



Comparing Features for Ranking Relationships Between Financial Entities based on Text

Tim Repke, Michael Loster, Ralf Krestel
Hasso Plattner Institute
Potsdam, Germany

What is the FEIII Challenge?

■ Given:

- set of tuples (financial entity A, role, financial entity B)
- text snippet the tuple was extracted from (10-K, 10-Q filings)
- tuples labelled by experts as (highly) relevant, neutral, irrelevant

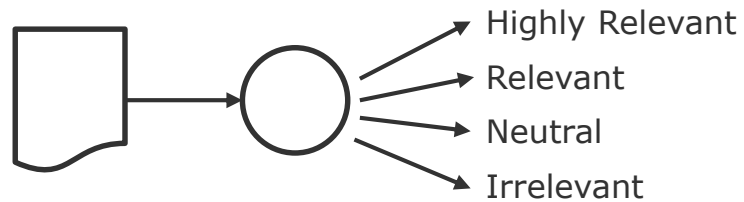
■ Challenge:

per role, rank relationship tuples by relevance

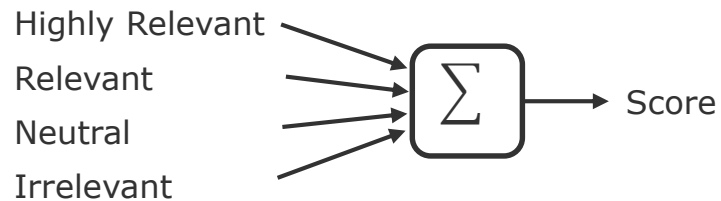
**Text Features for
Relationship
Ranking**

12/05/2017

- Restrict input to **context snippets** (no external information)
- Train **classifiers** to distinguish between levels of relevance

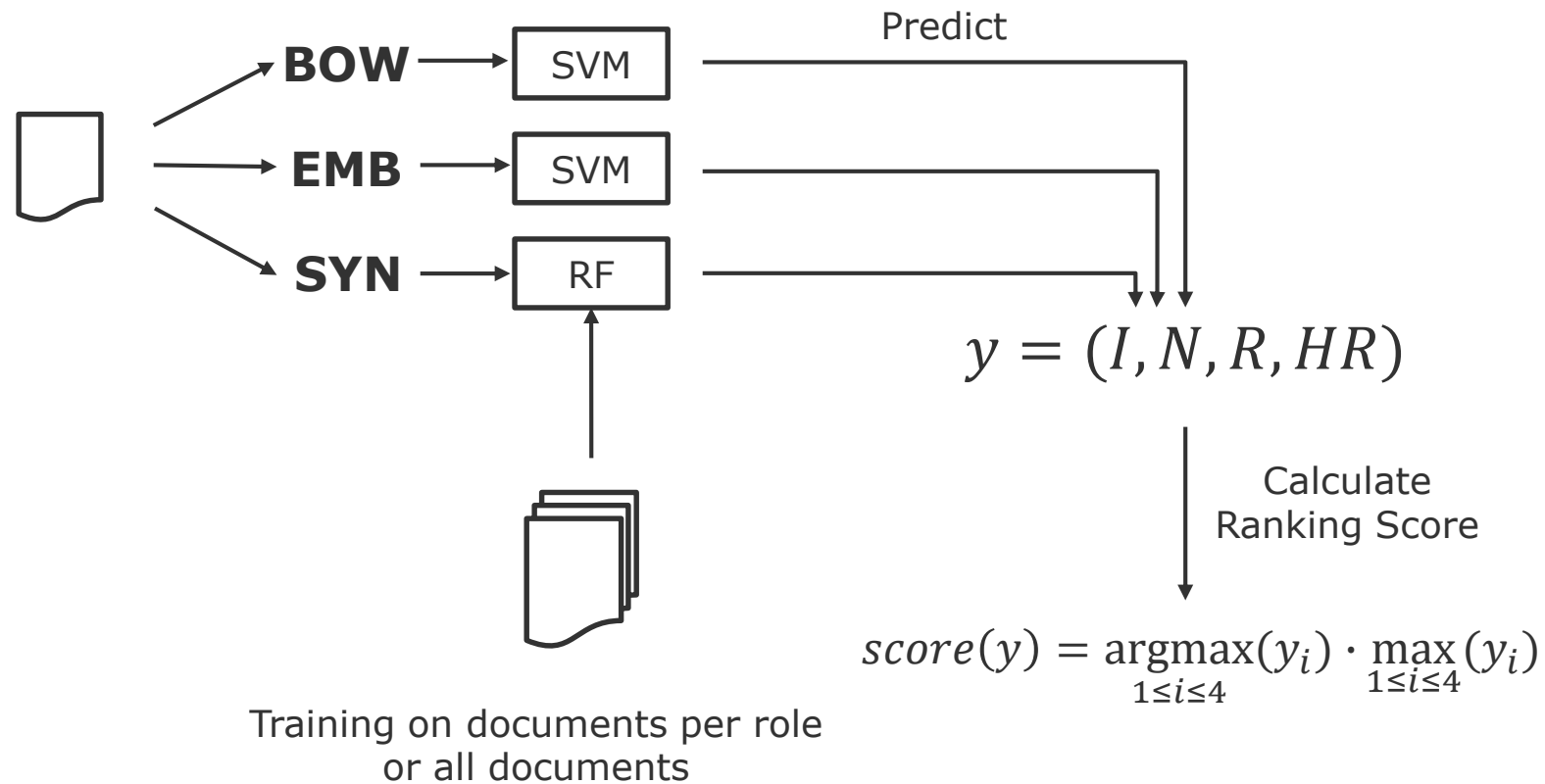


- Use classifiers' prediction to calculate ranking score



**Text Features for
Relationship
Ranking**

12/05/2017



**Text Features for
Relationship
Ranking**

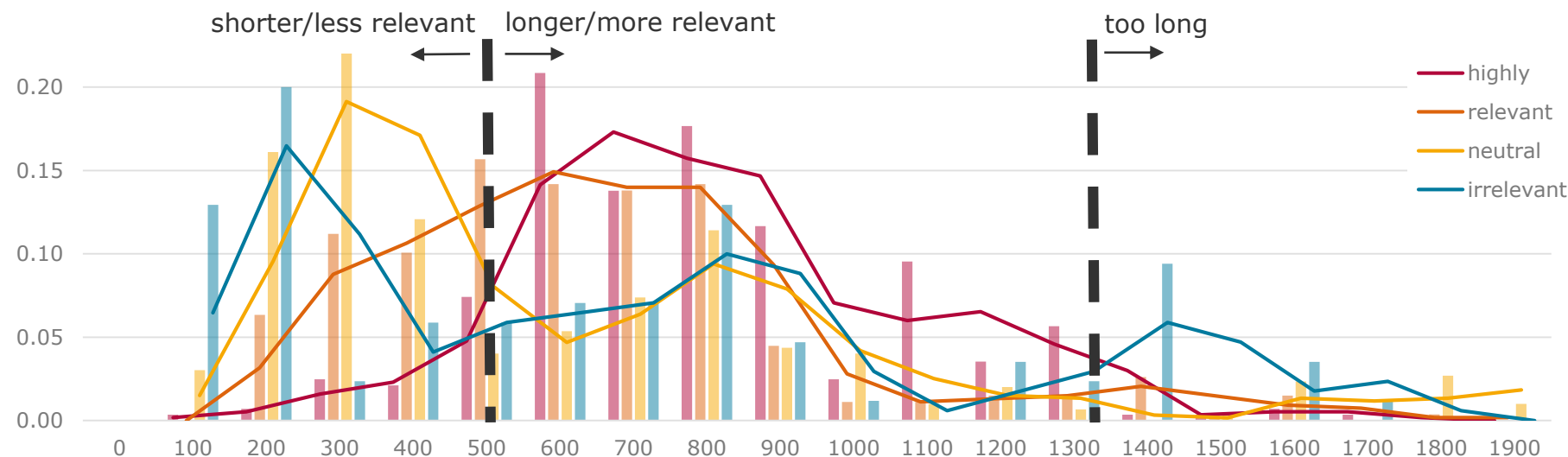
12/05/2017

Features

SYN (Syntax Features)

- Language independent, fewer training samples needed
- Character based counts/ratios
- Token based counts/ratios
- POS-tag based features

Example: Histogram of snippet length (characters)



Text Features for Relationship Ranking

12/05/2017

- Baseline (**random**): average NDCG of 100 random rankings per pass
- Baseline (**worst**): NDCG of inverse perfect order

| Approach | NDCG (std) | F1-Score |
|-------------------|-------------|----------|
| Baseline (random) | 0.88 (0.03) | -- |
| Baseline (worst) | 0.72 (0.06) | -- |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

- 5-fold cross-validation
- Trained on ~900 training samples (per pass, leave out samples from 5 documents)
- Tested on all testing samples, scores calculated per role and aggregated

| | | | |
|---|-------|-------|---|
| ↓ | ideal | worst | ↓ |
| | 4 | 1 | |
| | 4 | 2 | |
| | 3 | 2 | |
| | 3 | 3 | |
| | 3 | 3 | |
| | 2 | 3 | |
| | 2 | 4 | |
| | 1 | 4 | |

**Text Features for
Relationship
Ranking**

12/05/2017

| Approach | NDCG (std) | F1-Score |
|-------------------|-------------|-------------|
| Baseline (random) | 0.88 (0.03) | -- |
| Baseline (worst) | 0.72 (0.06) | -- |
| BOW | 0.88 (0.05) | 0.34 (0.13) |
| EMB | 0.89 (0.04) | 0.24 (0.18) |
| | | |
| | | |

- 5-fold cross-validation
- Trained on ~900 training samples (per pass, leave out samples from 5 documents)
- Tested on all testing samples, scores calculated per role and aggregated

**Text Features for
Relationship
Ranking**

12/05/2017

| Approach | NDCG (std) | F1-Score |
|-------------------|-------------|-------------|
| Baseline (random) | 0.88 (0.03) | -- |
| Baseline (worst) | 0.72 (0.06) | -- |
| BOW | 0.88 (0.05) | 0.34 (0.13) |
| EMB | 0.89 (0.04) | 0.24 (0.18) |
| SYN | 0.94 (0.04) | 0.44 (0.11) |
| | | |

- 5-fold cross-validation
- Trained on ~900 training samples (per pass, leave out samples from 5 documents)
- Tested on all testing samples, scores calculated per role and aggregated

Text Features for Relationship Ranking

12/05/2017

- **BOW+EMB+SYN**: soft vote of BOW, EMB, and SYN classifier

| Approach | NDCG (std) | F1-Score |
|-------------------|--------------------|--------------------|
| Baseline (random) | 0.88 (0.03) | -- |
| Baseline (worst) | 0.72 (0.06) | -- |
| BOW | 0.88 (0.05) | 0.34 (0.13) |
| EMB | 0.89 (0.04) | 0.24 (0.18) |
| SYN | 0.94 (0.04) | 0.44 (0.11) |
| BOW+EMB+SYN | 0.95 (0.04) | 0.43 (0.12) |

- 5-fold cross-validation
- Trained on ~900 training samples (per pass, leave out samples from 5 documents)
- Tested on all testing samples, scores calculated per role and aggregated

Text Features for
Relationship
Ranking

12/05/2017

Results

Confusion Matrix

| <i>predicted annotated</i> | Highly Relevant | Relevant | Neutral | Irrelevant |
|--------------------------------|-----------------|----------|---------|------------|
| Highly Relevant | 0.73 | 0.23 | 0.04 | 0.00 |
| Relevant | 0.25 | 0.39 | 0.34 | 0.02 |
| Neutral | 0.21 | 0.56 | 0.17 | 0.06 |
| Irrelevant | 0.09 | 0.54 | 0.21 | 0.16 |

**Text Features for
Relationship
Ranking**

12/05/2017

- Shallow features outperforms word based approaches
 - Word features seem over-fitted, require many labelled samples
 - Semi-supervised approach doesn't generalise much better
 - Manually tailored syntax features avoid this
- Ensemble takes best of both worlds
- Classifiers trained per role are too sparse
 - Although role specific wording could be beneficial

**Text Features for
Relationship
Ranking**

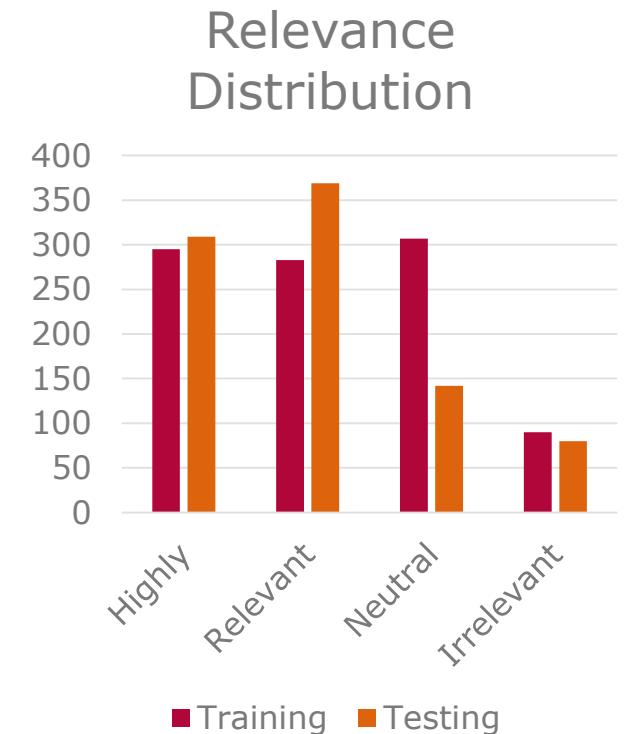
12/05/2017

Thanks for your attention!
Do you have questions?

Supplementary Slides

Dataset Distribution

| | Training | Testing |
|-------------|------------|------------|
| Affiliate | 186 | 129 |
| Agent | 61 | 40 |
| Counterpart | 64 | 108 |
| Guarantor | 34 | 28 |
| Insurer | 19 | 47 |
| Issuer | 129 | 98 |
| Seller | 20 | 49 |
| Servicer | 21 | 57 |
| Trustee | 420 | 304 |
| Underwriter | 21 | 40 |
| All | 975 | 900 |



Inter Annotator Agreement (Training Set)

| | | | | | | | | | | |
|-----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|-----------|
| Expert 1 | 409 | 0.26 | 0.94 | 0.53 | | | 0.40 | | | 0.06 |
| Expert 2 | 128 | 366 | | 0.06 | 0.91 | | 1.00 | 1.00 | | 0.03 |
| Expert 3 | 24 | | 60 | | 0.03 | | 0.03 | 0.09 | | 0.09 |
| Expert 4 | 35 | 1 | | 35 | | 0.09 | | 0.09 | | 0.09 |
| Expert 5 | 11 | 36 | 1 | | 80 | | 0.09 | 0.54 | | 0.54 |
| Expert 6 | | | | 1 | | 60 | | 0.54 | | 0.54 |
| Expert 7 | 3 | 41 | 1 | | 1 | | 75 | 0.64 | | 0.64 |
| Expert 8 | 23 | 12 | 39 | 1 | 44 | 1 | 35 | 226 | | 0.07 |
| Expert 9 | | | | | | | | | 74 | |
| Expert 10 | 8 | 13 | 1 | 1 | 1 | 1 | 1 | 14 | | 20 |
| | Expert 1 | Expert 2 | Expert 3 | Expert 4 | Expert 5 | Expert 6 | Expert 7 | Expert 8 | Expert 9 | Expert 10 |

- Lower left: overlapping annotations
- Upper right: Cohens kappa
- Diagonal: number of annotations

Syntax Features

```
{
    'num_chars': len(raw),
    'num_words': len(raw.split()),
    'num_upper_chars': sum(1 for c in raw if c.isupper()),
    'num_upper_words': sum(1 for w in raw.split() if w[0].isupper()),
    'ratio_upper_chars': sum(1 for c in raw if c.isupper()) / len(raw),
    'ratio_upper_words': sum(1 for w in raw.split() if w[0].isupper()) / max(1, len(raw.split())),
    'mean_word_len': np.mean([len(w) for w in raw.split()]),
    'num_word_repetitions_raw': len({k: v for k, v in counts_r.items() if v > 1}),
    'num_word_repetitions_clean': len({k: v for k, v in counts_c.items() if v > 1}),
    'ratio_word_repetitions_raw': len({k: v for k, v in counts_r.items() if v > 1}) / max(1, len(raw.split())),
    'ratio_word_repetitions_clean': len({k: v for k, v in counts_c.items() if v > 1}) / max(1, len(clean.split())),
    'num_dollarsigns': len(raw) - len(raw.replace('$', '')),
    'num_numbers': len(re.findall(r'\d+', raw)),
    'num_digits': len(re.findall(r'\d', raw)),
    'ratio_numbers': len(re.findall(r'\d+', raw)) / max(1, len(raw.split())),
    'ratio_digits': len(re.findall(r'\d', raw)) / len(raw),
    'num_ents_org': len([ent for ent in n.ents if ent.root.ent_type_ == 'ORG']),
    'num_ents_person': len([ent for ent in n.ents if ent.root.ent_type_ == 'PERSON']),
    'num_verbs': tagc['VB'] + tagc['VBD'] + tagc['VBG'] + tagc['VBN'] + tagc['VBP'] + tagc['VBZ'],
    'num_adverbs': tagc['RB'] + tagc['RBR'] + tagc['RBS'],
    'num_adj': tagc['JJ'] + tagc['JJR'] + tagc['JJS'],
    'num_punct': posc['PUNCT'] + posc['SYM'],
    'num_1letter': len([tok for tok in n if len(tok) == 1 and tok.pos_ != 'PUNCT' and tok.pos_ != 'SYM']),
    'num_role_mentions': raw.lower().count(row['ROLE']),
    'role+ent': len([sent for sent in n.sents if row['MENTIONED_FINANCIAL_ENTITY'] in str(sent) and row['ROLE'] in str(sent)]),
    'dist_role_ent': dist_role_ent + (len(row['MENTIONED_FINANCIAL_ENTITY']) if dist_role_ent < 0 else len(row['ROLE'])),
    'dist_role_ent_abs': abs(dist_role_ent + (len(row['MENTIONED_FINANCIAL_ENTITY']) if dist_role_ent < 0 else len(row['ROLE'])))
}
```

Full results (participant rankings)

| weighted average | Name | gt1 | gt1 500 | gt2 | gt3 | gt4 | gt5 | SUM | Ranking |
|---------------------|------|-----|---------|-----|-----|-----|-----|-----|---------|
| | P1 | 17 | 17 | 17 | 17 | 17 | 17 | 102 | 17 |
| | P2 | 14 | 14 | 12 | 12 | 12 | 10 | 74 | 12 |
| P3 scored_full_all | P3 | 7 | 7 | 6 | 8 | 6 | 5 | 39 | 6 |
| P4 scored_full_bow | P4 | 15 | 15 | 14 | 16 | 16 | 9 | 85 | 16 |
| P5 scored_full_emb | P5 | 11 | 11 | 14 | 10 | 14 | 14 | 74 | 12 |
| P6 scored_full_syn | P6 | 3 | 3 | 3 | 2 | 2 | 3 | 16 | 2 |
| P7 scored_role_all | P7 | 13 | 13 | 11 | 13 | 15 | 15 | 80 | 15 |
| P8 scored_role_bow | P8 | 4 | 4 | 7 | 5 | 9 | 11 | 40 | 7 |
| P9 scored_role_emb | P9 | 12 | 12 | 13 | 11 | 13 | 15 | 76 | 14 |
| P10 scored_role_syn | P10 | 8 | 8 | 8 | 3 | 4 | 2 | 33 | 4 |
| | P11 | 16 | 16 | 16 | 15 | 8 | 1 | 72 | 11 |
| | P12 | 6 | 6 | 5 | 9 | 7 | 12 | 45 | 8 |
| | P13 | 10 | 10 | 10 | 14 | 11 | 13 | 68 | 10 |
| | P14 | 9 | 9 | 9 | 6 | 10 | 6 | 49 | 9 |
| | P15 | 2 | 2 | 2 | 4 | 2 | 7 | 19 | 3 |
| | P16 | 5 | 5 | 4 | 7 | 4 | 8 | 33 | 4 |
| | P17 | 1 | 1 | 1 | 1 | 1 | 4 | 9 | 1 |

Full results (participant rankings)

| average | Name | gt1 | gt1 500 | gt2 | gt3 | gt4 | gt5 | SUM | Ranking |
|---------------------|------|-----|---------|-----|-----|-----|-----|-----|---------|
| | P1 | 17 | 17 | 17 | 17 | 17 | 17 | 102 | 17 |
| | P2 | 14 | 14 | 11 | 12 | 12 | 10 | 73 | 12 |
| P3 scored_full_all | P3 | 5 | 5 | 5 | 4 | 5 | 5 | 29 | 4 |
| P4 scored_full_bow | P4 | 15 | 15 | 15 | 16 | 16 | 11 | 88 | 16 |
| P5 scored_full_emb | P5 | 12 | 12 | 14 | 11 | 13 | 13 | 75 | 14 |
| P6 scored_full_syn | P6 | 3 | 3 | 2 | 3 | 1 | 3 | 15 | 2 |
| P7 scored_role_all | P7 | 13 | 13 | 12 | 14 | 14 | 15 | 81 | 15 |
| P8 scored_role_bow | P8 | 8 | 8 | 8 | 6 | 9 | 9 | 48 | 7 |
| P9 scored_role_emb | P9 | 10 | 10 | 13 | 10 | 15 | 16 | 74 | 13 |
| P10 scored_role_syn | P10 | 6 | 6 | 7 | 5 | 4 | 2 | 30 | 5 |
| | P11 | 16 | 16 | 15 | 15 | 7 | 1 | 70 | 11 |
| | P12 | 7 | 7 | 4 | 9 | 8 | 14 | 49 | 8 |
| | P13 | 11 | 11 | 10 | 13 | 11 | 12 | 68 | 10 |
| | P14 | 9 | 9 | 9 | 8 | 10 | 6 | 51 | 9 |
| | P15 | 2 | 2 | 3 | 2 | 3 | 7 | 19 | 3 |
| | P16 | 4 | 4 | 6 | 7 | 6 | 8 | 35 | 6 |
| | P17 | 1 | 1 | 1 | 1 | 2 | 4 | 10 | 1 |

Highly Relevant Sentences: One type of highly relevant sentences will identify potential sources of significant (large) expenses and/or significant business opportunities. Examples of the source of the expenses or opportunities include litigation, spin-offs, acquisitions, etc. Most of these sentences describe a change from the status quo or current situation. Another type of highly relevant sentence will identify corporate character, e.g., the compensation of senior executives or commentary about business activities.

Relevant Sentences: One type of relevant sentences will identify existing assets, liabilities, revenues, or expenses. They may be very specific, e.g., interest rate expenses. Another type of relevant sentences will also identify the size and nature of current business activities, e.g., retail division, underwriting, investment banking, etc.

Neutral Sentences: These sentences may describe the type of business activity, the location of some business entity or activity. They are informative sentences but convey less information value compared to the highly relevant or relevant sentences.

Irrelevant: This is boilerplate text that is not informative. In some cases, the extracted sentences may be irrelevant to the filing financial entity or the mentioned entity or the role

**Text Features for
Relationship
Ranking**

12/05/2017

Example

(somewhat relevant)

Form: 10-K

Filing: Ally Financial Inc

Mentioned: Computershare Trust Company

Role: Agent

Rating: Neutral, Relevant, Highly Relevant

| Exhibit | Description | Method of Filing |
|---------|--|---|
| 10.11 | Form of Award Agreement related to the issuance of an Ally Leader Equity Participation Award | Filed herewith. |
| 10.12 | Tax Asset Protection Plan dated as of January 10, 2014 between Ally Financial Inc. and <u>Computershare Trust Company, N.A.</u> , as Rights <u>Agent</u> . | Filed as Exhibit 10.1 to the Company's Current Report on Form 8-K dated as of January 13, 2014 (File No. 1-3754) incorporated herein by reference |
| 10.13 | Amendment No. 1 to the Tax Asset Protection Plan, dated February 3, 2015 | Filed as Exhibit 10.18 to the Company's Annual Report for the period ended December 31, 2014, on Form 10-K (File No. 1-3754), incorporated herein by reference. |
| 10.14 | Consent Order Dated December 23, 2013 (Department of Justice) | Filed as Exhibit 10.34 to the Company's Annual Report for the period ended December 31, 2013, on Form 10-K (File No. 1-3754), incorporated herein by reference. |

[Excerpt]

Example (irrelevant)

Form: 10-K,

Filing: Ally Financial Inc

Mentioned: Queens University of Charlotte

Role: Trustee

Rating: 2x Irrelevant

~~Item 10. Directors, Executive Officers, and Corporate Governance~~ ~~Executive Officers and Other Significant Employees~~

~~Jeffrey J. Brown~~ — Chief Executive Officer of Ally since February 2015 [..] Mr. Brown received a bachelor's degree in economics from Clemson University and an executive master's degree in business from Queens University in Charlotte. He serves on the Trevillian Cabinet of the College of Business and Behavioral Sciences at Clemson University and is a Board of Trustees member of Queens University of Charlotte.

~~Christopher Halmy~~ — Chief Financial Officer of Ally since November 2013. [..]

(NORTHERN TRUST CORP - **Guarantor** - NTC Capital)

(RELEVANT) *Guarantee Agreement, dated as of January 16, 1997, relating to NTC Capital I, by and between Northern Trust Corporation, as Guarantor, and The First National Bank of Chicago, as Guarantee Trustee (incorporated herein by reference to Exhibit 4(j) to the Corporation's Current Report on Form 8-K dated January 16, 1997).10.16*

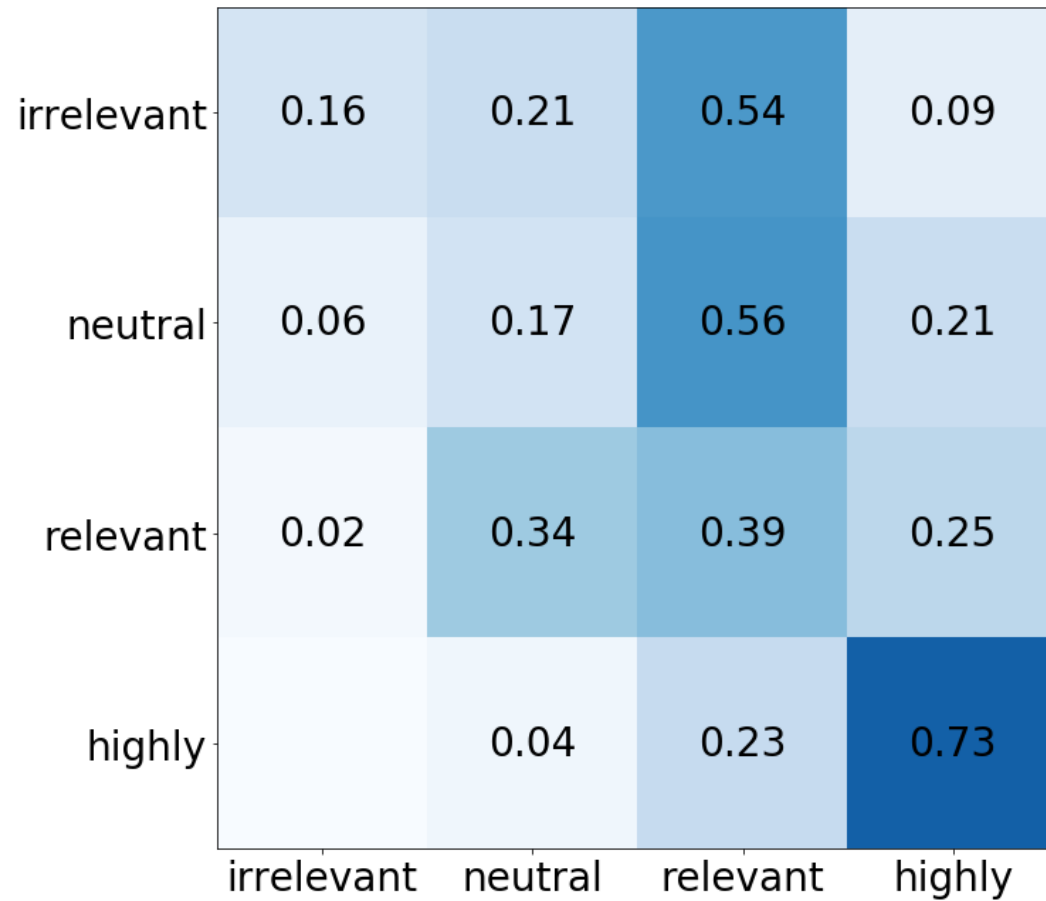
(MORGAN STANLEY - **Trustee** - Morgan Stanley ABS Capital)

(HIGHLY RELEVANT) *On November 6, 2013, Deutsche Bank, in its capacity as trustee, became the named plaintiff in Federal Housing Finance Agency, as Conservator for the Federal Home Loan Mortgage Corporation, on behalf of the Trustee of the Morgan Stanley ABS Capital I Inc. Trust, Series 2007-NC3 (MSAC 2007-NC3) v. Morgan Stanley Mortgage Capital Holdings LLC, and filed a complaint in the Supreme Court of NY under the caption Deutsche Bank National Trust Company, solely in its capacity as Trustee for Morgan Stanley ABS Capital I Inc.*

(STATE STREET CORP - **Issuer** - Equity Securities)

(Discover Financial Services - **Counterparty** - DFS Services LLC)

(PNC FINANCIAL SERVICES GROUP INC - **Affiliates** - Deferred Compensation Plan)



Text Features for Relationship Ranking

12/05/2017