

学生生活満足度調査アンケート自由記述欄の解析 II (2016 年度データを用いて)

佐藤友哉¹・塚越 清²

¹ 情報科学センター

² システム情報分野

Text Mining Analysis of Student life questionnaire Data II

Tomoya SATO, Kiyoshi TSUKAGOSHI

Abstract

This paper is a comparison of the free description column of the student life satisfaction investigation questionnaire that Ashikaga Institute of Technology carried out in 2015 and 2016. Analysis of the free description column of the student life satisfaction investigation questionnaire of 2015 is just what We announced it in Research Reports Ashikaga Institute of Technology No.51 March 2017. We clarify hope and a request to Ashikaga Institute of Technology of the student by comparing the free description column of the questionnaire for 2 years. In this paper, association analysis and a co-occurrence network is used.

Keywords: *Text Mining Analysis, Student life questionnaire Data, association analysis, co-occurrence network*

1. はじめに

足利工業大学では、2015 年から工学部の全学生を対象に後期授業事前ガイダンスにおいて学生生活に関する「学生生活満足度調査」アンケートを実施している。

2015 年の同アンケートの自由記述欄については、テキストマイニング¹⁾で抽出した名詞と形容詞から共起ネットワークと主成分分析を用いて解析を実施し、その結果は足利工業大学研究集録第 51 号でも発表した²⁾。

本研究では、昨年度とは違う解析方法であるアソシエーション分析と共起ネットワークを用いて、2015 年と 2016 年の 2 年間のアンケート自由記述欄を比較することで、学生からの要望や不満

の変化を捉えることを試みた。

2. 解析に用いたデータについて

2 度実施した同アンケートは、学生自身に関する基本項目、学生生活に関する項目、課外活動に関する項目、大学施設に関する項目、授業に関する項目、大学全体に関する項目からなる 50 問の選択問題、さらに 9 項目に関する意見・要望の自由記述欄と多岐にわたるものである。

今回は、昨年の解析結果を考慮して“学生生活についての要望や不満”の自由記述欄に焦点を絞り解析を試みた。

2015 年と 2016 年のデータ量から単純比較すると全体では、2016 年のデータが 3 割程度減少した(表 1)。

表 1 データ量の比較

	2015年度	2016年度
1年生	46	20
2年生	51	37
3年生	35	32
4年生	34	30
総記述数	166	119

これらの記述内容から名詞と形容詞について、R パッケージ「RMeCab」, 「arules」, 「arulesViz」^{3) 4)} で可視化を行い、さらに「KH Coder」⁵⁾ でも可視化を行った。

3. 分析について

今回利用したアソシエーション分析とは、トランザクションデータをもとに「もし A ならば B である」という因果関係を見つける手法である。別名バスケット分析と呼ばれ、POS システムのビックデータから「A 商品を購入した人は一緒に B 商品も購入する」といった解析に用いられる。

共起ネットワーク分析とは、「複数の語の繋がりをネットワーク図化」したもので、感覚的に語と語の繋がりを確認することが可能である。

3. 1 名詞と形容詞の出現頻度

2015 年のデータとしては、記述数は 166 メッセージで名詞と形容詞の数は 350 アイテムであり、2016 年のデータとしては、記述数は 119 メッセージで名詞と形容詞の数は 242 アイテムあり、名詞と形容詞の出現頻度トップ 10 を示す(図 1, 図 2)。

2016 年では、「欲しい」が 31 件で最も多く、続いて「スクールバス」が 17 件、「学生食堂」が 13 件、「ない」が 10 件、「タバコ」・「時間」・「本数」がともに 8 件となる。

2015 年では、「欲しい」が 51 件で最も多く、続いて「スクールバス」が 28 件、「時間」が 19 件、「学生食堂」が 16 件、「少ない」が 15 件となる。

3. 2 アソシエーション分析

アソシエーション分析は、単語の出現回数の比率、単語間の共起回数の比率、全単語の出現回数に対する単語間の共起回数の比率により重要なキーワードを探す手法である。

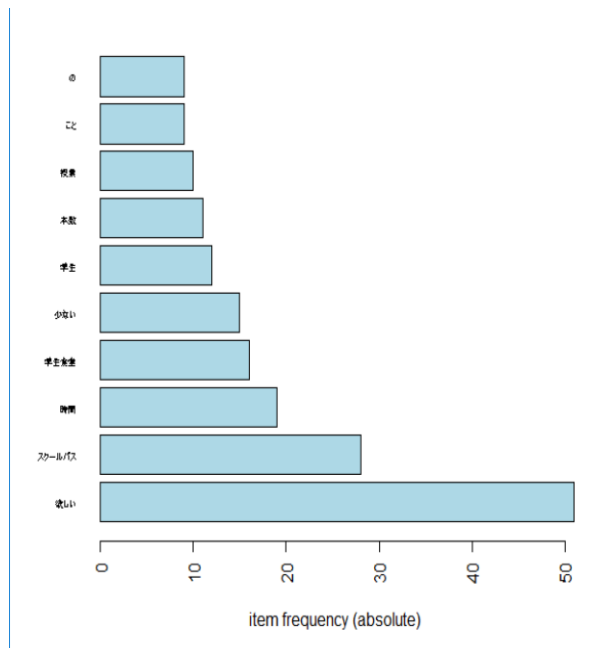


図 1 名詞・形容詞出現トップ 10 (2015 年)

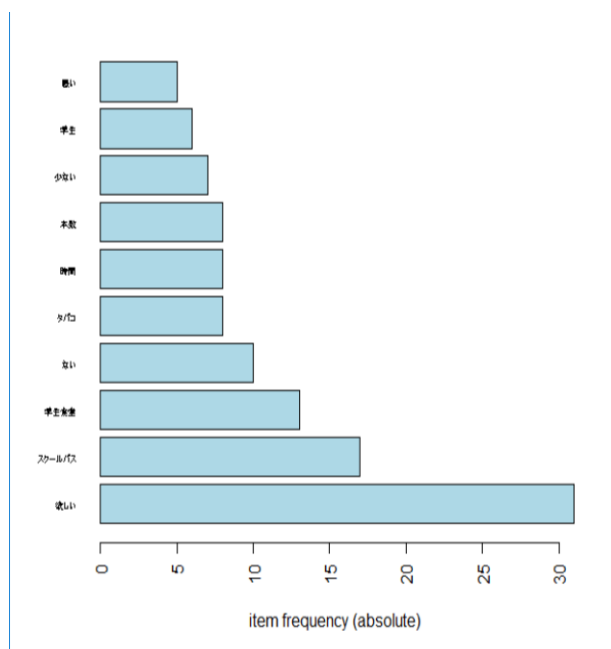


図 2 名詞・形容詞出現トップ 10 (2016 年)

ルール数は、支持度 (support), 信頼度 (confidence), maxlen (頻出アイテムの最大数) のパラメータを低く設定すると抽出過多となり、あまり意味を持たないルールまで含まれてしまう可能性があり、逆にパラメータを高く設定してしまうと抽出減となり、重要なルールが欠損してしまう可能性がある。そのため調整し、support=0.01, confidence=0.5, maxlen=10 に設定し、その結果からは 2015 年のデータでは 272 個のルール

が抽出され、2016 年のデータでは 50 個のルールが抽出された。

3. 2. 1 支持度 (support)

支持度 (support) は条件部・文字列と結論部・文字列を含む事例が全事例中に占める比率である。2015 年と 2016 年の支持度をそれぞれ表 2 と表 3 に示す。

2015 年支持度の上位からは、1 位が「時間」と「スクールバス」で 0.072, 2 位が「本数」と「スクールバス」で 0.066, 3 位が「本数」と「欲しい」で 0.036, 4 位が「スクールバス+本数」と「欲しい」で 0.036, 5 位が「本数+欲しい」と「スクールバス」で 0.036 となる。さらに、上位 15 ルールからは「スクールバス」関連、「授業時間」関連、「学生食堂」関連に占められていることが確認できる。

表 2 支持度上位 15 ルール (2015 年)

条件部		結論部	支持度
時間	⇒	スクールバス	0.072
本数	⇒	スクールバス	0.066
本数	⇒	欲しい	0.036
スクールバス+本数	⇒	欲しい	0.036
本数+欲しい	⇒	スクールバス	0.036
スクールバス+欲しい	⇒	本数	0.036
授業	⇒	時間	0.030
駅	⇒	スクールバス	0.024
所	⇒	欲しい	0.024
時間+欲しい	⇒	スクールバス	0.024
種類	⇒	学生食堂	0.018
メニュー	⇒	学生食堂	0.018
メニュー	⇒	欲しい	0.018

2016 年支持度の上位からは、1 位が「本数」と「スクールバス」で 0.067, 2 位が「時間」と「スクールバス」で 0.050, 3 位が「本数」と「欲しい」で 0.033, 4 位が「スクールバス+本数」と「欲しい」で 0.033, 5 位が「本数+欲しい」と「スクールバス」で 0.033 となる。さらに、上位 15 ルールからは「スクールバス」関連、「学生食堂」関連、「規制緩和」関連に占められていることが確認できる。

表 3 支持度上位 15 ルール (2016 年)

条件部		結論部	支持度
本数	⇒	スクールバス	0.067
時間	⇒	スクールバス	0.050
本数	⇒	欲しい	0.033
スクールバス+本数	⇒	欲しい	0.033
本数+欲しい	⇒	スクールバス	0.033
スクールバス+欲しい	⇒	本数	0.033
電車	⇒	時間	0.025
電車	⇒	スクールバス	0.025
よう	⇒	欲しい	0.025
学生	⇒	ない	0.025
時間+電車	⇒	スクールバス	0.025
スクールバス+電車	⇒	時間	0.025
スクールバス+時間	⇒	電車	0.025
まずい	⇒	学生食堂	0.017
規制	⇒	緩和	0.017

3. 2. 2 確信度 (confidence)

確信度 (confidence) は条件部・文字列、結論部・文字列を含む事例と条件部・文字列を含む事例との比率である。2015 年と 2016 年の確信度をそれぞれ表 4 と表 5 に示す。

表 4 確信度上位 15 ルール (2015 年)

条件部		結論部	確信度
自習	⇒	室	1.000
室	⇒	自習	1.000
工学部	⇒	ロッカー	1.000
種類	⇒	学生食堂	1.000
始業時間	⇒	欲しい	1.000
電車	⇒	時間	1.000
電車	⇒	スクールバス	1.000
調整	⇒	スクールバス	1.000
調整	⇒	欲しい	1.000
用	⇒	設置	1.000
ゲーム	⇒	ホール	1.000
ゲーム	⇒	学生	1.000
メニュー	⇒	学生食堂	1.000
メニュー	⇒	欲しい	1.000
ラウンジ	⇒	女子	1.000

2015 年確信度の上位 15 ルールからは、「スクールバス」関連のルール、「授業の開始時間」関連のルール、「学生食堂」関連のルール、「女子ラウンジ」関連のルール、「自習室」関連に占められていることが確認できる。

2016 年確信度の上位 15 ルールからは、「学生食堂」関連、「禁煙の規制緩和」関連、「学生ホール」関連、「部室棟」関連に占められていることが確認できる。

表 5 確信度上位 15 ルール (2016 年)

条件部		結論部	確信度
まずい	⇒	学生食堂	1.000
規則	⇒	緩和	1.000
緩和	⇒	規則	1.000
メニュー	⇒	学生食堂	1.000
種類	⇒	学生食堂	1.000
種類	⇒	欲しい	1.000
ホール	⇒	学生	1.000
あと	⇒	欲しい	1.000
スペース	⇒	タバコ	1.000
他	⇒	ない	1.000
電車	⇒	時間	1.000
電車	⇒	スクールバス	1.000
汚い	⇒	大学	1.000
部室	⇒	棟	1.000
棟	⇒	部室	1.000

表 7 リフト上位 15 ルール (2016 年)

条件部		結論部	リフト
規制	⇒	緩和	60.000
緩和	⇒	規制	60.000
部屋	⇒	棟	60.000
棟	⇒	部室	60.000
学生食堂+欲しい	⇒	種類	40.000
時間+本数	⇒	電車	40.000
スクールバス+時間+本数	⇒	電車	40.000
所	⇒	喫煙	26.667
喫煙	⇒	所	26.667
汚い	⇒	大学	24.000
ホール	⇒	学生	20.000
スクールバス+時間	⇒	電車	20.000
スペース	⇒	タバコ	15.000
電車	⇒	時間	15.000
電車+本数	⇒	時間	15.000

3. 2. 3 リフト (lift)

リフト (lift) は条件部・文字列, 結論部・文字列の同時発生確率と条件部・文字列・結論部・文字列の発生確率の積の比率である. 2015 年と 2016 年の確信度をそれぞれ表 6 と表 7 に示す.

3. 2. 4 上位 50 ルールの関連について

2015 年の支持度 (support) とリフト (lift) の上位 50 ルールの関連性を図 3 に示す.

表 6 リフト上位 15 ルール (2015 年)

条件部		結論部	リフト
自習	⇒	室	83.500
室	⇒	自習	83.500
科	⇒	看護	83.500
看護	⇒	科	83.500
市	⇒	足利	83.500
足利	⇒	市	83.500
の+女子	⇒	ラウンジ	83.500
駅+市	⇒	足利	83.500
駅+足利	⇒	市	83.500
スクールバス+市	⇒	足利	83.500
スクールバス+足利	⇒	市	83.500
スクールバス+駅+市	⇒	足利	83.500
スクールバス+駅+足利	⇒	市	83.500
用	⇒	設置	55.667
配置	⇒	用	55.667

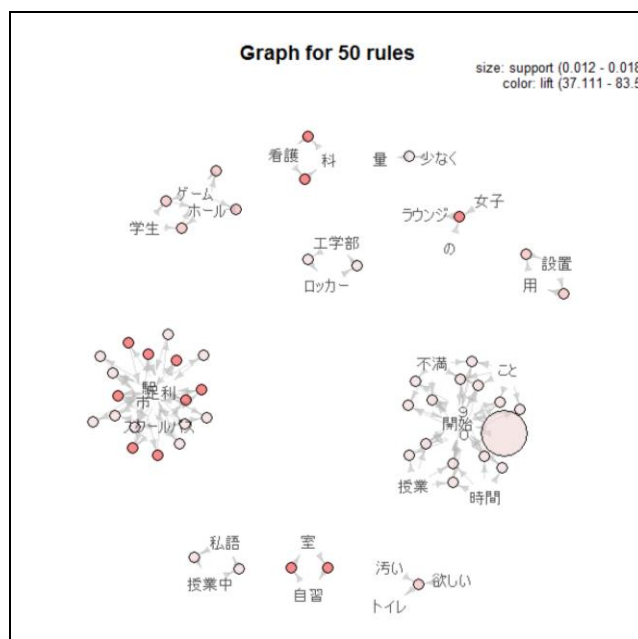


図 3 上位 50 ルールのグラフ (2015 年)

2015 年リフトの上位 15 ルールからはスクールバスに関するルール, 女子ラウンジに関するルール, 看護学科に関するルール, 自習室に関するルールが占めることが確認できる.

2016 年リフトの上位 15 ルールからは禁煙の規制緩和に関するルール, 学生食堂に関するルール, スクールバスに関するルールが占めることが確認できる.

2015 年の同アンケートからアソシエーション分析を実施した際の特徴は次のようになる.

- (1) 1 コマ目の授業開始時間が 9 時 30 分から 9 時に早まったため, 1 コマ目の開始時間を 9 時 30 分に戻してほしいとの改善要望
- (2) スクールバスの運行が授業の開始・終了に合わせられているため, 時間帯によっては JR 山前駅・東武足利市駅での待ち時間が長くなり, それを解消したいがための改善要

望

- (3) 学生ホールがみんなの共有スペースであることを認識していない学生が引き起こす雑音（ゲーム音）への不満に対する改善要望
- (4) 洋式トイレの追加要望やきれいにして欲しいとの改善不満
- (5) 女子専用ラウンジが設置されていることによる、男子専用のラウンジ設置を望む要望
- (6) 看護学科との交流を望む要望，その反面，女子が増えたことによる不満
- (7) 授業中に私語がうるさいという不満
- (8) 工学部の学生が利用できるロッカーがないという不満
- (9) 女子学生からランチのボリュームを少ない量に変更してほしいという不満

2016 年の支持度（support）とリフト（lift）の上位 50 ルールの関連性を図 4 に示す。

2016 年の同アンケートからアソシエーション分析を実施した際の特徴は次のようになる。

- (1) スクールバスの運行が授業の開始・終了に合わせられているため，時間帯によっては JR 山前駅・東武足利市駅での待ち時間が長くなり，それを解消したいがための改善要望
- (2) 学生食堂のメニューや品質の改善を望む要望と不満
- (3) 学生ホールがみんなの共有スペースであることを認識していない学生が引き起こす様々な問題に対する改善要望
- (4) 学内完全喫煙禁止による喫煙スペースに関する要望と不満
- (5) 工学部の学生が利用できるロッカーがないとの不満
- (6) クラブ活動を行うための部室に関する要望と不満
- (7) 大学内の清掃が行き届いていないとの不満とゴミを自ら片付けない学生への不満

2015・2016 年ともにスクールバスの運行についての記述が上位に挙がっている。

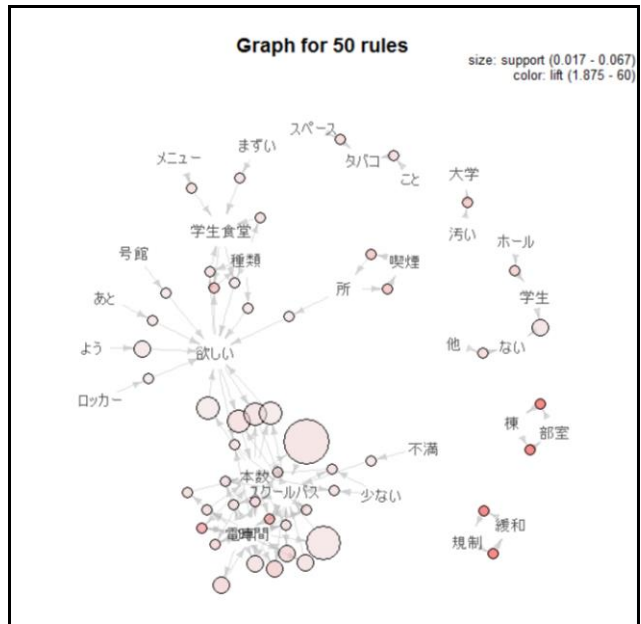


図 4 上位 50 ルールのグラフ（2016 年）

3. 3 共起ネットワーク分析

名詞・サ変名詞・形容詞からなる共起ネットワークを作成し可視化を試みると，2015 年のデータでは図 5 となる。

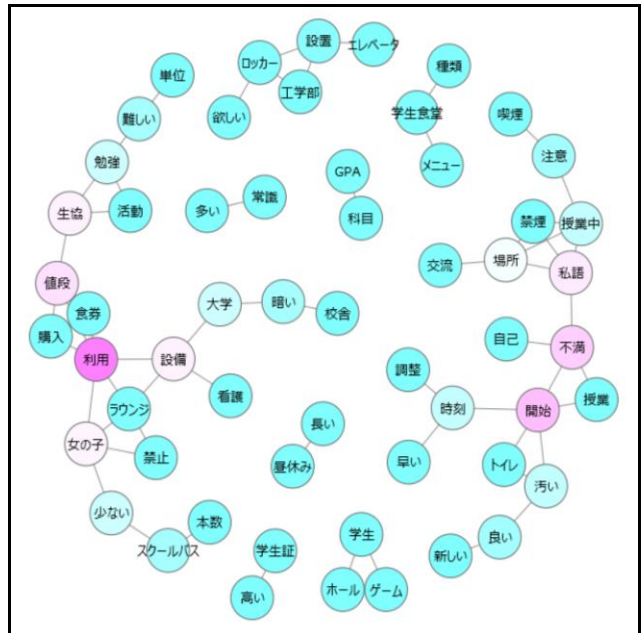


図 5 2015 年データにおける共起ネットワーク

2015 年データの特徴としては，最大のネットワークでは“利用”を中心とし“女の子”・“設備”・“値段”と“生協”からで，この中には，“スクールバス関連”・“女子ラウンジ関連”・“看護学部関連”・“大学施設関連”・“学生食堂の券売機関連”・“大学生協活動関連”・“大学での勉強関連”

で構成されている。

2 番目のネットワークでは“開始”を中心とし“不満”・“私語”からで“トイレ関連”・“授業の開始時刻関連”・“授業中の私語関連”・“喫煙関連”で構成されている。

3 番目のネットワークでは“ロッカー関連”・“エレベータ関連”で構成され、それ以外では、“学生ホール関連”・“学生証の再発行”・“学生食堂のメニュー関連”・“講義課目”になっているのがわかる。

また、2016 年のデータでは図 6 となる。

2016 年データの特徴としては、最大のネットワークでは“欲しい”を中心とし“少ない”・“スクールバス”と“思い”からで、この中には、“学生食堂関連”・“スクールバス関連”・“大学施設関連”で構成されている。

2 番目のネットワークでは“部室”・“クーラー”・“温度”・“駐車場”で構成された“空調やクラブ棟に関する内容”で、それ以外では、“ネット”・“規制”・“緩和”で構成された“学内 LAN の利用に関する内容”になっているのがわかる。

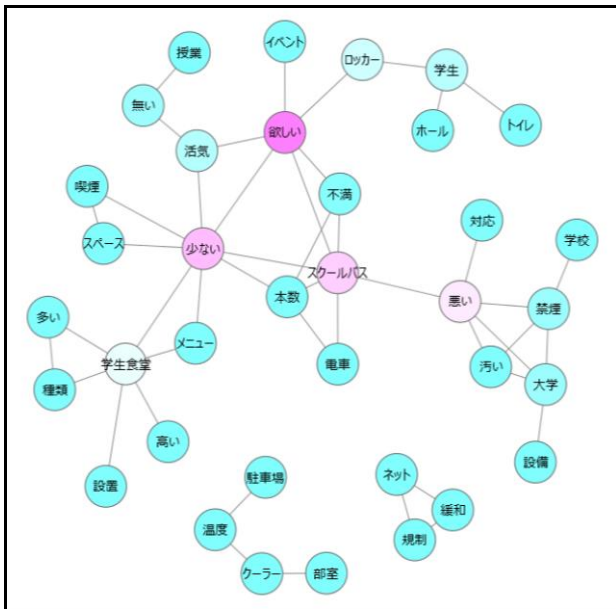


図 6 2016 年データにおける共起ネットワーク

4. 解析結果

2015 年度だけの特徴としては、“1 コマ目の授業開始時間を 9 時 30 分に戻してほしい”と“女子学生からは学生食堂のボリュームを少なくしてほしい”という内容になる。

また、2016 年度だけの特徴としては、“学生食堂のメニューを増やしてほしい”という内容になる。

さらに、アソシエーション分析・共起ネットワーク分析のどちらの手法でも、2015 年度と 2016 年度ともに、“学生ホールにおける騒音”と“スクールバスの運行”に関する内容が学生にとっての重要な関心事になっている。

参考文献

- 1) 石田基広, テキストマイニング入門, 森北出版(2008), 173pp.
- 2) 佐藤友哉, 菅野久幸, 塚越 清: 学生生活満足度調査アンケート自由記述欄の解析 第 51 号足利工業大学研究集録(2017), pp. 76-83.
- 3) アソシエーション分析(1), <http://www1.doshisha.ac.jp/~mjn/R/40/40.html>
- 4) アソシエーション分析の結果を可視化するパッケージの arulesViz, <http://d.hatena.ne.jp/EulerDijkstra/20121215/1355564356>
- 5) KH Coder, <http://khc.sourceforge.net/>
- 6) 佐藤友哉, 塚越 清: 日本経営工学会 2017 年春季大会予稿集, pp. 20-21.

原稿受付日 平成 30 年 3 月 1 日