

令和 6 年 6 月 17 日現在

機関番号：32515

研究種目：挑戦的研究（萌芽）

研究期間：2021～2023

課題番号：21K18420

研究課題名（和文）政治活動における公的資金の透明性・公正性の確保に向けた人工知能技術の活用

研究課題名（英文）Utilization of artificial intelligence for transparency and fairness in political funds

研究代表者

藤原 丈史（FUJIWARA, Takeshi）

東京情報大学・総合情報学部・准教授

研究者番号：60348456

交付決定額（研究期間全体）：（直接経費） 2,000,000円

研究成果の概要（和文）：本研究の目的は、政治活動における透明性の確保に向けたシステム化を念頭に、人工知能技術を活用した公的資金の適正使用の監視、チェック体制を実現することにある。具体的には、2020年度の各都道府県議会における政務活動費を対象としてデータの整備を行い、分析として各都道府県および議員の年度データおよび日次データを対象に、支出項目の比率を特徴量として支出パターンについてクラスタリング手法をはじめとしたさまざまな分析を行った。研究結果として、人工知能技術を活用することにより特徴ある支出を行っている議員の抽出が可能であり、今後のシステム化への道筋が示されたと言える。

研究成果の学術的意義や社会的意義

政治活動における不正な問題は多種にわたるが、その中心となるのは金銭をとまなうものであり大きな社会問題となっている。特に公金が投入される政党や議員への交付金はその用途や対象を含め透明性が強く求められる。一方、その情報公開は依然として非公開や限定的であるなどの問題もあるが、実際に公開されている情報であっても適正な使用かどうかの確認やチェックはデータ量や統一されていない形式であるなど種々問題があり非常に困難である。本研究での人工知能技術を活用した分析結果の知見を活かすことにより、定量的なチェックや監視体制を確立することが可能となり、政治の透明性の実現に寄与できるといえる。

研究成果の概要（英文）：The purpose of this research is to develop a system that monitors and verifies the proper use of public funds by politicians using artificial intelligence technology, ensuring transparency in political activities. In detail, this research analyzes annual and daily data on political activity expenses for each prefectural assembly in FY2020 using clustering methods and various analytical techniques. The results show that artificial intelligence techniques can identify assembly members with distinctive expenditure patterns. This will contribute to the development of a system that ensures the reasonable and fair use of funds by politicians.

研究分野：データサイエンス

キーワード：政治 可視化 情報公開 データサイエンス 都道府県議会 政務活動費 地方自治体 議員

1. 研究開始当初の背景

政治は国家において根幹をなすものであり、そのあり方については公明正大でなければならない。しかしながら、昨今においては政治家の不祥事は後を絶たず、社会的に大きな問題となっている。その問題の中心となるのは収賄、裏金、脱税等さまざまであるが中心となるのはやはり金銭が絡んだものである。その中でも税金が投入されている政党や議員への交付金は、特に透明性をもって公平で公正な利用が求められる。現在は情報公開の必要性が叫ばれており、その要請に答える形で、ある程度は報告書や証明書類等が議会へ提出され、一般にも公開されるようになってきてはいる。しかし、それらの情報から正しく使用されているかどうかの確認および評価はさまざまな問題から非常に困難であり、その対策としてチェック体制も含めた体系的な対策が必要となっている。

2. 研究の目的

本研究の目的は、政治活動の透明性、特に政治資金や政務活動費といった公金における公正で公平な適正使用を促進するための監視およびチェック体制を、人工知能技術(AI)を活用して実現することにある。政治家の活動はさまざまであり、その中で適正な使用であるかの評価は難しい。そのため本研究では、機械学習や統計学など種々の人工知能技術を活用しその支援を行うことで客観的で効率的な評価を実施できるシステムの構築を行い、政治活動における透明性および公平性を確保することを目標とする。

3. 研究の方法

本研究では議員の適正な公金の利用について分析するにあたり、地方議会において議員および会派に交付されている政務活動費、より具体的には各都道府県議会における政務活動費を対象とする。政務活動費は地方自治法第100条で規定されているが、交付対象や額、経費の範囲、およびその収入および支出の報告書等は条例で定めることになっている。つまりは各都道府県によって細かい規定や運用は異なっている。また透明性の確保という観点から、各都道府県は政務活動費において情報公開を行っているものの、その公開のレベルもさまざまである。したがって、まずは政務活動費全般についての各都道府県の状況の調査から行った。次にその結果を受け、各都道府県における年度単位の政務活動費の決算である収支報告書について取りまとめた。これには Web 上で公開されている電子データ、公文書公開請求を個別に行い入手した紙書類からのデータなどさまざまな形式があり、多大な時間を費やしこれらのデータを分析可能な形式へと前処理等を行った上で分析データとしての整理を行った。さらに分析対象データとしては、一部の都道府県については年度単位の収支報告書単位だけではなく、日次の支出を対象とした時系列データの整備も行っている。

具体的な分析としては、各都道府県ごとにまとめた基礎集計を行い、収支報告書における支出の分類である支出項目ごとの額を比率に変換したものを特徴量として各都道府県の特徴をクラスタリング手法、異常値検出手法などを適用し分析を行った。さらに議員ごとについても同様な分析を行った。より詳細な分析として、一部の都道府県を対象に日次データによる議員ごとの支出パターンに着目した分析を時系列クラスタリング手法を用いて行った。これらの結果から支出パターンに着目した分析による特徴的な支出パターンを行っている都道府県、および議員について抽出の可能性を考察した。

4. 研究成果

・都道府県単位の分析

各都道府県議会の2020年度における政務活動費について基礎集計を行った。会派および議員を総計した支出総額(残余除く)を図1に示す。総額で見れば人口や県内総生産額等が大きい都道府県が比例して大きくなっている。これを各都道府県人口一人当たりで表したのが図2、それを地図上で表したものが図3であり、産業規模が小さく人口が少ない都道府県においては県内人口一人当たりの負担は大きくなっている。

各支出項目の比率を都道府県ごとに示したのが図4である。支出項目は会派および議員の各支出についてどのような用途で使用したかを分類したもので、10分類(調査研究費、研修費、広聴広報費、要請陳情等活動費、会議費、資料作成費、資料購入費、事務所費、事務費、人件費)となっている¹。この支出項目ごとの比率を特徴量として分析することで、支出パターンにおける特徴を抽出することを行った。具体的には各クラスタリング手法を用いて支出パターンが類

¹ この支出項目の分類については各都道府県ごとに条例で規定されており、多くの都道府県ではこの10分類となっている。しかしながら、いくつかの都道府県では異なる分類になっているため、本研究ではそれらの都道府県については各都道府県の政務活動費についての要項や手引きを考慮し比例配分等の前処理を行っている。

似している都道府県をクラスタとして分類を行った。階層型クラスタリングを用いた分析結果を図5、各クラスタの都道府県数と項目比率を図6、図7に示す。この結果からは「広聴広報費」「人件費」の支出割合の違いにより大きく2つのクラスタに分類されることがわかる。一方、山梨県と徳島県の2県のみのクラスタが形成されている。この2つの都道府県の支出パターンは他と大きく異なり、「広聴広報費」の割合が非常に大きいことが特徴となっている。

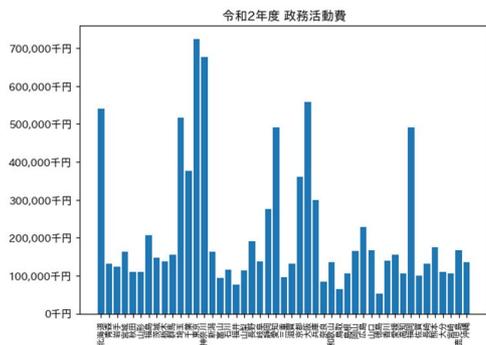


図1 各都道府県支出総額

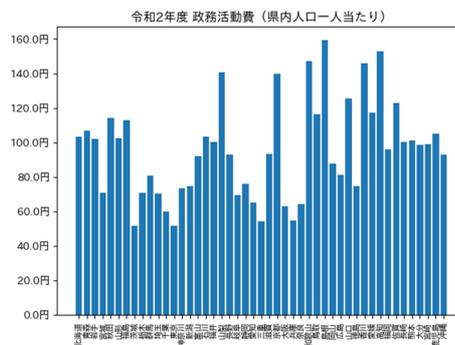


図2 各都道府県支出総額県内人口一人当たり

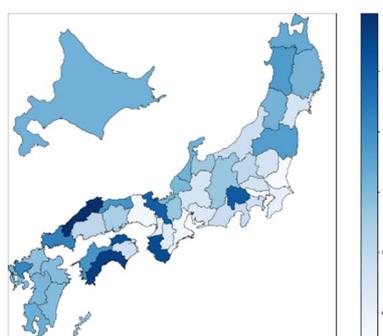


図3 各都道府県支出総額県内人口一人当たり

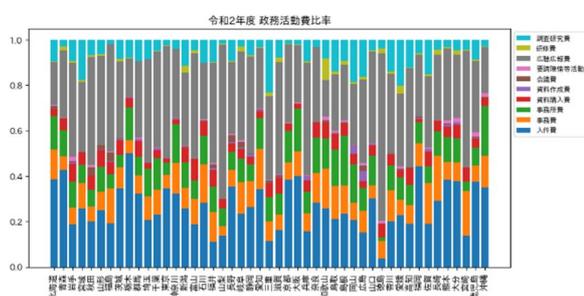


図4 各都道府県支出項目比率

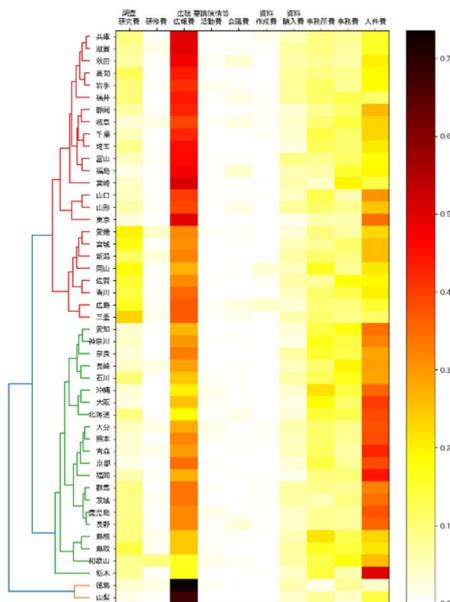


図5 各都道府県議会の階層型クラスタリ

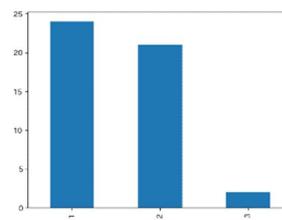


図6 各クラスタの都道府県数

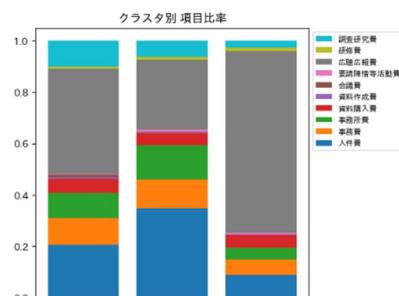


図7 各クラスタの項目比率

・議員単位の分析

2020年度における政務活動費収支報告書の提出されている全国の議員1582名について、支出項目の比率を特徴量とした分析を行った²。これには都道府県の分析と同様にさまざまなクラス

² 政務活動費の交付対象としては各都道府県ごとに、会派、議員、会派および議員、の3つのパターンがあり、それぞれ政務活動費の収支報告書の作成単位が異なっている。例えば交付が会派であれば、収支報告書も会派全体で1つで各

タリング手法、さらに異常検知手法を用いた。階層型クラスタリングの結果を図8、各クラスタの議員数および項目比率を図9、図10に示す。大きなクラスタとしては、「広聴広報費」が支出の中心であるクラスタ1、比較的バランスよく各項目を支出しているクラスタ3、そして「人件費」を中心に支出しているクラスタ7の3つである。注目すべきは極端に支出が偏った小さなクラスタが見られることであり、これらに分類された議員は他の多くの議員と比較して特徴ある支出を行っており、より詳細な支出の確認が必要となる可能性をもっているといえる(クラスタ2,4,5,6)。

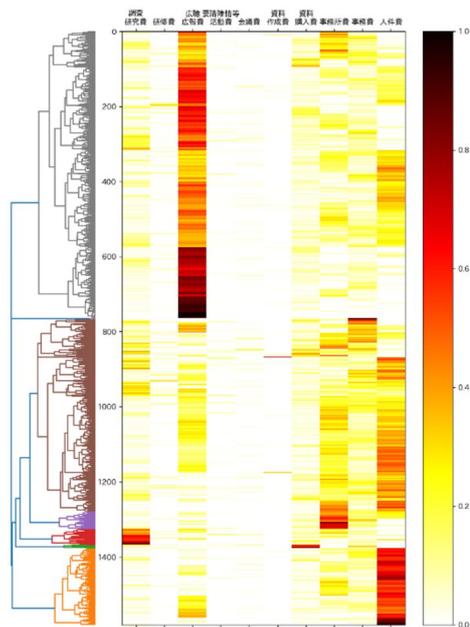


図8 各議員の階層型クラスタリング

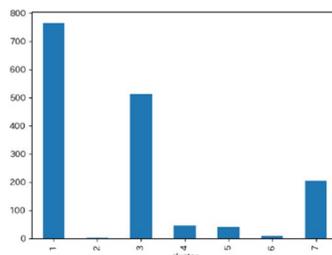


図9 各クラスタの議員数

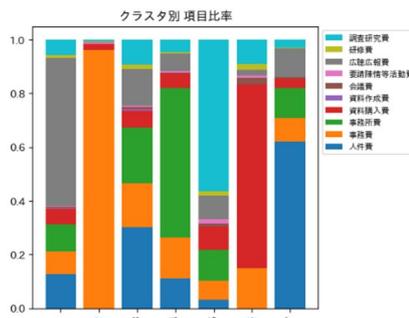


図10 各クラスタの項目比率

・日次データによる議員単位の分析

大阪府議会 2020 年度の全議員 89 名について、日ごとの支出データである時系列データを対象に分析を行った。まずは年度としてまとめたデータについて上記と同様な分析を行い大阪府議会での特徴を確認した。階層型クラスタリングの結果を図11、各クラスタの議員数、項目比率を図12、図13に示す。この結果では大きく2つのクラスタに分類できることがわかるが、特徴的な議員の抽出という観点では分類は困難である。

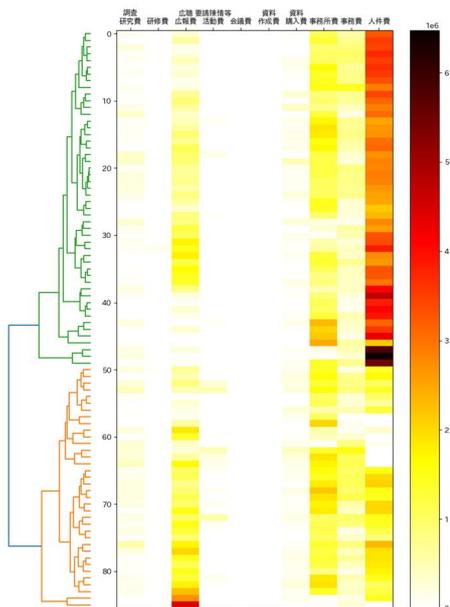


図11 大阪府議会各議員の階層型クラスタリング

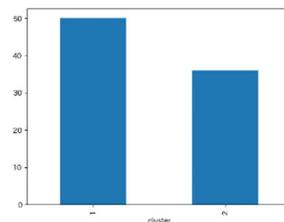


図12 各クラスタの議員数

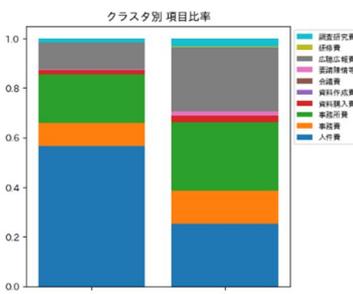


図13 各クラスタの項目比率

次に日ごとの各項目の支出額である日次データでの分析を行った。実際には今回の支出については日次データとしてはスパースであるため、月ごとの各項目の支出額を集計した月次デー

議員の報告書は無い場合もある。ここでは会派分は除いた議員ごとの報告書が提出されているものだけを対象とした。

タを分析対象としている。分析には時系列クラスタリング手法であるDTW(Dynamic Time Warping)を類似度としたk-means法を適用したUMAP(Uniform Manifold Approximation and Projection)を使った2次元マッピングを図14、各クラスタの議員数、年度でまとめた支出項目比率、支出項目ごとの月次平均支出額をそれぞれ図15、図16、図17に示す。少数のクラスタであるクラスタ4(議員数1)とクラスタ5(議員数2)は、他の議員とは異なる特徴的な支出パターンを持っているといえる。特にクラスタ4は、年度でまとめた支出項目比率(図16)では最大規模のクラスタ3と比較してもその特徴はほぼ変わらず識別できない。しかしながら、支出項目ごとの月次平均支出額(図17)ではクラスタ3とクラスタ4では大きく異なることが明らかになっている。クラスタ4は「事務所費」の支出が12月に突出しており、他の議員とは大きく異なる支出パターンであり、より詳細な確認の必要性を示している。

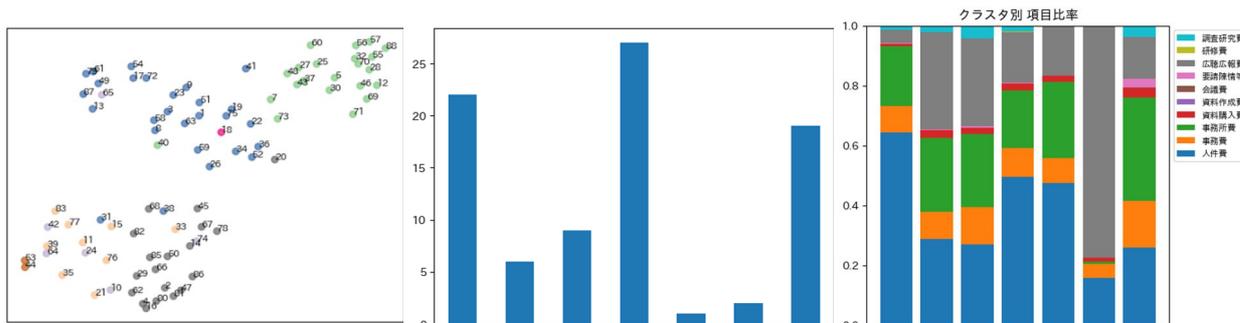


図14 UMAPを用いた各議員のマッピングとクラスタ

図15 各クラスタの議員数

図16 各クラスタの項目比率(年度)

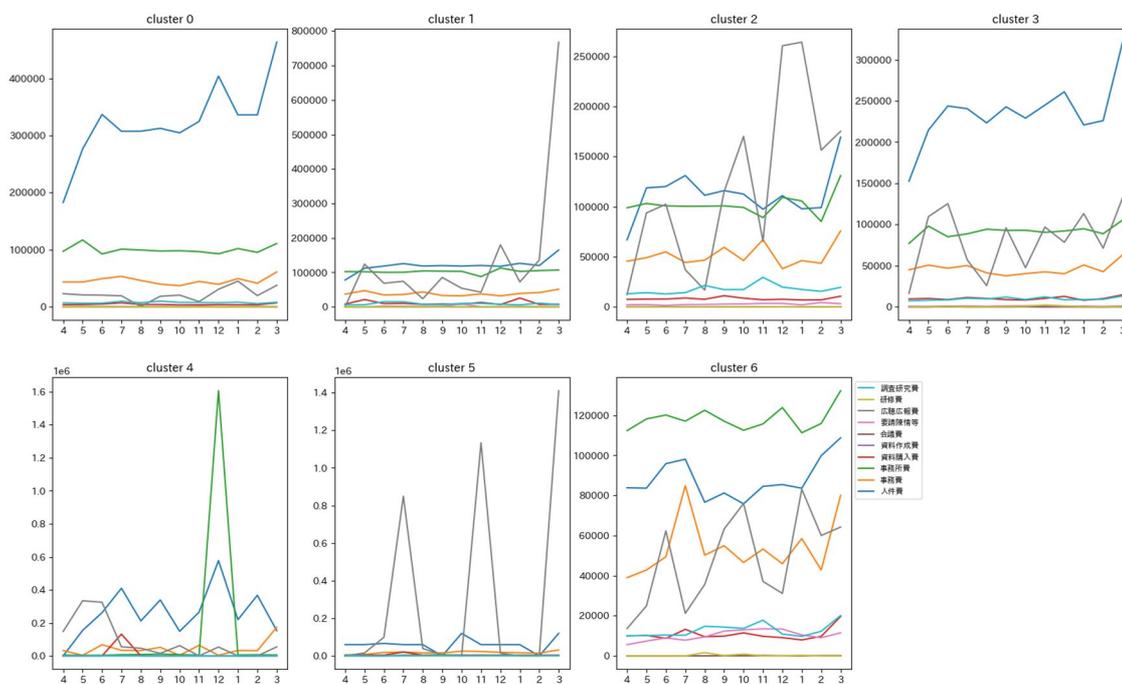


図17 各クラスタの支出項目ごと月次平均支出額

・まとめと今後

本研究では、都道府県議会における政務活動費についてデータの整備と、支出パターンから各都道府県および各議員の特徴の抽出を人工知能技術を活用して行った。この結果、他の多数の都道府県および議員とは異なる支出パターンをもつ特徴的な都道府県、議員を識別できる可能性を示せた。この知見を活用することにより、公正で公平な公金の適正使用について客観的な評価を行い、政治活動における透明性および公平性の確保に向けたシステム化への道筋が見えたといえる。

今後の課題としては、対象年度および対象範囲が挙げられる。本研究での対象年度は2020年度の1年度分だけであり、他の年度と比較した時系列での分析が必要である。また対象についても都道府県議会だけでなく、市区町村議会まで対象を広げて今回の分析結果の有用性を確認するべきである。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------