

判例の構造を利用した判例文書検索

江越 裕紀[†] 片上 大輔[†] 新田 克己[†]

あらまし 本研究は、判例の特徴を利用し、事件を記述した文書をクエリとする類似判例検索を提案する。判例は1つの文書ではなく役割の異なる複数の文書からなるため、ベクトル空間法を用いた検索だけでは精度が悪い。そのため本研究では、法的な専門用語が偏って出現する判例の特徴を利用する。まず、キーワード検索のクエリとして用いられ易い専門用語の検索有用性を検証するため、TF*IDF法に専門用語を応用した検索とTF*IDF法のみを用いた検索の比較実験を行った。また、適合性フィードバックによる精度改善の結果、文書ベクトルに法的な専門用語が用いられやすくなることの検証を行った。さらに判例構造を利用して、判例を争点ごとに分割した文章に対する検索実験を行った結果についても報告する。

Judicial precedent retrieval using the structure of documents

Hiroki EGOSHI[†] Daisuke KATAGAMI[†] Katsumi NITTA[†]

Abstract We propose a judicial precedent retrieval system using the characteristic of documents and Vector Space Model. The system uses a judicial precedent document as a query. Judicial precedent retrieval using only TF*IDF is not good in accuracy, because of the structure of a judicial precedent which consists of some documents. We use judicial terms that are unevenly distributed among issues. First, we compare the system using TF*IDF and judicial terms, and the system using only TF*IDF to verify value of judicial terms for judicial precedent retrieval. Then, we confirm the number of judicial terms increase in relevance feedback retrieval. And, to use the structure of judicial precedents, we experiment the effect of issues retrieval.

1. はじめに

判例とは、過去の裁判の内容を文書にまとめたものである。判例DBの文書量は膨大であるため、判例検索は、専門家にとって重要な役割を果たしている。現状の判例DBの検索には、キーワード検索が用いられているが、絞込みに有効なクエリの選択が難しく、また、検索結果がランキングされないため、重要な判例が見つかりにくいという問題がある。また、事件を記述した文書から、関連の判例を検索したいというニーズがある。そのため、近年、事件の記述を入力し、TF*IDF法等を利用することで、類

似度ランキングを行う判例検索の試みがなされ始めた。例えば、督永らは判例データベースのアブストラクトを用いて、その類似性に着目した類似検索に一定の効果があることを報告している[1]。

しかし、判例の検索にはベクトル空間法を用いるだけでは検索精度が良くない問題がある。その理由は、判例は1つの文書ではなく、役割の異なる複数の文書に分けることが出来、どの部分に着目するかによって語の出現頻度が変わるためである。すなわち、当事者名や裁判官名といった書誌的な情報部分や、判決のみを記した主文、事実を記した部分、争いを記した原告及び被告の主張等の構造を用いずに、1つの文書として扱うことが、精度の低下につながっている。

そこで我々は、ベクトル空間法による検索に加え、判例構造の特徴を利用することで精度改善を目的とした研究を行っている。

[†]東京工業大学 大学院総合理工学研究科知能システム科学専攻

[†] Department of Computational Intelligence and Systems Science Interdisciplinary Graduate School of Science and Engineering Tokyo Institute of Technology

[†] E-mail: {egoshi,katagami,nitta}@ntt.dis.titech.ac.jp

判例特有の法律判断に用いられる語の利用で精度改善が期待できるか、また、判例の構造を利用した部分別検索を行い、事件事実に関する類似検索、法的な内容に関する類似検索といったように目的を分けた検索が精度改善に有効か検証する。また、別の視点から、検索結果に対するユーザの評価を再検索に利用する適合性フィードバックによる類似判例検索が有効であるか、フィードバックを繰り返した結果、どのような特徴語が文書ベクトルに用いられるようになるかを評価する。

本論文では、まず第2章で判例構造及び特徴を述べ、第3章ではTF*IDF法による判例類似検索と法律判断用語を利用した検索の比較実験について述べる。第4章では、適合性フィードバックを争点別に分割した判例に対して用いた場合の評価について述べ、第5章では判例要旨やニュース記事からの類似判例検索の評価について、第6章ではまとめを述べる。

2. 判例構造を利用した判例検索

文書ベクトルを利用した検索は、キーワード検索に比べて文書の類似度を測ることが容易であり、結果のランキングの指標が立てやすい。しかしながら、一般文書と異なり、判例構造は内容の異なる複数部分の組み合わせで成っているため、全文検索では完全に類似した判例を検索することは難しい。また、ユーザもどの部分に対する類似性を求めているかが異なるため、全文よりも部分に対する検索の必要性がある。以下の節に構造とその利用について述べる。

2.1. 従来の判例検索

従来の判例検索には、判例本文検索、要旨検索共に主にキーワード検索が用いられている。例として、図に判例体系[2]の操作画面を示す。キーワードは、民刑区分、本文中の出現語、参照法条、裁判年月日、裁判所名といったカテゴリに分けられた入力が可能であり、これらは判例の先頭に示されている裁判記録の記述を利用している。他に、最高裁判所ホームページ[3]でもこの形式による最高裁判決が参照可能であり前

述のものに加え、事件番号、原審といったカテゴリが用いられている。これらの検索は、検索目的とする判例が明確に定まっている場合に最も有効であるが、事件番号や裁判年月日が入力されなかった場合は、複数のクエリを用いた通常のキーワード検索を行う。このため、有効なクエリを複数用いなければ絞込みがされ難しく、また入力語の有無によるブール検索であるため、検索結果から目的の判例を見つけ出すことが難しい問題がある。

2.2. 判例構造分析

判例は1つの文書ではなく、役割の異なる複数の文書に分けることが出来る。以下の節に判例構造を述べる。

2.2.1. 判例構造

判例の多くは、図1のような構成になっており、各段落はそれぞれ内容によって段落を分けた階層構造になっている。

この構造のうち、「争いのない事実」では事件の概要が記述され、「原告の主張」、「被告の主張」、及び、「裁判官の判断」には、法的解釈の記述を含んでいる。文章量は、法的な解釈を含む記述部分が判例の大半の部分を占めている。法的な解釈の記述には、やや専門的な法律判断用語を含んでいる。ここで、法律判断用語とは、法的判断部で用いられる語の内、事例の絞込みに有効と思われる特徴語と定義する。これらは、判例中で争点の内容と密接に結びつく重要な語であるが、出現頻度は一般用語に比べて非常に少ない。また、それぞれが争点の内容を示唆していることが多く、その

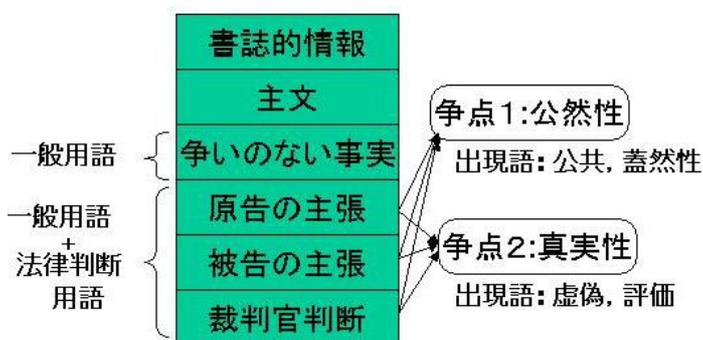


図1:判例文書構造

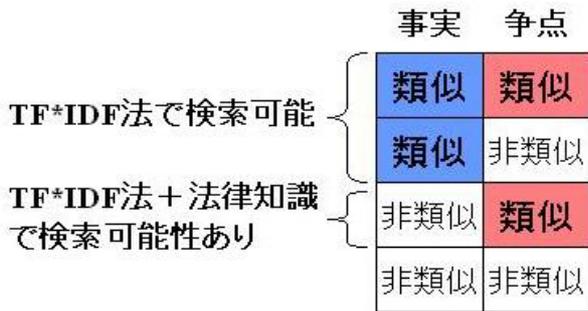


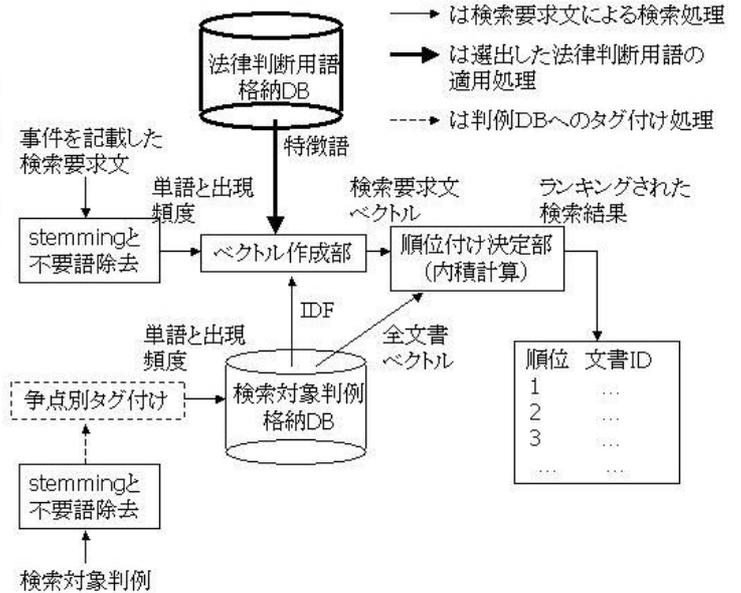
図 2: 検索目的とする判例の種類

ため争点毎に出現する語に偏りがあるという特徴がある。例えば図 1 は名誉毀損に関する判例の構造であるが、問題とされる情報が公に公開されたものであるかという公然性についての争いと、その情報が真実を述べたものであるかという真実性についての争いの 2 争点で争われている。この公然性の争いで出現する法律判断用語は、不特定多数の人の目に触れる情報であるかといった内容の記述に公共、蓋然性等が用いられており、真実性では虚偽、評価といった語が用いられている。本研究では、このように部分、または争点によって異なる法律判断用語が現れる特徴を利用する。

2.2.2. 部分による語の偏りの利用

文書ベクトルによる検索を判例全文に対して行った場合、適切なクエリを用いたキーワード検索ほどの精度が得られないため、本研究では上記判例の特徴を利用して改善を図る。まず、法律判断が成される主張以下の部分に出現する法律判断用語は、キーワード検索のクエリとして選ばれ易い特徴語を含んでいるが、出現頻度が少ないために頻度を尺度とした文書ベクトルに採用されない問題がある。そのため TF*IDF 等を用いる場合は、これらの語に対する重みの修正が必要になる。法律判断用語の検索有用性を確認した予備実験については 3 章で述べる。

また、類似検索では、事件事実についての類似と法律判断についての類似というように、長文である判例全文に対して、ある部分に着目した検索を行うことが出来れば有用である。図 1 の争点別に現れる出現語



の偏りに見られるように、内容に関する部分別に検索を行うことは精度の面でも改善が期待できる。例えば、図 2 に示すように判例が事実の類似、争点（法的解釈）の類似と大きく 2 つの部分から成るとした場合、事実を記述した一般用語の頻度が大きいために TF*IDF 法のみを用いて事実が類似した判例を検索することが出来るが、事実が異なり法的判断が類似した事例を検索するためには低頻度であるが重要な法律判断用語を利用する必要がある。

2.3. 構造自動抽出

後に述べる争点別検索や、特徴語選出への利用のために、判例が 2.2.1 節で述べたような階層構造であることを利用して、本文を段落別に分割する。本研究で補助的に使用している自動タグ付けシステムは、次の要素を決定する。

3. 類似判例検索

本研究で用いる文書ベクトルには、TF*IDF 法による重み付けを用いる。

検索の流れを図 3 に示す。TF*IDF のみを用いた検索は実線の流れ、我々が拡張を行った部分は破線、太線の流れで示される。検索要求文、及び、検索対象文書群から文書ベクトルを作成し、類似度を算出して

る。結果は類似度の降順で示される。ここで、法律判断用語 DB とは、法律判断用語の有用性確認のために人手で選出したものであり、キーワード検索において絞込みに有効と思われる法律用語に代表される語を選んでいる。詳細は次節の評価実験で説明する。

3.1. 比較実験方法

TF*IDF 法を用いた検索は、図 3 実線の流れになる。検索対象は民事、地方裁判所での初審に関する判例 200 件を用いた。

民事判例を用いた理由は、刑事に比べて裁判での判断が難しく、キーワードでの絞込みが行い難いといった特徴のためであり、初審のものを用いたのは、事件概要、及び、判決に関して省略されずに全容が書かれているためである。

検索要求文には、プライバシー侵害を含む名誉毀損、医療における過失の有無、株式に関するトラブルの 3 種類の判例を用いた。予め総争点のうち正解判例を人手により定め、それらを検索目標とする。評価は非常に類似、類似、やや類似、非類似の 4 段階に分け、各判例内の類似している争点数を評価基準とした。以下の争点のうち全てを含むものを非常に類似した事例と判断している。

- 名誉毀損
 - 肖像権、プライバシー権、公共の利害と成りうるか
- 医療
 - 治療に関する過失の有無、治療に関する義務不履行の有無、診療契約の締結等の事実
- 株式
 - 勧誘行為自体の違法性、過当取引にあたるか、先物及びワラント取引契約の締結等の事実

図 3 に示される方法で判例における類似検索を行う。ベクトル作成部では、TF*IDF 検索は TF*IDF 値上位のベクトルを用い、法律判断用語を用いた検索では、人手により選出したベクトルを固定で用いる。なお文書ベクトルの次元数は 200 とする。

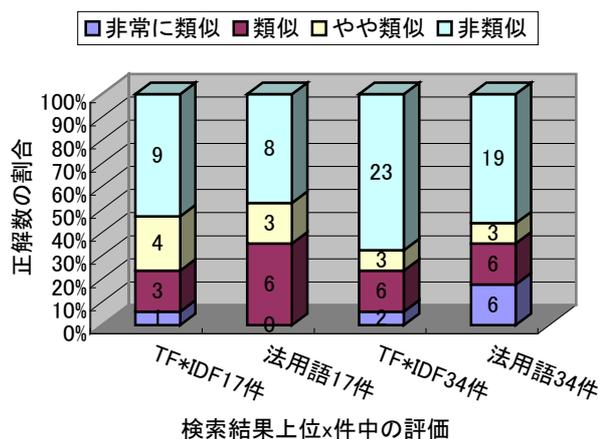


図 4: 名誉毀損判例をクエリとした実験結果

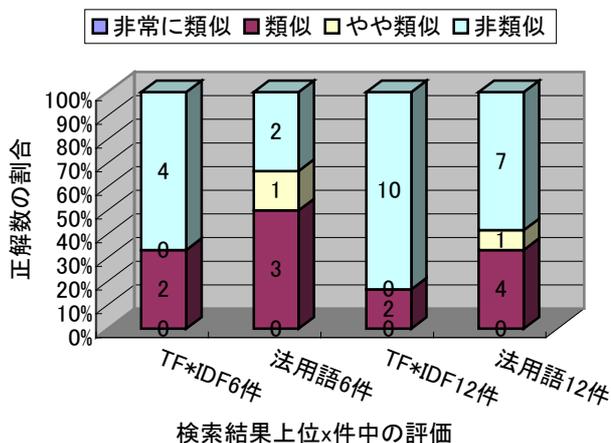


図 5: 医療判例をクエリとした実験結果

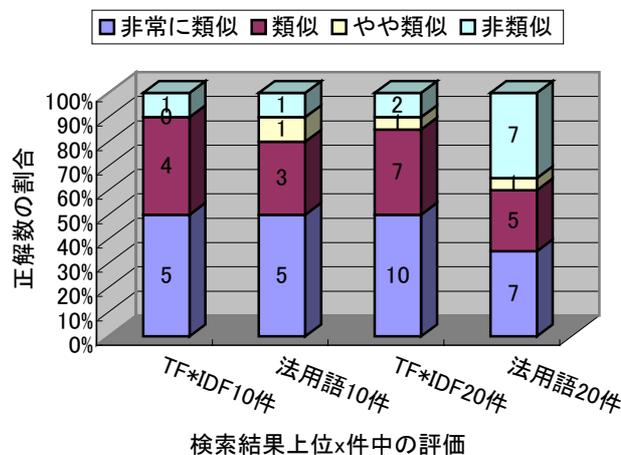


図 6: 株式判例をクエリとした実験結果

3.2. 実験結果

両方のベクトルで3種類の事例について評価を行った実験結果を図4,5,6に示す。グラフ左の2本は、検索結果上位から正解数までの順位中の正解数の比率、また右の2本はその倍の順位までの正解数の比率の評価を行っている。上位X件の数が異なるのは、クエリとする判例によって正解数が異なるためである。判断用語を用いた場合に正解数に改善が見られ、株式については正解数をやや落とす結果となった。これは、名誉毀損、医療に関する判例は、法的な争点が明確にされている傾向の記述が多いが、株式では事件事実から判断を導く記述が多く、定義した専門的な法律判断用語の数も十分でなかったためと考えられる。このため、他2件と異なり、株式の事例同士の類似を見るためには、事実を記述した語を多く用いる必要がある。この問題は、TF*IDFから算出される事実部に関するベクトルを併用することで解消できると考えられ、また、次元数を上げることで精度の改善が期待できると思われる。

しかしながらこれらの結果から、法律判断用語をはじめとする、法的な解釈に関して特徴的な低頻度語を利用することで、法律判断部に関する検索精度が改善される可能性があることを確認した。この予備実験では法律判断用語を手手で選出しているが、それらの処理をシステムにより自動化することで判例検索に有用であると考えられる。

4. 適合性フィードバックと語の偏りを利用した検索

検索有効性の確認された法律判断用語をベクトルに用いた検索に対し、重要な低頻度語の選出の半自動化、及び、詳細な内容に関する類似事例を得るための争点別検索を行った結果を述べる。

4.1. 適合性フィードバックを利用した検索

適合性フィードバックとは、検索結果に対するユーザの評価をシステムに反映させる手法であり、再現率向上に有効とされるものである。図7に適合性フィードバックを用いたシステムの概要を示す。本研究で

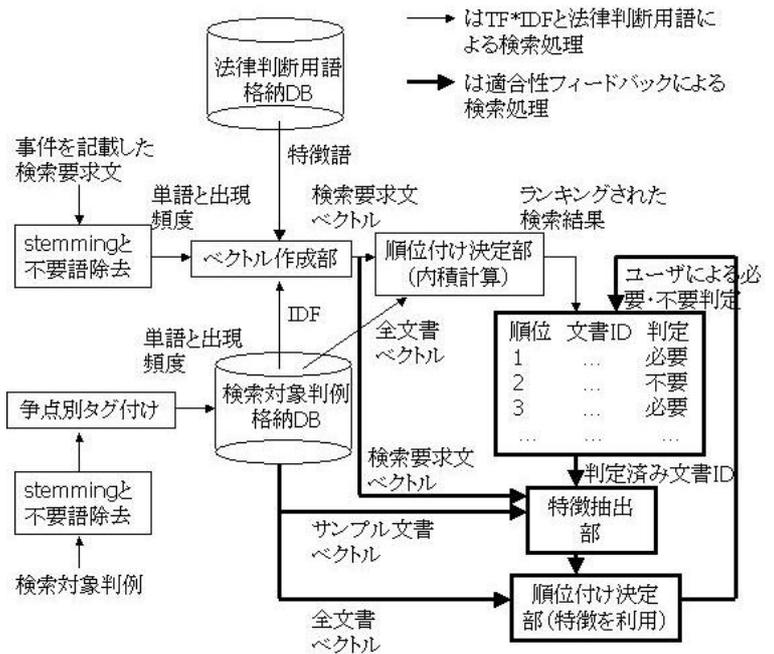


図7: 適合性フィードバックを用いた検索

評価	順位	判例ファイル	類似度
<input type="radio"/>	1	C:\tmp\upload14276.cha 争点1	1.0000000000000002
<input type="radio"/>	2	C:\tmp\upload14276.cha 争点2	0.9948334014989143
<input checked="" type="radio"/>	3	C:\tmp\upload14276.cha 争点3	0.9863008555153429
none	4	C:\tmp\upload14276.cha 争点4	0.9856244152201663
<input type="radio"/>	5	C:\tmp\upload14276.cha 争点5	0.978816921040703
<input checked="" type="radio"/>	6	C:/testcha/28042593.cha 争点1	0.07658081593180625
<input type="radio"/>	7	C:/testcha/28042593.cha 争点2	0.05299226775011101
none	8	C:/testcha/28042593.cha 争点3	0.05215565297464182
none	9	C:/testcha/28041977.cha 争点1	0.04719446838148202

図8: 検索結果画面

は、Rocchio アルゴリズム[4]を使用する。

3.1 節の予備実験で有用性が確認された人手による選出ベクトルを、システムに自動選出させるためにこれを利用した検索を行う。適合性フィードバックは高次元ベクトルを扱うことができる SVM による検索[5]と組み合わせた手法が注目されている。しかし、本研究は、検索対象を判例、さらに民事事件のカテゴリに絞っているため、SVM を用いずに法律判断用語等、特有の語の数が限られている特徴を利用した検索を行う。中島は適合性フィードバックを実現する要素技術の改良の研究において、文書フォーマットを利用し、文書の特定範囲内の単語出現頻度を語の重要度に反映させ重要語を抽出する手法を報告している[6]。すなわち、判例で述べられる事件、争点の

種類に限りがあり，それらを良く特徴付ける語の数にも限りがある，という特徴の元でそれらの語を抽出することを最終的な目標としている．この方法で全判例における事件，争点の種類を網羅するようなベクトルを作成することが出来れば，新しい判例に対してもそのベクトルでの検索が可能であるはずである．

本研究の文書ベクトルを用いた検索結果は，図 8 の示すように，類似度評価によりランキングされて表示される．これらの結果に対し，例えばランキング上位にある判例が，質問判例に類似したものか否かという評価をユーザが指定することで，再検索の際にそれら正・負例のベクトルを反映させる．この評価を繰り返すことにより検索精度の改善を図る．

4.2. 争点別検索

判例中の法的判断部分が複数の争点から成っていることから，法律判断に関する検索には争点別の検索が有効であると考えられる．住田は，文書構造に基づく文書検索において，文書中の部分の意味役割を利用することが検索結果の絞込みやランキング付けに有効であることを報告している [7]．

本研究では，検索要求文，及び，検索対象文を XML タグにより争点毎に分けた判例に対し，法律判断用語を用いたベクトルにより予備実験として類似検索を行ったところ，法律判断等，内容の近い判例争点が検索上位に得られ，同一の判例であっても争点毎に順位に差が見られる，という結果を得た．また，争点には事実の記述も含まれているため，事件の類似した事例も上位に得られた．この結果から，争点別の検索が精度改善につながる可能性が期待できる．

4.3. 類似判例検索実験

Rocchio フィードバック，及び，争点別検索を組み合わせた判例検索の評価実験を行った．対象判例データは，タグで争点別に分割された民事判例 200 件，総争点数 538 件を用い，判例検索実験と同様に，プライバシー侵害を含む名誉毀損，医療における過失の有無，株式に関するトラブルの 3 種

類の判例を検索質問とした．予め総争点のうち正解のものを人手により定め，それらを検索目標とする．文書ベクトルの次元数は 1000 とする．ユーザは検索結果上位 50 件に対して正解，不正解の評価を行い，再検索を繰り返すものとする．フィードバックを繰り返すと類似度自体は下がるため，ランキング上位から総正解数の順位までに含まれる正解数，すなわち適合率で評価を行っている．

フィードバック時にユーザが行う評価は，求める内容であるか否かの 2 値であるため，検索結果の評価も正解数で行った結果を図 9 に示す．いずれも 5 回前後のフィードバックで結果が収束している．名誉毀損や医療に関する事例については 5 回のフィードバックで 30%程の正解率の改善が見られた．

また，3.2 節と同様に正解争点を類似の度合いにより 3 段階に分けた評価を行ったところ，名誉毀損と医療では検索結果上位に得られる「非常に類似」した事例の数が 20%ほど増えたが，株式では不変であった．検索要求文との類似度評価に影響していた文書ベクトルは，フィードバック前には 100 語程度，後には 250 語程度である．それぞれ上位 100 語中に含まれる法律判断用語の数は，名誉毀損，医療の事例では共に約 5 語から約 15 語に増えていた．株式では約 3 語のままだった．また，検索対象文書を争点別に分けてあるため，判例全文に対する検索よりも高い正解率が得られている．

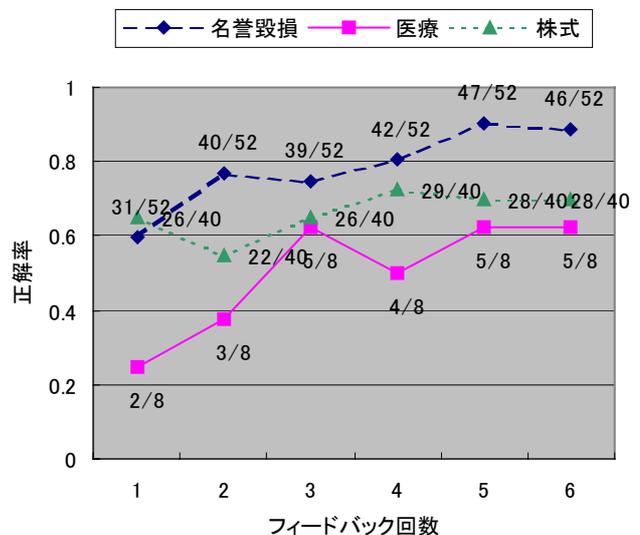


図 9:判例をクエリとした結果

この結果から、精度が改善された事例の検索には 3.2 節の結果と同様に、法律判断用語が特徴語に用いられやすいことが確認され、争点別に内容の分けられた判例に対して Rocchio フィードバックが有効であることがわかった。また、あまり結果が改善されなかった株式では、もともと用いられる法律判断用語が少ないことから、判例中の争点を分けることが難しく、フィードバックによってもあまり精度が上がらなかったと考えられる。

5. 判例以外の文書からの類似判例検索

実用的な検索システムを作るためには、判例同士の類似検索の他に、判例要旨からの類似判例の検索、または一般のニュース記事からの類似判例の検索が行えると便利である。ユーザは要旨にまとめられた内容から、法律判断の類似した判例を導くことが出来、ニュースから事実部の類似した事例がどのような判断に帰結するかといった判例検索の利用が可能になる。また、法学教育の分野で、ある判例について複数の学生に作らせた要約が、システムにより元の判例を検索できるかといった要約のチェックにも用いることが出来る。

判例からの判例類似検索と同様に、要旨、ニュース検索についても名誉毀損、株式トラブル、医療ミスの 3 種類の事例について争点別検索を行い、Rocchio フィードバックによる検索要求文文書ベクトルの重み修正を行っている。

5.1. 要旨からの類似判例検索

検索要求文に用いる要旨は、判例要旨 DB 判例マスター[6]から用い、文章量は 0.5～1 頁程度のものを使用した。検索要求文に類似した判例を検索した結果を図 10 に示す。

名誉毀損では TF*IDF による初期検索及びフィードバック後ともに最も高い検索精度が得られた。名誉毀損、医療ともに明確に争点が分けられる判例が多く、事実等の、法的に重要ではない記述が少ないために、要旨に述べられる内容と近いものが得られていると考えられ、使用ベクトルのうち

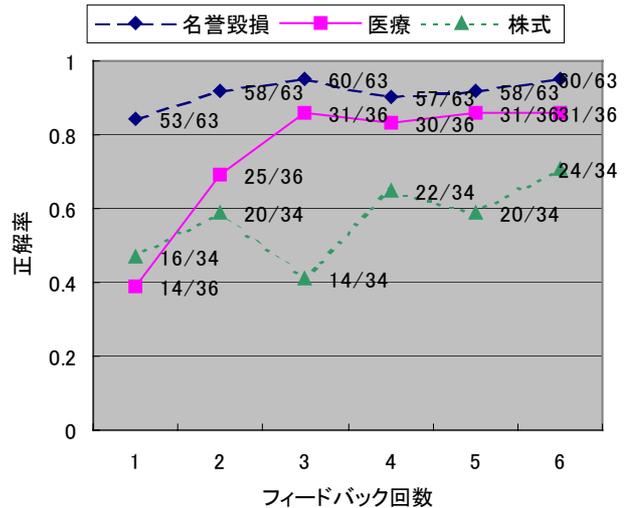


図 10:要旨をクエリとした結果

重みの大きいものに法律判断用語が多く見られている。判例類似検索ではフィードバック後も精度が良くなかった株式についても要旨からの検索ではフィードバックを繰り返すことで徐々に精度の改善が見られたことが特徴的な結果である。要旨 DB では、事実の記述よりも主に法的な内容に関する記述が多く、文章量も少ないため、フィードバックにより絞られた上位の低頻度語が改善に影響していると思われる。株式の正解率が他事例よりも低いのは、同一判例内の争点を特徴付け、区別する法律判断用語が少ないため、他事例にも記述されることが多い争点「賠償額」等、不正解とした争点も上位に検索してしまっていることが原因である。

5.2. ニュース記事からの類似判例検索

記事の文章量は要旨と同様に 0.5～1 頁程度のものを選んでいく。各カテゴリに関する記事は以下のサイトから引用している。

- 名誉毀損
MYCOM PCWEB サイト総合/ニュースカテゴリ「プロバイダ責任制限法施行に伴うガイドラインの修正」
- 医療
日経 BP 社 MedWave サイト「注目される大阪カルテ改ざん裁判、慰謝料請求に司法の判断は？」
- 株式
YOMIURI ON-LINE サイト @Money

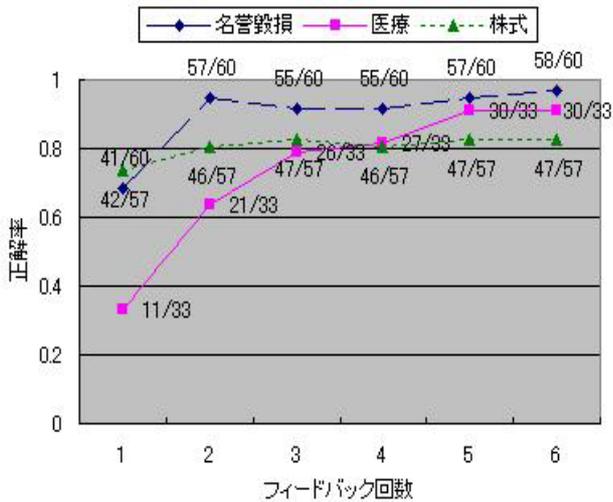


図 11: ニュース記事をクエリとした結果

マネーニュース「きんゆうファイル 金融商品販売法施行」

検索要求文に類似した判例を検索した結果を図 11 に示す。

ニュース記事をクエリとする類似判例検索では、正解率はクエリとした記事に大きく依存するが、5.1 節の実験と類似した結果が得られた。実験では名誉毀損、株式の事例が TF*IDF による初期検索で高い正解率を示した。3 つの記事はいずれも法律判断に関する内容であったため、特徴語には専門用語が多く含まれ、フィードバック後に用いられる特徴語に含まれる法律判断用語の数が多くなる傾向が見られた。

要旨、ニュース記事をクエリとした検索では、共に検索要求文の文章量が少ないため、事実よりも法的な内容に関する記述が固まって現れ、そのため判例をクエリとする場合よりも高い正解率が得られたと考えられる。

6. おわりに

本研究では、判例の特徴とベクトル空間法を利用した、判例をクエリとする類似判例検索を提案した。

比較実験と適合性フィードバックによる特徴語の変化から、判例の特徴である法律判断用語の検索有用性を検証した。また、判例構造を用いて争点別の検索を行うことで、目的に合わせた検索が可能であり、精

度が改善されることを検証した。さらに、争点別検索と適合性フィードバックを用いた類似判例検索が、判例要旨やニュース記事をクエリとした場合にも有効であり、再検索を繰り返すことで、類似評価に用いられる特徴語に法律判断用語の比率が高くなることを確認した。

今後は、専門家によるキーワード検索との比較評価を行う。また、専門家に類似判断を行ってもらい、より正確な評価と重要語の選出を行っていく予定である。

文 献

- [1] 督永涼平, 樋口茂樹, 若木利子, 新田克己: 判例データベースからの類似文書検索システムの開発と評価, 電子情報通信学会, 2003.
- [2] (参考資料) 判例体系. 第一法規出版, 2003.
- [3] 最高裁判所ホームページ.
(<http://courtdomino2.courts.go.jp/home.ns> f).
- [4] Rocchio, J: Relevance feedback in information retrieval, The SMART Retrieval System: Experiments in Automatic Document Processing, pp. 313-323, 1971.
- [5] Thorsten Joachims: Text Categorization with Support Vector Machine: learning with many relevant features, Proceedings of ECML-98, 10th European Conference on Machine Learning, pp.137-142, 1998.
- [6] 中島浩之: 適合性フィードバックによる高精度文書検索の研究, 東京工業大学学位論文, 2003.
- [7] 住田一男: 自然言語処理システムにおける文脈情報の利用に関する研究, 東京工業大学学位論文, 1998.
- [8] (参考資料) 判例マスター. 大日本法規, 2001.