

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ  
им. Н.Э. Баумана

Факультет «Информатика и системы управления»  
Кафедра «Систем обработки информации и управления»

ОТЧЕТ

**Лабораторная работа № 4**  
по дисциплине «Методы машинного обучения»

ИСПОЛНИТЕЛЬ:

группа ИУ5-23М

Морозенков О.Н.

ФИО

подпись

" " 2022 г.

ПРЕПОДАВАТЕЛЬ:

Гапанюк Ю.Е.

ФИО

подпись

" " 2022 г.

Москва - 2022

---

**Цель лабораторной работы:** изучение разработки рекомендательных моделей.

**Задание:**

1. Выбрать произвольный набор данных (датасет), предназначенный для построения рекомендательных моделей.
2. Опираясь на материалы лекции, сформировать рекомендации для одного пользователя (объекта) двумя произвольными способами.
3. Сравнить полученные рекомендации (если это возможно, то с применением метрик).

```

import numpy as np
import pandas as pd
from typing import Dict, Tuple
from scipy import stats
from IPython.display import Image
from IPython.display import Image
from sklearn.feature_extraction.text import CountVectorizer, TfidfVectorizer
from sklearn.datasets import load_iris, load_boston
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsRegressor, KNeighborsClassifier
from sklearn.model_selection import GridSearchCV, RandomizedSearchCV
from sklearn.metrics import accuracy_score, balanced_accuracy_score
from sklearn.metrics import precision_score, recall_score, f1_score,
classification_report
from sklearn.metrics import confusion_matrix
from sklearn.tree import DecisionTreeClassifier, DecisionTreeRegressor,
export_graphviz
from sklearn.ensemble import RandomForestClassifier, RandomForestRegressor
from sklearn.ensemble import ExtraTreesClassifier, ExtraTreesRegressor
from sklearn.ensemble import GradientBoostingClassifier,
GradientBoostingRegressor
from sklearn.ensemble import BaggingClassifier
from sklearn.ensemble import AdaBoostClassifier
from sklearn.metrics import mean_absolute_error, mean_squared_error,
mean_squared_log_error, median_absolute_error, r2_score
from sklearn.metrics import roc_curve, roc_auc_score
from sklearn.metrics.pairwise import cosine_similarity, euclidean_distances,
manhattan_distances
from collections import defaultdict
import seaborn as sns
import matplotlib.pyplot as plt
from matplotlib_venn import venn2
%matplotlib inline
sns.set(style="ticks")

```

## Чтение и обработка данных

```

data = pd.read_csv('./winemag-data-130k-v2.csv.zip')
data.head()

```

	Unnamed: 0	country	description	\
0	0	Italy	Aromas include tropical fruit, broom, brimston...	
1	1	Portugal	This is ripe and fruity, a wine that is smooth...	
2	2	US	Tart and snappy, the flavors of lime flesh and...	
3	3	US	Pineapple rind, lemon pith and orange blossom ...	
4	4	US	Much like the regular bottling from 2012, this...	

  

	designation	points	price	province	\
0	Vulkà Bianco	87	NaN	Sicily & Sardinia	
1	Avidagos	87	15.0	Douro	

2		NaN	87	14.0	Oregon
3	Reserve Late Harvest		87	13.0	Michigan
4	Vintner's Reserve Wild Child Block		87	65.0	Oregon

	region_1	region_2	taster_name \
0	Etna	NaN	Kerin O'Keefe
1	NaN	NaN	Roger Voss
2	Willamette Valley	Willamette Valley	Paul Gregutt
3	Lake Michigan Shore	NaN	Alexander Peartree
4	Willamette Valley	Willamette Valley	Paul Gregutt

	taster_twitter_handle		title \
0	@kerinokeefe	Nicosia 2013 Vulkà Bianco	(Etna)
1	@vossroger	Quinta dos Avidagos 2011 Avidagos Red	(Douro)
2	@paulgwine	Rainstorm 2013 Pinot Gris	(Willamette Valley)
3	NaN	St. Julian 2013 Reserve Late Harvest Riesling ...	
4	@paulgwine	Sweet Cheeks 2012 Vintner's Reserve Wild Child...	

	variety	winery
0	White Blend	Nicosia
1	Portuguese Red	Quinta dos Avidagos
2	Pinot Gris	Rainstorm
3	Riesling	St. Julian
4	Pinot Noir	Sweet Cheeks

```
data.shape
```

```
(129971, 14)
```

```
description_data = data[data['description'].notnull()]
```

```
description_data.shape
```

```
(129971, 14)
```

```
title = description_data['title'].values
```

```
title[0:5]
```

```
array(['Nicosia 2013 Vulkà Bianco (Etna)',
      'Quinta dos Avidagos 2011 Avidagos Red (Douro)',
      'Rainstorm 2013 Pinot Gris (Willamette Valley)',
      'St. Julian 2013 Reserve Late Harvest Riesling (Lake Michigan Shore)',
      "Sweet Cheeks 2012 Vintner's Reserve Wild Child Block Pinot Noir
      (Willamette Valley)"],
      dtype=object)
```

```
descriptions = description_data['description'].values
```

```
descriptions[0:5]
```

```
array(["Aromas include tropical fruit, broom, brimstone and dried herb. The
palate isn't overly expressive, offering unripened apple, citrus and dried
sage alongside brisk acidity.",
```

"This is ripe and fruity, a wine that is smooth while still structured. Firm tannins are filled out with juicy red berry fruits and freshened with acidity. It's already drinkable, although it will certainly be better from 2016.",

'Tart and snappy, the flavors of lime flesh and rind dominate. Some green pineapple pokes through, with crisp acidity underscoring the flavors. The wine was all stainless-steel fermented.',

'Pineapple rind, lemon pith and orange blossom start off the aromas. The palate is a bit more opulent, with notes of honey-drizzled guava and mango giving way to a slightly astringent, semidry finish.',

"Much like the regular bottling from 2012, this comes across as rather rough and tannic, with rustic, earthy, herbal characteristics. Nonetheless, if you think of it as a pleasantly unfussy country wine, it's a good companion to a hearty winter stew."],

dtype=object)

description\_data.keys()

Index(['Unnamed: 0', 'country', 'description', 'designation', 'points',  
 'price', 'province', 'region\_1', 'region\_2', 'taster\_name',  
 'taster\_twitter\_handle', 'title', 'variety', 'winery'],  
 dtype='object')

wine\_ids = description\_data['Unnamed: 0'].values  
wine\_ids

array([ 0, 1, 2, ..., 129968, 129969, 129970])

%%time

tfidf = TfidfVectorizer()

description\_matrix = tfidf.fit\_transform(descriptions)

description\_matrix

CPU times: user 3.15 s, sys: 40 ms, total: 3.19 s

Wall time: 3.2 s

<129971x31275 sparse matrix of type '<class 'numpy.float64'>'  
with 4475479 stored elements in Compressed Sparse Row format>

description\_matrix

<129971x31275 sparse matrix of type '<class 'numpy.float64'>'  
with 4475479 stored elements in Compressed Sparse Row format>

##Фильтрация на основе содержания. Метод k-ближайших соседей

**class** SimplerKnnRecomender:

**def** \_\_init\_\_(self, X\_matrix, X\_ids, X\_title, X\_overview):

"""

*Входные параметры:*

*X\_matrix - обучающая выборка (матрица объект-признак)*

*X\_ids - массив идентификаторов объектов*

```

X_title - массив названий объектов
X_overview - массив описаний объектов
"""

#Сохраняем параметры в переменных объекта
self._X_matrix = X_matrix
self.df = pd.DataFrame(
    {'id': pd.Series(X_ids, dtype='int'),
     'title': pd.Series(X_title, dtype='str'),
     'overview': pd.Series(X_overview, dtype='str'),
     'dist': pd.Series([], dtype='float')})

def recommend_for_single_object(self, K: int, \
                                X_matrix_object, cos_flag = True, manh_flag = False):
    """
    Метод формирования рекомендаций для одного объекта.
    Входные параметры:
    K - количество рекомендуемых соседей
    X_matrix_object - строка матрицы объект-признак, соответствующая
объекту
    cos_flag - флаг вычисления косинусного расстояния
    manh_flag - флаг вычисления манхэттенского расстояния
    Возвращаемое значение: K найденных соседей
    """

    scale = 1000000
    # Вычисляем косинусную близость
    if cos_flag:
        dist = cosine_similarity(self._X_matrix, X_matrix_object)
        self.df['dist'] = dist * scale
        res = self.df.sort_values(by='dist', ascending=False)
        # Не учитываем рекомендации с единичным расстоянием,
        # так как это искомый объект
        res = res[res['dist'] < scale]

    else:
        if manh_flag:
            dist = manhattan_distances(self._X_matrix, X_matrix_object)
        else:
            dist = euclidean_distances(self._X_matrix, X_matrix_object)
        self.df['dist'] = dist * scale
        res = self.df.sort_values(by='dist', ascending=True)
        # Не учитываем рекомендации с единичным расстоянием,
        # так как это искомый объект
        res = res[res['dist'] > 0.0]

    # Оставляем K первых рекомендаций
    res = res.head(K)
    return res

```

```

test_id = 11
print(title[test_id])
print(descriptions[test_id])

Leon Beyer 2012 Gewurztraminer (Alsace)
This is a dry wine, very spicy, with a tight, taut texture and strongly
mineral character layered with citrus as well as pepper. It's a food wine
with its almost crisp aftertaste.

test_matrix = description_matrix[test_id]
test_matrix

<1x31275 sparse matrix of type '<class 'numpy.float64'>'
  with 25 stored elements in Compressed Sparse Row format>

skr1 = SimplerKnnRecomender(description_matrix, wine_ids, title,
descriptions)

# 15 вин, наиболее похожих на Leon Beyer 2012 Gewurztraminer (Alsace)
# в порядке убывания схожести на основе косинусного сходства
rec1 = skr1.recommend_for_single_object(15, test_matrix)
rec1

```

	id	title \
24045	24045	Domaine Michel Thomas et Fils 2015 Rosé (Sance...
90700	90700	Henri de Villamont 2014 Morgeot Premier Cru (...)
58330	58330	Schröder & Schöyler 2013 Chartron la Fleur (Bo...
66081	66081	Maison Champy 2014 Viré-Clessé
78572	78572	Domaine Olivier Merlin 2014 Mâcon La Roche Vi...
105230	105230	Domaine Nigri 2013 Pierre de Lune (Jurançon Sec)
25907	25907	Louis Max 2014 Mâcon-Villages
99011	99011	Joseph Drouhin 2013 Les Clos (Macon-Bussières)
5406	5406	Aveleda 2015 Alvarinho (Vinho Verde)
22652	22652	Maison Malet Roquefort 2012 Léo de la Gaffeliè...
129715	129715	Boeckel 2012 Vieilles Vignes Sylvaner (Alsace)
119482	119482	Boeckel 2012 Vieilles Vignes Sylvaner (Alsace)
21920	21920	Moncigale 2014 Frais et Délicat Rosé (Coteaux ...)
96505	96505	Domaine Alban Roblin 2014 Rosé (Sancerre)
92292	92292	Domaine Alban Roblin 2014 Rosé (Sancerre)

  

	overview	dist
24045	The wine is textured and tight with crisp acid...	633624.990866
90700	This wine is still tight and crisp. It has ple...	442624.176096
58330	The wine is tight and nervy, very fresh, crisp...	432556.705703
66081	This taut and structured wine has weight as we...	430242.028148
78572	This wine is tight, structured and taut. Still...	428504.458538
105230	This rich and ripe wine is full of apricot and...	425886.605501
25907	Tight and structured, this wine has minerality...	424385.444731
99011	This crisp wine offers plenty of acidity as we...	423757.525560
5406	Ripe Alvarinho gives a wine that is rich as we...	421592.529700
22652	Very herbaceous in character, this is a wine t...	418388.507228

```

129715 Intensely peppery as well as fruity, this is a... 416866.789965
119482 Intensely peppery as well as fruity, this is a... 416866.789965
21920 This is crisp, fruity with apple and citrus fl... 411434.544994
96505 This is a fresh wine with caramel as well as r... 408987.116976
92292 This is a fresh wine with caramel as well as r... 408987.116976

```

```

# При поиске с помощью Евклидова расстояния получаем такой же результат
rec2 = skr1.recommend_for_single_object(15, test_matrix, cos_flag = False)
rec2

```

	id	title \
24045	24045	Domaine Michel Thomas et Fils 2015 Rosé (Sance...
90700	90700	Henri de Villamont 2014 Morgeot Premier Cru (...)
58330	58330	Schröder & Schöler 2013 Chartron la Fleur (Bo...
66081	66081	Maison Champy 2014 Viré-Clessé
78572	78572	Domaine Olivier Merlin 2014 Mâcon La Roche Vi...
105230	105230	Domaine Nigri 2013 Pierre de Lune (Jurançon Sec)
25907	25907	Louis Max 2014 Mâcon-Villages
99011	99011	Joseph Drouhin 2013 Les Clos (Macon-Bussières)
5406	5406	Aveleda 2015 Alvarinho (Vinho Verde)
22652	22652	Maison Malet Roquefort 2012 Léo de la Gaffeliè...
119482	119482	Boeckel 2012 Vieilles Vignes Sylvaner (Alsace)
129715	129715	Boeckel 2012 Vieilles Vignes Sylvaner (Alsace)
21920	21920	Moncigale 2014 Frais et Délicat Rosé (Coteaux ...)
92292	92292	Domaine Alban Roblin 2014 Rosé (Sancerre)
96505	96505	Domaine Alban Roblin 2014 Rosé (Sancerre)

	overview	dist
24045	The wine is textured and tight with crisp acid...	8.560082e+05
90700	This wine is still tight and crisp. It has ple...	1.055818e+06
58330	The wine is tight and nervy, very fresh, crisp...	1.065311e+06
66081	This taut and structured wine has weight as we...	1.067481e+06
78572	This wine is tight, structured and taut. Still...	1.069108e+06
105230	This rich and ripe wine is full of apricot and...	1.071553e+06
25907	Tight and structured, this wine has minerality...	1.072953e+06
99011	This crisp wine offers plenty of acidity as we...	1.073539e+06
5406	Ripe Alvarinho gives a wine that is rich as we...	1.075553e+06
22652	Very herbaceous in character, this is a wine t...	1.078528e+06
119482	Intensely peppery as well as fruity, this is a...	1.079938e+06
129715	Intensely peppery as well as fruity, this is a...	1.079938e+06
21920	This is crisp, fruity with apple and citrus fl...	1.084957e+06
92292	This is a fresh wine with caramel as well as r...	1.087210e+06
96505	This is a fresh wine with caramel as well as r...	1.087210e+06

```

# Манхэттенское расстояние дает несколько иные результаты поиска
rec3 = skr1.recommend_for_single_object(15, test_matrix,
                                         cos_flag = False, manh_flag = True)
rec3

```

	id	title \
24045	24045	Domaine Michel Thomas et Fils 2015 Rosé (Sance...



22652	22652	Maison Malet Roquefort 2012 Léo de la Gaffeliè...
35502	35502	Château de Piote 2012 Perles (Crémant de Bord...
58330	58330	Schröder & Schöyler 2013 Chartron la Fleur (Bo...
25907	25907	Louis Max 2014 Mâcon-Villages
21920	21920	Moncigale 2014 Frais et Délicat Rosé (Coteaux ...
97201	97201	Ravoire et Fils 2013 Domaine la Rabiote Rosé ...
70762	70762	Château du Seuil 2015 Domaine du Seuil (Borde...
128577	128577	Ravoire et Fils 2014 Domaine Bel Eouve Rosé (C...
78572	78572	Domaine Olivier Merlin 2014 Mâcon La Roche Vi...
92292	92292	Domaine Alban Roblin 2014 Rosé (Sancerre)
96505	96505	Domaine Alban Roblin 2014 Rosé (Sancerre)
108912	108912	Quinta do Portal 2012 Colheita Rosé (Douro)
66081	66081	Maison Champy 2014 Viré-Clessé
88898	88898	Markus Huber 2009 Hugo Grüner Veltliner (Niede...

		overview	dist
24045	The wine is textured and tight with crisp acid...	3.865262e+06	
22652	Very herbaceous in character, this is a wine t...	5.251729e+06	
35502	Tight and sharp, this is an herbaceous wine wi...	5.312967e+06	
58330	The wine is tight and nervy, very fresh, crisp...	5.316624e+06	
25907	Tight and structured, this wine has minerality...	5.354298e+06	
21920	This is crisp, fruity with apple and citrus fl...	5.452536e+06	
97201	Tight, zingy and crisp, this wine has fresh, c...	5.535851e+06	
70762	The wine is tight and mineral in character. It...	5.564448e+06	
128577	This is a tangy, spicy wine, a character that ...	5.628584e+06	
78572	This wine is tight, structured and taut. Still...	5.644448e+06	
92292	This is a fresh wine with caramel as well as r...	5.653916e+06	
96505	This is a fresh wine with caramel as well as r...	5.653916e+06	
108912	This rosé is almost as rich as a red wine, the...	5.701024e+06	
66081	This taut and structured wine has weight as we...	5.734040e+06	
88898	Very crisp fruit, with light acidity and a tau...	5.751297e+06	

##Коллаборативная фильтрация. Метод на основе сингулярного разложения

data.head()

	Unnamed: 0	country	description \
0	0	Italy	Aromas include tropical fruit, broom, brimston...
1	1	Portugal	This is ripe and fruity, a wine that is smooth...
2	2	US	Tart and snappy, the flavors of lime flesh and...
3	3	US	Pineapple rind, lemon pith and orange blossom ...
4	4	US	Much like the regular bottling from 2012, this...

	designation	points	price	province \
0	Vulkà Bianco	87	NaN	Sicily & Sardinia
1	Avidagos	87	15.0	Douro
2	NaN	87	14.0	Oregon
3	Reserve Late Harvest	87	13.0	Michigan
4	Vintner's Reserve Wild Child Block	87	65.0	Oregon

	region_1	region_2	taster_name \
0	Etna	NaN	Kerin O'Keefe
1	NaN	NaN	Roger Voss
2	Willamette Valley	Willamette Valley	Paul Gregutt
3	Lake Michigan Shore	NaN	Alexander Peartree
4	Willamette Valley	Willamette Valley	Paul Gregutt

  

	taster_twitter_handle		title \
0	@kerinokeefe	Nicosia 2013 Vulkà Bianco	(Etna)
1	@vossroger	Quinta dos Avidagos 2011 Avidagos Red	(Douro)
2	@paulgwine	Rainstorm 2013 Pinot Gris	(Willamette Valley)
3	NaN	St. Julian 2013 Reserve Late Harvest Riesling ...	
4	@paulgwine	Sweet Cheeks 2012 Vintner's Reserve Wild Child...	

	variety	winery
0	White Blend	Nicosia
1	Portuguese Red	Quinta dos Avidagos
2	Pinot Gris	Rainstorm
3	Riesling	St. Julian
4	Pinot Noir	Sweet Cheeks

```
data3 = data[30000:55000]
```

```
# Количество уникальных дегустаторов
```

```
len(data3['taster_name'].unique())
```

```
20
```

```
# Количество уникальных вин
```

```
len(data3['title'].unique())
```

```
24517
```

```
# Сформируем матрицу взаимодействий на основе рейтингов
```

```
# Используется идея из статьи - https://towardsdatascience.com/beginners-guide-to-creating-an-svd-recommender-system-1fd7326d1f65
```

```
def create_utility_matrix(data):
```

```
    itemField = 'title'
```

```
    userField = 'taster_name'
```

```
    valueField = 'points'
```

```
    userList = data[userField].tolist()
```

```
    itemList = data[itemField].tolist()
```

```
    valueList = data[valueField].tolist()
```

```
    users = list(set(userList))
```

```
    items = list(set(itemList))
```

```
    users_index = {users[i]: i for i in range(len(users))}
```

```
    pd_dict = {item: [0.0 for i in range(len(users))] for item in items}
```

```

for i in range(0,data.shape[0]):
    item = itemList[i]
    user = userList[i]
    value = valueList[i]
    pd_dict[item][users_index[user]] = value

X = pd.DataFrame(pd_dict)
X.index = users

itemcols = list(X.columns)
items_index = {itemcols[i]: i for i in range(len(itemcols))}

return X, users_index, items_index

%%time
user_item_matrix, users_index, items_index = create_utility_matrix(data3)

CPU times: user 745 ms, sys: 4.02 ms, total: 749 ms
Wall time: 747 ms

user_item_matrix

```

	Château Roc Meynard 2008	Bordeaux Supérieur \
NaN		0.0
Roger Voss		87.0
Mike DeSimone		0.0
Anne Krebiehl MW		0.0
Matt Kettmann		0.0
Carrie Dykes		0.0
Alexander Peartree		0.0
Kerin O'Keefe		0.0
Joe Czerwinski		0.0
Jeff Jenssen		0.0
Lauren Buzzeo		0.0
Michael Schachner		0.0
Christina Pickard		0.0
Fiona Adams		0.0
Anna Lee C. Iijima		0.0
Virginie Boone		0.0
Paul Gregutt		0.0
Jim Gordon		0.0
Sean P. Sullivan		0.0
Susan Kostrzewa		0.0

  

	Sannino 2014 Bianca Rosé Merlot (North Fork of Long
Island) \	
NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0

Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	84.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Pascal Aufranc 2014 Juliéнас \

NaN	0.0
Roger Voss	88.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Vignoble des 2 Lunes 2012 Comète Brut Pinot Blanc

(Alsace) \	
NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	92.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0

Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

JCB 2012 No. 8 Pinot Noir Rosé (Sonoma Coast) \

NaN	87.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Lantieri de Paratico NV Rosé Sparkling (Franciacorta) \

NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	85.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0

Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Falua 2006 Conde de Vimioso Red (Ribatejano) \

NaN	0.0
Roger Voss	89.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Grove Mill 2013 Pinot Noir (Wairau Valley) \

NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	86.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Chanson Père et Fils 2010 Les Clos Grand Cru (Chablis)

\	
NaN	0.0
Roger Voss	92.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Swiftwater Cellars 2010 Zephyr Ridge Estate Bottled

Cabernet Sauvignon (Horse Heaven Hills) \	
NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	87.0
Susan Kostrzewa	0.0

...	\
NaN	...
Roger Voss	...
Mike DeSimone	...

Anne Krebiehl MW	...
Matt Kettmann	...
Carrie Dykes	...
Alexander Peartree	...
Kerin O'Keefe	...
Joe Czerwinski	...
Jeff Jenssen	...
Lauren Buzzeo	...
Michael Schachner	...
Christina Pickard	...
Fiona Adams	...
Anna Lee C. Iijima	...
Virginie Boone	...
Paul Gregutt	...
Jim Gordon	...
Sean P. Sullivan	...
Susan Kostrzewa	...

Douglas Green 2008 Vineyard Creations Shiraz (Western

Cape) \	
NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	87.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Château Roubine 2016 Inspire Rosé (Côtes de Provence) \

NaN	0.0
Roger Voss	93.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0



Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Ironstone 2006 Cabernet Franc (Lodi) \

NaN	83.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Reininger 2012 Seven Hills Vineyard Carmenère (Walla

Walla Valley (WA)) \

NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0

Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	91.0
Susan Kostrzewa	0.0

Omero 2015 Chardonnay (Willamette Valley) \

NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	88.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

A.A. Badenhorst Family Wines 2013 Red (Swartland) \

NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	94.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0

Susan Kostrzewa	0.0
-----------------	-----

Black Ink 2015 Red (California) \

NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	85.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Herdade Grande 2014 Reserva Branco White (Alentejano) \

NaN	0.0
Roger Voss	90.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Darcie Kent Vineyards 2007 Picazo Vineyard Merlot  
(Livermore Valley) \

NaN	85.0
Roger Voss	0.0

Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	0.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

Chilcas 2010 Sauvignon Blanc (Central Valley)

NaN	0.0
Roger Voss	0.0
Mike DeSimone	0.0
Anne Krebiehl MW	0.0
Matt Kettmann	0.0
Carrie Dykes	0.0
Alexander Peartree	0.0
Kerin O'Keefe	0.0
Joe Czerwinski	0.0
Jeff Jenssen	0.0
Lauren Buzzeo	0.0
Michael Schachner	83.0
Christina Pickard	0.0
Fiona Adams	0.0
Anna Lee C. Iijima	0.0
Virginie Boone	0.0
Paul Gregutt	0.0
Jim Gordon	0.0
Sean P. Sullivan	0.0
Susan Kostrzewa	0.0

[20 rows x 24517 columns]

*# Выделение тестовой строки*

```
user_item_matrix__test = user_item_matrix.loc[['Kerin O'Keefe']]
user_item_matrix__test
```

	Château Roc Meynard 2008	Bordeaux Supérieur	\
Kerin O'Keefe			0.0

Kerin O'Keefe	Sannino 2014 Bianca Rosé Merlot (North Fork of Long Island) \	0.0
Kerin O'Keefe	Pascal Aufranc 2014 Juliéñas \	0.0
Kerin O'Keefe	Vignoble des 2 Lunes 2012 Comète Brut Pinot Blanc (Alsace) \	0.0
Kerin O'Keefe	JCB 2012 No. 8 Pinot Noir Rosé (Sonoma Coast) \	0.0
Kerin O'Keefe	Lantieri de Paratico NV Rosé Sparkling (Franciacorta) \	85.0
Kerin O'Keefe	Falua 2006 Conde de Vimioso Red (Ribatejano) \	0.0
Kerin O'Keefe	Grove Mill 2013 Pinot Noir (Wairau Valley) \	0.0
Kerin O'Keefe	Chanson Père et Fils 2010 Les Clos Grand Cru (Chablis) \	0.0
Kerin O'Keefe	Swiftwater Cellars 2010 Zephyr Ridge Estate Bottled Cabernet Sauvignon (Horse Heaven Hills) \	0.0
Kerin O'Keefe	... \	...
Kerin O'Keefe	Douglas Green 2008 Vineyard Creations Shiraz (Western Cape) \	0.0
Kerin O'Keefe	Château Roubine 2016 Inspire Rosé (Côtes de Provence) \	0.0
Kerin O'Keefe	Ironstone 2006 Cabernet Franc (Lodi) \	0.0
Kerin O'Keefe	Reininger 2012 Seven Hills Vineyard Carmenère (Walla Walla Valley (WA)) \	0.0
Kerin O'Keefe	Omero 2015 Chardonnay (Willamette Valley) \	0.0
	A.A. Badenhorst Family Wines 2013 Red (Swartland) \	

Kerin O'Keefe	0.0
Black Ink 2015 Red (California) \	
Kerin O'Keefe	0.0
Herdade Grande 2014 Reserva Branco White (Alentejano) \	
Kerin O'Keefe	0.0
Darcie Kent Vineyards 2007 Picazo Vineyard Merlot (Livermore Valley) \	
Kerin O'Keefe	0.0
Chilcas 2010 Sauvignon Blanc (Central Valley)	
Kerin O'Keefe	0.0

[1 rows x 24517 columns]

```
#taster_names = description_data['taster_name'].unique()
taster_names = np.delete(data3['taster_name'].unique(), 0)
taster_names = np.delete(taster_names, 7)
taster_names

array(['Jim Gordon', 'Michael Schachner', 'Matt Kettmann',
       'Sean P. Sullivan', 'Roger Voss', 'Virginie Boone',
       'Joe Czerwinski', 'Paul Gregutt', 'Mike DeSimone', 'Jeff Jenssen',
       nan, 'Anna Lee C. Iijima', 'Susan Kostrzewa', 'Lauren Buzzeo',
       'Alexander Peartree', 'Fiona Adams', 'Carrie Dykes',
       'Christina Pickard'], dtype=object)
```

```
# Оставшаяся часть матрицы для обучения
user_item_matrix__train = user_item_matrix.loc[taster_names]
user_item_matrix__train
```

	Château Roc Meynard 2008 Bordeaux Supérieur \
Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	87.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0

Carrie Dykes	0.0
Christina Pickard	0.0

Sannino 2014 Bianca Rosé Merlot (North Fork of Long

Island) \	
Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	84.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Pascal Aufranc 2014 Juliéñas \

Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	88.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Vignoble des 2 Lunes 2012 Comète Brut Pinot Blanc

(Alsace) \	
Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0

Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

JCB 2012 No. 8 Pinot Noir Rosé (Sonoma Coast) \

Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	87.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Lantieri de Paratico NV Rosé Sparkling (Franciacorta) \

Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0



Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Falua 2006 Conde de Vimioso Red (Ribatejano) \

Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	89.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Grove Mill 2013 Pinot Noir (Wairau Valley) \

Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	86.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Chanson Père et Fils 2010 Les Clos Grand Cru (Chablis)

\	
Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0

Sean P. Sullivan	0.0
Roger Voss	92.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Swiftwater Cellars 2010 Zephyr Ridge Estate Bottled

Cabernet Sauvignon (Horse Heaven Hills) \	
Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	87.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

...	\
Jim Gordon	...
Michael Schachner	...
Matt Kettmann	...
Sean P. Sullivan	...
Roger Voss	...
Virginie Boone	...
Joe Czerwinski	...
Paul Gregutt	...
Mike DeSimone	...
Jeff Jenssen	...
NaN	...
Anna Lee C. Iijima	...

Susan Kostrzewa	...
Lauren Buzzeo	...
Alexander Peartree	...
Fiona Adams	...
Carrie Dykes	...
Christina Pickard	...

Douglas Green 2008 Vineyard Creations Shiraz (Western

Cape) \	
Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	87.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Château Roubine 2016 Inspire Rosé (Côtes de Provence) \

Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	93.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Ironstone 2006 Cabernet Franc (Lodi) \

Jim Gordon	0.0
------------	-----

Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	83.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Reininger 2012 Seven Hills Vineyard Carmenère (Walla Walla Valley (WA)) \	
Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	91.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Omero 2015 Chardonnay (Willamette Valley) \	
Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	88.0
Mike DeSimone	0.0
Jeff Jenssen	0.0

NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

A.A. Badenhorst Family Wines 2013 Red (Swartland) \

Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	94.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Black Ink 2015 Red (California) \

Jim Gordon	85.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Herdade Grande 2014 Reserva Branco White (Alentejano) \

Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	90.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	0.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Darcie Kent Vineyards 2007 Picazo Vineyard Merlot

(Livermore Valley) \

Jim Gordon	0.0
Michael Schachner	0.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0
Jeff Jenssen	0.0
NaN	85.0
Anna Lee C. Iijima	0.0
Susan Kostrzewa	0.0
Lauren Buzzeo	0.0
Alexander Peartree	0.0
Fiona Adams	0.0
Carrie Dykes	0.0
Christina Pickard	0.0

Chilcas 2010 Sauvignon Blanc (Central Valley)

Jim Gordon	0.0
Michael Schachner	83.0
Matt Kettmann	0.0
Sean P. Sullivan	0.0
Roger Voss	0.0
Virginie Boone	0.0
Joe Czerwinski	0.0
Paul Gregutt	0.0
Mike DeSimone	0.0

```

Jeff Jenssen                                0.0
NaN                                           0.0
Anna Lee C. Iijima                          0.0
Susan Kostrzewa                             0.0
Lauren Buzzeo                              0.0
Alexander Peartree                          0.0
Fiona Adams                                0.0
Carrie Dykes                               0.0
Christina Pickard                           0.0

```

```
[18 rows x 24517 columns]
```

```
%%time
```

```
U, S, VT = np.linalg.svd(user_item_matrix__train.T)
```

```
V = VT.T
```

```
CPU times: user 52.7 s, sys: 10.8 s, total: 1min 3s
```

```
Wall time: 18.6 s
```

```
# Матрица соотношения между дегустаторами и латентными факторами
```

```
U.shape
```

```
(24517, 24517)
```

```
# Матрица соотношения между объектами и латентными факторами
```

```
V.shape
```

```
(18, 18)
```

```
S.shape
```

```
(18,)
```

```
Sigma = np.diag(S)
```

```
Sigma.shape
```

```
(18, 18)
```

```
# Диагональная матрица сингулярных значений
```

```
Sigma
```

```

array([[6328.37615756,    0.,    0.,    0.,    0.,    0.,    0.,    0.,
        0.,    0.,    0.,    0.,    0.,    0.,    0.,    0.,    0.],
       [    0., 6214.00788753,    0.,    0.,    0.,    0.,    0.,    0.,
        0.,    0.,    0.,    0.,    0.,    0.,    0.,    0.,    0.],
       [    0.,    0., 4603.41568838,    0.,    0.,    0.,    0.,    0.,
        0.,    0.,    0.,    0.,    0.,    0.,    0.,    0.,    0.]])

```

0.	,	0.	,	0.	,	0.	,
0.	,	0.	,	0.	,	0.	,
0.	,	0.	,	0.	,	0.	,
0.	,	0.	],				
[	0.	,	0.	,	0.	,	3880.90866797,
	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
	0.	,	0.	],			
[	0.	,	0.	,	0.	,	0.
3683.95055254,		0.	,	0.	,	0.	,
	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
	0.	,	0.	],			
[	0.	,	0.	,	0.	,	0.
	0.	,	3004.09187609,	0.	,	0.	,
	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
	0.	,	0.	],			
[	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	2830.86117639,	0.	,
	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
	0.	,	0.	],			
[	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	2714.05324192,
	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
	0.	,	0.	],			
[	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
2451.15054617,		0.	,	0.	,	0.	,
	0.	,	0.	,	0.	,	0.
	0.	,	0.	],			
[	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
	0.	,	2416.46258391,	0.	,	0.	,
	0.	,	0.	,	0.	,	0.
	0.	,	0.	],			
[	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	1559.43034471,	0.	,
	0.	,	0.	,	0.	,	0.
	0.	,	0.	],			
[	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	0.
	0.	,	0.	,	0.	,	1283.41926119,
	0.	,	0.	,	0.	,	0.
	0.	,	0.	],			
[	0.	,	0.	,	0.	,	0.



```

0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
922.26406197, 0.      , 0.      , 0.      ,
0.      , 0.      ],
[ 0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 825.2545062 , 0.      , 0.      ,
0.      , 0.      ],
[ 0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 745.06040024, 0.      ,
0.      , 0.      ],
[ 0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 390.0179483 ,
0.      , 0.      ],
[ 0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
177.03107072, 0.      ],
[ 0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 0.      , 0.      , 0.      ,
0.      , 124.45882853]]))

```

*# Используем 3 первых сингулярных значения*

*r=3*

Ur = U[:, :r]

Sr = Sigma[:, :r]

Vr = V[:, :r]

*# Матрица соотношения между новым дегустатором и латентными факторами*

test\_user = np.mat(user\_item\_matrix\_\_test.values)

test\_user.shape, test\_user

((1, 24517), matrix([[0., 0., 0., ..., 0., 0., 0.])))

tmp = test\_user \* Ur \* np.linalg.inv(Sr)

tmp

matrix([[ 3.78394162e-04, -4.35827216e-06, 2.92218350e-18]])

test\_user\_result = np.array([tmp[0,0], tmp[0,1], tmp[0,2]])

test\_user\_result

array([ 3.78394162e-04, -4.35827216e-06, 2.92218350e-18])

```

# Вычисляем косинусную близость между текущим дегустатором
# и остальными дегустаторами
cos_sim = cosine_similarity(Vr, test_user_result.reshape(1, -1))
cos_sim[:10]

array([[ 9.99999728e-01],
       [-1.44540659e-18],
       [ 1.61406973e-32],
       [-3.33224730e-35],
       [-4.12491330e-04],
       [ 9.9999975e-01],
       [ 1.02200867e-37],
       [-1.04994959e-03],
       [ 0.00000000e+00],
       [ 0.00000000e+00]])

# Преобразуем размерность массива
cos_sim_list = cos_sim.reshape(-1, cos_sim.shape[0])[0]
cos_sim_list[:10]

array([ 9.99999728e-01, -1.44540659e-18,  1.61406973e-32, -3.33224730e-35,
        -4.12491330e-04,  9.9999975e-01,  1.02200867e-37, -1.04994959e-03,
         0.00000000e+00,  0.00000000e+00])

# Находим наиболее близкого дегустатора
recommended_user_id = np.argsort(-cos_sim_list)[0]
recommended_user_id

5

test_user

matrix([[0., 0., 0., ..., 0., 0., 0.]])

# Получение названия вина
wine_list = list(user_item_matrix.columns)
def wine_name(ind):
    try:
        wine = wine_list[ind]
        #print(wineId)
        #flt_links = data3[data['movieId'] == wineId]
        #tmdbId = int(flt_links['tmdbId'].values[0])
        #md_links = df_md[df_md['id'] == tmdbId]
        #res = md_links['title'].values[0]
        return wine
    except:
        return ''

# Вина, которые оценивал текущий дегустатор:
i=1
for idx, item in enumerate(np.ndarray.flatten(np.array(test_user))):
    if item > 0:

```

```

film_title = wine_name(idx)
print('{} - {} - {}'.format(idx, film_title, item))
if i==20:
    break
else:
    i+=1

```

```

5 - Lantieri de Paratico NV Rosé Sparkling (Franciacorta) - 85.0
15 - Ronco Blanchis 2015 Friulano (Collio) - 89.0
17 - Italo Cescon 2014 Pinot Grigio (Friuli Grave) - 85.0
26 - Cantina Valpolicella Negrar 2013 Amarone della Valpolicella Classico -
87.0
46 - Nino Franco 2016 Vigneto della Riva di San Floriano Brut (Valdobbiadene
Prosecco Superiore) - 93.0
51 - Feudi di San Gregorio 2011 Taurasi - 91.0
84 - Abbazia di Novacella 2011 Praepositus Riserva Lagrein (Alto Adige) -
90.0
98 - Ca' Marcanda 2010 Camarcanda (Bolgheri) - 93.0
99 - Tenuta La Badiola 2013 642° Il Canapone (Maremma Toscana) - 88.0
109 - Abbadia Ardenga 2010 Vigna Piaggia (Brunello di Montalcino) - 94.0
124 - Castello San Donato in Perano 2009 Riserva (Chianti Classico) - 90.0
129 - Proprietà Sperino 2009 Lessona - 92.0
134 - Colutta 2013 Pinot Grigio (Colli Orientali del Friuli) - 88.0
144 - Col d'Orcia 2013 Banditella (Rosso di Montalcino) - 87.0
158 - La Fornace 2012 Origini (Brunello di Montalcino) - 86.0
159 - Fattoria Fibbiano 2012 Casalini (Chianti Superiore) - 85.0
175 - Corte Moschina 2013 Amarone della Valpolicella - 87.0
189 - Benanti 2012 Rovittello (Etna) - 90.0
232 - Terrazze dell'Etna 2012 Ciuri (Etna) - 89.0
234 - Teanum 2012 Alta Nero di Troia (Puglia) - 86.0

```

*# Вина, которые оценивал наиболее схожий дегустатор:*

```

i=1
recommended_user_item_matrix = user_item_matrix.loc[['Roger Voss']]
for idx, item in
enumerate(np.ndarray.flatten(np.array(recommended_user_item_matrix))):
    if item > 0:
        _wine = wine_name(idx)
        print('{} - {} - {}'.format(idx, _wine, item))
        if i==20:
            break
        else:
            i+=1

```

```

0 - Château Roc Meynard 2008 Bordeaux Supérieur - 87.0
2 - Pascal Aufranc 2014 Juliéna - 88.0
6 - Falua 2006 Conde de Vimioso Red (Ribatejano) - 89.0
8 - Chanson Père et Fils 2010 Les Clos Grand Cru (Chablis) - 92.0
14 - Jeaux-Robin NV Extra Brut Grande Tradition (Champagne) - 90.0
22 - Wolfberger 2013 Eichberg Grand Cru Riesling (Alsace) - 90.0

```

31 - Fonseca 2005 Quinta do Panascal (Port) - 91.0  
38 - Clos des Terrasses 2008 Cuvée le Clos Red (Côtes de Bergerac) - 89.0  
43 - Château la Fleur-Pétrus 2000 Pomerol - 94.0  
44 - Château d'Aydie 2013 Aydie l'Origine Tannat-Cabernet (Madiran) - 90.0  
60 - Château Duhart-Milon 2014 Pauillac - 93.0  
64 - Domaine de la Prébende 2014 Vieilles Vignes (Beaujolais) - 86.0  
68 - Florent Descombe NV Syrah Rosé (Vin de France) - 87.0  
87 - Château Margaux 2007 Pavillon Rouge de Château Margaux (Margaux) - 89.0  
89 - Helfrich 2009 Riesling (Alsace) - 85.0  
91 - Roux Père et Fils 2010 La Perrière (Mercurey) - 87.0  
92 - Lombard et Cie NV Brut Nature Grand Cru Blanc de Blancs Chardonnay (Champagne) - 93.0  
110 - Château de Fuissé 2014 Tête de Cru (Pouilly-Fuissé) - 94.0  
128 - Adega Cooperativa de Borba 2014 Montes Claros Reserva Branco White (Alentejo) - 88.0  
130 - Domaine Parent 2003 Les Epenots Premier Cru (Pommard) - 93.0

*Как видно, фильтрация на основе содержания и коллаборативная фильтрация показывают различные результаты работы в рамках рекомендательных систем*