

ECONOMIC EVENT DETECTION IN COMPANY-SPECIFIC NEWS

Gilles Jacobs

gillesm.jacobs@ugent.be

Els Lefever

els.lefever@ugent.be

Véronique Hoste

veronique.hoste@ugent.be

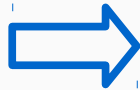
Language & Translation Technology Team (LT3), Ghent University
ATILA 2018, Ghent, Belgium

1. Introduction
2. SentiFM Dataset and Experiments
3. SENTiVENT Dataset
4. Conclusion



1. INTRODUCTION

- 2 annotated resources for event processing in economic news.
- 1 pilot study on economic event detection.
- Research goal: structured factual and sentiment data from economic news text.
 - **Event extraction:** Scenario templates of real-world economic events.
 - **Sentiment analysis:**
 - **Aspect-based** approach: Which opinion is expressed about what? Informed by EE?
 - **Implicit investor** sentiment: Common-sense subjectivity connotations, “polar facts”.
Economic news text contains less explicit sentiment expressions than other genres (Van de Kauter 2015).

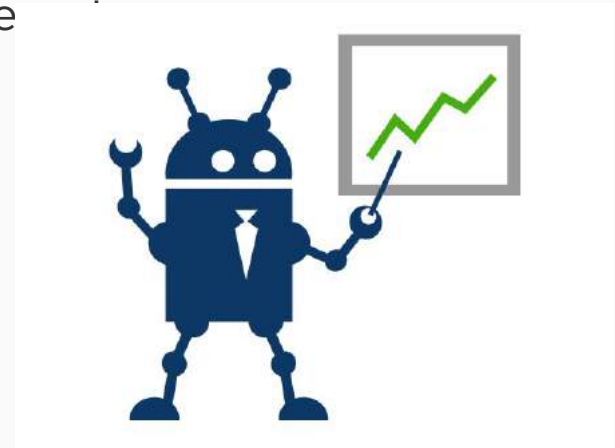


"Unilever's underlying sales rose by some 7 per cent."

Revenue report event with positive sentiment			
Event data		Sentiment data	
Type	Revenue report	Polarity	Positive +
Event expression	"sales rose"	Aspect	Revenue
Company	Unilever	Aspect expression	"sales rose"
Movement	Increase		
Amount	+7%		
Cause	NA		
Money expression	NA		

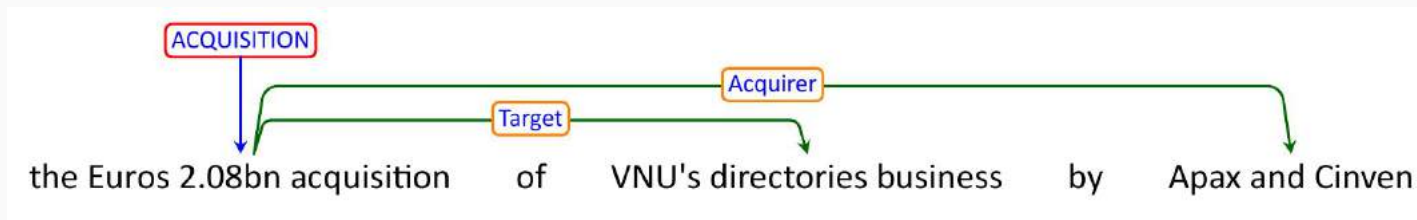
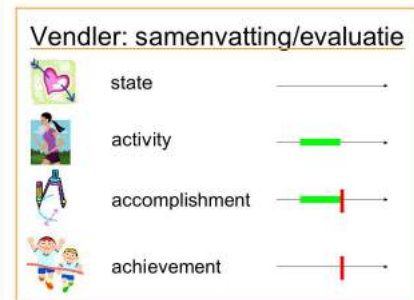
1. INTRODUCTION: APPLICATIONS

- Applications:
 - Economics: academic event studies (MacKinlay, 1997); assessing impact of news events (Boudoukh et al., 2016)
 - Downstream financial applications: Security price prediction, business intelligence, (semi)-automated trading strategies
cf. NLP APIs for market sentiment and (financial) IE: SentDex, OpenCalais, TheySay)



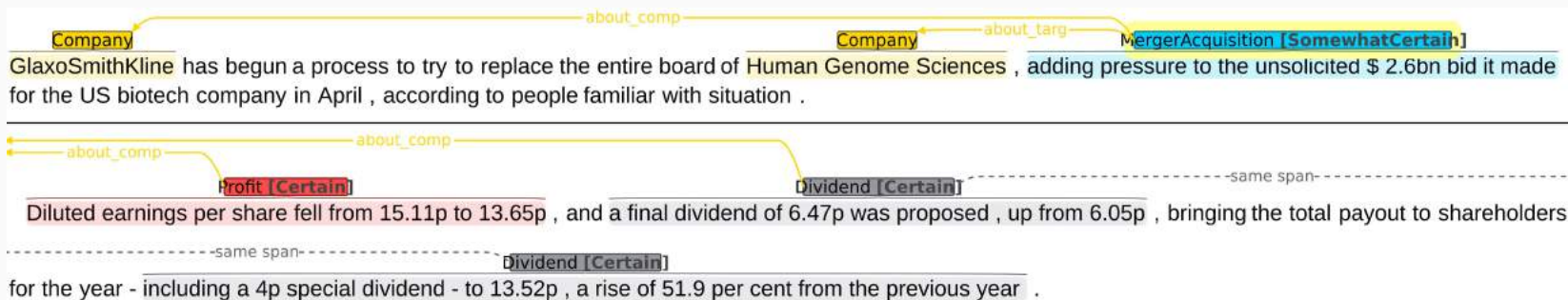
1. INTRODUCTION: EVENT PROCESSING

- What are events?
 - Linguistic and ontological: only dynamic actions or also static situations?
 - No one-to-one linguistic mapping: verbs → eventive, nominals → objects?
 - Vendler (1967) classification based on durativity, dynamicity, and telicity.
- Approaches to event processing in NLP:
 - As topics in document stream (conflated with topic modeling).
 - Temporal processing: time relations (Time-ML, TempEval)
 - **Scenario templates: prototypical semantic schemata of events that capture the content. (MUC-3, ACE, ERE)**



2. SENTIFM PILOT STUDY: DATASET DESCRIPTION

- SentiFM dataset enables **event mention detection and typing** (!= event extraction with argument slots).
- Event conceptualization: “*textually reported real-world occurrences, actions, situations involving companies*”
- **7 companies** selected for sector diversification.
- **2522 event mentions** in **497 news articles** from the Financial Times (2004-2013) in English.



2. SENTIFM PILOT STUDY: DATASET DESCRIPTION

- Dutch counterpart available (Lefever and Hoste, 2016).
- Validity of the annotation scheme was evaluated on Dutch subset:
 - 78.41% Inter-annotator F1-score.
- Event types:
 - 10 event types, typology constructed iteratively on corpus subsample.
 - Type overlap with independently created economic event typologies: StockSonar (Feldman et al., 2011) and SPEED ontology (Hogenboom, 2013).

Event type	Type ratio	# mentions
BuyRating	9.00%	227
Debt	2.38%	60
Dividend	7.22%	182
MergerAcquisition	10.03%	253
Profit	25.81%	651
QuarterlyResults	10.59%	267
SalesVolume	19.31%	487
ShareRepurchase	2.42%	61
TargetPrice	3.73%	94
Turnover	9.52%	240
total	100%	2522

Event type distribution and mention count.

2. SENTIFM PILOT STUDY: INTRODUCTION

- Current methods are pattern- or knowledge-based (Feldman et al., 2011; Arendarenko and Kakkonen, 2012; Hogenboom et al., 2013; Du et al., 2016): largely handmade ontologies.
 - Manual work, captures lexical variation poorly.
- Supervised, data-driven methods: potential to generalize over lexical variation.
 - requires annotated **gold-standard dataset**.
- Currently no resources exist for supervised event detection in economic domain.
 - In general-domain event detection: large amount of resources (e.g. ACE/ERE (TAC-KBP) (Aguilar et al., 2014)).

2. SENTIFM PILOT STUDY: TASK

- Goal: provide a baseline for the dataset.
- Task: Sentence-level event typing, multi-label classification.

“However, revenues from voice and text fell in the period.”

→ Turn-over event

“So far, free cash flow has been used to finance share buybacks and dividend increases.”

→ Share buyback event

“It will increase the number of Barclays’ customers in France by 25 per cent.”

→ Sales Volume event

2. SENTIFM PILOT STUDY: CLASSIFIERS

1) **Feature-engineering + SVM.**

- Lexical features: bag-of-n-gram (token, char, lemma), special token presence.
- Syntactic features: PoS and NER-tags.
- Kernels: linear and RBF.

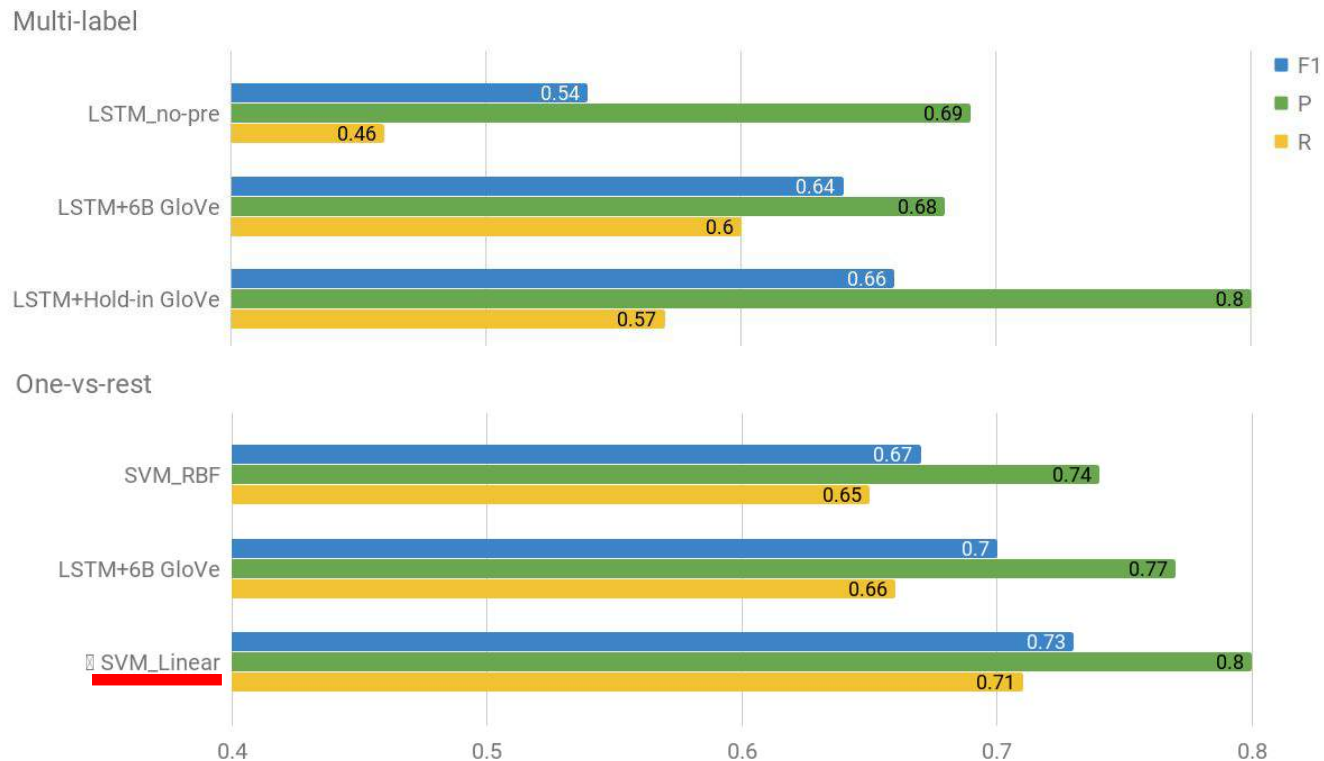
2) **wvec + LSTM:**

- Pre-trained, hold-in set GloVe.
- Pre-trained, 6B corpus GloVe.
- No pre-trained, token-sequence input.
- Pre-trained vectors chosen from multiple candidates by quality evaluation on analogy task.

2. SENTIFM EXPERIMENTS: SET-UP

- Evaluation: support-weighted macro-F1 on 10% random hold-out test.
- Hyper-parameter optimization:
 - SVM-RBF: 5-fold cross-validation, grid-search.
 - SVM-Linear: No optimization, default LibSVM hyper-parameters.
 - LSTM: 3-fold cross-validation, randomized-search (32 it.).
- SVM: one-vs-rest.
- LSTM: multi-label & one-vs-rest (for best input).

2. SENTIFM EXPERIMENT RESULTS: ALL SYSTEMS



2. SENTIFM EXPERIMENT RESULTS: BEST SVM & LSTM SYSTEM, SCORES BY TYPE

Event type	Precision	Recall	F_1 -score
Linear kernel one-vs-rest			
BuyRating	<u>0.95</u>	<u>0.91</u>	<u>0.93</u>
Debt	<u>0.50</u>	<u>1.00</u>	<u>0.67</u>
Dividend	<u>0.62</u>	<u>0.73</u>	<u>0.67</u>
MergerAcquisition	<u>0.56</u>	<u>0.40</u>	<u>0.47</u>
Profit	0.75	0.74	0.75
QuarterlyResults	0.82	0.53	0.64
SalesVolume	0.88	<u>0.75</u>	<u>0.81</u>
ShareRepurchase	<u>1.00</u>	<u>0.50</u>	<u>0.67</u>
TargetPrice	<u>1.00</u>	<u>0.75</u>	<u>0.86</u>
Turnover	<u>0.91</u>	<u>0.77</u>	<u>0.83</u>
avg	<u>0.80</u>	<u>0.71</u>	<u>0.73</u>

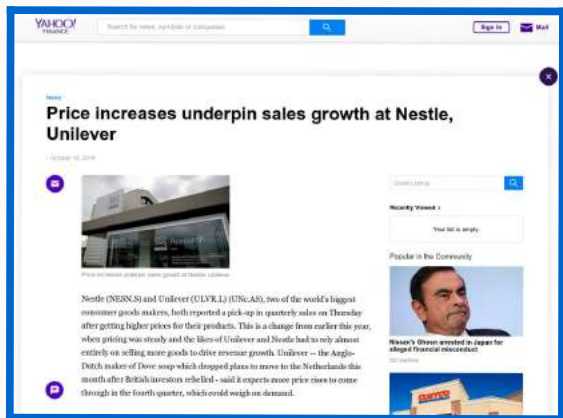
Event type	Precision	Recall	F_1 -score
6B corpus GloVe one-vs-rest			
BuyRating	0.88	<u>0.95</u>	0.91
Debt	0.50	0.50	0.50
Dividend	0.55	0.55	0.55
MergerAcquisition	<u>0.58</u>	<u>0.44</u>	<u>0.50</u>
Profit	0.81	0.74	0.77
QuarterlyResults	0.84	0.47	0.60
SalesVolume	0.81	<u>0.76</u>	0.79
ShareRepurchase	0.75	0.50	0.60
TargetPrice	<u>1.00</u>	<u>1.00</u>	<u>1.00</u>
Turnover	0.94	0.65	0.77
avg	0.77	0.66	0.70

2. SENTIFM EXPERIMENTS: LIMITATIONS

- Corpus collection:
 - Keyword-search to retrieve articles per event type: introduces lexical bias.
 - Type coverage could be improved: 18 event types in Boudhouk et al. (2016) vs. our 10 types.
- Baseline classification too elementary:
 - Task: Our data-set allows for token-span event mention detection instead of sentence-level.
 - Straight-forward classifiers: much room for improvement and more advanced approaches.

3. SENTIMENT ENGLISH EVENT CORPUS

- New corpus for event extraction with participants:
 - Yahoo Finance news aggregator: Randomly crawled for all companies in the S&P 500.
 - Selected 30 companies for based on sector diversification and reporting frequency.
 - Corpus article time-span: 11 months (April 2017 – March 2018).



"Unilever's underlying sales rose by some 7 per cent."

Revenue report event with positive sentiment			
Event data		Sentiment data	
Type	Revenue report	Polarity	Positive +
Event expression	"sales rose"	Aspect	Revenue
Company	Unilever	Aspect expression	"sales rose"
Movement	Increase		
Amount	+7%		
Cause	NA		
Money expression	NA		

1. EVENT EXTRACTION: AS SEMANTIC FRAMES

- Events are prototypical scenarios conceptualized similar to semantic frames.
 - Trigger: lexical unit that evokes the event.
 - Type, subtype: domain-specific event typology of economic events.
 - Arguments: participants in the event: who/what is involved with what prototypical semantic role?
 - Attributes: modality (certain/other) and polarity (negated/asserted).



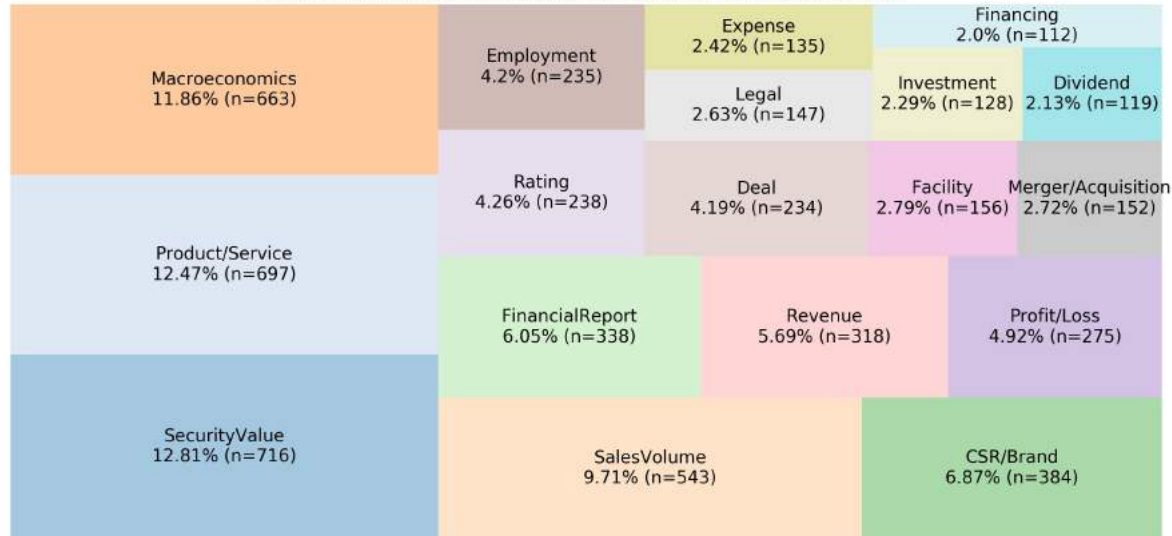
4. SENTIVENT ENGLISH: EVENT ANNOTATION

- Event annotation guidelines:
 - Adapted from Rich ERE guidelines.
 - Company-specific, business news event typology: Developed in cooperation with financial domain expert.
- 18 types, 43 subtypes: Balance coverage, distinctiveness, and descriptiveness.
- Annotation done in WebAnno web-based annotation tool.

The screenshot displays the WebAnno web-based annotation tool interface. At the top, there is a red navigation bar with the text "Annotation" and "Home". Below this is a toolbar with various icons for document, page, script, help, and workflow management. The main area shows a news article snippet: "AMAZON EXPECTS TO HIRE AT LEAST 1,000 AT ITS NEW TROUTDALE FACILITY - PORTLAND BUSINESS JOURNAL". The text is annotated with various event types and roles, such as "Employment Start", "Facility Open", and "CoreEvent". The interface also includes a sidebar on the right with a "Layer" dropdown set to "d_Discontiguous" and a "Forward annotation" checkbox. The bottom right corner of the interface shows "Showing 1-16 of 16 sentences [document 2 of 334]" and "No annotation selected!".

4. SENTIVENT ENGLISH: CORPUS STATISTICS

Distribution of event categories in SENTIVENT English corpus.



- 300 annotated articles containing 6000 events (approx. 170.000 tokens).
- Macroeconomics: largest types is a left-over class: re-conceptualize this class, split this into multiple classes.
- Other type frequency along the line one would expect in news reporting.

4. SENTIMENT ENGLISH: INTERANNOTATOR STUDY

- 3 annotators annotated 30 randomly selected documents.
- Event trigger span detection:
 - Token-based agreement:
 - Fleiss' kappa: 0.57 κ : moderate agreement (Landis & Koch 1977)
 - Pairwise F1-score: 0.60
- Pairwise event nugget score (Liu et al. 2015):
 - Incorporates dice coefficient of partial span matches.
 - **Pairwise Span F1-score: 0.61**
 - Attributes (type, subtype, modality, polarity): All attribute F1-score: 0.44
 - **Type accuracy: 0.88**
 - **Subtype accuracy: 0.82**
 - Modality accuracy: 0.88
 - Polarity accuracy: 0.99

Event trigger span detection is difficult: unsurprising in semantic annotation.

Typology and attributes report good accuracy.

4. CONCLUSION

- We presented 2 of the first annotated for event processing in the business news-domain.
 - SentiFM economic event typing dataset.
 - SENTiVENT economic extraction dataset.
- A pilot study on event typing with satisfactory results.
 - SVM 73% F1-score
- Future research will focus on full event extraction with participants and attributes.
 - Lexical variation in event types: Knowledge-base and semi-supervised extensions to supervised methods.

REFERENCES

- Gilles Jacobs, Els Lefever, and Véronique Hoste. 2018. Economic event detection in company-specific news text. In Proceedings of the 1st Workshop on Economics and NLP (ECONLP). ACL 2018, Melbourne, AUS, 1-10.
- Craig MacKinlay. 1997. Event Studies in Economics and Finance. *Journal of Economic Literature*, 35(1):13–39.
- Jacob Boudoukh, Ronen Feldman, Shimon Kogan, and Matthew P Richardson. 2016. Information, trading, and volatility: Evidence from firm-specific news. SSRN Working Paper 2193667.
- Ronen Feldman, Benjamin Rosenfeld, Roy Bar-Haim, and Moshe Fresko. 2011. The stock sonarsentiment analysis of stocks based on a hybrid approach. In Twenty-Third IAAI Conference.
- Ernest Arendarenko and Tuomo Kakkonen. 2012. Ontology-Based Information and Event Extraction for Business Intelligence. In *Artificial Intelligence: Methodology, Systems, and Applications*, volume 7557 of Lecture Notes in Computer Science, pages 89–102. Springer.
- Alexander Hogenboom, Frederik Hogenboom, Flavius Frasinca, Kim Schouten, and Otto van der Meer. 2013. Semantics-Based Information Extraction for Detecting Economic Events. *Multimedia Tools and Applications*, 64(1):27–52.
- Mian Du, Lidia Pivovarova, and Roman Yangarber. 2016. PULS: natural language processing for business intelligence. In Proceedings of the 2016 Workshop on Human Language Technology, pages 1–8.
- Jacqueline Aguilar, Charley Beller, Paul McNamee, Benjamin Van Durme, Stephanie Strassel, Zhiyi Song, and Joe Ellis. 2014. A comparison of the events and relations across ACE, ERE, TAC-KBP, and FrameNet annotation standards. In Proceedings of the Second Workshop on EVENTS: Definition, Detection, Coreference, and Representation, pages 45–53.
- Els Lefever and Veronique Hoste. 2016. A classification-based approach to economic event detection in dutch news text. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC '16), pages 330–335. European Language Resources Association (ELRA).