

大学webサイトのテキストから大学らしさの抽出

白井ゼミ 3年

17161278 佐藤功基

説明・目的

- 各大学webサイトのテキストに登場する単語の出現頻度からその大学の特徴が突き止められるかを検証する。
- その大学らしさを形成、特徴づける単語を突き止められる。
- 強み(力を入れていること)、アピールしたいポイント、各大学の共通点
- 大学経営(広報部)の観点から、大学が受験生、企業に対してどういったイメージを押し出したらよいのか。

分析した大学

- 東京大学
- 東京理科大学
- 日本体育大学
- 多摩美術大学
- 東京家政大学
- 大東文化大学

大学らしさの計算

- $$\frac{\text{ある大学ででてきた単語の出現頻度}}{\text{他の大学の出現頻度の平均}} = \text{大学らしさ}$$

Ex: 東京都

大東(0.879566)

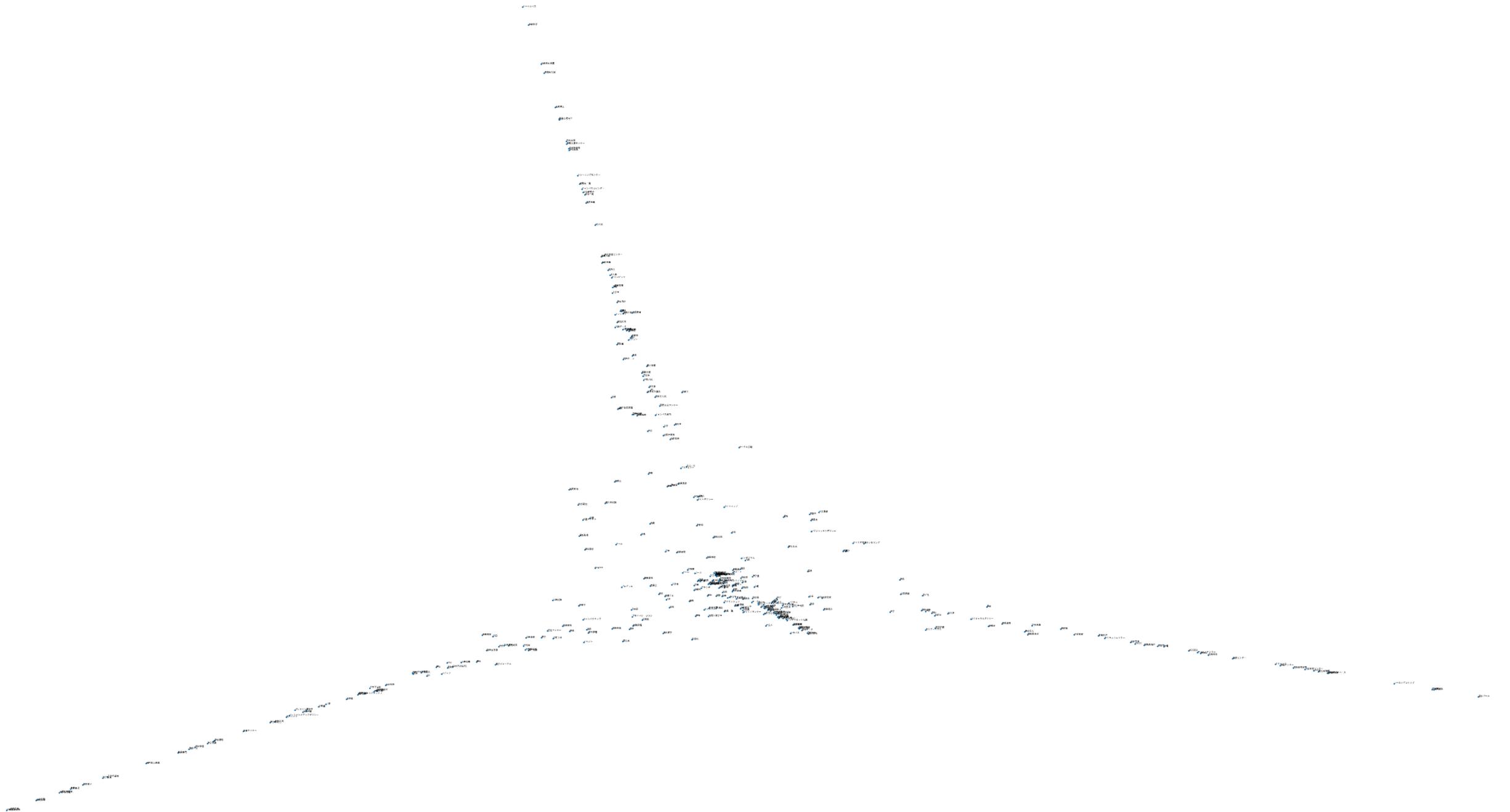
他大平均 (0.032516 + 0.15028 + 1.025436 + 0.106107 + 0.81345)/5

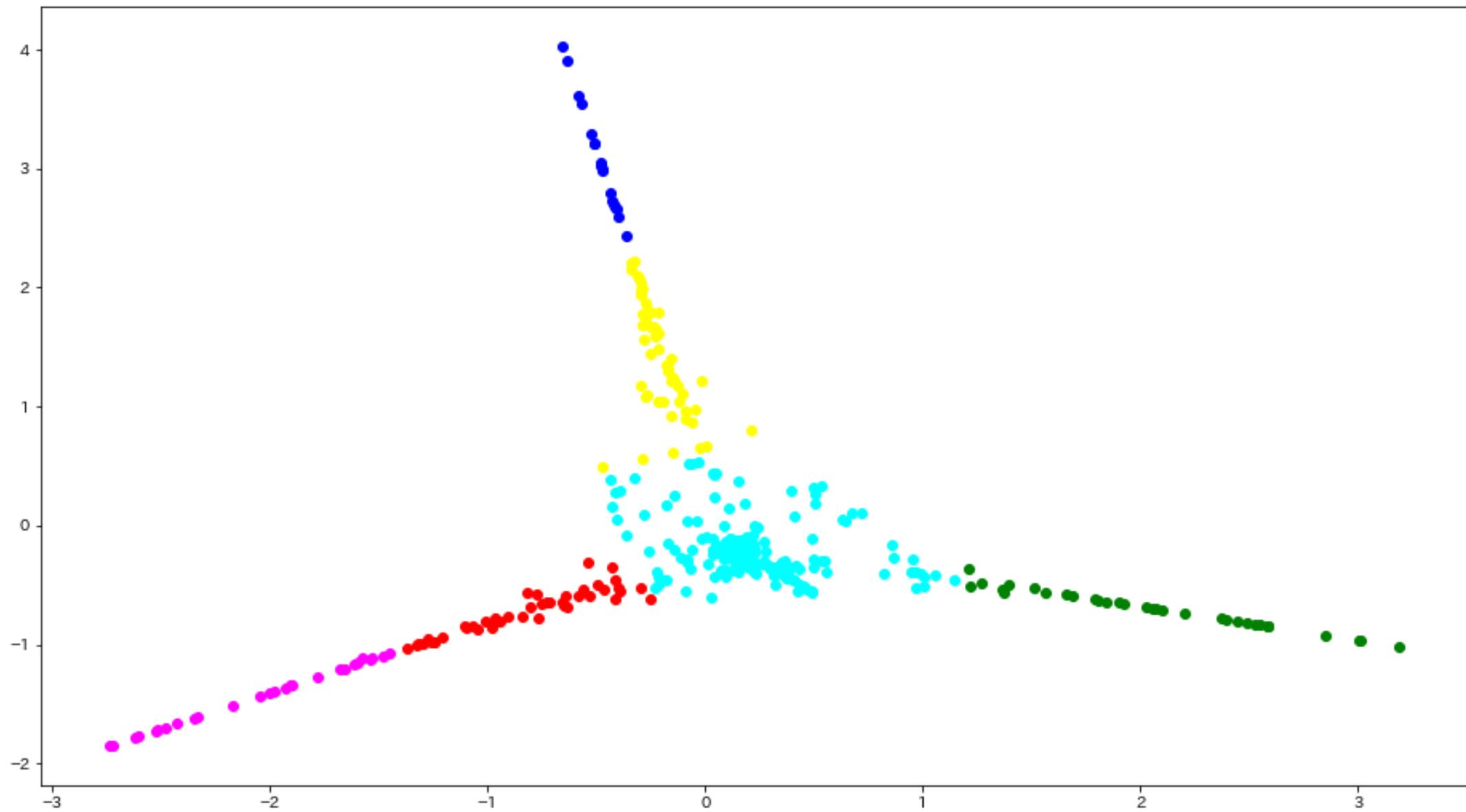
index	word	Daito	U_Tokyo	Tus	Nittai	Kasei	tamabi
1	東京都	2.066845	0.054651	0.262995	2.586967	0.182864	1.853908
2	キャリア	47.84643	0.159045	0.217539	0.009459	0.039984	0.062112
3	先生	20.6974	0.112733	0.128809	0.053015	0.628252	0.129739
4	カリキュ ラム	8.630332	0.016787	0.740517	0.314341	0.709451	0.269125
5	板橋区	3.902704	0.000726	0	0	6.402042	0

可視化してみる

主成分分析 (6次元 => 2次元)

クラスタリング (k-means)





分析

6つの大学共通でよく出てくる単語

各大学でよく出てくる単語

WordCloud

東京家政大学

臨床相談センター SCROLLDOWN 附属みどりヶ丘幼稚園 グローバル教育センター—教員養成教育
単位互換 中学校高等 大学教育 サロン 生涯学習 包括協定 教育内容
東京家政大学大学院
抱負 140周年 英会話 共通教育科目
家政学部 箱根仙石 緑窓会
渡邊辰五郎賞 東京家政大学 学長裁量 北区 個人情報等 海外提携 健康センター
子ども学部 UNIVERSITYAII
専門研修 TJUP埼玉 学長便り 放課後等ディサービス
短大 渡辺学園 スリ 狭山 津浦
イングリッシュ情報 財務関係 教育研究情報 入間市 リセット 担当科目 取組 生涯学習センター
SD 入間市 推進センター 文学部
東上地域 生活信条
せい森SEARCHrightsReserved 附属女子 研究者情報 総合研究プロジェクト短期大学部 ピンクリボン運動 放課後ディサービス
管理運営 Copyrightc search ENGLISH
東京家政大学短期大学部
放課後等 生活科学研究所 女性未来研究所 ヒューマンライフ支援センター 研究関連

多摩美術大学

多摩美

searchButton
UNIVERSITYALL
グラフィックデザイン学科

情報デザイン

学科プロダクト
1920394

学科テキストスタイル
統合デザイン学科

優秀作品

上野毛3
上野毛3
0426768611
デザイン学科

油画

卒業制作
工芸学科
多摩美術大学Twitter
特別支援

学科メディア

デザインコース
サポーター募金
RIGHTSRESERVED
searchWindow
彫刻学科
非常勤講師

デザイン専攻
版画専攻

集

2018

1723
2019
1534

劇場美術

絵画学科

0337021141
生産デザイン

芸術コース

PBL科目

取引業者様
公益通報窓口
1588558
サイトご利用

多摩美術大学

TAMAART
激甚災害等

TAMABINNEWS

大学院美術
EnglishSite
八王子市

グラフィックデザイン

環境デザイン学科
美術学部
演劇舞踊
COPYRIGHT2019

多摩美術大学Facebookページ

八王子キャンパス

上野毛キャンパス

非常時

教育情報

デザイン

芸術情報

担当教員

三つ

技法

黒材

丸

考察

- 各大学の特徴が現れた。
- 単語の相対的な出現確率(大学の出現頻度 / 他大学の出現頻度平均)を
求めることで大学を特徴づける言葉一応見出すことができた。
- 今回は特徴がでそうな大学で比較したが、大学に重要なのは似ている
ような大学で比べたときにより目立つこと。
- 特徴をアピールする手法は出現頻度以外にもある。
ビジュアル(デザイン、配置、カラー、フォント)

課題

- 同系統の大学での分析。
- 今回の大学間のベクトルから似ている大学、まったく違う大学を計算する。
- 大学間での分析に加えて、大学に対する評判に対しても同様の分析をすることで、大学の経営戦略とリアルギャップがあきらかになるのでは。