

# キャリア支援のための自由記述文を 用いた計量テキスト分析

——キャリアセンター資料の企業担当者コメントに基づいて——

柴 田 淳 子

神戸学院経済学論集

第49巻 第4号 抜刷

平成30年3月発行

# キャリア支援のための自由記述文を 用いた計量テキスト分析

——キャリアセンター資料の企業担当者コメントに基づいて——

柴 田 淳 子

## 1. はじめに

厚生労働省と文部科学省が共同で実施している大学生等卒業者の就職状況調査によると、平成28年3月卒業の大学生の就職率は97.6%であった。平成23年3月卒業の大学生の就職率91.0%から年々増加傾向にあり、ここ数年は特に学生優位の売り手市場が続いている。神戸学院大学において、平成28年3月卒業の大学生の就職率は94.6%であった。学部により多少の違いはあるが、近年は高い数値を保持している。その背景には、本学キャリアセンターが年間延べ約7,000人も学生の個別相談を行うなど、学生の就職を全面的にサポートしていることがある。

一方、情報通信技術の発展により、日々たくさんのデータが生み出されている。事業に役立つ知見を導出するためのデータはビッグデータと呼ばれ、私たちの生活を豊かにするために役立っている。ビッグデータの解析方法の一つにテキストマイニングがあり、実際に教育や看護などの分野において活用されている。

本研究では、本学キャリアセンターが実施した採用活動に関するアンケート調査によって得られた自由記述文に対して、テキストマイニング技術を用いて分析し、学生の就職活動に役立つ知識発見を目的とする。さらに、業種ごとの特徴を捉えることで、より効率的な就職支援が可能になると考えられる。

## 2. 調査データ

本研究で使用するデータは、2017年求人紹介会で学生に採用活動を継続して行っている企業を紹介するために、キャリアセンターが実施したアンケート調査の一部である。この調査は、2017年2月から6月末までに本学学生に対して求人票を送付した企業を対象としている。求人紹介会とは、主に就職活動中の4年生を対象とした毎年9月頃にキャリアセンターがその時点で応募可能な求

表1：業種ごとの回答件数

| 順位 | 職種              | 件数  | 順位 | 職種               | 件数 |
|----|-----------------|-----|----|------------------|----|
| 1  | 化学・薬品・化粧品小売業    | 109 | 25 | 精密機械器具卸売業        | 10 |
| 2  | 情報・ソフトウェアサービス業  | 68  | 26 | 精密機械器具製造業        | 9  |
| 3  | その他のサービス業       | 62  | 27 | 新聞・放送・通信・広告業     | 9  |
| 4  | 建設業             | 45  | 28 | 教育               | 9  |
| 5  | 自動車小売業          | 42  | 29 | 出版・印刷・関連産業       | 8  |
| 6  | 観光業・興業・娯楽       | 36  | 30 | 薬品・化粧品製造業        | 8  |
| 7  | 運輸・倉庫業          | 32  | 31 | 木製品・紙製品製造業       | 7  |
| 8  | 食料品小売業          | 30  | 32 | その他の製造業          | 7  |
| 9  | 一般機械・輸送用機械器具卸売業 | 29  | 33 | 各種商品小売業          | 7  |
| 10 | 食料品卸売業          | 21  | 34 | 石油・石炭・ゴム・皮革製品製造業 | 6  |
| 11 | 百貨店・スーパー        | 21  | 35 | 木製品・紙製品卸売業       | 6  |
| 12 | 一般機械・輸送用機械器具製造業 | 20  | 36 | 繊維工業・繊維製品製造業     | 5  |
| 13 | その他の小売業         | 20  | 37 | 繊維製品等小売業         | 5  |
| 14 | 鉄鋼・非鉄金属・金属製品製造業 | 18  | 38 | 信用金庫・組合・労働金庫     | 5  |
| 15 | その他の卸売業         | 16  | 39 | 各種商品卸売業          | 4  |
| 16 | 不動産業            | 15  | 40 | 電気・機械製品小売業       | 4  |
| 17 | 繊維製品卸売業         | 13  | 41 | 証券業              | 4  |
| 18 | 鉄鋼・非鉄金属・金属製品卸売業 | 13  | 42 | 農業・林業・水産・鉱業      | 3  |
| 19 | 調査・専門サービス業      | 12  | 43 | その他金融業           | 3  |
| 20 | 食料品製造業          | 11  | 44 | 保険業              | 3  |
| 21 | 化学工業・化学製品製造業    | 11  | 45 | 化学製品卸売業          | 2  |
| 22 | 電気機械器具製造業       | 11  | 46 | 薬品・化粧品卸売業        | 2  |
| 23 | 電気機械器具卸売業       | 11  | 47 | 石油・石炭・ゴム・皮革製品卸売業 | 2  |
| 24 | 公益業・非営利団体       | 11  |    |                  |    |

合計：805件

人を公開するものである。具体的には、複数個の質問項目の中の一つである「今後の参考のために本学の学生に限らず、現在の就職活動生を人事担当者様からはどのように見えているのか、どういったところが不足しているのかご教示ください」の質問項目に対して、企業の人事担当者が回答した自由記述文を使用する。本報告では、2017年9月15日までに回答が得られた805件のアンケート調査結果に対して分析を行う。ここで、業種ごとの分類を表1に示す。

次に、その回答文の一部を図1に示す。

|  |
|--|
| 農業・林業・水産・鉱業  |
| 当社は牛という生き物相手の仕事ですので、定時定刻の勤務時間にはなかなかないのが実情です。しっかり牛を観察しながら、自分で仕事を組み立てて取り組んで行くかが求める人物像になります。働きかけるとその分結果になってかえって来るのが「農業」「酪農」だと思っています。少しの障害に当たっても負けずに、みんなと協力しながら対処し、定着していただける力が欲しいと思います。                  |
| 酪農の場合は生き物と生乳に携わるため、強い信念（原点に立ち返る力をきちんともっている人）が必要と思われまます。その為、憧れだけでは生き物相手の仕事はできませんが、素晴らしい職場経験ができ、あらゆる面での糧となることと考えております。   |
| 自分にあった仕事は何か？きつい仕事は無理？楽しい仕事がいい。といった観点で仕事を選ぶ方が多く感じます。仕事は生きる為の手段で有り、どの仕事も世の中に必要です。だからこそ、その仕事でのやりがいを見つけられない人は転職を繰り返す傾向が強く感じます。まずは30歳の自分がどうなっていたいかの目標を明確にして、職業を選んでほしいと思います。まずは、当社で農業体験をしてみませんか？お待ちしております。 |
| 建設業  |
| どこの企業様も実感されることがあると考えますが、面談または企業説明会において、男性より女性の方が   |

図1：入力データ

出典：キャリアセンターの資料「調査結果報告－現在の就活生を企業の人事担当はどのように見えているか、どういったところが不足しているか」

企業の人事担当が学生をどのように見ているか、学生のどのようところが不足しているかという情報は、これから就職活動を行う学生にとって非常に重要な情報である。さらに、教員やキャリアセンターにとって有益な情報でもある。しかし、対象とする調査結果報告の資料は68ページで構成されているため、必要な情報の抽出が困難である。

### 3. 分析結果とその考察

本論文では、図1に示すテキストデータを用いて計量テキスト分析を行う。分析手法としては、KHcorderを用いる。本報告で用いるテキストデータが就職活動に関するデータであることから、前処理として、表2に示す「売り手市場」、「企業研究」など12個の言葉を強制抽出する語として設定した。

表2：設定したキーワード

|       |       |      |
|-------|-------|------|
| 売り手市場 | ガッツ   | 就活   |
| 企業研究  | 連絡なし  | 就職活動 |
| ドタキャン | 合同説明会 | 男子学生 |
| 無断欠席  | 会社説明会 | 女子学生 |

表3は、これらの設定により得られた抽出語リストを示している。

表3：抽出語リスト（上位50位まで）

| 抽出語  | 出現回数 | 抽出語 | 出現回数 | 抽出語  | 出現回数 | 抽出語   | 出現回数 | 抽出語 | 出現回数 |
|------|------|-----|------|------|------|-------|------|-----|------|
| 学生   | 781  | 会社  | 134  | 内定   | 98   | 売り手市場 | 84   | 受ける | 59   |
| 思う   | 548  | 仕事  | 132  | 企業研究 | 97   | 働く    | 80   | 必要  | 58   |
| 感じる  | 471  | 印象  | 129  | 自身   | 97   | 見る    | 75   | 自己  | 57   |
| 多い   | 387  | 持つ  | 118  | 質問   | 91   | 入社    | 73   | 増える | 56   |
| 企業   | 333  | 就職  | 116  | 説明   | 91   | 出来る   | 69   | 聞く  | 56   |
| 自分   | 300  | 人   | 113  | 選考   | 88   | 連絡    | 66   | 言う  | 55   |
| 就職活動 | 220  | 業界  | 112  | 良い   | 87   | 就活    | 65   | 行動  | 53   |
| 面接   | 172  | 少ない | 108  | 弊社   | 86   | 研究    | 62   | 志望  | 53   |
| 考える  | 141  | 採用  | 106  | 活動   | 85   | 社会    | 61   | 特に  | 52   |
| 不足   | 140  | 積極  | 105  | 情報   | 84   | 参加    | 60   | 意欲  | 51   |

アンケートの質問項目が学生に関することであるため、「学生」、「思う」、及び「感じる」の単語が上位に位置している。次に、どのような単語が共起しているかを明らかにするために、Jaccard 係数に基づく共起ネットワークを図2

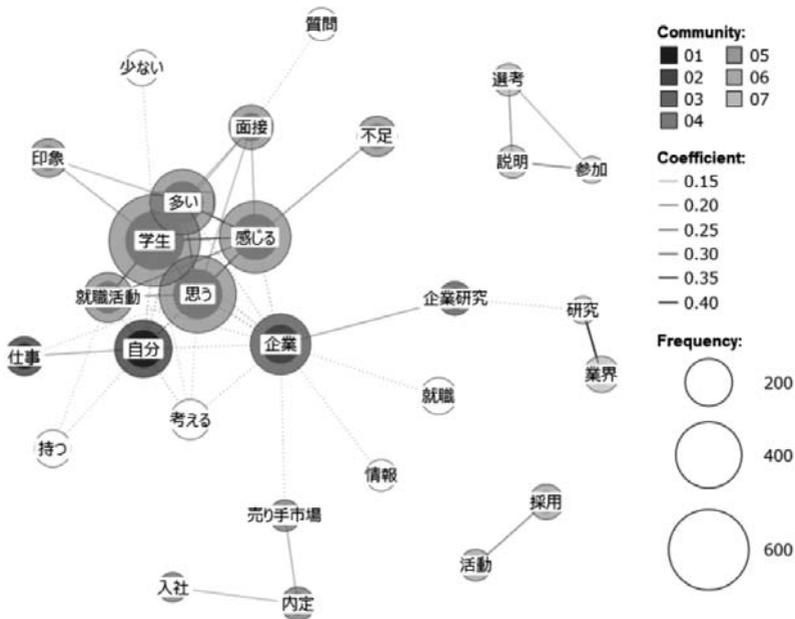


図2：共起ネットワーク

に示す。図中、語句を囲む円の大きさは出現回数が多いほど大きい円で表されている。そして、円の配色によってサブグラフを表している。また、語句と語句を結ぶ線の太さはそれらの語句の共起関係を表しており、共起関係が強いほど太い線で表現されている。

質問文の「どのように見えているのか」という点から、企業担当が学生の就職活動に対して思うことや感じる内容が回答されていることが図2より分かる。ただし、太字の語句は図2の共起ネットワーク中に存在する語句であることを表している。自分に合う仕事を探すことや、企業研究によって企業を知ることが、就職活動の中でも重要な項目である。企業担当が面接に関して思うことや感じる内容は、面接受け答えやその準備など多岐に渡っている。また、近年の売り手市場の影響もあり、企業としては、内定を出しても確実に入社の子測が困難であること、学生が当社の説明会や選考への参加を望んでいることが読み

キャリア支援のための自由記述文を用いた計量テキスト分析

取れる。さらに、質問文の「どういったところが不足しているか」に対して、企業担当は何かしら不足だと感じることもあることが分かる。特に、表2の結果から不足の頻出回数は140回であった。企業担当が不足していると感じる内容は、就職活動を行う学生にとって有益な情報であると考えられる。

次に、不足の内容を明らかにするために、ここでは不足の語句の前後に出現する語句を調べるコロケーションによる分析を行う。語句の重要度を表す指標としては、直前や直後に出現する語句  $s$  を重要視するスコア  $f(s)$  を用いる。スコア  $f(s)$  は以下の式によって計算される。

$$f(s) = \sum_{k=1}^K \frac{(L_k + R_k)}{k}$$

ただし、 $L_k$  は不足の  $k$  個前に出現した回数であり、 $R_k$  は不足の  $k$  個後に出現した回数である。本報告では  $K=5$  とし、不足の直前の5語と直後の5語に出現する語句を調べ、その結果を表4に示す。

表4：コロケーション統計の結果

| N  | 抽出語  | 品詞   | 合計 | 左合計 | 右合計 | 左5 | 左4 | 左3 | 左2 | 左1 | 右1 | 右2 | 右3 | 右4 | 右5 | スコア   |
|----|------|------|----|-----|-----|----|----|----|----|----|----|----|----|----|----|-------|
| 1  | 感じる  | 動詞   | 31 | 4   | 27  | 1  | 0  | 2  | 1  | 0  | 0  | 9  | 1  | 1  | 16 | 9.650 |
| 2  | 企業研究 | タグ   | 15 | 14  | 1   | 1  | 0  | 0  | 9  | 4  | 0  | 1  | 0  | 0  | 0  | 9.200 |
| 3  | 研究   | サ変名詞 | 10 | 10  | 0   | 0  | 1  | 1  | 4  | 4  | 0  | 0  | 0  | 0  | 0  | 6.583 |
| 4  | 特に   | 副詞   | 8  | 8   | 0   | 0  | 1  | 1  | 0  | 6  | 0  | 0  | 0  | 0  | 0  | 6.583 |
| 5  | 準備   | サ変名詞 | 7  | 7   | 0   | 1  | 0  | 0  | 1  | 5  | 0  | 0  | 0  | 0  | 0  | 5.700 |
| 6  | 理解   | サ変名詞 | 9  | 8   | 1   | 0  | 0  | 1  | 5  | 2  | 0  | 0  | 0  | 1  | 0  | 5.083 |
| 7  | 学生   | 名詞   | 12 | 4   | 8   | 0  | 3  | 1  | 0  | 0  | 0  | 3  | 0  | 5  | 0  | 3.833 |
| 8  | 分析   | サ変名詞 | 7  | 7   | 0   | 2  | 2  | 0  | 1  | 2  | 0  | 0  | 0  | 0  | 0  | 3.400 |
| 9  | 思う   | 動詞   | 11 | 4   | 7   | 1  | 0  | 3  | 0  | 0  | 0  | 1  | 1  | 0  | 5  | 3.033 |
| 10 | 少し   | 副詞   | 3  | 3   | 0   | 0  | 0  | 0  | 0  | 3  | 0  | 0  | 0  | 0  | 0  | 3.000 |
| 11 | 力    | 名詞C  | 6  | 6   | 0   | 0  | 0  | 0  | 6  | 0  | 0  | 0  | 0  | 0  | 0  | 3.000 |
| 12 | 業界   | 名詞   | 8  | 8   | 0   | 1  | 3  | 2  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 2.617 |
| 13 | 時間   | 副詞可能 | 3  | 3   | 0   | 0  | 0  | 0  | 1  | 2  | 0  | 0  | 0  | 0  | 0  | 2.500 |
| 14 | 志望   | サ変名詞 | 7  | 2   | 5   | 2  | 0  | 0  | 0  | 0  | 0  | 2  | 1  | 0  | 2  | 2.133 |
| 15 | 自己   | 名詞   | 6  | 6   | 0   | 2  | 0  | 2  | 2  | 0  | 0  | 0  | 0  | 0  | 0  | 2.067 |

表4のスコア上位の抽出語から、不足している具体的な内容は企業研究、準備、志望動機や自己分析などであることが分かる。特に、企業研究と研究の順位が高く、就職活動を行う前の準備が不足していると言える。

次に、表1に示す回答数の上位10位までの業種の特徴を捉えるために、頻出回数上位30位までの抽出語を用いて対応分析を行い、その結果を図3に示す。

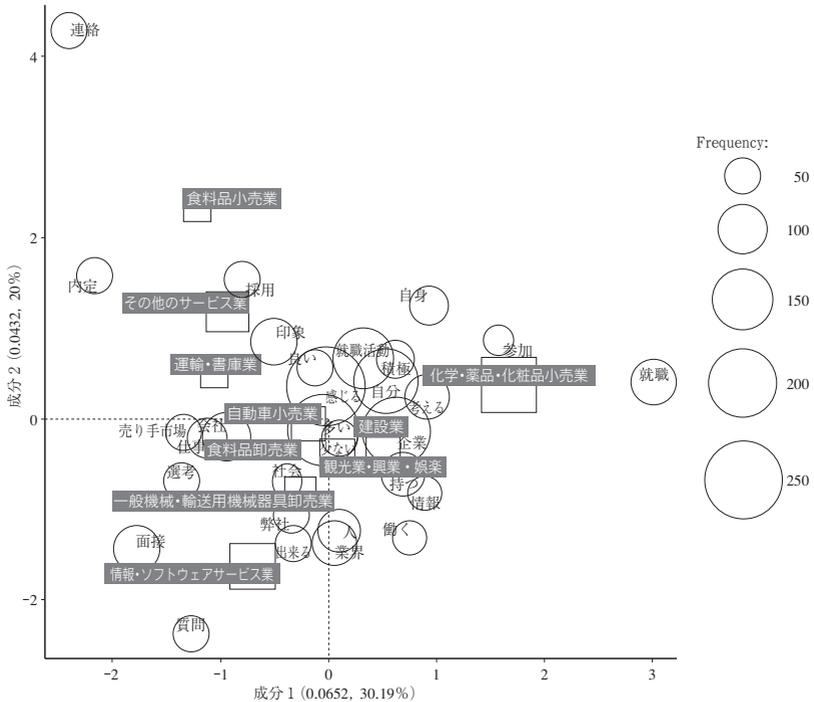


図3：業界と頻出後の対応分析

業種ごとの回答数が一番多かった化学・薬品・化粧品小売業は、考えるや自分といった語句の近くに布置している。実際に、「これまでお会いした学生の傾向から申し上げます、自ら考え、それを自分のことばで表現または主張できればなお良いと思います。」という意見があった。また、左下に位置する情報・ソフトウェアサービス業の近くに布置している語句は、面接や質問などであり、

キャリア支援のための自由記述文を用いた計量テキスト分析

「面接官の質問の意図を汲み取り、適切に回答する力（が不足している）」という意見があった。この業種は技術職を希望する学生が多く、面接に対して苦手意識を持っている学生が多いことが推察される。

#### 4. おわりに

本論文では、就職活動に関して企業担当が回答した自由記述型のテキストデータを用いて、学生が就職活動を行う際に役立つ情報の抽出を試みた。具体的には、2017年9月に開催された求人紹介会のためのアンケート調査結果の一部を用いてテキストマイニングを行い、それによって得られた頻出語の特徴から全体的な内容を明らかにし、さらに業種ごとの傾向を知るために対応分析を行った。

今回の目的は、805件もあるテキストデータの全体像を掴むことであった。今後は、学生の就職に対する意識調査など学生側の意見を取り扱うデータを併せて分析することで、双方向型のキャリア教育に役立つ知識を発見できるように継続して研究を行う予定である。

#### 謝辞

本論文を作成するにあたり、アンケート資料を提供して頂いた神戸学院大学キャリアセンターと、コメントやアドバイスを頂いた同センターの吉川彰一氏に深謝する。

#### 参 考 文 献

- [1] “大学，短期大学，高等専門学校及び専修学校卒業予定者の就職内定状況等調査”，文部科学省，2018-03-16.
- [2] 菰田文男，那須川哲哉，“ビッグデータを活かす 技術戦略としてのテキストマイニング”，中央経済社，2014.
- [3] 樋口耕一，“テキスト型データの計量的分析：2つのアプローチの峻別と統合”，理論と方法，19(1)，101-115，2004.

- [4] 樋口耕一, “社会調査のための計量テキスト分析”, ナカニシヤ出版, 2014.
- [5] 市村由美, 長谷川隆明, 渡部勇, 佐藤光弘, “テキストマイニング: 事例紹介”, 人工知能学会誌, 16(2), 192-200, 2001.
- [6] 松河秀哉, 齊藤貴浩, “データ・テキストマイニングを活用した授業評価アンケートフィードバックシステムの開発と評価”, 日本教育工学会論文誌 35(3), 217-226, 2011.
- [7] 越中康治ほか, “テキストマイニングによる授業評価アンケートの分析—共起ネットワークによる自由記述の可視化の試み—”, 宮城教育大学情報処理センター研究紀要, COMMUE (22), 67-74, 2015.
- [8] KH Corder (<http://khc.sourceforge.net/>)