Media Framing in *Fraud Scandals*

Author

Nipun Angkavichai

A Classifier Model Approach to Understanding Media Portrayal in Fraud Scandals

1 Introduction

Media framing in fraud scandals refers to how the media shapes public perception by *selecting* certain aspects of a story and making them *salient* through rhetoric and stylistic devices.

Building on prior research that identified **6 frames** in media coverage of the Wirecard fraud scandal, this study develops **a classifier model to automate media framing analysis** focusing on a fraud scandal.

2 Objective

- **Develop an accurate framing classifier:** to provide tools to investigate the media's role in shaping public opinion and constructing narratives surrounding scandals
- **Build a multilingual and generalizable model:** to handle multiple languages and generalize to fraud cases across linguistic contexts, support research in low-resource languages.

3 Dataset

Using 795 newspaper articles on the Wirecard (Germany, June 2020) fraud scandal, annotated with 6 frames, to develop the model.
Based on selection and salience, the 6 frames include:

Failure Attribution

Focus on <u>oversight bodies' malpractice</u> in accounting fraud.

Factualization

Focus on <u>key facts</u> about the fraud.

Personification

Focus on <u>top managers' criminal behavior</u>.

Treatment Recommendation

Focus on <u>needed reforms</u> for oversight bodies and auditors.

Consequences

Focus on the <u>financial and reputational impact</u> on the company and stakeholders.

Moral Evaluation

Focus on the <u>lack of morality</u> among stakeholders.

4 Methodology

Standard Approach

- **Baseline methods:** majority baseline & MaxEnt.
- **Pre-trained language models:** mBERT for fine-tuning.
- **LLMs:** GPT-4o for in-context learning with techniques such as zero-shot and few-shot learning.

Translation-Based Data Augmentation:
- Generate additional training data by translating German articles into English and back to the original.

Ignoring the First Paragraph:
- Exclude the first paragraph, often dominated by informative content, from training data.

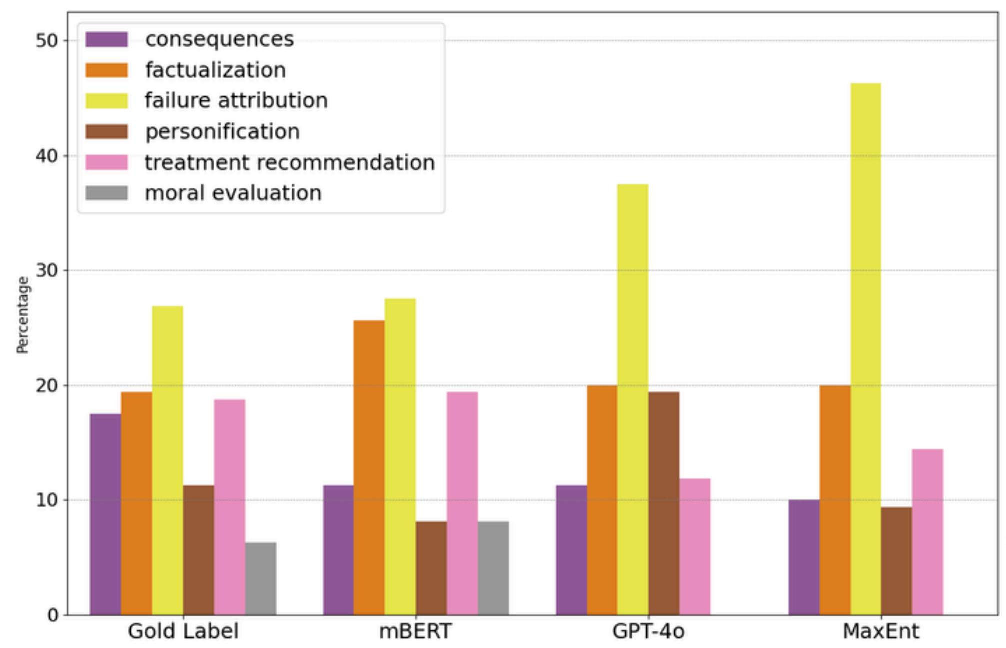
Paragraph/Sentence Level Classification:
- Breaking down articles into smaller segments

Multi-Label Classification:
- Allow articles to have multiple frame labels to better capture their complexity.

5 Results

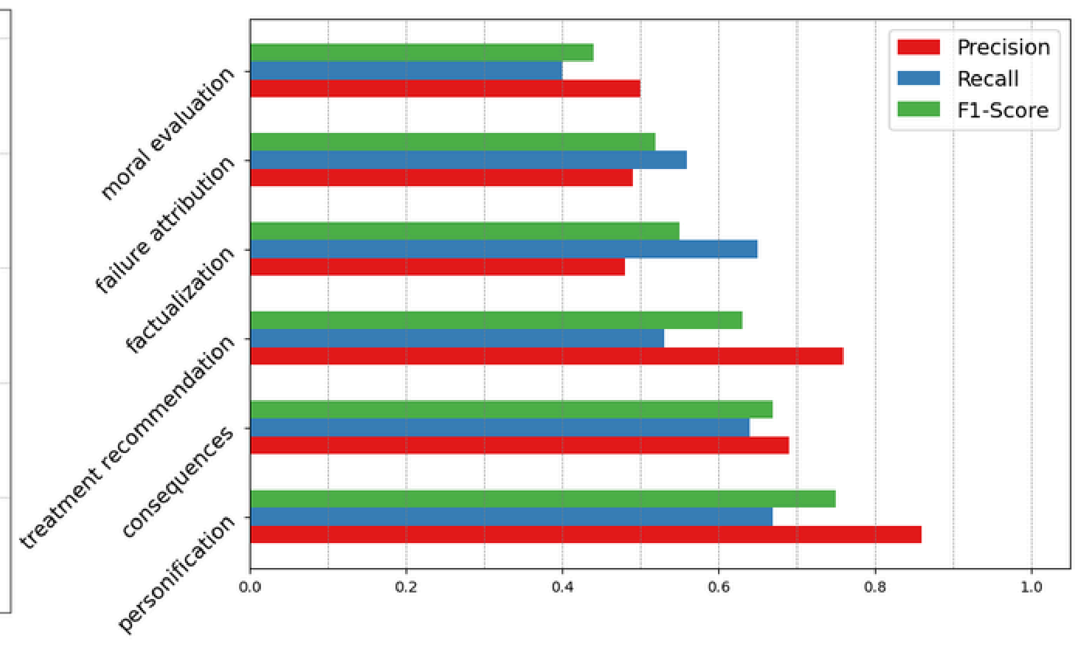
	Approach	Model	Accuracy
Baseline	Majority Baseline	-	0.27
	Standard	MaxEnt (TF-IDF)	<u>0.59</u>
	Translated data (DE to EN)	MaxEnt (TF-IDF)	0.29
	Translated data augmentation	MaxEnt (TF-IDF)	0.55
	Paragraph-level segment	MaxEnt (TF-IDF)	All 'no signal'
	One-vs-All	MaxEnt (TF-IDF)	0.59
	Ignore first paragraph (11 lines) only in training data	MaxEnt (TF-IDF)	0.54
	Ignore first paragraph (11 lines)	MaxEnt (TF-IDF)	0.46
	Standard (Top 2 Prob.)	MaxEnt (TF-IDF)	0.75
	Standard (Top 3 Prob.)	MaxEnt (TF-IDF)	0.90
Transformer	Standard	mBERT	<u>0.59</u>
	context window = 1024	mBERT	0.27
	Standard	G-BERT	0.56
	Translated data augmentation	mBERT	0.55
	context window = 1024	XLM-Long	0.31
	context window = 2824	XLM-Long	0.27
	Paragraph-level segment	mBERT	All 'no signal'
	One-vs-All	mBERT	0.24
	Ignore first paragraph (11 lines) only in training data	mBERT	0.45
	Ignore first paragraph (11 lines)	mBERT	0.40
	Ignore first paragraph (8 lines) only in training data	mBERT	0.49
	Ignore first paragraph (8 lines)	mBERT	0.51
	Standard (Top 2 Prob.)	mBERT	0.76
	Standard (Top 3 Prob.)	mBERT	0.86
In-context Learning	Zero-shot (Top 1 Prob.)	GPT-4o	<u>0.53</u>
	Zero-shot (Top 2 Prob.)	GPT-4o	0.76
	Zero-shot (Top 3 Prob.)	GPT-4o	0.90
	Zero-shot (Top 1 Prob.)	GPT-3.5	0.49
	Zero-shot (Top 2 Prob.)	GPT-3.5	0.63
	Zero-shot (Top 3 Prob.)	GPT-3.5	0.76
	One-shot	GPT-3.5	0.32

- **Paragraph/Sentence Level:** Revealed annotation misalignment, hindering accuracy.
- **Multi-Label Classification:** Top-2 accuracy (75%) exceeded top-1 (59%), indicating multi-label data and need for better annotations.



Predicted Frame Distribution

mBERT demonstrated the most accurate frame distribution, closely aligning with the gold labels in the test dataset.



mBERT Evaluation

The mBERT model achieved **0.59 accuracy**, with strong performance on the Personification frame (F1 = 0.75) but weaker on Moral Evaluation (F1 = 0.44), reflecting uneven frame handling.

6 Conclusion

Our investigation reveals that the data and task exhibit multi-label characteristics, yet the current dataset is annotated with a single label per instance.

This mismatch causes confusion for the model, impacting its reliability and accuracy. Redefining the dataset with multi-label annotations will be essential to develop a more robust and high-performing model.

Based on

Oelrich, S., & Siebold, N. (2024). Media framing in Wirecard's fraud scandal: Facts, failures, and spying fraudster fantasies. Critical Perspectives on Accounting, 100, 102755.
https://doi.org/10.1016/j.cpa.2024.102755