

# Implicit ALS

Андрей Москаленко

# Implicit факторы

Штрафуем фильм в инициализации, если пользователь его не смотрел  
т.е. отсутствует и в трейне и в тесте

```
for u in range(user_size):  
    for i in range(item_size):  
        if train[u][i]:  
            matrix_full[u][i] = train[u][i]  
        elif (u, i) in watched:  
            matrix_full[u][i] = mean_score  
        else:  
            matrix_full[u][i] = mean_score * self.neg_sampling
```

# Implicit факторы

Штрафуем фильм в инициализации, если пользователь его не смотрел  
т.е. отсутствует и в трейне и в тесте

```
for u in range(user_size):  
    for i in range(item_size):  
        if train[u][i]:  
            matrix_full[u][i] = train[u][i]  
        elif (u, i) in watched:  
            matrix_full[u][i] = mean_score  
        else:  
            matrix_full[u][i] = mean_score * self.neg_sampling
```

# Бэггинг

Фитим на валидации в 0.33 от датасета

```
sorted(GS, key=lambda x: x[0])[:20]
```

```
[(0.9139662442046143, 11, 5, 100, 15, 0.75, 1),  
(0.9144519547170014, 10, 5, 100, 15, 0.85, 1),  
(0.9147856216111326, 11, 5, 40, 15, 0.85, 10),  
(0.9149946763732804, 14, 5, 40, 10, 0.85, 5),  
(0.9155571578200764, 11, 5, 40, 20, 0.85, 20),  
(0.9158111653793563, 8, 10, 100, 20, 0.85, 1),  
(0.9158415430872886, 12, 5, 40, 15, 0.75, 10),  
(0.9158767554267505, 15, 5, 40, 10, 0.75, 5),  
(0.9158834603710929, 13, 5, 25, 10, 0.85, 20),  
(0.9159032701005957, 25, 5, 40, 15, 0.85, 20),  
(0.9159282695248893, 9, 5, 100, 20, 0.75, 1),  
(0.9160828203524338, 8, 5, 100, 20, 0.85, 1),  
(0.9163132854131656, 9, 5, 40, 15, 0.85, 5),  
(0.9163207303660508, 9, 10, 100, 20, 0.75, 1),  
(0.9165103356830069, 14, 5, 40, 20, 0.85, 50),  
(0.9165156567796929, 10, 5, 25, 10, 0.85, 10),  
(0.9167239313150672, 11, 5, 25, 15, 0.85, 50),  
(0.9167525893117328, 19, 5, 40, 15, 0.75, 20),  
(0.9167969535263346, 11, 5, 40, 20, 0.75, 20),  
(0.9169190198281769, 7, 15, 100, 20, 0.85, 1)]
```

# Бэггинг над беггингом

Фитим на ответы первого беггинга

```
sorted(GS, key=lambda x: x[0])[:20]
```

```
[(0.8695166666651777, 6, 10, 25, 5, 0.85, 1),  
(0.8706953771223594, 8, 10, 10, 5, 0.85, 20),  
(0.8708991969948306, 16, 15, 25, 5, 0.85, 1),  
(0.8712175719863978, 8, 5, 10, 5, 0.85, 20),  
(0.8712504037270057, 7, 10, 10, 5, 0.85, 10),  
(0.8716418715089047, 10, 5, 10, 5, 0.85, 50),  
(0.8716988531174237, 5, 5, 25, 5, 0.85, 1),  
(0.8721054678992153, 8, 5, 40, 5, 0.85, 1),  
(0.8724489406357814, 13, 15, 10, 5, 0.85, 10),  
(0.8730113091855535, 7, 5, 10, 5, 0.85, 10),  
(0.8732856292401278, 18, 15, 10, 5, 0.85, 20),  
(0.8738634358006234, 10, 10, 10, 5, 0.85, 50),  
(0.8744412628713872, 9, 10, 40, 5, 0.85, 1),  
(0.8751510233697914, 15, 5, 25, 5, 0.85, 5),  
(0.8759978203050359, 7, 10, 10, 5, 0.85, 5),  
(0.8763067254854525, 7, 15, 10, 5, 0.85, 5),  
(0.8777708254734176, 7, 5, 10, 5, 0.85, 5),  
(0.8783185335303358, 29, 15, 10, 5, 0.85, 50),  
(0.8787555177458003, 29, 15, 40, 5, 0.85, 1),  
(0.8807564999242398, 21, 5, 25, 5, 0.85, 10)]
```

# Бэггинг над беггингом x5

Фитим на ответы предыдущих беггингов  
PROFIT!

Минусы – качество ограничено числом попыток?