

事業セグメントに着目した有価証券報告書からの文脈抽出とキーワード生成による類似企業検索

國吉 房貴 井本 稔也

Japan Digital Design 株式会社

{fusataka.kuniyoshi, toshiya.imoto}@japan-d2.com

概要

事業セグメントに着目した類似企業の特定は、ポートフォリオの構築や、資産価格や債券金利の適切な設定など金融機関の主要業務に関わる重要な課題とされている。類似企業の特定は、各社の事業セグメントに基づき異なる粒度で各企業に独自のカテゴリーを割り当てた業種分類に基づいて行われる。しかし、複数の業種にまたがる事業をもつ企業の場合、ひとつの企業に複数の業種を割り当てることができないため、類似企業を適切に特定することが困難である。そこで本研究では、有価証券報告書で報告された事業セグメントの記述よりキーワードを抽出し、その周辺のコンテキストを含めてベクトル化することで、多様な意味や潜在的な文脈を抽出する手法と、大規模言語モデルの内部知識に基づき企業の一般的な業種イメージをキーワードとして表現する手法を組み合わせるハイブリッド検索手法を提案した。これにより、従来の業種分類では困難だった、事業セグメントを考慮した類似企業検索を実現した。評価実験により、業種抽出性能と株式投資のリターン相関の評価において、提案モデルがベースラインモデルを上回る性能を示した。

1 はじめに

類似企業の分析は、企業を公平に比較し、パフォーマンスを比較することにつながるため、金融機関において重要なタスクとされている。類似企業の分析が迅速かつ正確に実施可能になると、特定された類似の企業グループに対して金融政策や規制変更がどのような影響を与えるかを分析できるようになり、また、同一のセクターに属する企業間で財務データの異常値を検出し、これらのデータを用いてポートフォリオのリスクを最小化し、最適な投資ポートフォリオを構築できるようになるため、その

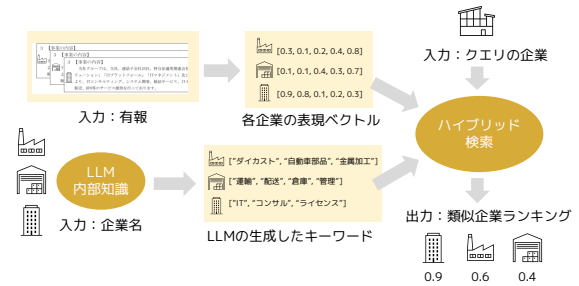


図1 提案モデルの概要

実現が期待されている。

類似企業の検索は、日本取引所グループ (Japan Exchange Group, JPX) によって作成された東証 33 業種の業種分類システムなどを用いることが一般的である¹⁾。一方で、業種分類システムを用いた類似企業の特定では、複数の業種にまたがって活動する企業を見逃してしまう可能性があり、また、流動的に変化する企業の事業セグメントを捉えられない可能性がある。例えば、小売業にはオンラインストアや実店舗も含まれていることがあり、事業再編により事業や企業の構成・編成が作り直されることもある。さらに、異なるデータベンダーは、同じビジネスに異なるコードを割り当てる可能性がありその定義に揺らぎが発生することがある [1]。

このように、従来の業種分類システムは、企業の業種を特定するために広く用いられている一方、類似企業の特定に用いるには限界があることは明らかである。このため、金融機関では、より解像度の高い企業分析を実現することを目的として、データベンダーから提供されるデータセットを超えて企業の類似性を比較可能な手法が求められている。

有価証券報告書 (有報) は、日本の上場企業が年次で提出する法定開示書類であり、事業の状況、財務情報、リスク情報、企業ガバナンス、株主情報などが記載されている。このため、有報は従来の手法

1) <https://www.jpx.co.jp/>

では捉えにくかった企業同士の関係を抽出するためのデータソースとして注目されており、様々な情報抽出手法が提案されている [2, 3, 4].

しかしながら、事業セグメントに関する情報は、数値データとしてではなく、テキストデータとして記載されていることが主なため、金融工学の分野で提案される財務情報への適用が想定された数理モデルを直接適用し分析を行うことが困難である。

そこで本研究では、自然言語処理を用いて有報に含まれるテキストデータより抽出した事業セグメントを用いて企業の類似性を比較可能な手法を提案する。本研究の貢献は次のとおりである。図 1 に提案モデルの概要を示す。

1. 有報全体より事業セグメントを表す文脈を含むチャンクを抽出する文脈ベクトル検索を提案。
2. 大規模言語モデル (LLM) を用いて企業のもつ業種イメージをキーワード (KW) として取得する生成 KW 検索を提案。
3. 業種抽出性能と株式投資におけるリターンの相関の評価において、提案モデルのハイブリッド検索モデルがベースラインモデルを上回る性能を示した。

2 データセット

本研究で用いるデータセットは有報から抽出したテキストで構成されている。本章では、有報の取得と前処理を説明する。

2.1 有報の取得

EDINET API ²⁾ を用いて有報を取得した。2021 年 4 月 1 日から 2024 年 3 月 31 日の期間で docID を取得し、EDINET コードを紐付けて証券コードとの対応をとった。最後に、docID を用いて EDINET から有報を zip でダウンロードし、抽出した情報を単一のテーブルデータとして結合した。

2.2 前処理

事業セグメントに着目した企業間比較を行うため、EDINET より取得した有報から「事業の内容」と「経営者による財政状態、経営成績及びキャッシュ・フローの状況の分析 (以下、MD&D)」の章を抽出した。MD&D に関しては、事業セグメントが記載されていることが直接明記されていないが、

実際には事業セグメントに関する記述が多く存在するため、これも抽出対象に含めた。また、「事業の内容」か「MD&D」あるいは「提出者業種」が空の企業は抽出の対象外とした。「上場区分」の列に「上場」と記載されている企業のみを抽出した。最後に、各企業ごとに「提出日」の列が最新の日付の有報のみに絞り込み、3,880 社分のテキストデータを抽出し、これをデータセットとした ³⁾。

3 有報からの文脈抽出と KW 生成

提案手法は、文脈抽出と KW 生成のハイブリッド検索モデルであり、企業名を入力として、類似企業をランキングとして出力する。本章では、要素技術について個別に説明する。

3.1 文脈ベクトル検索

有報の各章のテキスト部分を D_k と定義する。特に、「事業の内容」と「MD&D」の章はそれぞれ、 D_{biz}, D_{mdnd} と記載する。抽出した章に KeyBERT 法 [5] を適用する。KeyBERT 法は、文書 D の埋め込みベクトル $\mathbf{e}_D = \text{BERT}(D)$ を計算する。次に、文書 D より KW の候補リスト $\{K_1, K_2, \dots, K_n\}$ を抽出し、同様に、各候補 K_i の埋め込みベクトル $\mathbf{e}_{K_i} = \text{BERT}(K_i)$ $i = 1, 2, \dots, n$ を計算する。最後に、 \mathbf{e}_D と各候補 \mathbf{e}_{K_i} の類似度: $S(K_i) = \text{sim}(\mathbf{e}_D, \mathbf{e}_{K_i})$ の高さに基づいて適切な KW を抽出する。

文書 D_{biz} から KeyBERT 法で抽出された KW の集合を $\{K_1, K_2, \dots, K_m\}$ とする。また、 D_{biz}, D_{mdnd} のテキストをそれぞれチャンキングし、得られたチャンクの集合を $\{C_1, C_2, \dots, C_n\}$ とする。

$$\text{KeyBERT}(D_{biz}) = \{K_1, K_2, \dots, K_m\} \quad (1)$$

$$\text{Chunk}([D_{biz}, D_{mdnd}]) = \{C_1, C_2, \dots, C_n\} \quad (2)$$

ここで、事業セグメントに関連する KW が抽出されやすくなるため、KeyBERT 法を適用する章を D_{biz} のみとする。

チャンクの集合 $\{C_1, C_2, \dots, C_n\}$ のうち、KeyBERT で抽出された KW を含むチャンクの部分集合を Chunk' と定義する。

$$\text{Chunk}' = \{C_i \mid K_j \in C_i, \forall i \in \{1, 2, \dots, n\}, \exists j \in \{1, 2, \dots, m\}\} \quad (3)$$

3) 本研究で用いたデータセットは次のドライブよりダウンロード可能である:<https://drive.google.com/file/d/1yxt5QZ3cF0y300SE-SIHhNJeyJSI951S/view?usp=sharing>

2) <https://disclosure2.edinet-fsa.go.jp>

$C' \in \text{Chunk}'$ は K_j を含み、その KW 前後の文脈を含むチャンクである。さらに、得られたチャンク C' を Sentence BERT (SBERT) [6] に入力し、埋め込みベクトル $e_{C'}$ を得る。

$$e_{C'} = \text{SBERT}(C'), \quad C' \in \text{Chunk}' \quad (4)$$

本手法の重要な点は、文書から得られたチャンクの中から KW を含むチャンクに絞り込んで抽出したことにより、KW に関する文脈を効果的に取り入れた点にある。

3.2 生成 KW 検索

本研究では、有報から企業のもつ一般的な事業内容を KW として抽出するため、LLM を用いて KW を生成する手法を採用した。我々は生成した KW の類似度を計算するため、Word Mover's Distance (WMD) [7], BM25 Okapi [8], Jaccard 係数, 重み付き Jaccard 係数の 4 つの類似度関数を用いた。各関数を付録の章 A.3 で説明する。また、4 つの類似度関数の出力を融合するとき、生成 KW 検索と文脈ベクトル検索の出力結果を融合するときのランキング融合アルゴリズムとして Ranked Retrieval Fusion (RRF) を用いた。RRF を付録の章 A.4 で説明する。

4 評価実験

企業同士がどの程度類似しているかを定量的に評価する手法は確立されていない。我々は、業種分類と株式投資のリターンに基づく評価により提案手法の性能評価を試みた。

4.1 設定

D_{biz} に KeyBERT 法を適用して得られた KW をクエリに用いた通常 KW 検索 (k.1), 有報から抽出された章 ($[D_{\text{biz}}, D_{\text{mdnd}}]$) に Sentence BERT を適用して得られたベクトルをクエリに用いた通常ベクトル検索 (v.1) を用意し、これらをベースラインモデルとした。提案手法として生成 KW 検索 (k.2) と文脈ベクトル検索 (v.2) を用意し、これらのハイブリッド検索 (hyb) を評価した。KW を生成する際は gpt-4o⁴⁾ を用いた。文脈抽出の際は、KeyBERT 法で得られた類似度上位 7 位の KW を用いた。チャンキングには、RecursiveCharacterTextSplitter⁵⁾ を用いた。チャンクの最大サイズを 64 文字、チャ

4) <https://openai.com>

5) <https://python.langchain.com>

ンク間のオーバーラップを 4 文字に設定し、長さを測る関数として len 関数を用いた。KeyBERT と Sentence BERT の日本語埋め込みモデルとして unsup-simcse-ja-large [9] を、WMD の単語埋め込みモデルとして日本語 Wikipedia エンティティベクトル [10] を用いた。株価は Yahoo! Finance⁶⁾ を介して取得した。入力テキストは章 2.2 により取得した 3,880 社の有報より抽出した。評価の際は、3,880 社からクエリの 1 社を選択し、残りの 3,799 社から類似企業を検索した。一般的に、時価総額の大きい企業の類似企業は、非専門家でも容易に推測できる可能性が高いため、時価総額が 10 億円を下回る企業にクエリを絞り込み、10 の業種より 1 社ずつ選択し評価した。各検索モデルのクエリとした企業を付録の表 A.3 に示す。以降企業は証券コードで示す。

4.2 業種抽出性能に基づく評価

EDINET コードマスタの業種分類情報を基準として使用し、提案手法によって特定された類似企業と同じ業種に分類される割合を評価指標とした。クエリの企業に対して特定された上位 20 社の類似企業について、nDCG (Normalized Discounted Cumulative Gain) を比較した。

4.3 リターンの相関に基づく評価

本研究は先行研究 [11] に基づき、類似企業は通常、長期的な視点で株式投資のリターンの類似した動きが現れると仮定し、2023 年 4 月 1 日から 2024 年 3 月 31 日における各企業の株式リターンデータを用いて、2 銘柄 i と j の間の日次リターンのペアワイズ相関 ρ_{ij} を計算し、最も近い k 個の近傍企業のリターンとの平均ペアワイズ相関 $\bar{\rho}_i = \frac{1}{k} \sum_{j=1, j \neq i}^k \rho_{ij}$ を評価指標とした。ここで、クエリの企業と同一業種の 20 社をランダムに選択しそれらを類似企業とするナイーブな手法 (rand) をベースラインモデルに追加する。rand において、同一業種が 20 社に満たない場合は業種が同じ企業を類似企業とした。

4.4 分析

表 1 にクエリの企業と hyb によって特定された最も類似度の高い企業のペアを示す。生成 KW と通常 KW はそれぞれ LLM で得た KW, KeyBERT 法を「事業の内容」に適用して得た KW を指す。クエリの企業の業種が「水産農林」であるにも関わらず、

6) <https://finance.yahoo.com/>

表 1 提案手法 (hyb) によって特定された類似企業. 左がクエリの企業, 右が最も類似度の高い企業. 生成 KW を黄で示し, KeyBERT 法で得られた通常 KW を緑で示す.

証券コード	1380	2060
業種	水産農林	食料品
生成 KW	農業, 畜産, 食品 , 有機, 鶏肉, 卵 , 乳製品	飼料 , 畜産, 水産 , 農業, 養殖, ペットフード, 家畜
通常 KW	ファーム, 生協, 農場, 直販 , 牧場, 食品 , 加工	飼料 , 牧場, 委託, 仕入 , 農場, 畜産, ファーム
事業の内容	... 農薬・化学肥料・抗生物質等の薬物・化学添加物に頼らない安全な食肉, 加工食品 , 鶏卵 , 牛乳, 一般 食品 等に関連する事業活動を展開しております。(中略)(1) 生産卸売事業主要な製品は, 食肉, 加工食品 , 鶏卵 及び牛乳等であります。(中略)(2) 直販 事業 生産卸売事業等において製造された製品と、外部取引先より仕入れた商品を会員様に販売しております。	... 飼料 事業として配合 飼料 の製造・販売、食品事業として豚などの飼育・ 仕入 ・販売及び食肉・加工品の 仕入 ・販売、鶏卵の 仕入 ・生産・加工・販売、 水産 物の 仕入 ・販売等、その他の事業活動を展開しております。(中略) 水産 物事業関連... 当社は、 水産 物の 仕入 ・販売等を行っております。(中略) その他... 当社は、畜水産機材及び不動産賃貸等の事業を行っております。

表 2 業種抽出性能の評価結果. 企業ごとに最も nDCG の高い値を太字で示し, 2 番目に高い値に下線を引く.

企業	業種	k.1	k.2	v.1	v.2	hyb
8365	銀行業	1.00	1.00	1.00	1.00	1.00
4667	情報通信	.558	<u>.935</u>	.634	.929	.991
4579	医薬品	.998	.998	.932	.878	<u>.992</u>
3286	不動産業	<u>.993</u>	.918	.430	.965	.898
6408	機械	.886	.526	.330	.535	<u>.840</u>
9514	電気ガス	.853	<u>.774</u>	.431	.449	.724
9206	空運業	.922	1.00	0.00	<u>.967</u>	1.00
9849	卸売業	.917	.766	.720	<u>.790</u>	.562
1380	水産農林	.315	.518	0.00	0.00	<u>.337</u>
7709	精密機器	<u>.335</u>	0.00	0.00	.370	.250

表 3 日次リターンの相関値. 企業ごとに最も相関値の高い値を太字で示し, 2 番目に高い値に下線を引く.

企業	業種	rand	k.1	k.2	v.1	v.2	hyb
8365	銀行業	.547	.604	.601	<u>.653</u>	.644	.704
4667	情報通信	.214	.267	.218	.130	.387	<u>.286</u>
4579	医薬品	.226	.435	<u>.383</u>	.254	.364	.355
3286	不動産業	.141	.243	.219	.044	<u>.312</u>	.423
6408	機械	.062	-.093	<u>.113</u>	-.033	.086	.142
9514	電気ガス	.351	.239	<u>.335</u>	.014	.153	.181
9206	空運業	.159	.022	<u>.166</u>	.031	.083	.187
9849	卸売業	.010	.065	.106	-.087	.175	<u>.137</u>
1380	水産農林	.116	.185	.194	-.002	<u>.435</u>	.470
7709	精密機器	.149	.092	.154	.157	.023	.159

「食料品」の企業が類似企業となった. 生成 KW と通常 KW を観察すると, 通常 KW が「食品」や「加工」など事業に関わる単語が多く抽出され, 生成 KW では「卵」や「鶏肉」など企業の取扱品目を示す単語が抽出された. しかし, 通常 KW には「委託」や「生協」など事業から遠い単語が含まれており, 生成 KW の中には関連性の低い「ペットフード」の単語が含まれる. 表 2 に業種抽出性能の評価結果を示す. KW 検索 (k.1 と k.2) の性能がベクトル検索 (v.1 と v.2) より高い傾向があることがわかった. 企業ごとに観察すると, 「銀行業」においてはすべての手法で同一業種を検索することができ, 値が 1.00 になった. 一方で, 「空運業」, 「水産農林」, 「精密機器」において, 同一業種の企業を挙げることができずに値がゼロになった項目が k.2, v.1 と v.2 にあった. 表 3 に日次リターンの相関値を示す. 「医薬品」, 「電気ガス」業種の企業を除くすべての企業において hyb が最も高いか, 2 番目の相関値を示すことがわかった. 提案手法 (k.2, v.2, hyb) と rand を比較すると, 「電気ガス」を除くすべての企業で提案手法が rand の相関値を上回った. 表 2 におい

てすべての手法で nDCG が 1.00 であった「銀行業」は, 表 3 において, すべての手法で相関値に違いがあり, hyb が最も高い相関値を示した. また, 業種抽出性能比較において低い値 0.337 を示した「水産農林」は, hyb で相関値 0.470 を示した. これは, 業種分類では捉えられない類似企業をテキスト分析によって捉えることができた結果だと考えている. まとめると, 我々が提案する生成 KW 検索と文脈ベクトル検索は, ベースラインモデルを上回る性能を示した. 評価実験により, 生成 KW 検索は同一業種を抽出しやすく, 文脈ベクトル検索はリターンの類似した企業を抽出しやすい特長が備わっていることがわかった. このため, 提案手法のハイブリッド検索は, 両者を相互補完する有用な手法だと主張する.

5 おわりに

本研究では, 有報に記載された事業セグメントに着目し, 類似企業を検索する手法を提案した. 今後は, モデルの性能向上を目的として有報を用いたベクトル埋め込みモデルのファインチューニングと, 業種ごとの特徴を考慮したモデルを検討していく.

参考文献

- [1] David A Guenther and Andrew J Rosman. Differences between compustat and crsp sic codes and related effects on research. **In Journal of Accounting and Economics**, 1994.
- [2] 高野海斗, 酒井浩之, 北島良三. 有価証券報告書からの事業セグメント付与された業績要因文・業績結果文の抽出. 人工知能学会論文誌, 2019.
- [3] 奥山和樹, 木村泰知. 有価証券報告書を対象とした機械判読が困難な表構造の分析. 言語処理学会 第 30 回年次大会 発表論文集, 2024.
- [4] 伊藤友貴, 平松賢士. 有価証券報告書の活用による事業セグメント関連語の拡張. 言語処理学会 第 30 回年次大会 発表論文集, 2024.
- [5] Maarten Grootendorst. Keybert: Minimal keyword extraction with bert. 2020.
- [6] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. **In Conference on Empirical Methods in Natural Language Processing (EMNLP)**, 2019.
- [7] Matt J. Kusner, Yu Sun, Nicholas I. Kolkin, and Kilian Q. Weinberger. From word embeddings to document distances. **In International Conference on Machine Learning (ICML)**, 2015.
- [8] Stephen E. Robertson, Steve Walker, Susan Jones, Micheline Hancock-Beaulieu, and Mike Gatford. Okapi at trec-3. **In Text Retrieval Conference**, 1994.
- [9] Hayato Tsukagoshi, Ryohei Sasano, and Koichi Takeda. Japanese SimCSE Technical Report. **In arXiv**, 2023.
- [10] 鈴木正敏, 松田耕史, 関根聡, 岡崎直観, 乾健太郎. Wikipedia 記事に対する拡張固有表現ラベルの多重付与. 言語処理学会 第 22 回年次大会 発表論文集, 2016.
- [11] Dimitrios Vamvourellis, Máté Attila Tóth, Snigdha Bhagat, Dhruv Desai, Dhagash Mehta, and Stefano Pasquali. Company similarity using large language models. **In IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFER)**, 2023.
- [12] S. F. Dierk. The smart retrieval system: Experiments in automatic document processing. **In IEEE**

A 付録

A.1 ハイブリッド検索手法の比較

生成 KW 検索 (k.2) と文脈ベクトル検索 (v.2) の組み合わせ (hyb) と、通常 KW 検索 (k.1) と v.2 の組み合わせ (hyb') を章 4.3 と同様の手法で比較した。表 A.1 より、医薬品以外のすべてにおいて、提案手法 (hyb) が安定して高い相関値を示した。

表 A.1 日次リターンの相関値。企業ごとに最も相関値の高い値を太字で示し、2 番目に高い値に下線を引く。

企業	業種	手法	hyb	hyb'
8365	銀行業		.704	.606
4667	情報通信		.286	.207
4579	医薬品		.355	.460
3286	不動産業		.423	.263
1380	水産農林		.470	.249
6977	電気機器		.268	.015
9060	陸運業		.231	.018
1994	建設業		.365	.258

A.2 業種抽出性能の評価結果

章 4.2 の評価に F1-Score および MRR (Mean Reciprocal Rank), nDCG を追加。表 A.2 に結果を示す。表 A.3 の 30 社すべての平均値をとった。

A.3 類似度関数

Word Mover's Distance (WMD) [7]: WMD は一方の文書の単語をもう一方のドキュメントの単語に移動させるための最小コストを求める: $WMD(D_1, D_2) = \min_{T \geq 0} \sum_{i,j} T_{ij} c_{ij}$. ここで, T_{ij} は文書 D_1 の i 番目の単語から文書 D_2 の j 番目の単語への流れを示し, c_{ij} は単語 i と単語 j の間のコストを示す。本研究では, 生成 KW を連結したリストのペアを文書 D_1, D_2 とした。

BM25 Okapi [8]: TF-IDF [12] を拡張したランキング関数であり, 文書 D に含まれるそれぞれのワードを w と表現すると次のように定義される。

$$BM25(w, D) = \frac{IDF(w) \cdot TF(w, D) \cdot (k_1 + 1)}{TF(w, D) + k_1 \left(1 - b + b \frac{|D|}{D_{avg}}\right)} \quad (5)$$

ここで, $IDF(w)$ は単語 w の逆文書頻度を表す。本研究において, w を生成 KW, D を分ち書きされた「事業の内容」章として $BM25(\cdot)$ を計算した。

Jaccard 係数: 2 つの集合 A と B の類似度を測定する関数であり, 次のように定義される:

表 A.2 業種抽出性能の評価結果。上位 $k = \{20, 50\}$ 個の類似企業について最も高い値を太字で示し, 2 番目に高い値に下線を引く。

手法	k	MRR	nDCG	F1-Score	
				Micro	Macro
k.1. 通常 KW 検索		.298	.606	.285	.121
k.2. 生成 KW 検索		.510	.672	.420	.177
v.1. 通常ベクトル検索	20	.224	.298	.010	.035
v.2. 文脈ベクトル検索		.379	.584	.298	.127
hyb. ハイブリッド検索		.359	.682	.433	.182
k.1. 通常 KW 検索		.300	.610	.239	.170
k.2. 生成 KW 検索		.511	.680	.361	.252
v.1. 通常ベクトル検索	50	.228	.329	.100	.064
v.2. 文脈ベクトル検索		.382	.613	.262	.183
hyb. ハイブリッド検索		.362	.704	.367	.254

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

重み付き Jaccard 係数: 重み付き Jaccard 係数は Jaccard 係数を拡張した関数であり, 要素の重要度を考慮する点で異なる: $J_w(A, B) = \frac{\sum_i \min(a_i, b_i)}{\sum_i \max(a_i, b_i)}$. ここで, a_i と b_i は集合 A と B の i 番目の要素の重みになる。本研究では, 生成 KW をソートした時の順番に基づく重要度を用いた。

A.4 ランキング融合

Ranked Retrieval Fusion (RRF) は, 複数のランキングリストを結合して全体のランクを計算する手法である: $RRF(d) = \sum_{i=1}^N (1/(k + r_i(d)))$. ここで, N はランキングリストの数, $r_i(d)$ はアイテム d の i 番目のランキングリストにおける順位, k はスムージング定数。

A.5 評価実験に用いたクエリの企業

本文の結果は, 表 A.3 より選択した 10 社を用いた。章 A.2 で 30 社すべてを用いた。

表 A.3 評価実験に用いたクエリの企業一覧

企業	業種	企業	業種
5268	ガラス・土石製品	5010	石油・石炭製品
5103	ゴム製品	9206	空運業
9720	サービス業	7709	精密機器
3878	パルプ・紙	8040	繊維製品
3286	不動産業	8705	証券、商品先物取引業
9311	倉庫・運輸関連	7235	輸送用機器
4243	化学	5936	金属製品
4579	医薬品	5446	鉄鋼
9849	卸売業	8365	銀行業
3352	小売業	9060	陸運業
1994	建設業	9514	電気・ガス業
4667	情報・通信業	6977	電気機器
6408	機械	1491	非鉄金属
1380	水産・農林業	2916	食料品
9173	海運業	8769	保険業