

公務員試験問題の類似判定手法の提案

大谷 紀子 研究室

1132167 永井優充

1. 研究の背景・目的

公務員試験には、5つの選択肢から1つを選ぶ形式の試験と、指定された時間内に論文を書く形式の試験がある。前者の試験では120分で40問を解く必要があり、時間をかけずに素早く各問を解く必要がある。すなわち、本番の試験で素早く正確に問題を解けるようになるための訓練が必要である。

公務員試験対策としては、多くの問題を解くのが有効である。例えば、数的処理の科目では、多くの問題を解いて形式やパターンを覚え、憲法の科目では選択肢の中から間違っているものを探したり、判例を覚えたりする。また、解く問題を選ぶ際には、昨年出題された問題の類似問題が今年も出題されることはなく、一昨年出題された問題の類似問題が今年も出題される確率は低いという出題傾向を考慮する必要がある。

本研究では、公務員試験の対策をするにあたって、今年出題される可能性の低い問題を除いて勉強できるようにすることを目的として、公務員試験問題の類似判定手法を提案する。

2. 類似判定

本研究では、問題に含まれる出題パターン文と出題キーワード、およびベクトル空間モデルを使用して、問題間の類似性を判定する。ベクトル空間モデルとは、文書を多次元空間上のベクトルとして表現し、2つのベクトルを比較することにより類似性を判定する手法である。ベクトルの方向は文書の特徴を表し、2つのベクトルのなす角が小さいほど類似しているといえる。

類似判定に先立ち、「○○の記述のうち、妥当なのはどれか」など、問題文に高頻度で含まれる問題の出題パターン文を用意する。○○は問題の特徴を表す出題キーワードである。問題文の出題パターン文の有無を判定し、含まれていたら出題キーワードを抽出する。また、問題文をベクトルで表現する。

2つの問題の類似性を判定する際には、出題パターン文の有無により判定の方法を切り替える。2つの問題が出題パターン文を含む場合、出題キーワードが同じであれば類似していると判定する。2つの問題のうちいずれか、あるいは両者が出題パターン文を含まない場合、および出題キーワードが異なった場合には、 $tf \cdot idf$ 法により2つの問題をベクトルで表現する。両者のベクトルの余弦値を求め、あらかじめ設定した閾値以上の場合には類似している、閾値未満の場合には類似していないと判定する。以上の方法により、学習対象の問題と、昨年および一昨年に出版された過去の問題との類似性を判定し、類似していると判定された過去の問題が1つでもある場合には、出題可能性が低いと判断して学習対象から除外する。

3. 評価実験

公務員試験の科目のうち行政学、社会学、政治学を対象として、提案手法の有用性を検証する。学習対象の問題としては、TAC出版「14 公務員 V 問題集」に掲載されている問題を使用し、過去の問題に

は 2013 年度 2014 年度の問題を使用する。各科目の過去の問題数，学習対象の問題数を表 1 に示す。網掛けの部分が出題パターン文と出題キーワードによる類似判定で類似していると判定された問題であり，すべてが実際に類似している問題となっている。

表 1 問題数

科目	過去の問題			学習対象の問題				
	すべて	出題パターン文		すべて	出題パターン文		出題キーワード	
		含む	含まない		含む	含まない	一致	不一致
行政学	8	8	0	172	156	16	15	141
社会学	9	9	0	168	109	59	11	98
政治学	11	11	0	193	176	17	14	162

出題パターン文の含まれていない問題，および出題キーワードが一致していない問題に関して，ベクトル空間モデルによる類似判定の閾値を 0.3 から 0.9 まで 0.1 刻みで変化させて得られた再現率，適合率，F 値を表 2 に示す。再現率は実際に類似している問題の中で類似していると判定された問題が占める割合を示しており，適合率は類似していると判定された問題の中で実際に類似している問題が占める割合を示している。適合率が再現率を上回る閾値が，行政学と社会学では 0.5 以上であるのに対し，政治学では 0.4 以上となっている。

表 2 評価結果

4. 考察

tf・idf 法により行政学で 194 語，社会学で 282 語，政治学で 243 語の単語を抽出してベクトルを作成したが，抽出した単語が 1 つも含まれず，いずれの問題とも類似度が 0 となった問題が，行政学と政治学でそれぞれ 2 問ずつ存在した。政治学の学習対象問題には，過去問に類似している問題が他の 2 科目より多く含まれているため，類似問題を漏れなく抽出するのは困難であったと考えられる。

	科目	閾値						
		0.3	0.4	0.5	0.6	0.7	0.8	0.9
再現率	行政学	0.81	0.69	0.54	0.46	0.31	0.27	0.12
	社会学	0.88	0.60	0.48	0.36	0.28	0.20	0.16
	政治学	0.60	0.57	0.43	0.29	0.26	0.20	0.14
適合率	行政学	0.47	0.55	0.67	0.75	1.00	1.00	1.00
	社会学	0.42	0.50	0.50	0.53	0.70	0.63	0.67
	政治学	0.48	0.63	0.75	0.71	0.82	0.78	0.71
F 値	行政学	0.59	0.61	0.60	0.57	0.47	0.42	0.21
	社会学	0.57	0.55	0.49	0.43	0.40	0.30	0.26
	政治学	0.53	0.60	0.55	0.41	0.39	0.32	0.24

類似しているのに類似していないと判定された問題に，「我が国における行政改革に関する記述として、妥当なのはどれか。」と「わが国の行政改革に関する記述として、妥当なのはどれか。」がある。両問題の出題キーワードはそれぞれ「我が国における行政改革」と「わが国の行政改革」で，完全には一致していないため，ベクトル空間モデルにより類似性が判定された。しかし，「行政」と「改革」という単語が抽出されなかったため，低い類似度が算出された。今回は問題文のみを処理対象としたが，選択肢も処理対象に加えることを検討する必要がある。また，ベクトルを構成する単語には「次」，「記述」，「妥当」という 3 つの単語も含まれていたため，問題の特徴を表さないような単語を除いて類似判定することも検討する必要がある。