

進捗報告

平井 遥斗

富山県立大学 情報システム工学科

2023年12月15日

jaccard 係数

今までちゃんと計算できていると思っていたが、少しおかしいことに気づいたので修正を行った。

間違えていたこと

共起関係の分析は一般的には文章単位で行う。

- Python は、機械学習でよく利用されるプログラム言語であり、他の言語よりも優れている。
- Python は、機械学習だけではなく、Web アプリ開発などでも利用されている。

文章単位で見たときの 2 つの単語の出現回数などを用いる。

しかし、すべての文章をひとくくりに計算していた。

- Python は、機械学習でよく利用されるプログラム言語であり、他の言語よりも優れている。Python は、機械学習だけではなく、Web アプリ開発などでも利用されている。

専門用語の辞書作成

3/13

辞書に専門用語を登録

抽出した重要語を Janome の辞書の書式に合わせて csv を作成した。ここで、辞書の大きさを大きくしすぎるとプログラムが正常に動作しないことが分かった。

制限しないと 50 万行から 100 万行の辞書が作成されてしまい。時間がかかる上にたまに正常にプログラムが動作しなくなる。

ログ記録	-1	-1	1000 名詞	固有名詞	*	*	*	*	ログ記録*	*		
カード	-1	-1	1000 名詞	固有名詞	*	*	*	*	カード*	*		
前記スクラッチ隠蔽層	-1	-1	1000 名詞	固有名詞	*	*	*	*	前記スクラ*	*		
隠蔽	-1	-1	1000 名詞	固有名詞	*	*	*	*	隠蔽*	*		
頬料	-1	-1	1000 名詞	固有名詞	*	*	*	*	頬料*	*		
屈折率層	-1	-1	1000 名詞	固有名詞	*	*	*	*	屈折率層*	*		
表面	-1	-1	1000 名詞	固有名詞	*	*	*	*	表面*	*		
スクラッチカード	-1	-1	1000 名詞	固有名詞	*	*	*	*	スクラッチ*	*		
組成物	-1	-1	1000 名詞	固有名詞	*	*	*	*	組成物*	*		
重合	-1	-1	1000 名詞	固有名詞	*	*	*	*	重合*	*		
印刷	-1	-1	1000 名詞	固有名詞	*	*	*	*	印刷*	*		
偽造防止性	-1	-1	1000 名詞	固有名詞	*	*	*	*	偽造防止性*	*		
用インキ組成物	-1	-1	1000 名詞	固有名詞	*	*	*	*	用インキ組*	*		
光ラジカル重合開始剤系	-1	-1	1000 名詞	固有名詞	*	*	*	*	光ラジカル*	*		
光カチオン重合開始剤系	-1	-1	1000 名詞	固有名詞	*	*	*	*	光カチオン*	*		
透明基材	-1	-1	1000 名詞	固有名詞	*	*	*	*	透明基材*	*		

図 1: 辞書 (csv)

ネットワークの可視化

4/13

2D グラフ

表示する要素が少ないとときは 2D の方が見やすいかもしれない、もっと要素が多くなるとごちゃごちゃするので 3D の方が見やすくなる。2D グラフも可視化の一つとして取り入れてもいい。

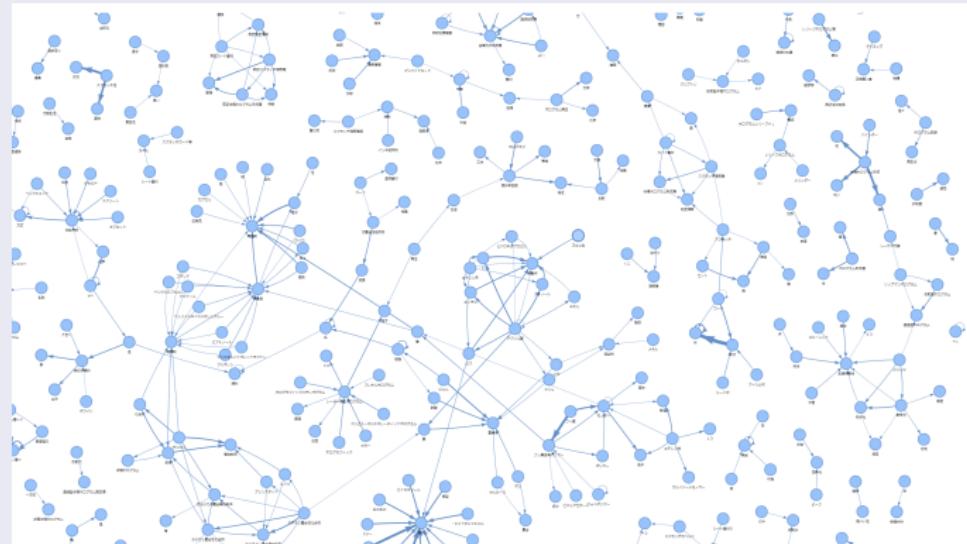


図 2: 2D グラフ

ネットワークの可視化

5/13

3D グラフ

自分のデータに適応してみた。
Json ファイルを作成してグラフを作成した

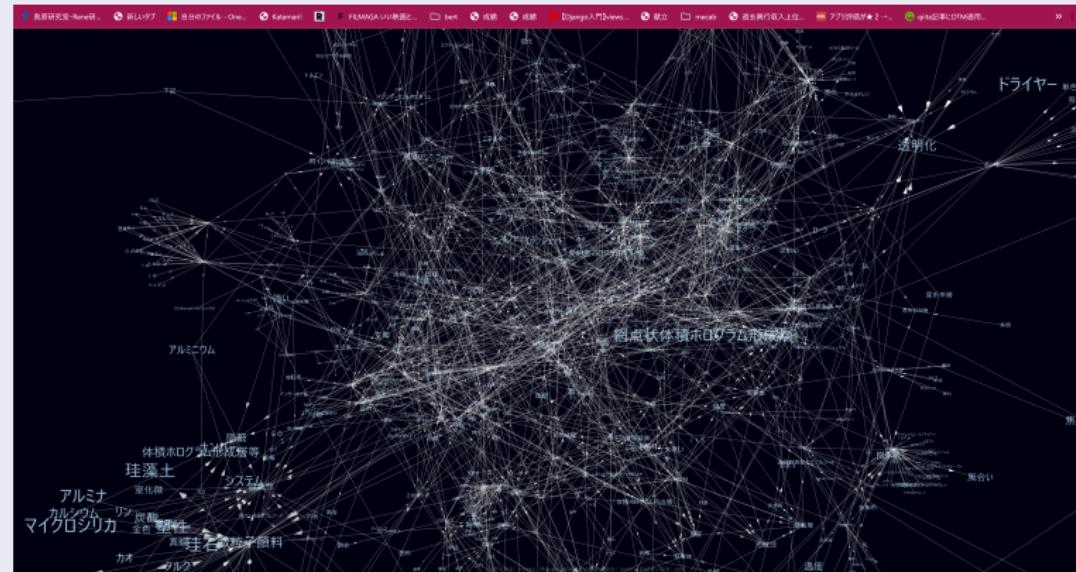


図 3: 3D グラフ

クラスターの解釈

6/13

クラスターの解釈

クラスター内の重要語を計算して表示することは可能だが、1つのクラスターの重要語を計算するのに結構時間がかかるので、10以上のクラスターを計算しようとすると、時間がかかりすぎる。

クラスターの代表的な文章だけを取り出して、重要度を計算すれば時間短縮になるのでは。

各クラスターの中心に近い文章ほどそのクラスターの中心的な文章である可能性が高い。

各クラスターの中心にある文章から重要語を計算すれば短時間で済むかもしれない。

フロントページ

ここに検索したい単語を入力する。
複数入力したい場合は間にスペースを空ける。

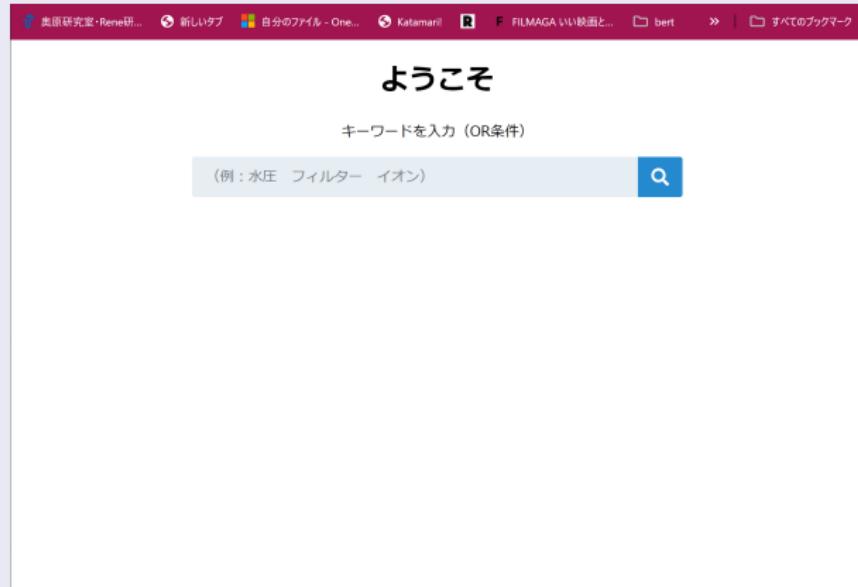


図 4: フロントページ

ロード画面

グラフを作るまで時間がかかるのでロード画面を作成した。
最終的には進捗バーを追加する予定。



図 5: ロード画面

ロード画面

クラスを選択できるようにした.

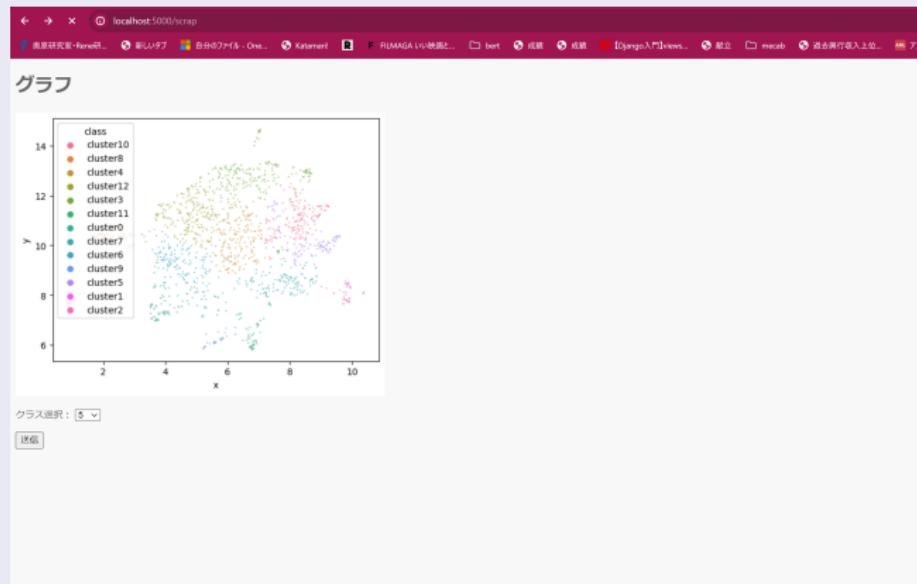


図 6: グラフ表示画面

特許以外の情報

10/13

顧客などの情報

今まで特許だけを分析するに固執していた。
特許だけではなくほかの情報も分析すればいいのでは
経営戦略の分析例

3C分析

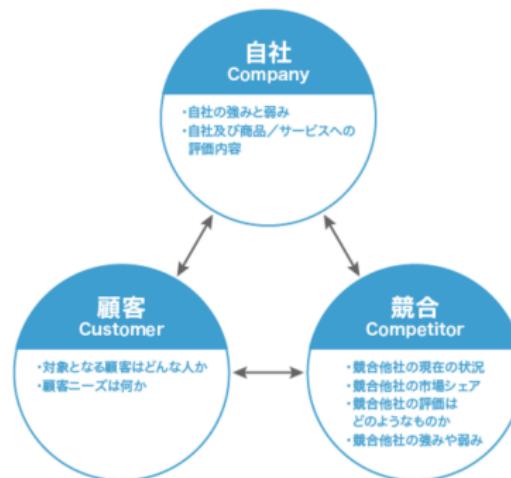


図 7: 3C 分析

特許以外の情報

11/13

環境的要因

経営戦略の分析例

SWOT分析

SWOT分析の例

		プラス面	マイナス面	
内部環境	S trong	強み	W eakness	弱み
	・メニュー数 ・配達システム ・顧客ロイヤルティの高さ ・業界No.1 ・味への評価		・調理の手間	
外部環境	O pportunity	機会	T hreat	脅威
	・安全思考の高まり ・人口減		・競合の対等 ・コンビニの中食ビジネスの激化 ・人手不足	

図 8: SWOT 分析

経営資源

経営戦略の分析例

VRIO分析



図 9: VRIO 分析

今まで、IP ランドスケープは「特許情報を用いて経営戦略を行うこと」という解釈だったが、「経営戦略の分析の中に特許情報を取り入れること」、と考えればこっちに寄せていった方がやっていることがわかりやすいんじゃないかな。

まとめ

13/13

まとめ

- まだまだ時間がかかる処理が多いので少しでも早く処理できるように工夫を凝らしたい.
- グラフから先の部分をアプリケーションに組み込む.
- ユーザーインターフェースを整える.
- 最終的にどういう実験をするのかを考えていきたい.