1766-2016
RUTGERS
250

The sentiment features of MD&As and financial misstatement prediction

A comparison of deep learning and text mining approach for textual analysis

Ting Sun, Yue Liu, Miklos A. Vasarhelyi

Presented by
Ting Sun
Rutgers Business School

Motivations

- Deep learning is able to effectively and automatically extract features from data, especially the unstructured or semi-structured data such as videos, audios, and text.
- It has achieved great success in speech recognition, object(face) recognition, and textual analysis.
- With deep learning approach, the sentiment features of the text can be extracted without human intervention
- Few prior literature has applied deep learning based textual analysis approach to auditing

Objective

- Demonstrate that deep learning technology can be applied to analyze finance-related text document to obtain the sentiment feature, which is an additional attribute to support audit judgement
- Provide evidence for the effectiveness of the sentiment features obtained by deep learning by comparing its prediction power to that of the sentiment features obtained by “bag of words”.

Research Questions

(1) Does the sentiment feature of 10-K MD&As extracted by deep learning approach provide essential information for financial misstatement prediction?

(2) How effective does the deep learning approach perform as compared to “bag of words” approach in terms of prediction accuracy?

What we did…

- We analyzed 30,239 MD&As of 10-K filings for fiscal years from 2006 to 2015 using deep learning and “bag of words” approach and obtained two sets of sentiment scores, Sentiment_DL and Sentiment_TM, respectively.
- Utilizing CHAID (CHI-square Adjusted Interaction Detection) algorithm, we established two classification models and compared their predictive performance.
- The results showed that both model 1 and model 2 performed better than previous prediction models for the financial misstatement.
- The sentiment feature extracted by Deep Learning approach generally performed as effectively as that obtained by “bag of words” approach.

Financial misstatement prediction

Distinguish financial misstatement (FM) from fraud:

FM: annual reports which contain misstatement and have been restated.

Fraud: An accounting misstatement is fraudulent if committed with intention.

FM can be seen as a superset of fraud. It is harder to predict than fraud

Prior literature for FM prediction

The misstatement literature, specifically those related to prediction with Machine Learning algorithms is limited as compared to fraud.

There is even less research involving content features of text (like sentiment): Cecchini, 2005; Larcker and Zakolyukina, 2012

the sample size is relative small and the predictive performance is modest

larcker and Zakolyukina, 2012: best AUC=0.597, total sample size=17,150

Cecchini, 2005: accuracy=55.84% ,total sample size=800

Sentiment analysis approaches

	Deep learning approach	Bag of words approach
Description of the technique	Emerging technique employing deep hierarchical neural network and trained with a large amount of text files	Prevalent technique using various pre-defined word lists, with each one representing a particular sentiment feature
Rationale	“understand” the meaning of a text file	count the frequency of the words originated from a specific dictionary
Output sentiment feature	Sentiment scores: Sentiment_DL	sentiment scores: Sentiment_TX
Is there prior literature in accounting and auditing domain	No	Yes
Tool	Alchemy language API	Loughran and McDonald (2011)
Is it a finance-specific tool	No	Yes
Required text document	HTML/text document and webpage	HTML/text document
Does it need data preprocessing	No	Yes

Sample

MD&As filed from 2007 through 2015 as provided by Seek iNF	58882
Less: fiscal year missing from Compustat	(27872)
Less: CIK missing from Compustat	(590)
Less: auditor missing from AuditAnalytics	(181)
Final Sample	30,239

Distribution of misstatements over fiscal years

Fiscal year	Normal Statements	Misstatements	Total
2006	2,461	367	2,828
2007	3,351	488	3,839
2008	3,451	514	3,965
2009	3,590	563	4,153
2010	3,424	596	4,020
2011	2,015	380	2,395
2012	2,020	387	2,407
2013	3,234	536	3,770
2014	1,961	216	2,177
2015	637	48	685
Total	26,144	4,095	30,239

Sentiment scores

	Obs.	Min	25% percentile	Median	75% percentile	Max
Sentiment_DL	30239	-.5606	-.0289	.0170	.0658	.7487
Sentiment_TM	30239	-.0721	-.0105	-.0062	-.0024	.0307

Classification models

		Model 1	**Model 2**
Dependent variable		Misstatement	Misstatement
Independent variables	Sentiment measures	SENTIMENT_TM	SENTIMENT_DL
	Other predictors following prior research	35 variables related to misstatement	35 variables related to misstatement

Prediction results of testing data

	Model 1	Model 2
Accuracy	64.23%	**65.7%**
Type 1 error rate	35.54%	**33.32%**
Type 2 error rate	**37.24%**	40.66%
Precision	0.2139	**0.2168**
Sensitivity	**0.6276**	0.5934
specificity	0.6446	**0.6668**
F1 score	**0.3191**	0.3176
AUC	**0.68**	**0.68**

Conclusions

The results show that

(1) the sentiment features generated by both approaches exhibit relatively high predictive accuracy in the two prediction models as compared with prior literature of similar sample size;

(2) With deep learning approach, we are less likely to have type one errors

(3) With “bag of words” approach, we are less likely to have type two errors. Possible reason is that it is a finance-specific approach.

(4) Generally speaking, deep learning approach performs as effectively as “bag of words” approach