Detect & Describe:
Deep learning of bank stress in the news

Samuel Rönnqvist[1] and Peter Sarlin[2]

[1] Turku Centre for Computer Science / Åbo Akademi Uni. / Goethe Uni. Frankfurt
[2] Hanken School of Economics / RiskLab Finland at Arcada

Systemic Risk Analytics, Helsinki

September 24, 2015

Detect & Describe

Financial text as a source of information on risk:

- Descriptive indicators from text
- In this paper, focus on bank distress

Key questions:

- Can we *detect* distress based on text?
- Can we obtain *descriptions* of events?

Background

- Fore-/nowcasting bank distress
  - Long term: Accounting data to measure imbalances
  - Short term: Market data to indicate volatility/jumps
  - Data challenges, e.g.:
    - information content / noise
    - availability
    - immediacy
- Identifying sources of financial risk
  - Accounting data indicate risk categories
  - Market data lacks descriptive information per se

Text as an information source

Alternative: Text data as a source of information on risk

Case: Bank distress (defined by event data)

1. To detect distress:
   - e.g., news articles as coinciding predictors
     Articles $\rightarrow$ *Bank distress measure*
2. To describe distress:
   - Extract text descriptions related to detected distress
     Extracts $\leftarrow$ *Bank distress measure*

Text as an information source

1. To detect distress
2. To describe distress

- Basis: predictive modeling of events based on text
- Problem: text data is very sparse, relatively few distress events
  - How to model?

A semantic deep learning set-up

- Deep learning with neural networks:
  - Learning abstractions of data, rather than handcrafting them

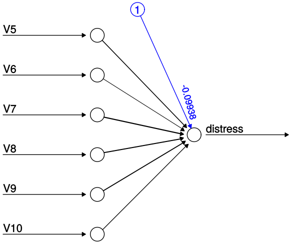
    $\rightarrow$ Multiple layers of learning
    $\rightarrow$ Data-driven, flexible

  - State-of-the-art performance in computer vision etc.

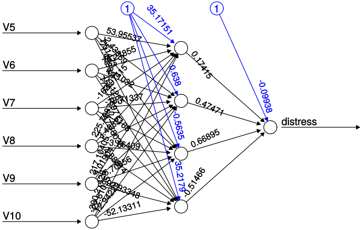
- Semantic deep learning:
  - Learning semantic abstractions of text (Le & Mikolov, 2014), as features in risk modeling
  - Dimensionality reduction: millions to hundreds

A semantic deep learning set-up

- Recap: Artificial Neural Networks



Single-layer network (Logistic regression)



Multi-layer network

- Hidden layer & logistic function $\rightarrow$ a non-parametric model

A semantic deep learning set-up

- Deep learning set-up: two-step learning with neural networks

Step 1: Unsupervised semantic modeling

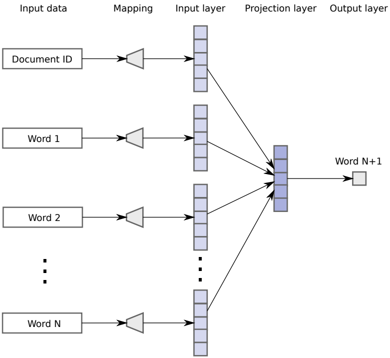
- In: Large set of news articles (~0.1-1M)
- Out: Semantic vectors (representations of article semantics)

Step 2: Supervised predictive modeling

- In: Semantic vectors
- Out: Distress signal (based on hundreds of events)

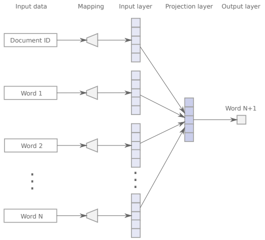
A semantic deep learning set-up

Step 1: Learning semantic vectors, through word prediction

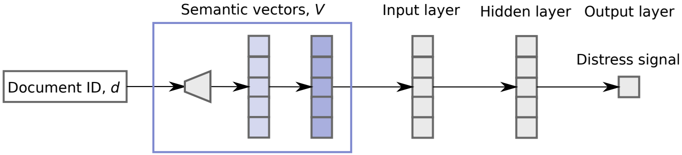


A semantic deep learning set-up

Step 1: Learning semantic vectors, through word prediction



Step 2: Learning to predict distress, based on semantic vectors of articles



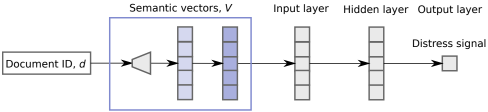
Experiments

- Data:
  - Text: 6.6M news articles from Reuters, 2007-2014
  - 243 reference events (govt. intervention, state aid, etc.), 2007Q2-2012Q2, involving 101 European banks (Betz et al., 2014)
- 262k articles mention any target bank
  - 8.6% of articles coincide with distress of bank
- Tasks:
  1. Learn to nowcast bank distress, based on mentioning articles
  2. Extract descriptions, based on predictive model

Experiments: predictive modeling

Step 1: 262k articles (210M words) → 262k semantic vectors

Step 2: Vectors → 400 input nodes → 20 hidden nodes → distress signal of article



Experiments: predictive evaluation

- We rather give false alarms than miss events (imbalanced loss function, preference $\mu = 0.9$)
- *Usefulness* score to evaluate gain of the model
- *Relative Usefulness*: no gain $\rightarrow$ perfect model (0$\rightarrow$100%)

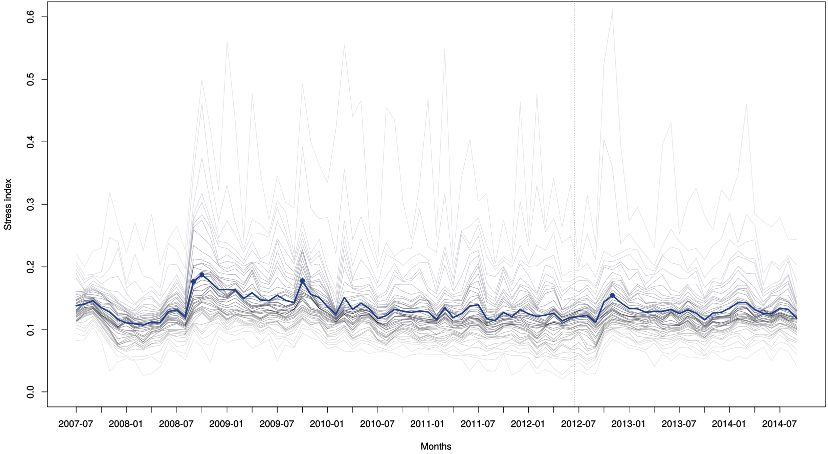
Experiments: predictive evaluation

Out-of-sample, relative Usefulness 27% (AUC 0.71)

μ	$\hat{U}_{\boldsymbol{r}}(\mu)$	$\sigma_{\boldsymbol{U}}$	$\hat{F}_\beta$	$\sigma_{\boldsymbol{F}}$	$\hat{TN}$	$\hat{FN}$	$\hat{FP}$	$\hat{TP}$
0.1	-33.34	2.47	0.070	0.03	18823	2249	0	2
0.2	-14.41	1.08	0.016	0.01	18823	2249	1	2
0.3	-8.094	0.62	0.023	0.01	18819	2242	4	8
0.4	-4.937	0.39	0.036	0.01	18807	2225	16	26
0.5	-3.044	0.25	0.037	0.02	18787	2208	36	43
0.6	-1.781	0.16	0.066	0.02	18694	2142	130	109
0.7	-0.879	0.09	0.114	0.02	18455	2023	368	228
0.8	-0.203	0.04	0.229	0.03	17503	1740	1321	511
0.85	0.147	0.02	0.397	0.05	15462	1331	3362	919
0.9	$\underline{0.271}$	0.01	$\underline{0.713}$	0.03	10306	577	8517	1673
0.95	0.146	0.01	0.934	0.01	4873	112	13950	2139

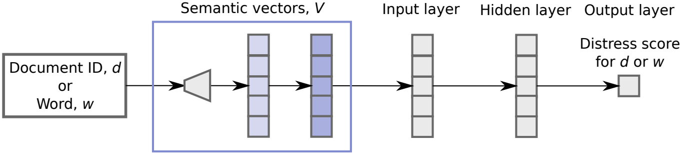
Experiments: distress indices

- Distress index: article distress scores aggregated by bank and period
- Overview: European-level average and percentiles of indices



Experiments: descriptions

- Key points identifiable through the index
  - How to describe the events?



- Use the model to score relevance of text excerpts

$$\sum \textit{article score} \times \textit{word score} \times \textit{word frequency}$$

Experiments: descriptions

Sep. 2008 *“...crisis has kept markets on tenterhooks by forcing European authorities to rescue troubled banks. Belgian-Dutch group Fortis FOR.BR underwent nationalisation on Sunday after emergency...”*

Oct. 2008 *“...banks booked losses on the U.S. housing market. On Sunday, the Dutch government agreed a 10 billion euro cash injection to ING to shore...”*

Oct. 2012 *“...high. Expectations that Spain will apply for a bailout, prompting the European Central Bank to start buying its bonds, have helped support the euro...”*

Conclusions

- We built a semantic model, and predictive model for distress
  - The link between text and events (defined by data)
  - Provides stress index to *detect* events of relevance
  - Provides text excerpts to *describe* the events
- Future work:
  - Extend coverage of event data
  - Explore alternative deep set-ups for the task
  - Improve method for description extraction

Thank you for your attention.