raised with IAG over its postBrexit ownership structure, the FT quoted a second Brussels official familiar with the conversations as saying.IAG was not immediately available.

BOW




W2V




- Analyze the context surrounding sentiment words is important

# Q2: LLMs VS Word_based models–Impact of Negation Words

$$r_{h,i,t+1}(s_{i,t}^{LLM} - s_{i,t}^{\text{word-based model}}) = \alpha + \beta \text{Negation}_{i,t} + \text{Fixed Effect} + \text{Control Variables}_{i,t} + \epsilon_{i,t}$$

| | SESTM | | | | Word2Vec | | | |
|---|---|---|---|---|---|---|---|---|
| | LLaMA2 | LLaMA | ROBERTa | BERT | LLaMA2 | LLaMA | ROBERTa | BERT |
| neg words count | 0.0137*** | 0.0134*** | 0.0077* | 0.0076* | 0.0189*** | 0.0186*** | 0.0129*** | 0.0128*** |
| | (0.0046) | (0.0047) | (0.0046) | (0.0045) | (0.0045) | (0.0046) | (0.0042) | (0.0041) |
| size | -0.0890*** | -0.0793*** | -0.0667*** | -0.0888*** | -0.0620*** | -0.0523*** | -0.0397*** | -0.0617*** |
| | (0.0169) | (0.0171) | (0.0168) | (0.0166) | (0.0166) | (0.0168) | (0.0152) | (0.0150) |
| BM | 0.0046 | 0.0061 | -0.0030 | -0.0012 | 0.0020 | 0.0035 | -0.0055 | -0.0038 |
| | (0.0075) | (0.0076) | (0.0075) | (0.0074) | (0.0074) | (0.0075) | (0.0068) | (0.0067) |
| liquidity | 0.0420*** | 0.0423*** | 0.0355*** | 0.0272*** | 0.0630*** | 0.0633*** | 0.0565*** | 0.0482*** |
| | (0.0105) | (0.0106) | (0.0104) | (0.0103) | (0.0104) | (0.0105) | (0.0095) | (0.0094) |
| IdioRisk | 0.0173*** | 0.0114* | 0.0204*** | 0.0086 | 0.0381*** | 0.0322*** | 0.0412*** | 0.0294*** |
| | (0.0065) | (0.0065) | (0.0064) | (0.0064) | (0.0064) | (0.0065) | (0.0058) | (0.0058) |
| sic2D | -0.0393*** | -0.0268* | -0.0328** | -0.0114 | -0.0376** | -0.0250* | -0.0311** | -0.0097 |
| | (0.0151) | (0.0152) | (0.0149) | (0.0147) | (0.0148) | (0.0150) | (0.0135) | (0.0133) |
| Constant | 0.0259*** | 0.0207*** | 0.0191*** | 0.0221*** | 0.0186*** | 0.0134** | 0.0118** | 0.0148*** |
| | (0.0053) | (0.0054) | (0.0053) | (0.0052) | (0.0052) | (0.0053) | (0.0048) | (0.0047) |
| Stock FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Date FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of obs | 1,552,769 | 1,552,769 | 1,552,769 | 1,552,769 | 1,552,769 | 1,552,769 | 1,552,769 | 1,552,769 |
| Adj R-squared | 0.0029 | 0.0035 | 0.0046 | 0.0047 | 0.0048 | 0.0036 | 0.0032 | 0.0030 |

# Q2: LLMs VS Word_based models–Impact of Context Complexity

- Model performance based on the headline and body of the same article

| | ChatGPT | | | | | | LLaMA2 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Headline | | | Body | | | Headline | | | Body | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.33 | -0.09 | 0.42 | 0.34 | -0.14 | 0.48 | 0.32 | -0.04 | 0.36 | 0.35 | -0.10 | 0.45 |
| Std | 0.20 | 0.22 | 0.10 | 0.20 | 0.22 | 0.10 | 0.20 | 0.22 | 0.10 | 0.20 | 0.23 | 0.11 |
| SR | 1.65 | -0.41 | 4.12 | 1.71 | -0.62 | 4.62 | 1.58 | -0.19 | 3.51 | 1.75 | -0.43 | 4.16 |

| | LLaMA | | | | | | RoBERTa | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Headline | | | Body | | | Headline | | | Body | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.33 | -0.02 | 0.35 | 0.34 | -0.07 | 0.41 | 0.35 | -0.00 | 0.35 | 0.33 | -0.06 | 0.39 |
| Std | 0.20 | 0.22 | 0.10 | 0.20 | 0.23 | 0.11 | 0.21 | 0.22 | 0.10 | 0.20 | 0.22 | 0.10 |
| SR | 1.62 | -0.11 | 3.51 | 1.67 | -0.33 | 3.89 | 1.69 | -0.02 | 3.47 | 1.62 | -0.29 | 3.75 |

| | BERT | | | | | | Word2vec | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Headline | | | Body | | | Headline | | | Body | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.31 | -0.02 | 0.33 | 0.32 | -0.04 | 0.36 | 0.35 | -0.01 | 0.37 | 0.29 | -0.01 | 0.30 |
| Std | 0.20 | 0.21 | 0.10 | 0.20 | 0.22 | 0.10 | 0.21 | 0.23 | 0.11 | 0.21 | 0.22 | 0.10 |
| SR | 1.55 | -0.09 | 3.48 | 1.59 | -0.19 | 3.60 | 1.70 | -0.06 | 3.32 | 1.41 | -0.05 | 3.06 |

| | SESTM | | | | | | LMMD | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Headline | | | Body | | | Headline | | | Body | | |
| | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S | Long | Short | L-S |
| Ret | 0.38 | -0.02 | 0.40 | 0.31 | -0.03 | 0.34 | 0.17 | -0.09 | 0.25 | 0.24 | 0.01 | 0.22 |
| Std | 0.21 | 0.22 | 0.10 | 0.20 | 0.22 | 0.10 | 0.23 | 0.25 | 0.15 | 0.20 | 0.23 | 0.10 |
| SR | 1.84 | -0.07 | 4.05 | 1.53 | -0.14 | 3.43 | 0.74 | -0.35 | 1.66 | 1.18 | 0.06 | 2.29 |

# Q3: The Virtue of Complexity

| | EW | | | | | VW | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Article | Alert | | | | Article | Alert | | | |
| | | All | TS1 | TS2 | Rest | | All | TS1 | TS2 | Rest |
| **Portfolios based on Sentiment Analysis** | | | | | | | | | | |
| LLAMA7B | 3.93 | 5.40 | 5.27 | 3.45 | 2.68 | 1.12 | 2.39 | 2.46 | 0.52 | 0.79 |
| LLAMA13B | 3.89 | 5.48 | 5.32 | 3.58 | 2.93 | 1.04 | 2.39 | 2.42 | 0.60 | 0.94 |
| LLAMA33B | 3.54 | 5.51 | 4.77 | 3.60 | 3.02 | 1.03 | 2.29 | 2.39 | 0.64 | 0.56 |
| LLAMA65B | 2.73 | 4.63 | 4.03 | 2.80 | 2.07 | 0.43 | 1.99 | 1.95 | 0.87 | 0.31 |
| LLAMA2_7B | 4.07 | 5.29 | 5.42 | 3.60 | 3.16 | 1.06 | 2.41 | 2.87 | 0.48 | 0.61 |
| LLAMA2_13B | 4.16 | 5.77 | 5.60 | 3.92 | 3.20 | 0.98 | 2.54 | 2.74 | 0.55 | 0.91 |
| LLAMA2_70B | 4.24 | 5.95 | 5.63 | 3.78 | 3.60 | 1.14 | 2.37 | 2.75 | 0.57 | 0.73 |
| **Portfolios based on Return Prediction** | | | | | | | | | | |
| LLAMA7B | 4.90 | 4.13 | 4.08 | 2.51 | 1.96 | 1.00 | 1.47 | 1.37 | 0.93 | 0.95 |
| LLAMA13B | 5.17 | 4.52 | 4.53 | 3.33 | 2.23 | 1.35 | 1.87 | 1.95 | 1.38 | 1.05 |
| LLAMA33B | 4.37 | 3.78 | 3.14 | 2.50 | 2.22 | 0.86 | 1.37 | 1.01 | 0.66 | 0.83 |
| LLAMA65B | 3.05 | 2.18 | 1.96 | 1.34 | 1.17 | 0.82 | 0.92 | 0.59 | 0.59 | 0.43 |
| LLAMA2_7B | 5.15 | 4.03 | 4.18 | 3.06 | 2.36 | 1.05 | 1.72 | 1.61 | 1.19 | 1.16 |
| LLAMA2_13B | 5.31 | 3.99 | 3.75 | 2.63 | 2.63 | 1.32 | 1.17 | 1.38 | 1.02 | 1.04 |
| LLAMA2_70B | 4.61 | 3.91 | 3.75 | 3.00 | 2.54 | 1.07 | 0.96 | 1.10 | 0.66 | 0.67 |

- More complex models perform better, but with diminishing marginal returns.

# Q4: Global Evidence

| | LLaMA2 | | LLaMA | | RoBERTa | | BERT | | Word2vec | | SESTM | | LMMD | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | EW | VW | EW | VW | EW | VW | EW | VW | EW | VW | EW | VW | EW | VW |
| US | 5.31 | 1.32 | 5.17 | 1.35 | 4.36 | 1.17 | 3.94 | 0.68 | 3.20 | 0.74 | 3.43 | 0.86 | 2.29 | 0.41 |
| UK | 3.10 | 1.64 | 2.96 | 1.34 | 2.30 | 0.60 | 2.19 | 1.22 | 2.04 | 0.81 | 2.05 | 0.73 | 0.82 | 0.34 |
| Australia | 0.30 | 0.25 | -0.02 | 0.02 | 0.04 | -0.01 | 0.21 | 0.07 | 0.01 | -0.02 | -0.16 | -0.11 | 0.37 | 0.02 |
| Canada | 2.01 | 1.27 | 2.07 | 0.76 | 1.64 | 0.60 | 1.92 | 0.89 | 1.49 | 0.79 | 0.62 | 0.33 | 0.76 | 0.37 |
| China (HK) | 1.05 | 0.54 | 1.37 | 1.07 | 0.76 | 0.55 | 1.05 | 0.87 | 0.46 | 0.33 | 1.03 | 0.76 | | |
| Japan | 1.52 | 0.56 | 1.29 | 0.65 | 0.87 | 0.39 | 1.09 | 0.54 | 0.68 | 0.45 | -0.54 | -0.29 | | |
| Germany | 1.31 | 0.63 | 1.18 | 0.40 | 0.51 | 0.23 | 0.63 | 0.34 | 0.47 | 0.20 | 0.92 | 0.70 | | |
| Italy | 0.39 | 0.08 | 0.38 | 0.14 | 0.55 | 0.13 | 0.63 | 0.06 | 0.39 | -0.04 | 0.12 | 0.21 | | |
| France | 1.49 | 0.63 | 1.09 | 0.67 | 0.79 | 0.40 | 1.35 | 0.72 | 0.74 | 0.19 | 1.06 | 0.14 | | |
| Sweden | 1.27 | 0.76 | 1.18 | 0.67 | 0.95 | 0.58 | 0.89 | 0.21 | 0.57 | 0.59 | 0.01 | 0.53 | | |
| Denmark | 0.16 | 0.02 | 0.04 | -0.06 | 0.58 | 0.49 | 0.58 | 0.53 | 0.44 | 0.31 | -0.01 | -0.16 | | |
| Spain | -0.17 | -0.15 | -0.11 | 0.05 | 0.08 | 0.02 | 0.03 | 0.07 | -0.02 | 0.14 | -0.26 | -0.43 | | |
| Finland | 0.35 | 0.09 | 0.23 | 0.01 | -0.06 | -0.21 | 0.01 | 0.01 | 0.11 | 0.11 | 0.18 | -0.06 | | |
| Portugal | -1.99 | -2.00 | -0.61 | -0.62 | 0.33 | 0.34 | -1.01 | -1.04 | 0.29 | 0.29 | 3.88 | 3.86 | | |
| Greece | -0.39 | -0.39 | 0.85 | 0.85 | 1.82 | 1.82 | 0.02 | 0.02 | -2.14 | -2.14 | 0.12 | 0.12 | | |
| Netherlands | -0.55 | -0.55 | -0.14 | -0.14 | 1.60 | 1.60 | 0.53 | 0.53 | -0.36 | -0.36 | 1.14 | 1.14 | | |
| Mean | 0.95 | 0.29 | 1.06 | 0.45 | 1.07 | 0.54 | 0.88 | 0.36 | 0.52 | 0.15 | 0.85 | 0.52 | 1.06 | 0.29 |
| Mean (Excluding US) | 0.66 | 0.23 | 0.79 | 0.39 | 0.85 | 0.50 | 0.67 | 0.34 | 0.34 | 0.11 | 0.68 | 0.50 | 0.65 | 0.24 |
| Median (Excluding US) | 0.39 | 0.25 | 0.85 | 0.40 | 0.76 | 0.40 | 0.63 | 0.34 | 0.44 | 0.20 | 0.18 | 0.21 | 0.76 | 0.34 |

- LLMs outperform word_based models(except:Portugal, Greece,Netherlands)

# New ideas

- Current focus is limited to news text; future work could integrate other text sources
  - Social Media...
- Use international news to model firm i's return
- Link text representations with investors' action