

コミュニティQAにおける文章の表層的特徴に基づく回答文の分析

朱 成敏^a 武田 英明^b

国立情報学研究所

a) *joo@nii.ac.jp* b) *takeda@nii.ac.jp*

概要 参加者達が質問と回答を投稿しながら知恵を共有するコミュニティQAでは、様々な理由によって適切な回答を得られない場合がある。こういった場合、投稿された回答文の中から最もベストアンサーに近い良質な回答を推薦することも考えられる。しかし、質問の中では様々なタイプがあり、それによって回答文の特徴も変わる可能性がある。質問のタイプはその質問が属しているカテゴリと関係があり、カテゴリの特徴に注目した分析を行う必要がある。本研究では、回答文が持つ表層的特徴を用いてベストアンサーの推定を行い、また、その結果がカテゴリ毎にどのように変化するのかが確認する。実際「Yahoo!知恵袋」に投稿された回答文のテキストを分析し、発見した特徴を用いて機械学習による判定結果について述べる。

キーワード コミュニティQA, ベストアンサー, 機械学習, SVM

1 はじめに

コミュニティQAは参加者が投稿した質問に対して他の参加者が回答を投稿するコミュニティサービスであり、新たな知恵を生み出す知識コミュニティの一種である。質問者は投稿した質問に対して集まった回答の中から最も適切だと思われる回答をベストアンサー(BA)として選択する仕組みである。参加者達は様々な質問を投稿し、疑問を解決しながら、または他の参加者の質問に回答をしながらお互い知恵を交換している。勿論、自分に必要な知恵を検索して活用する閲覧者も多い。

しかし、ベストアンサーを選択することを忘れて、質問者の都合によってベストアンサーの選択が行われないうちもある。一方、質問者が集まった回答文の中からベストアンサーを選択したが、より良質な回答が存在する場合もある。こういった場合、知恵を用いる立場である質問者や閲覧者のためにベストアンサー、または良質な回答をシステム側が推薦することによってコミュニティQAの質的向上が期待される。良質な回答の推薦は、質問文と回答文のテキストが持つ情報を分析し、機械分析などの手法を用いてベストアンサーや良質な回答を推定する方法が考えられる。

しかし、コミュニティQAには様々なカテゴリがあり、またその質問の中ではパソコンに関する質問や、社会問題、外国語の翻訳、恋愛相談など様々な話題がある。例えば、パソコンに関する質問はパソコンの機能や使用方法などの情報を、恋愛相談に関する質問は恋愛において他の参加者の経験や意見を聞く質問である。前者の場合は質問に対する客観的な正解が存在し、その情報を与えることで質問が解決されるが、後者の場合は決まった正解がないため多くの参加者が回答を投稿する場合もある。

このように質問の特徴は回答文に影響を与えており、回答文の分析やベストアンサーの推定にも関連があると考えられる。

本研究は、コミュニティQAサイトにおいてベストアンサーまたは良質な回答を推定に質問のタイプによる回答の特徴を分析することで、より質問者の希望を満たせる回答を推定することを目的とする。そして、類似な特徴を持つ質問の単位としてカテゴリ単位での分析と機械学習によるベストアンサーの推定をすることでその影響を確認する。

2 関連研究

コミュニティQAに関する研究は回答文と質問文のテキストが持つ様々な情報を用いてベストアンサーと良質な回答を推定する手法と回答の信頼性を判定する研究が活発に行われている。

石川ら [1] は判定者によるベストアンサーの推定実験を行い、その結果からベストアンサーの選択根拠となる要因を定義した。そして定義した要因の中で良質な回答に影響を与える要因を分類し、機械学習を用いて判定を行った。また、小林ら [2] は不適切な発言が持つ特徴を発見し、機械学習を用いて分類を行った。西原ら [3] はベストアンサーの基準は質問と回答の文体が持つ相性であると仮定し、文末表現の組み合わせに注目してベストアンサーになる可能性の高い回答の判定を行った。このように文章が持つ特徴要素を用いて機械学習による判定を行う研究が多く見られる。

また、回答文が持つ質的問題に関する研究は回答文に補完情報を提示した研究 [5] や、質問者に対して、より適切な回答を選択させるために回答者のメッセージ、外部参照、時間情報を用いた研究 [4] もある。そして、質問の種類を5W1Hで分類して、Webやコーパスでの検

表 1 カテゴリパスの詳細#2

順位	カテゴリパス	質問数	回答数 (平均回答数)	BA 数 (採択率)
1	インターネット, PC と家電 > パソコン > Windows 全般	248,811	500,084(2.01)	248,811(0.498)
2	インターネット, PC と家電 > インターネット	165,539	341,259(2.06)	165,539(0.485)
3	健康, 美容とファッション > 健康, 病気, 病院 > 病気, 病状, ヘルスケア	442,925	999,765(2.26)	442,925(0.443)
4	インターネット, PC と家電 > パソコン	292,657	716,011(2.45)	292,657(0.408)
5	健康, 美容とファッション > 健康, 病気, 病院	156,400	387,087(2.47)	156,400(0.404)

表 2 カテゴリパスの詳細#1

順位	カテゴリパス	質問数	回答数 (平均回答数)	BA 数 (採択率)
1	Yahoo! JAPAN > Yahoo!オークション	255,615	1,473,191(5.76)	255,615(0.174)
2	生き方と恋愛, 人間関係の悩み > 恋愛相談, 人間関係の悩み	444,521	2,276,502(5.12)	444,521(0.195)
3	Yahoo! JAPAN > Yahoo!知恵袋	380,200	1,668,888(4.39)	380,200(0.228)
4	ニュース, 政治, 国際情勢 > 政治, 社会問題	367,043	1,486,862(4.05)	367,043(0.247)
5	子育てと学校 > 子育て, 出産 > 妊娠, 出産	162,610	650,423(4.00)	162,610(0.250)

索結果を適用した研究もある [6][7] .

これらの研究に対して本研究は回答文のテキストが持つ表層的特徴を用いてベストアンサーの推定を行う。そして、カテゴリと推定結果との関係、カテゴリと回答文の特徴との関係も確認する。

3 コミュニティQAにおける質問の分類

本研究では、質問の特徴をカテゴリ毎に確認するために「Yahoo!知恵袋¹」のデータをカテゴリ単位で分類し、質問数と回答数から平均回答数を、ベストアンサーの数から平均採択率も求めた。質問数の多い上位 20 件のカテゴリパスの中から平均回答数を昇順と降順の 5 件を表 2 と表 1 に示す。この結果から見ると「PC と家電」と「健康, 病気, 病院」に関するカテゴリが平均回答数が少なく、「生き方と恋愛, 人間関係の悩み」、「ニュース, 政治, 国際情勢」、「子育て, 出産」に関するカテゴリの方が平均回答数が多かった。Yahoo!サービスに関する回答の場合は問い合わせシステムと併用しているからだと考えられる。このようにカテゴリ毎に回答における特徴があり、この特徴は質問の特徴でもある。

三浦ら [8] の知識共有コミュニティの分析に関する研究では、質問を「正解あり」と「正解なし」の二つのタイプが存在することを指摘した。「正解あり」タイプの場合は、信頼性の高い、客観的な情報を求める場合が多い。一方「正解なし」タイプの場合は他の人の意見や経験を求める場合である [9]。この分類から考えると平均回

答数が多い「生き方と恋愛, 人間関係の悩み」、「ニュース, 政治, 国際情勢」、「子育て, 出産」の場合、「正解なし」の質問が多いカテゴリだと考えられる。一方「PC と家電」と「健康, 病気, 病院」のカテゴリは「正解あり」タイプの質問が多いカテゴリだと予想される。このようにカテゴリによって質問タイプが分かれ、またそれに応じて回答文も異なる特徴を表すと考えられる。この特徴を用いるとよりベストアンサーに近い回答が推定できると考えられる。

4 予備実験

まず、カテゴリ毎に回答文が持つ特徴を確認するため予備実験を行った。回答文のテキストから特徴要素を抽出しベストアンサーとの関係性を確認する。ベストアンサーに影響を与える特徴要素とカテゴリによって与える影響の差について述べる。

4.1 実験概要

ベストアンサーとして選択された回答文の特徴をカテゴリごとに確認するために本研究では 5 つのカテゴリから回答文データを用意した。用意した回答文のカテゴリは「恋愛相談, 人間関係の悩み」、「生き方, 人生相談」、「プロ野球」、「Windows 全盤」、「一般教養」である。回答文はそれぞれ約 2 万件を基準として収集した。実験データの詳細は表 3 に示す。これらの回答文のテキストから特徴要素を抽出し、特徴要素がベストアンサーに与える影響を相関係数を用いて確認する。

¹Yahoo!知恵袋, <http://chiebukuro.yahoo.co.jp/>

表 3 実験データ

区分	カテゴリ	質問数	回答数	BA 数
A	恋愛相談, 人間関係の悩み	4,337	20,333	4,230
B	生き方, 人生相談	2,024	20,482	2,026
C	プロ野球	4,184	20,345	4,176
D	Windows 全般	9,618	21,023	9,490
E	一般教養	4,801	20,133	4,789
計		24,964	102,316	24,512

表 4 機械学習に与えたテキスト特徴要素

#	特徴要素	#	詳細
1	質問形語尾の出現数	10	主張動詞の出現数
2	勧誘形語尾の出現数	11	挨拶の出現数
3	敬語形語尾の出現数	12	参照記号の出現数
4	推測形語尾の出現数	13	外部リンクの出現数
5	命令希望形語尾の出現数	14	文数
6	意見要求の出現数	15	「？」の出現数
7	理由節の出現数	16	「！」の出現数
8	条件節の出現数	17	平均文字数
9	逆接語の出現数	18	Byte 数

1-16:文章別平均

4.2 回答文の特徴要素

本研究では回答文が持つ特徴とベストアンサーとの関係性を確認するために回答文のテキストから特徴要素を確定し、抽出を行った。回答文から文章表現と文法の特徴などの文法的要素、文数や外部リンク、参照記号などの構造的要素、そして挨拶のような意味的要素を用いた。今回、回答文から抽出したテキストの特徴要素は18種である(表4)。これらのテキストの特徴要素はWikipediaの議論テキストを分析した先行研究[10]を参考して用意したものである。

4.3 実験結果と考察

実験データから抽出した特徴要素とベストアンサーから求めた相関係数を表5に表す。一般的に弱い相関を見せた。その中で弱い正の相関(0.2~0.4)を見せた特徴要素を見ると、まず、すべてのカテゴリに対して平均文章数とByte数が影響を与えることがわかった。長い回答が持つ情報の量が関連していたと考えられる。そして、「恋愛相談, 人間関係の悩み」と「生き方, 人生相談」では敬語形語尾の出現数, 理由節の出現数, 主張動詞の出

現数, 文数が影響を与えていた。これらの特徴要素は質問者に対して敬意を表しながら丁寧に回答をしていたと考えられる。一方「Windows 全般」と「一般教養」では参照記号の出現数と外部リンクの出現数が影響を与えていたことが分かった。これらは質問に対して回答の根拠となる情報を提示することによって質問者の要求を満たす回答をしていたと考えられる。

また, 約 20,000 件に合わせて抽出した回答文に対して「生き方, 人生相談」は 2,024 件, 「恋愛相談, 人間関係の悩み」, 「プロ野球」, 「一般教養」は約 4,000 件, 「Windows 全般」の場合は 9,618 件であった。これは質問1つに対して恋愛相談の場合は約 3.88 件, 「生き方, 人生相談」は約 5.12 件, 「プロ野球」は 3.56 件, 「一般教養」は 3.01 件, そして「Windows 全般」は 2.01 件の回答文が投稿されたことを意味する。また, ベストアンサーの割合を確認すると, 「Windows 全般」の場合は 0.45, 「生き方, 人生相談」の場合は 5.12 であった。「Windows 全般」は投稿された回答文の数が最も少ないが, ベストアンサーの割合は最も高い結果となった。「生き方, 人生相談」は最も多い回答文が投稿された。これは「Windows 全般」は「生き方, 人生相談」に比べて質問者が求めた回答が限定されており, その回答以外に適切な回答が存在しない質問が多かったと考えられる。「生き方, 人生相談」, 「恋愛相談, 人間関係の悩み」の場合は, 決まった正解がなく, 質問者の相談内容に応じて誰もが自分の意見を投稿する形式で進行される場合が多く見られる。従って投稿された回答文の数が最も多かったと考えられる。

5 機械学習によるベストアンサーの推測

本章では前章で述べた回答文テキストの特徴要素を用いて機械学習による分類を行う。そして, カテゴリ単位で機械学習を行い, その結果について考察する。

5.1 概要

ベストアンサーを推測するために学習させる特徴量は前章で抽出した回答文テキストの特徴要素を用いた。今回用いる機械学習手法はサポートベクターマシン(SVM)手法である。抽出した特徴要素は学習させるために平均正規化を用いて特徴量正規化を行った。SVMはSVM light²を, カーネルは線形カーネルを用いた。実験データに対する検証は4分割交差検定を用いて行った。

5.2 実験結果と考察

カテゴリ毎に適合率, 再現率, F 値を以下の式に基いて算出し, F 値の平均を算出した。その結果を表6に示す。

$$\text{適合率} = \frac{\text{正解と判定された議論と正解の共通数}}{\text{正解と判定された議論}}$$

²SVM light, <http://svmlight.joachims.org/>

表 5 回答文の特徴要素とベストアンサーとの相関係数

データ	特徴要素								
	1	2	3	4	5	6	7	8	9
A	-0.005	0.119	0.239	0.027	0.019	0.194	0.247	0.017	0.032
B	-0.016	0.203	0.266	0.104	-0.002	0.142	0.242	0.025	0.031
C	-0.002	0.018	0.198	0.030	-0.002	-0.005	0.039	0.032	0.013
D	-0.021	-0.027	0.184	-0.008	0.023	-0.009	0.106	-0.014	0.000
E	-0.028	-0.003	0.227	0.210	0.010	-0.004	0.153	0.008	0.031
A+B+C+D+E	0.009	0.084	0.153	0.007	0.004	0.007	0.115	0.014	0.008
データ	特徴要素								
	10	11	12	13	14	15	16	17	18
A	0.226	0.033	0.105	0.102	0.286	0.014	-0.001	0.281	0.224
B	0.333	0.022	0.087	0.036	0.245	-0.007	-0.034	0.276	0.238
C	0.124	0.012	0.098	0.168	0.021	-0.049	-0.035	0.214	0.260
D	-0.005	0.006	0.229	0.368	0.096	-0.023	-0.035	0.341	0.256
E	0.013	0.001	0.321	0.375	0.004	-0.042	-0.071	0.341	0.305
A+B+C+D+E	0.015	-0.002	0.185	0.140	0.107	-0.006	0.005	0.221	0.247

$$\text{再現率} = \frac{\text{正解と判定された議論と正解の共通数}}{\text{正解数}}$$

$$F \text{ 値} = \frac{2 * \text{適合率} * \text{再現率}}{\text{適合率} + \text{再現率}}$$

実験結果、最も良い性能をみせたのは実験データ D 「Windows 全盤」のカテゴリである。4 つのデータセットの平均適合率は 0.587, 平均再現率は 0.900, 平均 F 値は 0.710 であった。一方、最も低い結果となったデータは「プロ野球」のカテゴリである。最も良い性能を見せたことはベストアンサーとして選択される基準が他のカテゴリに比べて最も明確であるからだと考えられる。予備実験では参照記号や外部リンクが影響を与えていることが分かった。このように質問に対して正解を出せる形の質問がこのカテゴリの特徴であり、推測においても良い性能を期待できると考えられる。「Windows 全版」と同様、参照記号や外部リンクがベストアンサーに影響を与えていた「一般教養」も比較的に良い性能を見せた。「恋愛相談, 人間関係の悩み」, 「生き方, 人生相談」に比べて、ベストアンサーの特徴が明確であり、ベストアンサーの推測においても良い結果が期待される。

6 投票によるベストアンサーの推定

「Yahoo!知恵袋」では質問者が適切な回答を選ばなかった場合、閲覧者の投票によってベストアンサーを決めることも可能である。質問者によるベストアンサーの選択は場合によっては適切ではない回答が選ばれる可能性もあり、多数の参加者が認めた投票によるベストアンサーの方が信憑性のある回答だと思われる場合もある。

本章では質問者によるベストアンサーと投票によるベストアンサー、それぞれが持つ特徴と推定結果について考察を行う。

6.1 分析

本研究で用いたカテゴリ A である「恋愛相談, 人間関係の悩み」の回答文 20,333 件中には 4,230 件のベストアンサーであり、4,230 件のベストアンサーの中には 454 件の投票によるベストアンサーが存在する。これらのベストアンサーを分けて回答文のテキストが持つ特徴要素を用いて相関係数を算出した。

表 7 の結果では、質問者によるベストアンサーに比べて「2. 勧誘形語尾の出現数」, 「8. 条件節の出現数」の特徴要素にも弱い相関があるのが分かった。条件節と理由節は説得における Toulmin モデルの構成要素であり、これは相談内容に対して説得のある回答が選ばれたことだと思われる。また、「18. Byte 数」が質問者によるベストアンサーより強い相関 (中間の強さ:0.4~0.6) をみせ、投票によるベストアンサーはより説得力がある、また丁寧に書かれた回答文であると考えられる。

6.2 機械学習による分類

今回は機械学習を用いて投票によるベストアンサーの分類を行う。前節の分析で質問者によるベストアンサーとは異なる特徴を発見した。本節では、この差が機械学習による分類に与える影響について述べる。

まず、質問者によるベストアンサーと投票によるベストアンサーを分け、それぞれ正解とする訓練データを作成した。そして、4 つのグループに分けて 4 分割交差検

表 7 質問者によるベストアンサーと投票によるベストアンサーの比較

選択方法	特徴要素								
	1	2	3	4	5	6	7	8	9
質問者	-0.007	0.117	0.228	0.025	0.019	0.102	0.241	0.011	0.029
投票	-0.005	0.319	0.239	0.027	0.019	0.001	0.447	0.217	0.032
選択方法	特徴要素								
	10	11	12	13	14	15	16	17	18
質問者	0.223	0.038	0.105	0.104	0.276	0.021	-0.007	0.267	0.201
投票	0.226	0.033	0.105	0.102	0.301	0.014	-0.001	0.384	0.414

定を用いて検定を行った。学習させる手法が前章と同様 SVM を用いた。

分類の結果を表 8 に示す。投票によるベストアンサーの分類が質問者によるベストアンサーの分類結果より高い F 値を見せた。質問者によるベストアンサーを用いて訓練させたデータの平均適合率は 0.528、平均再現率は 0.818、平均 F 値は 0.643 であったが、投票によるベストアンサーを用いて訓練させたデータは均適合率は 0.584、平均再現率は 0.957、平均 F 値は 0.726 であった。特に再現率においては高い結果となった。これは正解が少ないという原因も考えられるが、学習させる特徴要素が質問者によるベストアンサーの場合より比較的明確であると考えられる。

6.3 考察

本章では質問者によって選ばれたベストアンサーと投票によって選ばれたベストアンサーの分析と比較を行った。機械学習による分類では投票によるベストアンサーの分類の方がより良い結果を見せた。これは投票によるベストアンサーの回答文が良質な顔等が持つ特徴をより含んでいたからだと考えられる。そして、この特徴は投票を参加した参加者の多数から良質な回答だと判断された要素であるとも言える。今回の実験で確認した特徴要素を今後、良質な回答の判定における判定根拠として活用し、よりの確な良質な回答を推定する基準として期待される。

一方、質問者に選ばれたベストアンサーの分類実験では結果が投票による場合より低い結果となったのは、多数が正解だと判定した回答文ではなく、質問者の主観によって判定されたからだと考えられる。例えば、回答がベストアンサーとして正しいのかを判断する前に感情的に好む回答、または投稿の速さで選ばれた場合もある。「Yahoo!知恵袋」の仕組み上、ベストアンサーを選ぶ権限はまず質問者にあり、そのベストアンサーが適切ではないとは言いがたい。従って、ベストアンサーを推定する場合はこういったベストアンサーの選定基準について考慮しなければならないと考えられる。

7 考察と今後の課題

本研究ではコミュニティQA サイトである「Yahoo!知恵袋」の回答文データを分析し、ベストアンサーを推定するために考慮しなければならない要素について分析と機械学習を用いて推定を行った。

まず、カテゴリによってベストアンサーとして選ばれた回答文の特徴が異なることを実験を通じて確認した。決まった正解のない相談のような質問が多いカテゴリは知識な情報を求める質問が多いカテゴリより回答文の数が多かった。このように相談を求める質問の場合は、ベストアンサーの選択基準が質問者の主観となり、質問者の希望を満たせる回答が投稿されるまで回答を待つ状況が長くなったと考えられる。または知識や情報より回答者の経験や意見を聞く質問であるため誰もが回答することが出来る話題であることも一つの理由として挙げられる。このように質問者の主観的判断要素を導入して総合的な判定をする方法も考えられる。

今後の課題としてベストアンサー、または良質な回答を判定する際に、回答文の特徴以外に質問者と回答が持つ特徴に注目し、その関係性が与える影響を把握することが考えられる。これにより参加者の質問・回答における投稿活動を分析し、モデル化することによってコミュニティQA の質的状況や参加者の評価が可能になると思われる。

8 おわりに

本研究ではコミュニティQA での回答文が持つ特徴の分析と機械学習によるベストアンサーの推測を行った。また、カテゴリ毎に分析と実験を行うことによって回答文がカテゴリ毎にどういった特徴を持つのかを確認した。カテゴリが持つ質問文と回答文の傾向はベストアンサーまたは良質な回答の選択において大きい影響を与えていることが分かった。今回の結果を参考し、今後は質問・回答に対する参加者の参加姿勢に注目した質問・回答のモデル化とその検証を行いたい。

表 6 SVM を用いたベストアンサーの推測の結果

データ	回答数 (BA 数)	適合率	再現率	F 値
A#1	5,058(957)	0.532	0.879	0.663
A#2	5,059(1091)	0.547	0.843	0.664
A#3	5,058(1078)	0.510	0.816	0.628
A#4	5,058(1104)	0.569	0.867	0.687
(平均)	5120.5(1057.5)	0.540	0.851	0.661
B#1	5,120(532)	0.465	0.832	0.596
B#2	5,121(476)	0.521	0.760	0.618
B#3	5,121(508)	0.492	0.783	0.604
B#4	5,120(510)	0.475	0.796	0.595
(平均)	5120.5(506.5)	0.488	0.792	0.603
C#1	5,086(1037)	0.471	0.903	0.619
C#2	5,087(1061)	0.426	0.897	0.578
C#3	5,086(1041)	0.439	0.836	0.576
C#4	5,086(1037)	0.447	0.912	0.600
(平均)	5086.25(1044)	0.446	0.887	0.593
D#1	5,255(2,393)	0.573	0.897	0.699
D#2	5,256(2,401)	0.589	0.902	0.713
D#3	5,256(2,459)	0.601	0.916	0.726
D#4	5,256(2,237)	0.583	0.886	0.703
(平均)	5255.75(2372.5)	0.587	0.900	0.710
E#1	5,033(1,176)	0.581	0.842	0.688
E#2	5,034(1,181)	0.541	0.831	0.655
E#3	5,033(1,234)	0.563	0.855	0.678
E#4	5,033(1,198)	0.553	0.816	0.659
(平均)	5033.25(1197.25)	0.546	0.826	0.670

謝辞

本研究の実施にあたって、ヤフー株式会社が国立情報学研究所に提供した「Yahoo!知恵袋データ(第2版)」を利用させて頂きました。深く感謝いたします。

参考文献

- [1] 石川大介, 栗山和子, 酒井哲也, 間洋平, 神門典子: Q&A サイトにおけるベストアンサー推定の分析とその機械学習への応用, 情報知識学会誌, Vol.20 No2, pp.73-85, 2010.
- [2] 小林大祐, 松村真宏, 木戸冬子, 石塚満: 知識検索サイトにおける不適切な投稿の分類, 第 21 回人工知能学会全国大会, 2007.
- [3] 西原陽子, 松村真宏, 谷内田正彦: QA サイトにおける質問に適した回答の判定, 言語処理学会, 2007.
- [4] 瀧寛文, 森崎修司, 大平雅雄, 松本健一: Q&A コミュニティを対象とした回答の信頼性指標構築に向けた分析, 情報社会学会誌, Vol.4, No.1, pp.49-58, 2009.
- [5] 高田夏希, 山本裕補, 小山聡, 田中克己: 質問応答コンテンツに対する Web による回答補完, DEIM Forum

表 8 投票と質問者によるベストアンサーの分類結果

データ	回答数 (BA 数)	適合率	再現率	F 値
投票#1	5,058(131)	0.582	0.950	0.722
#2	5,059(100)	0.567	0.912	0.699
#3	5,058(111)	0.601	1.000	0.751
#4	5,058(112)	0.587	0.964	0.730
(平均)	5120.5(113.5)	0.584	0.957	0.726
質問者#1	5,058(957)	0.525	0.795	0.632
#2	5,059(991)	0.536	0.834	0.653
#3	5,058(967)	0.513	0.832	0.635
#4	5,058(992)	0.539	0.811	0.650
(平均)	5120.5(976.75)	0.528	0.818	0.643

C4-6, 2009.

- [6] Eugene Agichtein, Steve Lawrence and Luis Gravano, Learning search engine specific query transformations for question answering, The 10th International Conference on World Wide Web, pp.169-178, 2001.
- [7] Koji Eguchi: NTCIR-5 Query Expansion Experiments using Term Dependence Models, Proceedings of the Fifth NTCIR Workshop Meeting on Evaluation of Information Access Technologies: Information Retrieval, Question Answering and Cross-Lingual Information Access, 2005.
- [8] 三浦麻子, 川浦康至: 日とはなぜ知識共有コミュニティにさんかするのか: 質問行動と回答行動の分析, 社会心理学研究, Vol.23, No.3, 2008.
- [9] 渡邊直人, 島田諭, 関洋平, 神門典子, 佐藤哲司: QA コミュニティにおける質問者の期待に基づく質問分類に関する一検討, DEIM Forum 2011 B5-1, 2011.
- [10] 朱成敏, 武田英明: 参加者の議論能力に注目したオンライン議論のモデル化と分析 ~ Wikipedia の議論ページにおける分析 ~, 情報処理学会論文誌, Vol.55 No.1 pp.244-256, 2014.