

中期計画を対象としたロジックモデルの 各構成要素における特徴語分析

小柏香穂理（帝京大学），高田英一（神戸大学），大石哲也（東京工業大学），
森雅生（東京工業大学），廣川佐千男（九州大学）

1. はじめに

2004年の国立大学法人化以降，国立大学は現在，3巡目となる第3期中期目標期間に入っており，中期計画の策定も3回目である．大学の計画は，学内で共有されることが重要^[1]であり，そのためには，内容が明確で共有しやすいロジックモデルに基づいた中期計画であることが望ましいと思われるが，まだロジックモデルに基づいた計画策定^[2]に慣れていない大学が多数存在していることが先行研究により明らかにされている^{[3][4][5]}．これらの課題を解決するためには，まず初めに中期計画の内容をロジックモデルの各構成要素に分類し，計画内容の構造化やその対応関係を分析する必要がある．しかしながら，中期計画の文書を人手でロジックモデルの各構成要素に分類する作業はコストが膨大になる．そこで本研究は，中期計画の文書を，ロジックモデルの5つの構成要素^[3]に分類するために，機械学習の技術を活用した実験を行い，その識別性能を分析した．本発表ではその結果を報告する．

2. 研究方法

本研究の目的は，機械学習の技術を活用して，中期計画の文書をロジックモデルの5つの構成要素に分類し，その識別性能を分析することである．さらに各構成要素の特徴語も分析する．特徴語とは，単純な出現頻度ではなく，各構成要素において特徴的に出現する語句のことを示す．

2-1. 実験データ

本研究で使用したデータは，国立大学法人第3期中期計画^[6]の文書¹を，人手により，ロジックモデルの5つの構成要素（目的，投入（インプット），活動（アクティビティ），アウトプット，アウトカム）に分類したデータを用いた．データは国立大学法人86大学中84大学の全1132文書，2146単語からなる．「投入（インプット）」に関しては，データ数が少ないため実験対象から除外した．これらのデータを用いて，以下の2つのデータセットを作成した．

- (1) 全単語を使用したデータ
- (2) 属性選択^{[7][8]}された単語のみを使用したデータ

¹ 中期目標・中期計画は，大きく5つの項目（I 大学の教育研究等の質の向上に関する目標，II 業務運営の改善及び効率化に関する目標，III 財務内容の改善に関する目標，IV 自己点検・評価及び当該状況に係る情報の提供に関する目標，V その他業務運営に関する重要目標）に分かれている．本研究では，教育活動に関する分析が目的であるため，「I 大学の教育研究等の質の向上に関する目標」の「1. 教育に関する目標」に対応する中期計画のうち，「I 大学の教育研究等の質の向上に関する目標を達成するためにとるべき措置」「1 教育に関する目標を達成するための措置」「（1）教育内容及び教育の成果等に関する目標を達成するための措置」に該当する文書を対象とした．

2-2. 実験手順

2.1 の 2 つのデータセットに対して、機械学習モデルのひとつである SVM (Support Vector Machine) による実験を以下の方法で行った。評価指標として、Precision (適合率), Recall (再現率), F-measure (F 値), Accuracy (精度) の値を用いた。

- (1) SVM による識別性能の評価
- (2) F 値を最適とする属性選択と特徴語

3. 結果と考察

まず全単語での識別性能の結果を表 1 に示す。次に F 値を最適とする属性選択の識別性能 (表 2) と特徴語 (表 3) を示す。

識別性能は、属性選択された単語のみを使用した場合の方が、全単語を使用した場合と比較して性能が上がっている (表 1, 表 2)。属性選択された単語での識別性能について、5 つの構成要素の中では「活動 (アクティビティ)」についての文が一番識別しやすく、少数の単語だけで F 値 99%, accuracy 98% の識別ができる (表 2)。アウトプット, アウトカムの識別性能が悪い原因としては、「何を書けばいいかわからない」もしくは「書いてしまえば評価の対象となるので、不達成のリスクを避けるためになるべく書きたくない」という意識によると考えられる。この点、参考文献[3]によると、第 3 期中期計画への記載率は、活動 (アクティビティ) 99.0% に対して、アウトプット 6.4%, アウトカム 23.0% に留まっている。また、文科省も、繰り返し中期計画の策定の際に、成果 (アウトカム) を記載するよう指導していることから、今回の機械的分析による考察と符合する。

特徴語に関しては「活動 (アクティビティ)」の特徴語は、「する 行う 授業」の 3 つがあげられている (表 3)。文書中に「授業」が使われている例として、「授業内容」, 「授業科目」, 「〇〇の授業」等の使用例があった。本研究は教育に関する中期計画を対象としているため、活動 (アクティビティ) として、授業に関する内容が多く記載されていると考えられる。ただ、同じ単語であっても「目的, 活動 (アクティビティ), アウトプット, アウトカム」の各構成要素のうち、どこに現れているかで単語の意味が異なることが考えられる。

表 1. 全単語での識別性能

構成要素	precision	recall	F-measure	accuracy
目的	0.8630	0.5509	0.6710	0.7661
活動 (アクティビティ)	0.9922	0.9148	0.9518	0.9089
アウトプット	0.7577	0.2540	0.3778	0.8871
アウトカム	0.3333	0.1250	0.1814	0.5772

表 2. 属性選択された単語での識別性能

構成要素	N	precision	recall	F-measure	accuracy
目的	500	0.8324	0.9309	0.8785	0.8880
活動（アクティビティ）	3	0.9883	0.9925	0.9903	0.9809
アウトプット	400	0.5073	0.8740	0.6378	0.8677
アウトカム	200	0.6222	0.5100	0.5479	0.9706

表 3 特徴語

構成要素	特徴語
目的	ため 観点 向ける 学年 構成 受入 理論 リンク 卒業 活躍
活動（アクティビティ）	する 行う 授業 学修 科目 プログラム 的 カリキュラム 経験 改善
アウトプット	比す 輩出 県 促進 小中学校 向上 就職 脳 関連 もつ
アウトカム	名 % 年間 参加 増やす 教員 インターンシップ 専任 世界 採用

4. おわりに

本研究は、中期計画の文書を、ロジックモデルの5つの構成要素に分類するために、機械学習の技術を活用した実験を行い、その識別性能を分析した。活動（アクティビティ）の属性選択の識別性能はF値 99%、accuracy 98%であり、高い識別性能であった。今後は、各構成要素の対応関係や各大学の識別について実験を行う予定である。

【謝辞】

本研究は JSPS 科研費 JP18K02705 の助成を受けたものです。

【参考文献】

- [1] 両角亜希子, 「単年度計画への反映と学内共有が将来計画の実質化のカギ」『Between』2012年10-11月号, pp.3-5, 2012.
- [2] 田中啓, 政策体系と評価～政策の体系化による政策評価の体系的かつ合理的で的確な実施について～, 平成27年度 政策評価に関する統一研修(金沢会場), 2015年10月2日, http://www.soumu.go.jp/main_content/000420725.pdf, 2018年6月10日アクセス.
- [3] 高田英一, 国立大学における計画・評価のロジックモデルを用いた現況分析とIRによる支援のあり方, 2018年1月30日, IR集中講習会.
- [4] Tetsuya Oishi, Eiichi Takata, Noriko Kuwano, Takahiro Seki, Masao Mori, Masashi Sekiguchi. How can IR Support the Management of Japanese National Universities on the Mid-Term Plan Related to Globalization?, The 17th Annual SEAAIR Conference, Proc. of The 17th Annual SEAAIR Conference, pp. 303-309, Sep. 2017.
- [5] 大石哲也, 桑野典子, 高田英一, 関隆宏, 森雅生, 関口正司. 日本の国立大学における経営計画の課題とIRによる支援のあり方について—グローバル化に関する中期計画を中心に—, MJIR2017, 第6回 大学情報・機関調査研究会集会 論文集, pp. 8-13, 2017.

- [6] 文部科学省, 各国立大学法人・各大学共同利用機関法人の第3期中期目標・中期計画 (平成28年3月), http://www.mext.go.jp/a_menu/koutou/houjin/1368750.htm, 2018年6月10日アクセス.
- [7] Toshihiko Sakai, Sachio Hirokawa, Feature Words that Classify Problem Sentence in Scientific Article, Proc. iiWAS2012, pp.360-367, 2012.
- [8] Yusuke Adachi, Naoya Onimura, Takanori Yamashita, Sachio Hirokawa, Classification of Imbalanced Documents by Feature Selection, Proc. ICCDA 2017, pp.228-232, 2017.