

# 数学データからのモチベーション向上・キャリアパス支援のための情報推薦機構

平松 楓也

富山県立大学 情報基盤工学講座

1. はじめに
2. セクション名
3. おわりに

# はじめに

2/13

## 背景

近年の就職活動は売り手市場と言われていたが、コロナウイルスの影響で世界的に経済状況が悪化しており、買い手市場に推移していく可能性が考えられる。また、大手企業へ就職を考えた場合、応募人数が多く狭き門であることが多いため、企業は企業がより求めている人材を採用すると思われる。そのため、学生の間に、より効率的に企業が求める人材になるための勉強が必要になると思われる。

## 目的

過去の卒業生の就職先や、学業成績、野外活動のデータをクラスタリングし、在校生がより効率的に就職活動を行えるよう対話型の情報推薦機構の基礎技術を開発する。

はじめに

データ分析の種類

協調フィルタリング

進捗状況

おわりに

# 完成目標

3/13

はじめに

データ分析の種類

協調フィルタリ  
ング

進捗状況

おわりに



入学したばかりの大学1年生、卒業後就職したい企業が決まっている学生  
その企業に就職するためにはどうすれば効率的に動けるかわからない



協調フィルタリングで過去の卒業生の  
データからやるべきことを推薦



やるべき勉強、活動が明確になりモチベーション向上  
就職したかった企業への内定

# 説明的データ分析

4/13

## 説明

事実を説明する  
見つける

## 例

どんな人が何を買っているか？  
ある広告がどれだけ売り上げに貢献しているか？

## 手法

BI、クラスタリング、アソシエーション分析

はじめに

データ分析の種類

協調フィルタリ  
ング

進捗状況

おわりに

# 予測的データ分析

5/13

## 説明

未来や欠測値を予想する

## 例

ある商品群を閲覧した人の性別は？

広告を出稿したらどれだけ売り上げが上がるのか？

## 手法

分類・回帰、統計的機械学習、協調フィルタリング

はじめに

データ分析の種類

協調フィルタリング

進捗状況

おわりに

# 指示的データ分析

6/13

## 説明

### 最適解を探す

## 例

利益を最大化するための、最適な仕入れ量は？

売上を最大化するには、どこに広告を出稿するべきか？

## 手法

### 最適化、実験結果

はじめに

データ分析の種類

協調フィルタリング

進捗状況

おわりに

# 協調フィルタリング

7/13

協調フィルタリングとは、Amazon が開発したレコメンドエンジンで、多くのユーザの嗜好情報を蓄積し、あるユーザと嗜好の類似した他のユーザの情報を用いて自動的に推論を行う方法論である。

また、協調フィルタリングには二種類あり、ユーザベース協調フィルタリングとアイテムベース協調フィルタリングがある。

はじめに

データ分析の種類

協調フィルタリング

進捗状況

終わりに

# ユーザベース協調フィルタリング

8/13

ユーザベース協調フィルタリングでは「ユーザAは未評価アイテムIに対して、当該ユーザと似たような嗜好をしている他ユーザと同じような評価をするだろう」という仮定に基づいている。

## ユーザベース協調フィルタリング

履歴から  
類似ユーザ  
を見つける

	商品A	商品B	商品C	商品D
ユーザA	○	-	○	○
ユーザB	x	○	-	x
ユーザC	○	○	x	-
ユーザD	○	x	○	?

類似ユーザAはDを評価高くしているのでおすすめできそう

アイテムベース協調フィルタリングでは「アイテム同士の類似度とあるユーザ A の過去に評価したアイテムの評価点を用いて未評価アイテム I の評価点を予測する」というアプローチである。

## アイテム間協調フィルタリング

似た評価の商品を見つける。  
商品Aと商品Dは似た人に買われやすい

	ユーザA	ユーザB	ユーザC	ユーザD
商品A	○	-	○	○
商品B	x	○	-	x
商品C	○	○	x	-
商品D	○	x	○	?

商品Aの購買者評価と似た評価の商品Dを勧めよう

# 今後行うアイテムベース協調フィルタリングについて

10/13

はじめに

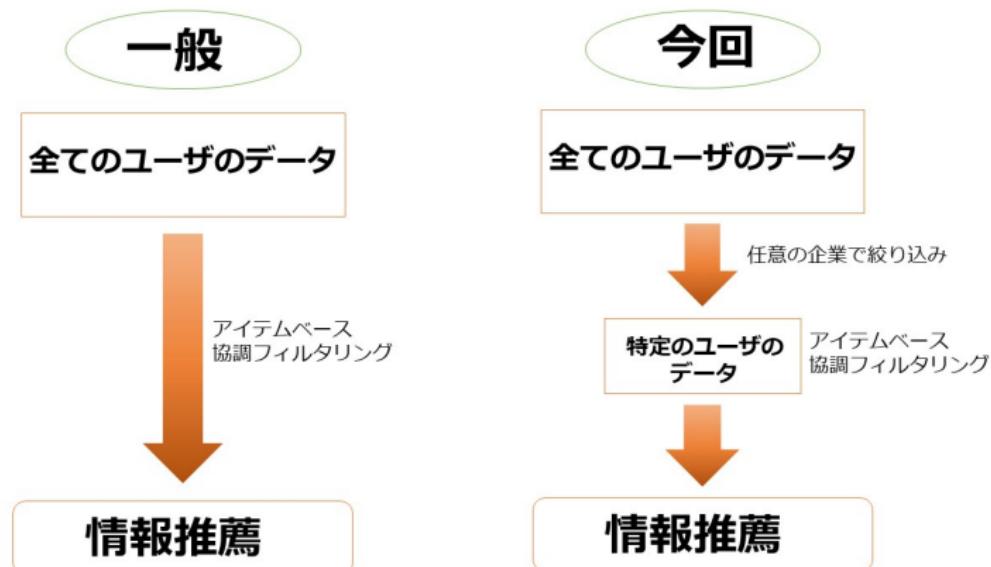
データ分析の種類

協調フィルタリング

進捗状況

おわりに

一般に使われる協調フィルタリングは全ユーザのデータを基にフィルタリングを行うのに対し、今回では、ユーザAが就職を希望している企業に就職したユーザのみでフィルタリングを行い情報推薦を行おうと考えている。



# デモデータ

11/13

- ・デモデータは3~5年分の卒業生データを作る予定
- ・デモデータはExcelを用いて作成
- ・項目は、学籍番号、工業数学などの科目の評価、卒業後の進路、留学経験の有無、資格の有無、サークル活動の有無、インターンの有無を考えている。
- ・現段階ではシンプルなもので検証したいので、学籍番号～留学経験の有無までの仮デモデータで行う。

	A	B	C	D	E	F	G	H
1	StudentN	RecordA	RecordB	RecordC	AfterGrad	Volunteer	StudyAbroad	
2	1713001	C	A	B	Employe	No		
3	1713002	S	D	A	Employe	No		
4	1713003	D	C	S	Employe	No		
5	1713004	S	C	C	Employe	No		
6	1713005	S	D	S	Continue	No		
7	1713006	C	A	B	Employe	No		
8	1713007	C	B	S	Employe	No		
9	1713008	C	C	S	Continue	No		
10	1713009	S	S	A	Employe	No		
11	1713010	S	B	A	Employe	No		
12	1713011	A	B	A	Continue	No		

はじめに

データ分析の種類

協調フィルタリング

進捗状況

おわりに

# mongoDBへのインポート

12/13

仮データを mongoDB へインポートした。

```
hiramatsu@zeta-desktop: ~
File Edit View Search Terminal Help
> db.students.find().limit(10)
{ "_id" : ObjectId("5e7993ca32f3fab4b1913047"), "StudentNumber" : 1713002, "RecordA" : "S", "RecordB" : "D", "RecordC" : "A", "AfterGraduating" : "Employment", "Volunteer" : "No", "StudyAbroad" : "No" }
{ "_id" : ObjectId("5e7993ca32f3fab4b1913048"), "StudentNumber" : 1713001, "RecordA" : "C", "RecordB" : "A", "RecordC" : "B", "AfterGraduating" : "Employment", "Volunteer" : "No", "StudyAbroad" : "No" }
{ "_id" : ObjectId("5e7993ca32f3fab4b1913049"), "StudentNumber" : 1713004, "RecordA" : "S", "RecordB" : "C", "RecordC" : "C", "AfterGraduating" : "Employment", "Volunteer" : "No", "StudyAbroad" : "No" }
{ "_id" : ObjectId("5e7993ca32f3fab4b191304a"), "StudentNumber" : 1713003, "RecordA" : "D", "RecordB" : "C", "RecordC" : "S", "AfterGraduating" : "Employment", "Volunteer" : "No", "StudyAbroad" : "No" }
{ "_id" : ObjectId("5e7993ca32f3fab4b191304b"), "StudentNumber" : 1713005, "RecordA" : "S", "RecordB" : "D", "RecordC" : "S", "AfterGraduating" : "Continue", "Volunteer" : "No", "StudyAbroad" : "No" }
{ "_id" : ObjectId("5e7993ca32f3fab4b191304c"), "StudentNumber" : 1713006, "RecordA" : "C", "RecordB" : "A", "RecordC" : "B", "AfterGraduating" : "Employment", "Volunteer" : "No", "StudyAbroad" : "No" }
```

はじめに

データ分析の種類

協調フィルタリング

進捗状況

おわりに

## まとめ

- ① 今後の研究の方向性と手法が決まった。
- ② 仮データを作成した。

## 今後の課題

- ① 実際に python でプログラムを作成し、どのような結果が得られるか検証を重ねていきたい。
- ② 本研究での有用性の評価をどう行うかも考えていきたい。

はじめに

データ分析の種類

協調フィルタリング

進捗状況

おわりに