

企業経営における意思決定支援のためのイベント抽出

Event Extraction for Decision Making Support in Corporate Management

柳瀬 利彦 *1
Toshihiko Yanase柳井 孝介 *1
Kohsuke Yanai丹羽 芳樹 *1
Yoshiki Niwa村上 聡一郎 *2
Soichiro Murakami渡邊 亮彦 *2
Akihiko Watanabe宮澤 彬 *3
Akira Miyazawa五島 圭一 *2
Keiichi Goshima高村 大也 *2
Hiroya Takamura宮尾 祐介 *3
Yusuke Miyao中田 亨 *4
Toru Nakata*1株式会社 日立製作所
Hitachi, Ltd.*2東京工業大学大学院
Tokyo Institute of Technology*3総合研究大学院大学/国立情報学研究所
The Graduate University for Advanced Studies/National Institute of Informatics*4産業技術総合研究所
The National Institute of Advanced Industrial Science and Technology

This paper discusses decision making support for corporate management. Our research goal is to develop a system that investigates past decisions as clues to solve current problems. In this paper, we conducted research on event extraction in order to collect such decisions from news articles. At first, we developed a dataset of events in a corporate management domain with two human annotators. Then, we showed results of preliminary experiments on two baseline methods.

1. はじめに

本研究では過去の経営判断の事例を現在の企業活動に活かすことを目標とする。日々のニュースには、様々な企業の経営判断が具体的なイベントとして記述されている。それらを抽出し、蓄積することで、過去の判断を自在に参照することができるようになる。本稿ではその要素技術となるニュース記事から経営判断を伴うイベントの抽出方法について論じる。

一般的なイベント抽出とは異なり、イベントに経営判断が含まれるかを見分ける必要がある。ユースケースから判定基準を考える。例えば、次の文は事業構造改革を検討する際に過去事例として利用できる：

Japanese electronics maker *Hitachi* will spin off its home appliance and industrial equipment divisions by April to become quicker in decision-making to respond to market changes.*1

一方で、次の文は単なる株価の説明であり、企業の意思決定の事例としては利用できない：

Hitachi gained 20 yen to 1,360, with 14.5 million shares traded.*2

一般に株価は市場により決められるものであり、企業が直接決めるものではないため、意思決定ではないと判断する。以上のように、本研究では、企業が能動的に影響力を発揮できる企業活動を抽出対象として定めた。そうした企業活動には、営業、法務、CSR、企業統治、知的財産、提携、調達、生産、広報、研究開発などがある。

本研究では、ニュース記事から抽出可能なイベントを明らかにするため、人手でデータセットを開発した。さらに、このデータセットを用いて、教師あり学習による自動抽出方法を評価し、完全な自動化に向けた課題を議論した。

連絡先: 柳瀬利彦, 株式会社 日立製作所 基礎研究センタ,
toshihiko.yanase.gm@hitachi.co.jp

*1 [Napoles 12] APW_ENG.20010927.0173, L3 より引用

*2 [Napoles 12] NYT_ENG.19970802.0223, L12 より引用

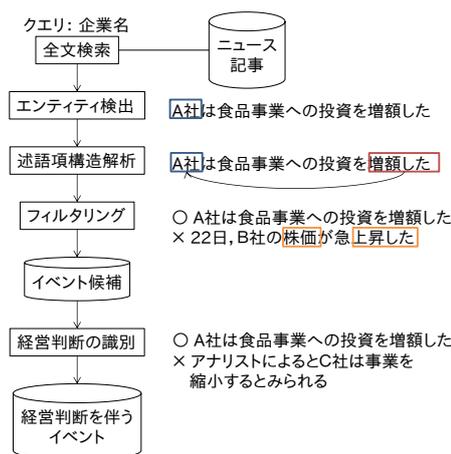


図 1: 経営判断を伴うイベント抽出の流れ

2. データセットの開発

これまで、人や組織のイベント [Tannier 13, Minard 15], 商品やサービスのイベント [阿部 09], タンパク質の反応に関するイベント [Kim 13]などを対象に、テキストからのイベント抽出が研究されてきた。抽出する目的やドメインが異なるため、それぞれ別のデータセットを構築し、研究に利用している。一方で、これらの研究に共通しているのは、イベントを、動作主となるエンティティと、何が起こったのかを表す述部の組として表現している点である。本研究でもこのイベントの表現に倣って、企業経営というドメインに適したデータセットを新たに開発する。

経営判断を伴うイベント抽出の流れを図 1 に示す。ルールベースの手法に基づくイベント候補の抽出と、その後の人手および機械による経営判断の有無の識別というふたつの部分に分けられる。

表 1: 自動抽出の結果

手法	Accuracy	Precision	Recall	F1
Majority Baseline	78.7	78.7	100	88.0
Linear SVM	82.3	86.0	92.7	89.2
fastText	85.0	87.9	93.8	90.7

2.1 イベント候補の抽出

はじめに、企業名で全文検索を行い、企業に関連するニュース記事を集めた。ニュース記事として Annotated English Gigaword [Napoles 12] を用いた。7 紙のニュース約 900 万文獻が含まれている。対象とする企業はインデックス銘柄 (ダウ工業株) に選ばれている企業から 5 社 (IBM, Caterpillar, Microsoft, Pfizer, Wal-Mart Stores) を選択した。有名企業であり、ニュース記事での言及が多いためである。次に、記事全文に対して DBPedia Spotlight [Daiber 13] を適用し、エンティティを同定した。この処理により異表記でのエンティティへの参照 (Big Blue と IBM) や多義性が解消される。続いて、Enju [Miyao 02] を用いて文の述語項構造解析を行い、企業が動作主として登場する文を抽出した。最後に、ルールベースのフィルタリングを行った。ニュースには、その日の株価のレポートなど、企業の自発的な判断とは明らかに関係のない記事がある。そうした記事に高い頻度で登場する語を辞書 (株価表現, 上昇下降表現) に登録した。そして、一文中に 2 つの辞書の表現がともに出現する文を候補から除いた。

2.2 経営判断の有無のアノテーション

イベントの各候補に対して、経営判断が含まれるか否かを人手により判定する。英語の母国語話者二人により、各企業につき 200 文、合計 1,000 文に対してアノテーションを付与した。999 文について二人から回答があり、そのうち判断が一致した文は 818 文、一致しなかった文は 181 文であった。一致した文のうち、経営判断を含むものは 644 文、含まないものは 174 文であった。アノテータ間のアグリーメントは、Fleiss の κ で 0.535 であった。

全体のうち半数以上に経営判断が含まれるため、イベント候補の段階でもアプリケーションを適切に選択すれば実用可能であると考えられる。例えば、情報推薦などユーザが最終的な判断をするようなアプリケーションには利用できる。一方で、質問応答のための知識ベースなど、一つの事例を回答する場合にはさらなる正確性が求められる。

3. 経営判断の有無の自動識別

判断の一致した文 818 文を用いて経営判断の有無の自動識別を評価した。Bag-of-words (BoW) による特徴量と線形 SVM [Fan 08] を用いた手法と、Neural Network に基づく手法 fastText を採用した [Joulin 16]。fastText は学習・識別の計算時間が短いことを特徴としており、かつ CNN や LSTM 等のより複雑なモデルと比べて複数のベンチマーク問題で同程度の精度を示している。fastText では、事前に Gigaword 全文を用いて Skip-gram により 100 次元の単語の埋め込み表現 [Mikolov 13] を求めた。ふたつの手法とも、単語の分かち書きには Stanford Core NLP [Manning 14] を利用した。

5-fold の交差検定によって評価した。結果を表 1 に示す。値はマクロ平均である。Accuracy でみると、Majority Baseline が 78.7% であるのに対し、線形 SVM は 4.1 ポイント、fastText は 6.3 ポイントの改善が見られる。すべてを真と識別する Majority Baseline と比べれば Recall が下がるのは当然であるため、改

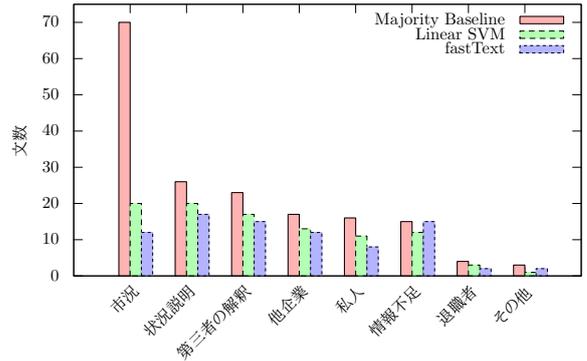


図 2: 誤抽出の内訳

善は Precision によってもたらされる。そこで、誤抽出 (偽陽性) の例を見ることで改善した点と課題を調査する。

アノテータによる識別理由のコメントをもとに、全ての負例の判定理由を 8 種類 (市況, 状況説明, 第三者の解釈, 他企業, 私人, 情報不足, 退職者, その他) に分類した。図 2 に各種法の誤抽出の内訳を示す。ここで、第三者の解釈とは、アナリストなど対象企業に属さない人による行動の予測や説明を意味する。また、私人とは、役員や従業員が個人としておこした行動であり、企業の代表が有名人 (例えば Microsoft の Bill Gates 氏) の場合に多い。

市況の誤抽出は、Majority Baseline では約 40% と大きな割合を占めている。SVM, fastText はそれぞれ約 2/3, 約 4/5 と大きな割合のエラーを削減している。以下のような文が市況の典型である:

Pharmaceutical issues advanced, led higher by Pfizer, which climbed 56 cents to \$45.25.*3

特徴的な単語 (climbed, cents や数値) が存在するため、その共起を見ることで識別できる。SVM や fastText は、ルールベースフィルタリングで漏れた表現を教師データから補ったのだと考えられる。特に、単語の埋め込み表現を用いる fastText は、教師データの類義の表現も識別できるため改善幅が大きくなっている。

一方で、状況説明は、市況ほど大きな改善は見られていない。例えば、SVM と fastText がともに誤抽出した状況説明の例として

Compaq Computer Corp. and International Business Machines Corp. both increased market share, the study said.*4

がある。“increase market share” は能動的な表現であるが、そこに判断は含まれていない。こうした文を見分けるには、述部の目的語 (market share) について企業が操作できるかどうかの知識が必要となる。

以上のように、精度の改善が見られたのは主に株価を解説した文の排除による。さらなる改善には個々の語の知識が必要になる。

4. おわりに

本研究では、ニュース記事から企業の経営判断を伴うイベントを抽出する方法について議論した。二人のアノテータにより 818 文の経営判断の有無を識別するデータセットを構築し、自動抽出に向けた評価を実施した。

開発したデータセットは教師あり学習のためには規模が小さいため、今後の課題はデータセットを拡充することである。

*3 [Napoles 12] APW_ENG.20001227.0677, L16 より引用

*4 [Napoles 12] NYT_ENG.19961028.0384, L25 より引用

参考文献

- [Daiber 13] Daiber, J., Jakob, M., Hokamp, C., and Mendes, P. N.: Improving Efficiency and Accuracy in Multilingual Entity Extraction, in *Proceedings of the 9th International Conference on Semantic Systems (I-Semantics)* (2013)
- [Fan 08] Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., and Lin, C.-J.: LIBLINEAR: A Library for Large Linear Classification, *J. Mach. Learn. Res.*, Vol. 9, pp. 1871–1874 (2008)
- [Joulin 16] Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T.: Bag of Tricks for Efficient Text Classification, *arXiv preprint arXiv:1607.01759* (2016)
- [Kim 13] Kim, J.-D., Wang, Y., and Yasunori, Y.: The Genia Event Extraction Shared Task, 2013 Edition - Overview, in *Proceedings of the BioNLP Shared Task 2013 Workshop*, pp. 8–15, Sofia, Bulgaria (2013)
- [Manning 14] Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., and McClosky, D.: The Stanford CoreNLP Natural Language Processing Toolkit, in *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55–60 (2014)
- [Mikolov 13] Mikolov, T., Chen, K., Corrado, G., and Dean, J.: Efficient Estimation of Word Representations in Vector Space, *Proceedings of Workshop at International Conference on Learning Representations* (2013)
- [Minard 15] Minard, A.-L., Speranza, M., Agirre, E., Aldabe, I., Erp, van M., Magnini, B., Rigau, G., and Urizar, R.: SemEval-2015 Task 4: TimeLine: Cross-Document Event Ordering, in *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, pp. 778–786, Denver, Colorado (2015)
- [Miyao 02] Miyao, Y. and Tsujii, J.: Maximum Entropy Estimation for Feature Forests, in *Proceedings of the Second International Conference on Human Language Technology Research, HLT '02*, pp. 292–297, San Francisco, CA, USA (2002)
- [Napoles 12] Napoles, C., Gormley, M., and Durme, B. V.: Annotated English Gigaword LDC2012T21, AKBC-WEKEX '12, pp. 95–100 (2012)
- [Tannier 13] Tannier, X. and Moriceau, V.: Building Event Threads out of Multiple News Articles, in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pp. 958–967, Seattle, Washington, USA (2013)
- [阿部 09] 阿部 修也, 江口 萌, 隅田 飛鳥, 大崎 梓, 乾 健太郎: みんなの経験: ブログから抽出したイベントおよびセンチメントのDB化, 言語処理学会 第15回年次大会 (2009)