Modeling and Detecting Company Risks from News

Jiaxin Pei[†*]   **Soumya Vadlamannati**[‡]   **Liang-Kang Huang**[‡]
Daniel Preoţiuc-Pietro[‡]   **Xinyu Hua**[‡]
[†]University of Michigan, Ann Arbor, MI, USA
[‡]Bloomberg, New York, NY, USA
pedropei@umich.edu
{svadlamanna1, lhuang214, dpreotiucpie, xhua22}@bloomberg.net

Abstract

Identifying risks associated with a company is important to investors and the wellbeing of the overall financial markets. In this study, we build a computational framework to automatically extract company risk factors from news articles. Our newly proposed schema comprises seven distinct aspects, such as supply chain, regulations, and competition. We annotate 666 news articles and benchmark various machine learning models. While large language models have achieved remarkable progress in various types of NLP tasks, our experiment shows that zero-shot and few-shot prompting state-of-the-art LLMs (e.g., Llama-2) can only achieve moderate to low performances in identifying risk factors. In contrast, fine-tuning pre-trained language models yields better results on most risk factors. Using this model, we analyze over 277K Bloomberg News articles and demonstrate that identifying risk factors from news could provide extensive insights into the operations of companies and industries.

1 Introduction

Risks are inherent and pervasive within companies' operations and our society (Stephany et al., 2022; Rausand, 2013; Albuquerque et al., 2019). Understanding and identifying corporate risk factors could benefit diverse stakeholders, including investors, regulators, and other relevant entities. Notably, publicly listed companies are mandated to disclose their risk factors, as these can inform shareholders and the public when making financial decisions (Beretta and Bozzolan, 2004). NLP models are also built to automatically extract company-related risk factors from public filings, providing consolidated and accessible insights for analysts to fathom and integrate these factors (Kogan et al., 2009; Yang et al., 2018).

While company filings offer a systematic view of company-related risks, they are beset by three

---

* Work done as an intern at Bloomberg.

principal issues: (1) Limited frequency—owing to mandatory quarterly reporting, risk analysis is confined to three-month intervals, disregarding the reality of swift, even daily, alterations in a company's risk profile. (2) Subjectivity—authored by internal personnel, company filings might inadvertently omit pivotal risk factors due to vested interests (Masson and Montariol, 2020; Klingebiel, 2018). (3) Bias towards public entities—only publicly listed companies are obligated to divulge risks through filings, neglecting the imperative to comprehend risks associated with private companies, which may be particularly pertinent when engaged in financial activities such as bond issuance (Abdel-Khalik, 1993; Vanstraelen and Schelleman, 2017).

To redress these limitations, we propose to model company-related risk factors from news articles. News articles offer the following merits in analyzing company-related risk factors: (1) High frequency—news updates occur in real time, providing a dynamic information stream conducive to measuring companies' risk factors in the ever-evolving market. (2) External perspective—news articles, devoid of company affiliations, proffer diverse viewpoints, shedding light on risk factors from external vantage points. (3) Coverage over both public and private companies—news articles encapsulate both publicly listed and private companies, thereby bridging the information gap present in public filings.

While there are existing studies on modeling risk factors, they typically focus on company filings (Zhu et al., 2016; Kravet and Muslu, 2013) and their categorization may not be directly applicable to news data. Combining existing literature and our manual examination of hundreds of news articles, we propose a novel theoretical framework to analyze company risk factors in news. Our taxonomy encompasses seven categories of risk factors: Supply Chain and Product, People and Management, Finance, Legal and

`Regulations`, `Macro`, `Competition`, and `Markets and Consumers`.

We annotate 666 news articles from Bloomberg News[1] and benchmark a series of models from feature-engineered baselines to prompting large language models (LLMs). Despite their impressive results on various other NLP tasks, LLMs perform worse than smaller transformer models (e.g., RoBERTa) fine-tuned on in-domain data. By applying the best-performing model to a large sample of 277, 112 news articles, we analyze the risk factor across companies in given industries, and across the entire macroeconomy. Our analysis shows that modeling company-related risk factors could reveal important signals of not only companies' operations but also can be used as the indicator of the macro-level risk for society.

2 Related Work

2.1 Risk Factors in Finance Domain

In finance and corporate operations, risks refer to the factors that may harm the company or may cause it to fail (Rausand, 2013). Existing research has identified many types of risk factors for companies, including financial risk (Malz, 2011; Fujii et al., 2022), credit risk (Kao, 2000), policy risk (Blyth et al., 2007), macro risk (Hiang Liow et al., 2006), operational risk (Fujii et al., 2022) and competition risk (Raith, 2003). In many countries, regulators require publicly listed companies to disclose risk factors in their quarterly and annual reports to inform the investors (Weil et al., 2006). Another line of research relevant to risk factors is financial and economic uncertainty (Moore, 2017). While both are forward-looking, risks specifically connotate those factors that may negatively affect the operations and market value of a company.

2.2 Risk Identification as an NLP Task

Natural Language Processing methods have long been used to analyze text documents in the finance domain (Loughran and McDonald, 2020), such as news reports (Day and Lee, 2016), social media posts (Souza et al., 2015), and company filings (Wang et al., 2013). Existing studies primarily focus on company filings (Wang et al., 2013; Yang et al., 2018; Kogan et al., 2009), which are issued by companies themselves and are subject to limited frequency. While researchers have also extracted risk factors from news articles (Lu et al., 2009;

Bhadani et al., 2020), these either focus only on specific types of risks (Bhadani et al., 2020) or only tried to identify the relevant claims (Lu et al., 2009) instead of providing a holistic view of the risks.

3 Theoretical Framework

In this study, we focus on modeling company-level risk factors in news articles. We survey existing literature (§ 2.1) and qualitatively examine hundreds of news articles. Our proposed framework is summarized as below.

Supply Chain and Product Risks associated with the company's supply chain, manufacturing, product or core technology. For example, "*Yum China Faces Challenges with Chicken Prices*" indicates risks regarding the supply chain, as chicken is an important ingredient in Yum China's products.

People and Management Risks regarding a company's internal operations such as layoffs, departures of top management, or specific operation strategies. For example, "*Tesla Pauses Hiring, Musk Says Need to Cut Staff by 10%.*"

Finance Risks related to the finances of a company such as cash flow, fund procurement, investments, and profits. For example, "*NIO Shares Soar as Loss Shrinks, Though Cash Concerns Linger.*"

Legal and Regulations Risks induced by potential policy changes, pressure from regulations or lawsuits. For example, "*Maple Leaf Plunges as China's Hog Suspension Impacts Profits.*"

Macro Risks caused by the macro socio-economic environment such as inflation, pandemics or a financial crisis. For example, "*Absa Drops on Profit Miss as South African Economy Struggles.*"

Markets and Consumers Risks or challenges from the market or consumer sales. For example, "*Hong Kong Protests Cut Demand for Hilton, Hyatt Hotel Rooms*" suggests that the demand for hotel space is shrinking, which indicates `Markets and Consumer` risk for both Hilton and Hyatt.

Competition Risks from a company's competitors in the market. For example, "*Apple Revamping Smart-Home Efforts to Challenge Amazon, Google.*"
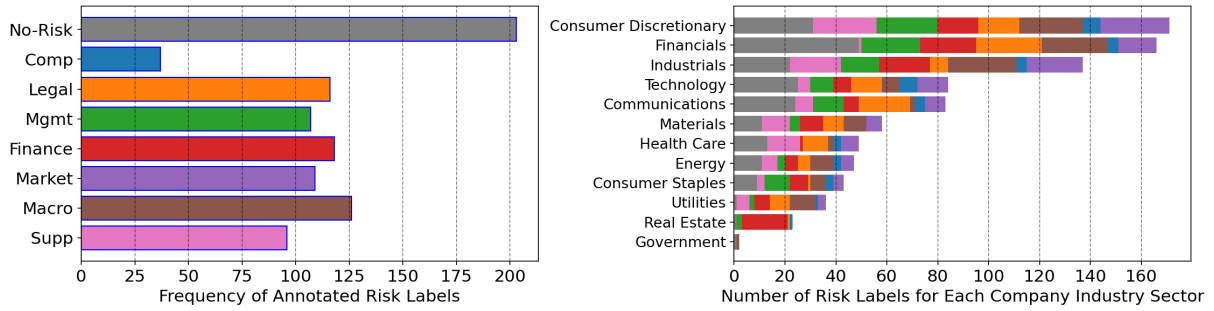
---

[1] https://www.bloomberg.com

Figure 1: [Left] Label distribution over the 716 annotated samples. We denote `Competition` as `Comp`, `Legal and Regulations` as `Legal`, `People and Management` as `Mgmt`, and `Supply Chain and Product` as `Supp`. All except `Competition` have an approximately 14% positive rate. In total, 71.6% of the samples are labeled with at least one risk factor. [Right] Distribution of risk labels for each company industry sector, based on the Bloomberg Industry Classification Standard (BICS). In total, 12 different sectors are covered in the annotated dataset.

4 Data and Annotation

BN Dataset We draw five years of articles published by Bloomberg News[2] (hereafter **BN**), covering diverse events and opinions pertaining to companies across the world. We choose news published between 2018 and 2022 to allow for COVID-related comparisons. This initial dataset is filtered by removing machine-generated and non-natural language content, and is preprocessed with a rule-based entity extraction pipeline. We further removed articles where no company is mentioned. This results in a collection of $277,112$ news articles covering $14,972$ public and $11,413$ private companies. For the sake of simplicity, we keep only the headline and first five sentences of each article for our study. These articles range from 20 to $4,430$ tokens long, with the average article being 151 tokens.

Year	2018	2019	2020	2021	2022	Total
# Docs.	56,741	62,862	57,421	53,906	46,182	277,112

Risk-Related Pre-Filtering Our pilot study shows that risk factors can be sparse in news. Directly annotating over a random subset of BN articles will therefore yield a very high negative ratio. We apply a lexicon-based filter before the sampling. Concretely, we iteratively curate 53 unigrams to capture various aspects of risk events, such as "*challenge*," "*layoff*," "*shrink*."[3] We consider annotating an article only if its headline matches at least one of the keywords. We also experimented with hedges similar to (Pei and Jurgens, 2021). However, we

found that while hedges are a good proxy for uncertainty, they are not able to reliably recall news articles regarding risks, highlighting the difference between uncertainty and risk factors.

Annotation We conduct an annotation study based on the seven risk factors mentioned in Section 3, using a multi-label classification setting. We hire three U.S.-based annotators who are experienced in the finance domain. They are first instructed to label 100 articles independently, followed by a discussion to resolve disagreements and make modifications to the annotation guideline. The adjudicated set is used as test data.

They further annotate 200 articles each. After removing samples with wrong mentions or low-quality text, the final dataset includes 716 samples from 666 unique news articles. In this dataset, 49% of the samples have exactly one label, while more than 20% mention multiple factors [4]. In Figure 1 we show the distribution of risk factors along with the number of news without any risk factors ("NO-RISK").

	SUPP	MGMT	FIN	LEGAL	MACRO	COMP	MRKT
SUPP	N/A	8	6	12	17	8	19
MGMT	8	N/A	15	16	12	1	12
FIN	6	15	N/A	8	22	0	9
LEGAL	12	16	8	N/A	11	4	8
MACRO	17	12	22	11	N/A	3	31
COMP	8	1	0	4	3	N/A	11
MRKT	19	12	9	8	31	11	N/A

Figure 2: Risk co-occurrence matrix (annotated dataset).

Statistics To better understand the characteristics of our newly annotated dataset, we first show the distribution of industry sectors to which the

---

[2]https://www.bloomberg.com
[3]Full list can be found in Appendix.

[4]Detailed distribution is in Appendix (§ A.1)

detected companies belong. We match each company to one of the 12 high-level industry sectors defined by the Bloomberg Industry Classification Standard (BICS) [5]. As shown on the right side of Figure 1, our dataset contains samples over all of these sectors. The distribution of risk types differs across industry sectors. For instance, there are more MARKET related risks for companies in "*Consumer Discretionary*," while more "*Legal*" risks are mentioned for companies in the "*Financials*" and "*Communications*" industries. For "*Real Estate*", the majority of the risks fall under FINANCE. In Figure 2 we further illustrate the co-occurrence of risk factors. Notably, we observe higher co-occurrence of (FINANCE, MACRO) and (MARKET and MACRO) pairs.

5 Benchmark

We formulate the risk prediction task as a multi-label classification problem: given a news article and a mentioned company, we aim to predict whether each of the seven risk factors is mentioned. We consider non-neural baseline models, fine-tuning pre-trained transformers, and large-language models (LLM) with in-context learning. We split the dataset into 484 samples for training, 126 for validation, and 106 for testing.

5.1 Models

We first experiment with non-neural baseline models: (1) **Random**: for each risk factor, randomly assign a binary label with equal probabilities. (2) **Logistic Regression**: we calculate TF-IDF (up to bigrams) features and run logistic regression models for each risk factor. Similarly, (3) **Support Vector Machine (SVM)** models are trained using the same TF-IDF features and linear kernel. (4) We further implement k-nearest neighbor (**KNN**) models using document embeddings calculated from a fine-tuned RoBERTa model (Liu et al., 2019) with SimCSE (Gao et al., 2021) objective.

Pre-trained Transformers with Fine-tuning
We benchmark common pre-trained transformer models as sequence classification tasks under a supervised fine-tuning setting: (1) **BERT-large** (Devlin et al., 2018), (2) **RoBERTa-base** and **RoBERTa-large** (Liu et al., 2019), (3) **RoBERTa-large-BB**: a RoBERTa model further pre-trained on 13 years of Bloomberg News data.
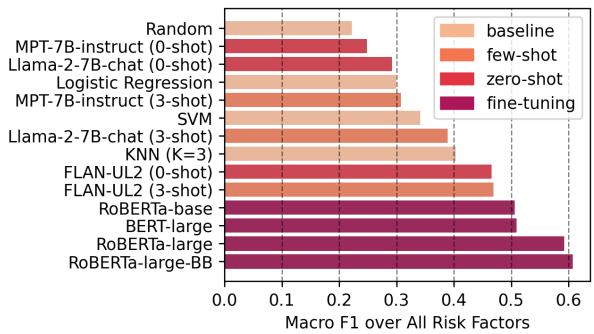
Figure 3: The overall performance of different models. The best result is achieved by fine-tuning the **RoBERTa-large-BB** model, which is pre-trained on domain-specific datasets. Zero-shot and few-shot prompting for LLM perform worse than the fine-tuned models by a large margin.
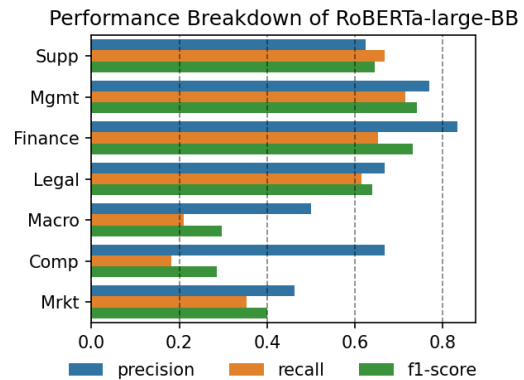

Figure 4: The best performance on each risk factor. Identifying company risks on `Macro`, `Markets` and `Consumer` and `Competition` remains hard.

LLM with Prompting We further compare with three open source, instruction-tuned large language models (LLM) under the in-context learning setting: (1) **FLAN-UL2**[6], which is an instruction-tuned version of the UL2 (Tay et al., 2022) model over the FLAN (Longpre et al., 2023) dataset. (2) **MPT-7B-instruct** (Team, 2023) is a decoder-only model with 7 billion parameters, trained on the `dolly-hhrlhf` [7] dataset. Lastly, (3) **Llama-2-7B-chat** (Touvron et al., 2023) is a decoder-only model optimized for dialogue tasks, achieving competitive performance on various NLP tasks against closed-source LLMs.

For each risk factor, we construct the following prompt template [8] with the input news `text`, the
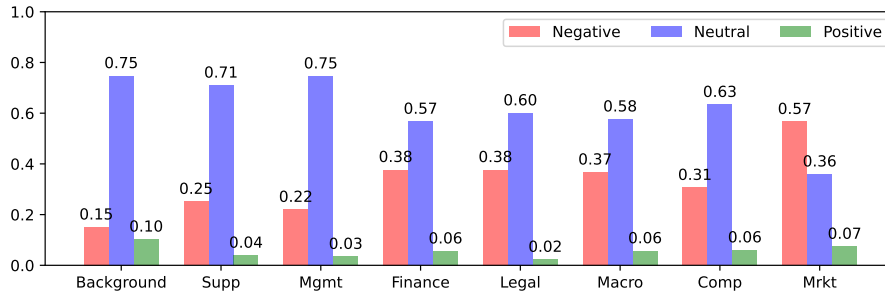
Figure 5: While news articles mentioning risk factors are more negative overall compared with the overall distribution (background), both positive and neutral news can mention risk factors for companies.

target company name, and a full description of the risk from the annotation guideline:

```
{news text}
For company {target}, does the above news
mention {risk} ?
Options: Yes, No
Your answer is (Please only use Yes or No):
```

We consider both the zero-shot and few-shot settings for all LLMs. The few-shot samples are chosen as the k-nearest neighbors ($k = 3$) from the training set, which are represented using the same template and directly prepended to the test sample.

5.2 Result

Figure 3 shows the overall model performance. Fine-tuning transformers yields the best performance, especially the **RoBERTa-large-BB** model that is trained on a domain-specific dataset. We breakdown the per-risk performance in Figure 4. The model achieves better results for Supply Chain and Product and Finance. However, identifying Macro, Competition and Markets and Consumers risks remain challenging.

6 Application and Analysis

Identifying company-related risk factors in news opens many potential applications. In this section, we explore the applications of our model over a large-scale Bloomberg News dataset.

6.1 Are risk factors just negative sentiment?

The term "risk" inherently carries negative connotations. In practice, are risk factors simply negative sentiment? In this study, we explore the connection between company-level sentiment and risk factors. We run an off-the-shelf sentiment analysis model[9]

---

[9]This model is based on DistilBERT (Sanh et al., 2019) and is fine-tuned on finance news with sentiment labels.

over the large 5-year **BN** dataset. For each company mentioned in a news article, a probabilistic distribution over "Positive," "Neutral," and "Negative" is estimated.

As shown in Figure 5, the overall sentiment for a company tends to be more negative when it faces risks. The largest gap of sentiment occurs for Legal and Regulations and Markets and Consumers, where risks are usually mentioned with negative sentiment. Nevertheless, risk factors can be mentioned even when the overall sentiment regarding a company is neutral or positive (See Table 1 in Appendix for examples). This suggests that risk factors are not just negative sentiment.

6.2 Company-level Study

Boeing In 2018, the first 737 MAX airplane crashed into the Java sea. As shown in Figure 6, Boeing faced high risks regarding its products (Supp) in 2018, while other types of risks generally remained low. In 2019, the second 737 MAX crashed, which immediately led to the involvement of the regulators (Legal). Risks related to the market, consumers, and management also rose in 2019.

Toyota Motor Corporation is the largest carmaker in the world. From 2018 to 2020, Toyota saw major Macro and Markets and Consumers risk. In 2020, the world was faced with a global chip shortage, which further led to a spike in Supply Chain and Product risk for Toyota.

Evergrande The Chinese real-estate company Evergrande Group has gone through various debt issues in recent years, which is reflected by the overwhelming percentage of Financial risks predicted by the model.

Huawei Unlike company filings which only include publicly traded companies, news articles also allow us to analyze private companies. Huawei
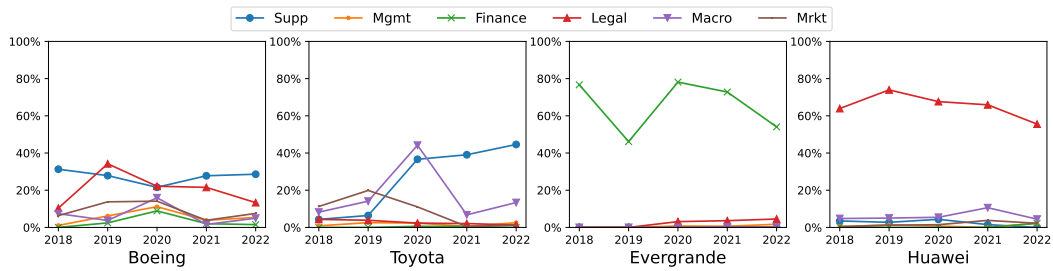
Figure 6: Percentage of news stories tagged by each risk factor type, for different companies.
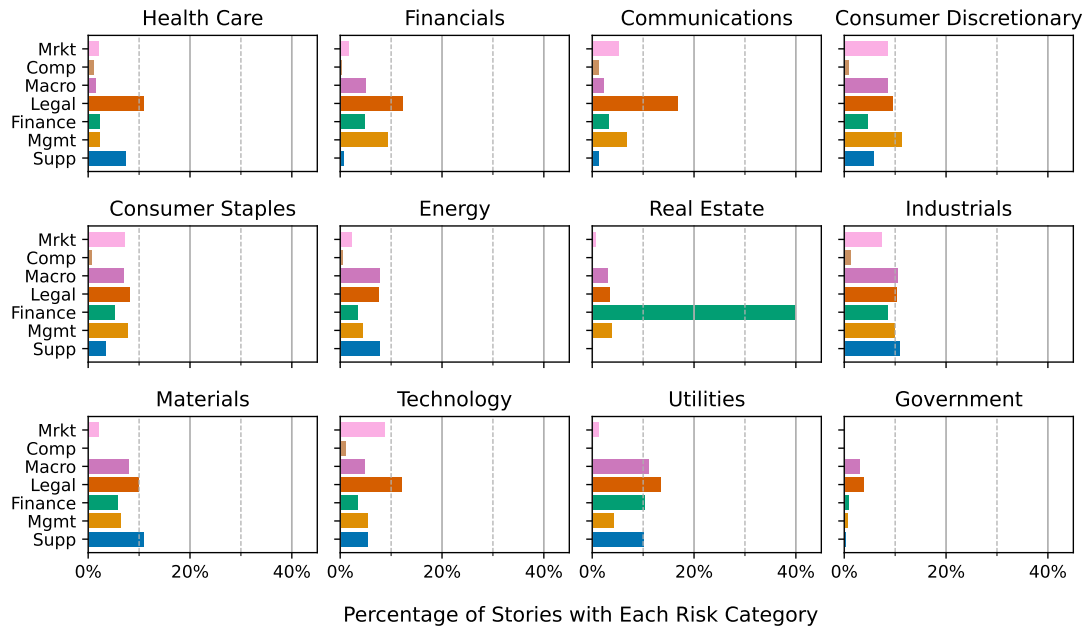

Figure 7: Risk distribution for companies in different industries.

Technologies Co., Ltd. is the world's leading communication technology and phone producer. Since 2018, Huawei has faced regulatory risks from the U.S. government. Figure 6 shows the overall change in Huawei's risk factors mentioned in Bloomberg News data. Huawei saw major regulatory risks from 2018 to 2020. Because of these regulations, Huawei's market and sales are also affected, and it has seen higher `Macro` and `Market` risks since 2021.

6.3 Industry-level Study

Companies in different industries are different in nature and therefore may face different types of risk factors. In this section, we explore the risk factors associated with companies in different industries. We use Bloomberg internal company categorization and map each company to one of the 12 top industry categories: Health Care, Financials, Technology, Energy, Consumer Discretionary, Utilities, Communications, Real Estate, Consumer Staples, Industrials, Materials, and Government. The results are displayed in Figure 7.

Financials Financial companies rarely see risks from the `Supply Chain and Product` side and are more likely to face risks from `People and Management` and `Legal and Regulations`.

Real Estate The real estate industry faced high `Finance` risk from 2018 to 2022, potentially due to the debt crisis of the real estate companies in China.

Health Care The Health Care industries are associated with high `Legal and Regulation` and `Supply Chain and Product` risks, potentially due to the production of and regulations surrounding the COVID-19 vaccines, in addition to providing other health care services in response to the COVID-19 pandemic.

Others Industries like Consumer Discretionary, Consumer Staples and Industrials generally see
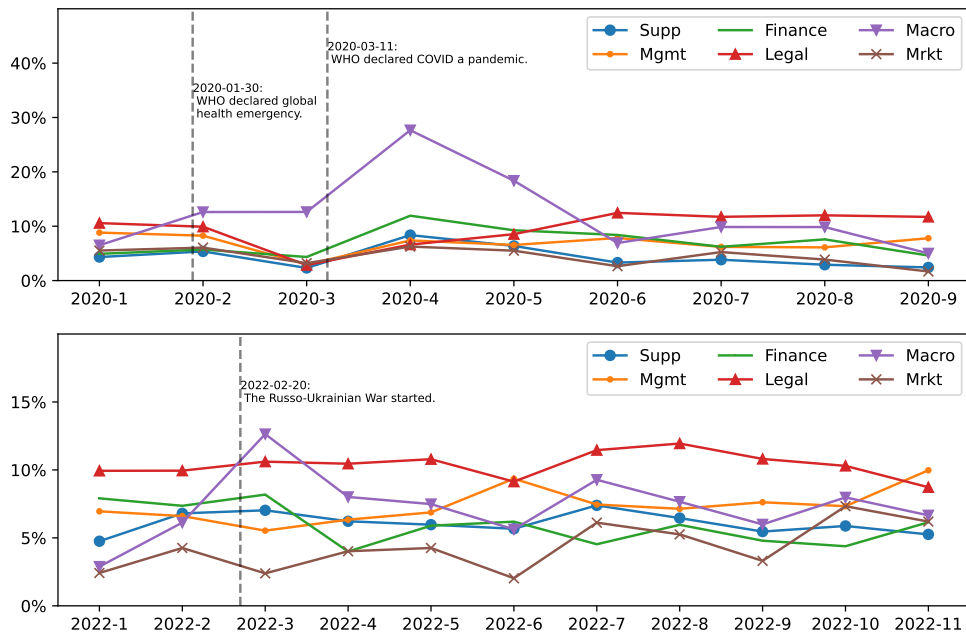
Figure 8: [Top] COVID-19 induces nearly all types of risk factors for companies. [Bottom] Russia's invasion of Ukraine sees increased Macro risks.

balanced risks across all factors.

6.4 Macro-level Study

COVID-19 Pandemic   Since early 2020, the COVID-19 pandemic posed huge global challenges. Figure 8 shows the aggregated risk factors in each month in 2020. The first COVID-19 case was identified in January 2020 and the World Health Organization (WHO) announced a global health emergency on January 31st in response to the rapid increase in infections and deaths worldwide. A new global health emergency led to a sharp rise of `Macro` risks in February. The world may still not have been fully aware of other types of risks, and therefore other risks remained stable in February. However, the situation was changing rapidly. In March, the WHO declared COVID-19 a pandemic and the United States officially issued a national emergency, which led to a sharp rise in all other risk factors in April.

Russia's Invasion of Ukraine   In February 2022, Russia invaded Ukraine and this event immediately led to an increase in `Macro` risks for companies. Similar to the beginning of COVID-19, other types of risks are not reflected at this early stage. However, in June, Russia cut natural gas supplies by more than half, which led to a rise in not only `Macro` risk, but also `Supply Chain and Product` and `Markets and Consumers` risks.

7 Conclusion

Risks are ubiquitous to all companies, industries, and society-at-large. Computational modeling of risk factors could better inform analysts, investors, and policymakers. However, how to systematically model risk factors at scale is a challenging question. In this study, we propose a new categorization framework for risks, and further annotate a new dataset over 666 news articles. We benchmark state-of-the-art NLP models, and analyze a large collection of Bloomberg News articles using the best model. Our analysis demonstrates that modeling risk factors from news could reveal important signals regarding the operations of a company. The aggregated data could further provide information regarding the risks to industries and society.

Acknowledgements

References

A Rashad Abdel-Khalik. 1993. Why do private companies demand auditing? a case for organizational loss of control. *Journal of accounting, auditing & finance*, 8(1):31–52.

Rui Albuquerque, Yrjö Koskinen, and Chendi Zhang. 2019. Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science*, 65(10):4451–4469.

Sergio Beretta and Saverio Bozzolan. 2004. A framework for the analysis of firm risk communication. *The International Journal of Accounting*, 39(3):265–288.

Saumya Bhadani, Ishan Verma, and Lipika Dey. 2020. Mining financial risk events from news and assessing their impact on stocks. In *Mining Data for Financial Applications: 4th ECML PKDD Workshop, MIDAS 2019, Würzburg, Germany, September 16, 2019, Revised Selected Papers 4*, pages 85–100. Springer.

William Blyth, Richard Bradley, Derek Bunn, Charlie Clarke, Tom Wilson, and Ming Yang. 2007. Investment risks under uncertain climate change policy. *Energy policy*, 35(11):5766–5773.

Min-Yuh Day and Chia-Chou Lee. 2016. Deep learning for financial sentiment analysis on finance news providers. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 1127–1134. IEEE.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Motomasa Fujii, Hiroki Sakaji, Shigeru Masuyama, and Hajime Sasaki. 2022. Extraction and classification of risk-related sentences from securities reports. *International Journal of Information Management Data Insights*, 2(2):100096.

Tianyu Gao, Xingcheng Yao, and Danqi Chen. 2021. SimCSE: Simple contrastive learning of sentence embeddings. In *Empirical Methods in Natural Language Processing (EMNLP)*.

Kim Hiang Liow, Muhammad Faishal Ibrahim, and Qiong Huang. 2006. Macroeconomic risk influences on the property stock market. *Journal of Property Investment & Finance*, 24(4):295–323.

Duen-Li Kao. 2000. Estimating and pricing credit risk: An overview. *Financial Analysts Journal*, 56(4):50–66.

Ronald Klingebiel. 2018. Risk-type preference shifts in response to performance feedback. *Strategic Organization*, 16(2):141–166.

Shimon Kogan, Dimitry Levin, Bryan R Routledge, Jacob S Sagi, and Noah A Smith. 2009. Predicting risk from financial reports with regression. In *Proceedings of human language technologies: the 2009 annual conference of the North American Chapter of the Association for Computational Linguistics*, pages 272–280.

Todd Kravet and Volkan Muslu. 2013. Textual risk disclosures and investors' risk perceptions. *Review of Accounting Studies*, 18:1088–1122.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized BERT pretraining approach.

Shayne Longpre, Le Hou, Tu Vu, Albert Webson, Hyung Won Chung, Yi Tay, Denny Zhou, Quoc V Le, Barret Zoph, Jason Wei, et al. 2023. The flan collection: Designing data and methods for effective instruction tuning. *arXiv preprint arXiv:2301.13688*.

Tim Loughran and Bill McDonald. 2020. Textual analysis in finance. *Annual Review of Financial Economics*, 12:357–375.

Hsin-Min Lu, Nina WanHsin Huang, Zhu Zhang, and Tsai-Jyh Chen. 2009. Identifying firm-specific risk statements in news articles. In *Intelligence and Security Informatics: Pacific Asia Workshop, PAISI 2009, Bangkok, Thailand, April 27, 2009. Proceedings*, pages 42–53. Springer.

Allan M Malz. 2011. *Financial risk management: Models, history, and institutions*, volume 538. John Wiley & Sons.

Corentin Masson and Syrielle Montariol. 2020. Detecting omissions of risk factors in company annual reports. In *Proceedings of the Second Workshop on Financial Technology and Natural Language Processing*, pages 15–21.

Angus Moore. 2017. Measuring economic uncertainty and its effects. *Economic record*, 93(303):550–575.

Jiaxin Pei and David Jurgens. 2021. Measuring sentence-level and aspect-level (un) certainty in science communications. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 9959–10011.

Michael Raith. 2003. Competition, risk, and managerial incentives. *American Economic Review*, 93(4):1425–1436.

Marvin Rausand. 2013. *Risk assessment: theory, methods, and applications*, volume 115. John Wiley & Sons.

Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. *ArXiv*, abs/1910.01108.

Thársis Tuani Pinto Souza, Olga Kolchyna, Philip C Treleaven, and Tomaso Aste. 2015. Twitter sentiment analysis applied to finance: A case study in the retail industry. *arXiv preprint arXiv:1507.00784*.

Fabian Stephany, Leonie Neuhäuser, Niklas Stoehr, Philipp Darius, Ole Teutloff, and Fabian Braesemann. 2022. The corisk-index: a data-mining approach to identify industry-specific risk perceptions related to covid-19. *Humanities and Social Sciences Communications*, 9(1):1–15.

Yi Tay, Mostafa Dehghani, Vinh Q Tran, Xavier Garcia, Jason Wei, Xuezhi Wang, Hyung Won Chung, Dara Bahri, Tal Schuster, Steven Zheng, et al. 2022. Ul2: Unifying language learning paradigms. In *The Eleventh International Conference on Learning Representations*.

MosaicML NLP Team. 2023. Introducing mpt-7b: A new standard for open-source, commercially usable llms.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Ann Vanstraelen and Caren Schelleman. 2017. Auditing private companies: what do we know? *Accounting and Business Research*, 47(5):565–584.

Chuan-Ju Wang, Ming-Feng Tsai, Tse Liu, and Chin-Ting Chang. 2013. Financial sentiment analysis for risk prediction. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, pages 802–808.

David Weil, Archon Fung, Mary Graham, and Elena Fagotto. 2006. The effectiveness of regulatory disclosure policies. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 25(1):155–181.

Rong Yang, Yang Yu, Manlu Liu, and Kean Wu. 2018. Corporate risk disclosure and audit fee: A text mining approach. *European Accounting Review*, 27(3):583–594.

Xiaodi Zhu, Steve Y Yang, and Somayeh Moazeni. 2016. Firm risk identification through topic analysis of textual financial disclosures. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–8. IEEE.

A Risk-Related Lexicon

In Section 4, we discuss a risk-related pre-filtering step to narrow down the dataset for annotation. We rely on a manually curated list of keywords by querying the entire dataset. They are listed below.

A.1 Label Count Distribution

We consider risk detection as a multi-label classification problem. In Figure 9, we show the distribution of positive labels per sample (news article) in our annotated dataset.

| affect	ban	cash
cashflow	challenge	competition
concern	crackdown	cut
debt	decline	decrease
delay	demand	downgrade
drop	fail	finance
harm	hit	impact
inflation	layoff	liable
limit	lose	loss
lowest	operation	plunge
pressure	protest	regulation
restriction	risk	rival
shortage	shrink	slump
strike	struggle	sue
suffer	supply	suspend
tension	unable	uncertain
volatile	warn	weak
worsen	worst	
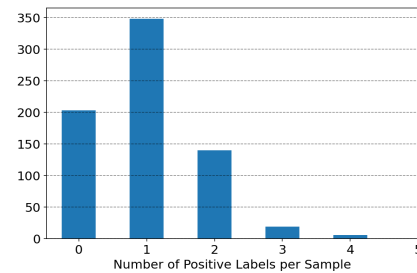
Figure 9: Distribution of number of positive labels per sample. Approximately half of the samples have exactly one risk label.

B Sample News Articles

In Table 1, we show sample articles where the sentiment analysis results are "Positive," while various risk factors are detected.

Tencent's Set for Fastest Growth Since 2018 After Outbreak
(Bloomberg) – Tencent Holdings Ltd. picked up millions of new gamers during the global coronavirus outbreak – yet that surge in mobile play may be slowing as the world's No. 2 economy goes back to work. [...]
Detected risks: *Market*
Wirecard Shares Surge After Statement on KPMG Audit
(Bloomberg) – Wirecard AG , the German payments company trying to move on from reports of alleged questionable accounting methods, said a special investigation has so far found no need to correct financial statements from 2016-2018. [...]
Detected risks: *Management*, *Legal*
SoftBank Soars After Unveiling $41 Billion Asset Sale Plan
(Bloomberg) – SoftBank Group Corp. surged the most in 11 years after unveiling a plan to raise as much as 4.5 trillion yen ($41 billion) over the coming year to buy back stock and slash debt, addressing concerns about its exposure to money-losing businesses during the coronavirus pandemic. [...]
Detected risks: *Finance*
Twitter Surges After Activists Seek to Replace CEO Dorsey
(Bloomberg) – Twitter Inc. shares rose in early trading Monday after Bloomberg reported that activist investors have built a sizable stake in the social media company and are pushing for changes, including possibly replacing co-founder and Chief Executive Officer Jack Dorsey. [...]
Detected risks: *Management*

Table 1: Sample news articles where the sentiment is *Positive* but company risks are detected. Due to space limitation, only the first paragraphs are shown.