

ラフ集合とクラスタリングを用いた就職活動者向け企業推薦システム

1180352 中井 祐貴 【知能情報学研究室】

1 はじめに

就職を支援することにおいて、求職者と企業のマッチングが最も重要なことである。しかし、インターネット上に様々な企業の情報が掲載されているものの、その量は膨大であり求職者一人で把握できる範囲に限られる他、中小企業では十分な情報が公開されていないこともしばしばある。また、近年は AI マッチング等の報道があるように、知能技術を用いたマッチングシステムも開発されてきている。

そこで、本研究では、部分的な情報から求職者と企業のマッチングを図るための企業推薦システムを提案する。本研究の対象者は高知工科大学情報学群の就職活動者とし、入社したいと思える会社と出会えることを目的とする。また、欠損値やテキスト情報等の数値化しにくい情報を扱うため、ラフ集合や階層的クラスタリング等の知能技術、データマイニング技術を用いる。

2 提案手法

提案手法は、ラフ集合による企業業績予測ルール抽出部と、業務内容・企業理念情報の階層的クラスタリング部からなる。

2.1 ラフ集合による近似予測ルール

ラフ集合は企業の所在地、従業員数、売上等の属性情報から、将来の業績上昇/下降等の予測したい情報(決定属性と呼ぶ)を導くルールを求める手法である [1]。

Company	Sales	Average Annual income	Ordinary income	Number of employees	listing	Type of business	Head office location	Operating income	Company performance
A	1000	8000	100	10000	Yes	Shipping	Tokyo	100	UP
B	4000	7000	200	100000	No	Manufacture	Tokyo	200	DOWN
C	20000	8000	2000	300000	No	Manufacture	Aichi	1000	UP
D	2000	8000	30	100000	No	Manufacture	Tokyo	30	DOWN

Co. performance up rule: Avg. annual income = 8000
 Co. performance down rule: Employees = 100000 and Business = Manufacture

図 1 決定ルール導出図 [中井, 2017]

図 1 はラフ集合の属性表の例である。右端の列が決定属性であり、これを予測するルールが矢印の先の業績上昇ルール、業績下降ルールである。本研究では、業績予測として「上がる」「下がる」「安定」の 3 要素の予測を行った。東証一部上場企業等、十分な情報が公開されているものに対して、ラフ集合を適用してルールを求め、求職者にも提示する。

2.2 業務内容と企業理念の階層的クラスタリング

各社が公開している業務内容と企業理念・経営理念のテキストデータに対して、TF-IDF での文章中の解析と

ベクトル化、コサイン距離 (文章間類似度) とワード法による階層的クラスタリングを行い、類似した企業のグルーピングを行う。その後、企業理念・経営理念のクラスに対し、単語頻出の多いものを理念のキーワードとし、企業情報と共に提示する。

3 実験と評価

データセットとして 537 社分の 2016 年度の決算情報を用意する。学習データは 203 社、残り 334 社のデータは欠損値のある属性の集合である。属性の要素として売上、営業利益の増減とその差分、単独又は連結従業員数、社員 1 人分の生産性、業種、所在地方、理念、純利益とした。被験者 3 名に、入力する事業内容を「システム」「ネットワーク」「アプリケーション」「IT コンサルティング」に限定し、関係ある会社の選択結果から適合率の算出を実施した。「システム」以外は 50%未滿とクラスタリングによる提示は良くなかった。また 4 つの事業内容の内、被験者が興味ある業務と近い要素を選んでもらい、興味を持った会社についてヒヤリングすると、業績予測よりも、企業理念を重視していた。

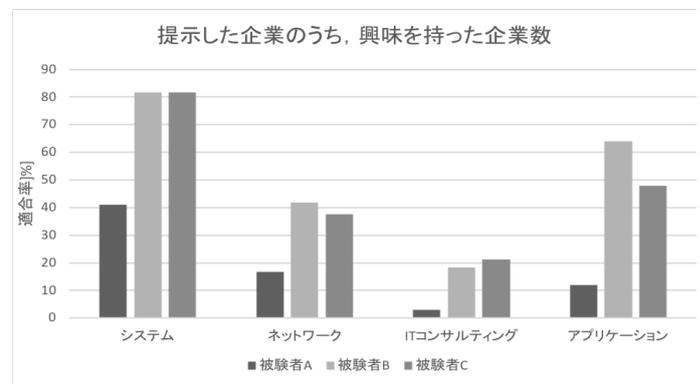


図 2 被験者実験の結果

4 まとめ

企業理念を 1 単語で提示することによって、各社が大切にしている事を把握が可能となり企業選びの有効な指標となった。ラフ集合は指標ではなく、会社選びの傾向に用いることで、各個人の性格や考え方に合致する会社を勧めることもできるため、将来のシステムでは個人の嗜好そのものをルールとして抽出することも考えている。

参考文献

[1] 森 典彦, 田中 英夫, 井上 勝雄, “データからの知識獲得と推論 ラフ集合と感性,” 海文堂出版株式会社, 2004.