Different models, recommend top 10 movies

In [1]:
```
import math
import copy
import pickle
from pathlib import Path
from itertools import zip_longest
from collections import defaultdict
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split, KFold
import torch
from torch import nn
from torch import optim
from torch.nn import functional as F
from torch.optim.lr_scheduler import _LRScheduler
from time import time
from collections import defaultdict
```

In [2]:
```
from model_archs import EmbeddingNet, EmbeddingNetGender, EmbeddingNetAge, EmbeddingNetGenderAge
```

In [26]:
```
def set_random_seed(state=1):
    gens = (np.random.seed, torch.manual_seed, torch.cuda.manual_seed)
    for set_state in gens:
        set_state(state)


RANDOM_STATE = 1
set_random_seed(RANDOM_STATE)
n_users = 6040
n_movies = 3706
dataset_sizes = {'train': 800167, 'val': 200042}
```

In [27]:
```
# batch-wise data iterator
class ReviewsIterator:
    def __init__(self, X, y, batch_size=32, shuffle=True):
        X, y = np.asarray(X), np.asarray(y)

        if shuffle:
            index = np.random.permutation(X.shape[0])
            X, y = X[index], y[index]

        self.X = X
        self.y = y
        self.batch_size = batch_size
        self.shuffle = shuffle
        self.n_batches = int(math.ceil(X.shape[0] // batch_size))
        self._current = 0

    def __iter__(self):
        return self

    def __next__(self):
        return self.next()

    def next(self):
        if self._current >= self.n_batches:
            raise StopIteration()
        k = self._current
        self._current += 1
        bs = self.batch_size
        return self.X[k*bs:(k + 1)*bs], self.y[k*bs:(k + 1)*bs]


def batches(X, y, bs=32, shuffle=True):
    for xb, yb in ReviewsIterator(X, y, bs, shuffle):
        xb = torch.LongTensor(xb)
        yb = torch.FloatTensor(yb)
        yield xb, yb.view(-1, 1)
```

In [38]:
```
def get_precision_recall(df_final, k=10, threshold=3.5):
    # map prediction to each user --> similar to top n
    # {id:(pred, truth)}
    user_pred_truth = defaultdict(list)
    for row in df_final.itertuples():
        _, user_id, movie_id, truth, pred = row
        user_pred_truth[user_id].append((pred, truth))

    precisions = dict()
    recalls = dict()

    for user_id, user_ratings in user_pred_truth.items():
        # Sort user ratings by estimated value
        user_ratings.sort(key=lambda x: x[0], reverse=True)

        # Number of relevant items
        n_rel = sum((true_r >= threshold) for (_, true_r) in user_ratings)

        # Number of recommended items in top k
        n_rec_k = sum((est >= threshold) for (est, _) in user_ratings[:k])

        # Number of relevant and recommended items in top k
        n_rel_and_rec_k = sum(((true_r >= threshold) and (est >= threshold))
                              for (est, true_r) in user_ratings[:k])

        # Precision@K: Proportion of recommended items that are relevant
        precisions[user_id] = n_rel_and_rec_k / n_rec_k if n_rec_k != 0 else 1

        # Recall@K: Proportion of relevant items that are recommended
        recalls[user_id] = n_rel_and_rec_k / n_rel if n_rel != 0 else 1

    # mean precision and recall
    mean_precision = sum(prec for prec in precisions.values()) / len(precisions)
    mean_recall = sum(rec for rec in recalls.values()) / len(recalls)
    print("Prec10 {:.4f}|Rec10 {:.4f}".format(mean_precision, mean_recall))

# get topn
def get_top_n(df_final, n=10):
    """
    key: user_id
    value: his top 10 highest movies as well as ratings
    """
    # map predictions to each user
    top_n = defaultdict(list)
    for row in df_final.itertuples():
        _, user_id, movie_id, truth, pred = row
        top_n[user_id].append((movie_id, pred))

    # sort the pred for each user
    for user_id, pred_ratings in top_n.items():
        pred_ratings.sort(key=lambda x: x[1], reverse=True)
        top_n[user_id] = pred_ratings[:n]
    return top_n

def get_train_pred_top10(df, df_final, datasets, user_id=1635, n=10):
    """
    df: the original df --> contain movie name
    df_final: the final df with predicted ratings
    datasets: the datasets with training and testing datasets
    user_id: the user id to be queried
    n: top_n
    """
    # step 1, get top n from df_final
    top_n = get_top_n(df_final, n)
    assert(user_id in top_n), "user_id {} is not in testing data, try another user such as 1635".format(user_
id)
    pred_ratings = top_n[user_id]

    # step 2: user information
    user = df[df['user_id'] == user_id]
    age = list(set(user['age']))[0]
    sex = list(set(user['sex']))[0]
    info = "User {}, age {}, {}, ".format(user_id, age, sex)

    # step 2, build df_train
    df_train = pd.DataFrame(datasets['train'][0])
    df_train['rating'] = datasets['train'][1]

    # step 3, find all movies user_id has been rated 5
    # df_refined = df_train[(df_train['user_id'] == user_id) & (df_train['rating'] == 5)]
    df_refined = df_train[df_train['user_id'] == user_id]
    movie_id_sets_train = set(df_refined['movie_id'])
    info = "{} has rated {} movies in training set\n".format(info, len(movie_id_sets_train))
    print(info)

    # step 4: get the top n
    # print("===== ===== ===== ===== ===== =====")
    # print("\nTop {} recommendations\n".format(n))
```

```
    for (movie_id, rating) in pred_ratings:
        movie_name = list(set(df[df['movie_id'] == movie_id]['movie_title']))[0]
        info = "ID {:<4d}|Rating {:2f}|{}".format(movie_id, rating, movie_name)
        if movie_id in movie_id_sets_train:
            info = "{}, but this movie has been rated during training!!!".format(info)
        print(info)
```

In [29]:

```
# load preprocessed df
df = pd.read_csv("ml-1m_dl.csv")
print(df.shape)
df.head()
```

```
(1000209, 9)
```

Out[29]:

	movie_id	movie_title	user_id	age	sex	occupation	rating	sex_index	age_index
0	1	Toy Story (1995)	1	1	F	10	5	0	0
1	48	Pocahontas (1995)	1	1	F	10	5	0	0
2	150	Apollo 13 (1995)	1	1	F	10	5	0	0
3	260	Star Wars: Episode IV - A New Hope (1977)	1	1	F	10	4	0	0
4	527	Schindler's List (1993)	1	1	F	10	5	0	0

In [30]:

```
# load dataset
datasets = pickle.load(open('ml-1m_dl.pkl','rb'))
datasets['val'][1]
```

Out[30]:

```
630120    4.0
229398    5.0
758377    3.0
159240    5.0
254252    4.0
         ...
875199    4.0
743921    4.0
527163    4.0
623363    3.0
120098    3.0
Name: rating, Length: 200042, dtype: float32
```

Compare following models

Arch	Index	n_factor	n_hidden	g_factor	a_factor	RMSE	Recall
EmbeddingNet	6	200	1000	-	-	0.8787	0.5528
EmbeddingNetGender	8	200	1000	25	-	0.8721	0.5637
EmbeddingNetAge	12	150	1000	-	15	0.8724	0.5647
EmbeddingNetGenderAge	5	200	1000	10	10	0.8688	0.5698

Compare following users

user_id	gender	age
1	F	1
19	M	1
176	F	18
181	M	18
253	F	25
261	M	25
704	F	35
749	M	35
1428	F	45
1565	M	45
1961	F	50
2088	M	50
5020	F	56
5583	M	56

In [54]:

```
user_id_list = [1, 19, 176, 181, 253, 261, 704, 749, 1428, 1565, 1961, 2088, 5020, 5583]
```

In [53]:
```
# set(df[(df['sex'] == 'M') & (df['age'] == 56)]['user_id'])
```

Now let's begin

1. No gender, no age

In [39]:
```
def get_result_df(datasets, model, best_weights, bs, save_path=None):
    '''
    model_parameter1_best.weights
    '''
    minmax = (1.0, 5.0)
    device = torch.device('cpu')
    model.load_state_dict(best_weights)
    groud_truth, predictions = [], []

    val_size = len(datasets['val'][0])
    # print("Total val size: {}".format(val_size))

    with torch.no_grad():
        for batch in batches(*datasets['val'], shuffle=False, bs=bs):
            x_batch, y_batch = [b.to(device) for b in batch]
            outputs = model(x_batch[:, 0], x_batch[:, 1], minmax)
            groud_truth.extend(y_batch.tolist())
            predictions.extend(outputs.tolist())

        last_num = val_size % bs
        # print("Last num: {}".format(last_num))
        dataset_last = (datasets['val'][0][-last_num:], datasets['val'][1][-last_num:])
        # print("Last dataset: {}".format(len(dataset_last[0])))
        for batch in batches(*dataset_last, shuffle=False, bs=1):
            x_batch, y_batch = [b.to(device) for b in batch]
            outputs = model(x_batch[:, 0], x_batch[:, 1], minmax)
            groud_truth.extend(y_batch.tolist())
            predictions.extend(outputs.tolist())

    groud_truth = np.asarray(groud_truth).ravel()
    predictions = np.asarray(predictions).ravel()

    assert(predictions.shape[0] == val_size)

    final_loss = np.sqrt(np.mean((predictions - groud_truth)**2))
    # print("RMSE: {:.4f}".format(final_loss))

    df_final = pd.DataFrame(datasets['val'][0])[['user_id','movie_id']]
    df_final['truth'] = datasets['val'][1]
    df_final['pred'] = predictions

    if save_path is not None:
        print("Save weight to:{}".format(save_path))
        with open(save_path, 'wb') as file:
            pickle.dump(best_weights, file)

    return df_final # note that here the sex and age is not included
```

In [32]:
```
# load model
n_factors = 200
hidden = [1000,1000,1000]
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
# training
lr = 1e-3
wd = 1e-5
bs =2000
n_epochs = 100
patience = 10

model = EmbeddingNet(n_users, n_movies,
                     n_factors=n_factors,
                     hidden=hidden,
                     embedding_dropout=embedding_dropout,
                     dropouts=dropouts)
best_weights = pickle.load(open("saved_models/noSex_noAge_para6.weights",'rb'))
```

In [55]:
```
df_final = get_result_df(datasets, model, best_weights, bs, save_path=None)
for user_id in user_id_list:
    get_train_pred_top10(df, df_final, datasets, user_id=user_id, n=10)
    print("===== ===== ===== ===== =====\n")
```

```
User 1, age 1, F,  has rated 41 movies in training set

ID 260 |Rating 4.290243|Star Wars: Episode IV - A New Hope (1977)
ID 1   |Rating 4.190775|Toy Story (1995)
ID 1961|Rating 4.174579|Rain Man (1988)
ID 595 |Rating 4.166494|Beauty and the Beast (1991)
ID 2804|Rating 4.131517|Christmas Story, A (1983)
ID 1022|Rating 4.043647|Cinderella (1950)
ID 2791|Rating 3.990579|Airplane! (1980)
ID 1207|Rating 3.976852|To Kill a Mockingbird (1962)
ID 2687|Rating 3.972957|Tarzan (1999)
ID 608 |Rating 3.905061|Fargo (1996)
===== ===== ===== ===== =====

User 19, age 1, M,  has rated 189 movies in training set

ID 2858|Rating 4.309461|American Beauty (1999)
ID 2571|Rating 4.237044|Matrix, The (1999)
ID 2028|Rating 4.189097|Saving Private Ryan (1998)
ID 1198|Rating 4.167509|Raiders of the Lost Ark (1981)
ID 457 |Rating 4.154977|Fugitive, The (1993)
ID 1   |Rating 4.143876|Toy Story (1995)
ID 1321|Rating 4.122013|American Werewolf in London, An (1981)
ID 1148|Rating 4.068268|Wrong Trousers, The (1993)
ID 1270|Rating 4.031488|Back to the Future (1985)
ID 2064|Rating 4.014668|Roger & Me (1989)
===== ===== ===== ===== =====

User 176, age 18, F,  has rated 58 movies in training set

ID 1196|Rating 4.510611|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 3552|Rating 4.434868|Caddyshack (1980)
ID 1197|Rating 4.257071|Princess Bride, The (1987)
ID 2378|Rating 4.249933|Police Academy (1984)
ID 1307|Rating 4.216447|When Harry Met Sally... (1989)
ID 2011|Rating 4.208655|Back to the Future Part II (1989)
ID 2795|Rating 4.164939|Vacation (1983)
ID 1544|Rating 4.015627|Lost World: Jurassic Park, The (1997)
ID 3715|Rating 3.991545|Burglar (1987)
ID 1968|Rating 3.975549|Breakfast Club, The (1985)
===== ===== ===== ===== =====

User 181, age 18, M,  has rated 242 movies in training set

ID 1136|Rating 4.839367|Monty Python and the Holy Grail (1974)
ID 1240|Rating 4.525747|Terminator, The (1984)
ID 1732|Rating 4.508242|Big Lebowski, The (1998)
ID 1213|Rating 4.504430|GoodFellas (1990)
ID 1704|Rating 4.382651|Good Will Hunting (1997)
ID 2000|Rating 4.331407|Lethal Weapon (1987)
ID 2683|Rating 4.302487|Austin Powers: The Spy Who Shagged Me (1999)
ID 1291|Rating 4.288813|Indiana Jones and the Last Crusade (1989)
ID 2028|Rating 4.288600|Saving Private Ryan (1998)
ID 2706|Rating 4.264375|American Pie (1999)
===== ===== ===== ===== =====

User 253, age 25, F,  has rated 41 movies in training set

ID 246 |Rating 4.240369|Hoop Dreams (1994)
ID 3083|Rating 4.081717|All About My Mother (Todo Sobre Mi Madre) (1999)
ID 3408|Rating 4.023458|Erin Brockovich (2000)
ID 3751|Rating 3.976006|Chicken Run (2000)
ID 1089|Rating 3.943969|Reservoir Dogs (1992)
ID 2599|Rating 3.909781|Election (1999)
ID 446 |Rating 3.868521|Farewell My Concubine (1993)
ID 296 |Rating 3.858854|Pulp Fiction (1994)
ID 1127|Rating 3.835043|Abyss, The (1989)
ID 1533|Rating 3.750873|Promise, The (La Promesse) (1996)
===== ===== ===== ===== =====

User 261, age 25, M,  has rated 297 movies in training set

ID 1221|Rating 4.675239|Godfather: Part II, The (1974)
ID 2692|Rating 4.622235|Run Lola Run (Lola rennt) (1998)
ID 3741|Rating 4.455191|Badlands (1973)
ID 1283|Rating 4.388023|High Noon (1952)
ID 1225|Rating 4.374327|Amadeus (1984)
ID 1307|Rating 4.212987|When Harry Met Sally... (1989)
ID 3424|Rating 4.159008|Do the Right Thing (1989)
ID 1265|Rating 4.144706|Groundhog Day (1993)
ID 2542|Rating 4.104301|Lock, Stock & Two Smoking Barrels (1998)
ID 3706|Rating 4.069049|Angel Heart (1987)
===== ===== ===== ===== =====

User 704, age 35, F,  has rated 61 movies in training set

ID 2959|Rating 3.921731|Fight Club (1999)
ID 2761|Rating 3.888439|Iron Giant, The (1999)
```

```
ID 2490|Rating 3.747899|Payback (1999)
ID 3911|Rating 3.701699|Best in Show (2000)
ID 3384|Rating 3.686259|Taking of Pelham One Two Three, The (1974)
ID 2390|Rating 3.568743|Little Voice (1998)
ID 3893|Rating 3.484054|Nurse Betty (2000)
ID 3896|Rating 3.462925|Way of the Gun, The (2000)
ID 3298|Rating 3.428260|Boiler Room (2000)
ID 3185|Rating 2.903764|Snow Falling on Cedars (1999)
===== ===== ===== ===== =====

User 749, age 35, M,  has rated 375 movies in training set

ID 2762|Rating 4.272712|Sixth Sense, The (1999)
ID 1276|Rating 4.179347|Cool Hand Luke (1967)
ID 1954|Rating 4.135741|Rocky (1976)
ID 912 |Rating 3.999698|Casablanca (1942)
ID 908 |Rating 3.956526|North by Northwest (1959)
ID 1196|Rating 3.902882|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 1090|Rating 3.894553|Platoon (1986)
ID 1204|Rating 3.869633|Lawrence of Arabia (1962)
ID 2423|Rating 3.821836|Christmas Vacation (1989)
ID 3072|Rating 3.791842|Moonstruck (1987)
===== ===== ===== ===== =====

User 1428, age 45, F,  has rated 41 movies in training set

ID 1206|Rating 5.187326|Clockwork Orange, A (1971)
ID 908 |Rating 5.048565|North by Northwest (1959)
ID 2227|Rating 4.868504|Lodger, The (1926)
ID 1282|Rating 4.369121|Fantasia (1940)
ID 175 |Rating 4.292841|Kids (1995)
ID 3930|Rating 3.958838|Creature From the Black Lagoon, The (1954)
ID 2291|Rating 3.911605|Edward Scissorhands (1990)
ID 3927|Rating 3.810028|Fantastic Voyage (1966)
ID 3926|Rating 3.601877|Voyage to the Bottom of the Sea (1961)
ID 3450|Rating 3.579489|Grumpy Old Men (1993)
===== ===== ===== ===== =====

User 1565, age 45, M,  has rated 37 movies in training set

ID 2997|Rating 3.997089|Being John Malkovich (1999)
ID 348 |Rating 3.948089|Bullets Over Broadway (1994)
ID 224 |Rating 3.928763|Don Juan DeMarco (1995)
ID 2396|Rating 3.893962|Shakespeare in Love (1998)
ID 34  |Rating 3.818663|Babe (1995)
ID 2599|Rating 3.816698|Election (1999)
ID 1732|Rating 3.598989|Big Lebowski, The (1998)
ID 357 |Rating 3.582452|Four Weddings and a Funeral (1994)
ID 3129|Rating 3.297520|Sweet and Lowdown (1999)
ID 619 |Rating 3.196335|Ed (1996)
===== ===== ===== ===== =====

User 1961, age 50, F,  has rated 71 movies in training set

ID 318 |Rating 4.328104|Shawshank Redemption, The (1994)
ID 1198|Rating 4.283365|Raiders of the Lost Ark (1981)
ID 1196|Rating 4.243160|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 858 |Rating 4.169449|Godfather, The (1972)
ID 457 |Rating 3.843350|Fugitive, The (1993)
ID 2951|Rating 3.820276|Fistful of Dollars, A (1964)
ID 2501|Rating 3.709575|October Sky (1999)
ID 1036|Rating 3.688698|Die Hard (1988)
ID 296 |Rating 3.680509|Pulp Fiction (1994)
ID 2966|Rating 3.662994|Straight Story, The (1999)
===== ===== ===== ===== =====

User 2088, age 50, M,  has rated 355 movies in training set

ID 589 |Rating 4.872320|Terminator 2: Judgment Day (1991)
ID 3469|Rating 4.719260|Inherit the Wind (1960)
ID 1304|Rating 4.704005|Butch Cassidy and the Sundance Kid (1969)
ID 480 |Rating 4.703722|Jurassic Park (1993)
ID 2396|Rating 4.701219|Shakespeare in Love (1998)
ID 1207|Rating 4.612859|To Kill a Mockingbird (1962)
ID 1222|Rating 4.605288|Full Metal Jacket (1987)
ID 2951|Rating 4.601109|Fistful of Dollars, A (1964)
ID 3737|Rating 4.582014|Lonely Are the Brave (1962)
ID 541 |Rating 4.555644|Blade Runner (1982)
===== ===== ===== ===== =====

User 5020, age 56, F,  has rated 84 movies in training set

ID 938 |Rating 4.152769|Gigi (1958)
ID 3792|Rating 4.143832|Duel in the Sun (1946)
ID 3171|Rating 3.939328|Room at the Top (1959)
ID 1210|Rating 3.882236|Star Wars: Episode VI - Return of the Jedi (1983)
ID 3723|Rating 3.874931|Hamlet (1990)
ID 592 |Rating 3.862677|Batman (1989)
ID 3712|Rating 3.822860|Soapdish (1991)
```

```
ID 3751|Rating 3.787452|Chicken Run (2000)
ID 2021|Rating 3.735971|Dune (1984)
ID 2784|Rating 3.683913|Masque of the Red Death, The (1964)
===== ===== ===== ===== =====

User 5583, age 56, M,  has rated 95 movies in training set

ID 2858|Rating 4.600380|American Beauty (1999)
ID 260 |Rating 4.597151|Star Wars: Episode IV - A New Hope (1977)
ID 2396|Rating 4.411871|Shakespeare in Love (1998)
ID 1198|Rating 4.354239|Raiders of the Lost Ark (1981)
ID 2571|Rating 4.199419|Matrix, The (1999)
ID 920 |Rating 4.155070|Gone with the Wind (1939)
ID 457 |Rating 4.062383|Fugitive, The (1993)
ID 714 |Rating 3.989401|Dead Man (1995)
ID 3176|Rating 3.921952|Talented Mr. Ripley, The (1999)
ID 2951|Rating 3.907883|Fistful of Dollars, A (1964)
===== ===== ===== ===== =====
```

2. With gender, no age

In [57]:

```python
def get_result_df(datasets, model, best_weights, bs, save_path=None):
    '''
    model_parameter1_best.weights
    '''
    minmax = (1.0, 5.0)
    device = torch.device('cpu')
    model.load_state_dict(best_weights)
    groud_truth, predictions = [], []

    val_size = len(datasets['val'][0])
    # print("Total val size: {}".format(val_size))

    with torch.no_grad():
        for batch in batches(*datasets['val'], shuffle=False, bs=bs):
            x_batch, y_batch = [b.to(device) for b in batch]
            outputs = model(x_batch[:, 0], x_batch[:, 1], x_batch[:, 2], minmax)
            groud_truth.extend(y_batch.tolist())
            predictions.extend(outputs.tolist())

        last_num = val_size % bs
        # print("Last num: {}".format(last_num))
        dataset_last = (datasets['val'][0][-last_num:], datasets['val'][1][-last_num:])
        # print("Last dataset: {}".format(len(dataset_last[0])))
        for batch in batches(*dataset_last, shuffle=False, bs=1):
            x_batch, y_batch = [b.to(device) for b in batch]
            outputs = model(x_batch[:, 0], x_batch[:, 1], x_batch[:, 2], minmax)
            groud_truth.extend(y_batch.tolist())
            predictions.extend(outputs.tolist())

    groud_truth = np.asarray(groud_truth).ravel()
    predictions = np.asarray(predictions).ravel()

    assert(predictions.shape[0] == val_size)

    final_loss = np.sqrt(np.mean((predictions - groud_truth)**2))
    # print("RMSE: {:.4f}".format(final_loss))

    df_final = pd.DataFrame(datasets['val'][0])[['user_id','movie_id']]
    df_final['truth'] = datasets['val'][1]
    df_final['pred'] = predictions

    if save_path is not None:
        print("Save weight to:{}".format(save_path))
        with open(save_path, 'wb') as file:
            pickle.dump(best_weights, file)

    return df_final # note that here the sex and age is not included
```

In [56]:
```
n_factors = 200
hidden = [1000] * 3
g_factor = 25

embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
# training
lr = 1e-3
wd = 1e-5
bs =2000
n_epochs = 100
patience = 10

model = EmbeddingNetGender(n_users, n_movies, n_factors=n_factors, hidden=hidden, dropouts=dropouts, g_factor
=g_factor)
best_weights = pickle.load(open("saved_models/withSex_noAge_para8.weights",'rb'))
```

In [58]:
```
df_final = get_result_df(datasets, model, best_weights, bs, save_path=None)
for user_id in user_id_list:
    get_train_pred_top10(df, df_final, datasets, user_id=user_id, n=10)
    print("===== ===== ===== ===== =====\n")
```

```
User 1, age 1, F,  has rated 41 movies in training set

ID 1207|Rating 4.472710|To Kill a Mockingbird (1962)
ID 595 |Rating 4.381732|Beauty and the Beast (1991)
ID 260 |Rating 4.348394|Star Wars: Episode IV - A New Hope (1977)
ID 1022|Rating 4.278702|Cinderella (1950)
ID 1   |Rating 4.227912|Toy Story (1995)
ID 1961|Rating 4.132434|Rain Man (1988)
ID 2804|Rating 4.101287|Christmas Story, A (1983)
ID 1287|Rating 4.100507|Ben-Hur (1959)
ID 608 |Rating 4.038425|Fargo (1996)
ID 2687|Rating 3.977126|Tarzan (1999)
===== ===== ===== ===== =====

User 19, age 1, M,  has rated 189 movies in training set

ID 1198|Rating 4.499864|Raiders of the Lost Ark (1981)
ID 2571|Rating 4.402059|Matrix, The (1999)
ID 110 |Rating 4.342436|Braveheart (1995)
ID 2028|Rating 4.332126|Saving Private Ryan (1998)
ID 2858|Rating 4.234061|American Beauty (1999)
ID 356 |Rating 4.213253|Forrest Gump (1994)
ID 1210|Rating 4.158370|Star Wars: Episode VI - Return of the Jedi (1983)
ID 1148|Rating 4.136994|Wrong Trousers, The (1993)
ID 1278|Rating 4.099858|Young Frankenstein (1974)
ID 1610|Rating 4.097485|Hunt for Red October, The (1990)
===== ===== ===== ===== =====

User 176, age 18, F,  has rated 58 movies in training set

ID 1307|Rating 4.286402|When Harry Met Sally... (1989)
ID 1197|Rating 4.262366|Princess Bride, The (1987)
ID 1196|Rating 4.262207|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 1544|Rating 4.239976|Lost World: Jurassic Park, The (1997)
ID 3715|Rating 4.155063|Burglar (1987)
ID 2011|Rating 4.072795|Back to the Future Part II (1989)
ID 2795|Rating 4.042375|Vacation (1983)
ID 3039|Rating 4.029671|Trading Places (1983)
ID 2378|Rating 3.985162|Police Academy (1984)
ID 1968|Rating 3.961801|Breakfast Club, The (1985)
===== ===== ===== ===== =====

User 181, age 18, M,  has rated 242 movies in training set

ID 1213|Rating 5.055884|GoodFellas (1990)
ID 1136|Rating 4.980361|Monty Python and the Holy Grail (1974)
ID 2028|Rating 4.515006|Saving Private Ryan (1998)
ID 1240|Rating 4.469388|Terminator, The (1984)
ID 1704|Rating 4.294208|Good Will Hunting (1997)
ID 1291|Rating 4.257468|Indiana Jones and the Last Crusade (1989)
ID 1732|Rating 4.150012|Big Lebowski, The (1998)
ID 1610|Rating 4.117481|Hunt for Red October, The (1990)
ID 3543|Rating 4.112954|Diner (1982)
ID 3360|Rating 4.094992|Hoosiers (1986)
===== ===== ===== ===== =====

User 253, age 25, F,  has rated 41 movies in training set

ID 3083|Rating 4.408238|All About My Mother (Todo Sobre Mi Madre) (1999)
ID 446 |Rating 4.388174|Farewell My Concubine (1993)
ID 246 |Rating 4.357580|Hoop Dreams (1994)
ID 3408|Rating 4.151003|Erin Brockovich (2000)
ID 2599|Rating 4.136911|Election (1999)
ID 3751|Rating 4.130851|Chicken Run (2000)
ID 1533|Rating 4.004681|Promise, The (La Promesse) (1996)
ID 296 |Rating 3.898038|Pulp Fiction (1994)
ID 2704|Rating 3.749000|Lovers on the Bridge, The (Les Amants du Pont-Neuf) (1991)
ID 1127|Rating 3.587751|Abyss, The (1989)
===== ===== ===== ===== =====

User 261, age 25, M,  has rated 297 movies in training set

ID 1221|Rating 4.586732|Godfather: Part II, The (1974)
ID 1283|Rating 4.574106|High Noon (1952)
ID 1225|Rating 4.431887|Amadeus (1984)
ID 2692|Rating 4.392269|Run Lola Run (Lola rennt) (1998)
ID 3741|Rating 4.255603|Badlands (1973)
ID 2542|Rating 4.240274|Lock, Stock & Two Smoking Barrels (1998)
ID 551 |Rating 4.138992|Nightmare Before Christmas, The (1993)
ID 1732|Rating 4.135163|Big Lebowski, The (1998)
ID 1265|Rating 4.116796|Groundhog Day (1993)
ID 3361|Rating 4.113025|Bull Durham (1988)
===== ===== ===== ===== =====

User 704, age 35, F,  has rated 61 movies in training set

ID 3911|Rating 4.095551|Best in Show (2000)
ID 2959|Rating 4.082156|Fight Club (1999)
```

```
ID 2761|Rating 3.953966|Iron Giant, The (1999)
ID 3384|Rating 3.823550|Taking of Pelham One Two Three, The (1974)
ID 2390|Rating 3.747732|Little Voice (1998)
ID 3298|Rating 3.505182|Boiler Room (2000)
ID 3185|Rating 3.443740|Snow Falling on Cedars (1999)
ID 3893|Rating 3.435017|Nurse Betty (2000)
ID 2490|Rating 3.340688|Payback (1999)
ID 3896|Rating 3.281293|Way of the Gun, The (2000)
===== ===== ===== ===== =====

User 749, age 35, M,  has rated 375 movies in training set

ID 2762|Rating 4.250895|Sixth Sense, The (1999)
ID 1196|Rating 4.077009|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 3114|Rating 3.852480|Toy Story 2 (1999)
ID 1204|Rating 3.821676|Lawrence of Arabia (1962)
ID 1210|Rating 3.818435|Star Wars: Episode VI - Return of the Jedi (1983)
ID 2791|Rating 3.815598|Airplane! (1980)
ID 1610|Rating 3.803171|Hunt for Red October, The (1990)
ID 969 |Rating 3.782071|African Queen, The (1951)
ID 1276|Rating 3.781012|Cool Hand Luke (1967)
ID 3897|Rating 3.757148|Almost Famous (2000)
===== ===== ===== ===== =====

User 1428, age 45, F,  has rated 41 movies in training set

ID 908 |Rating 4.991136|North by Northwest (1959)
ID 1282|Rating 4.807029|Fantasia (1940)
ID 1206|Rating 4.801775|Clockwork Orange, A (1971)
ID 2227|Rating 4.693388|Lodger, The (1926)
ID 175 |Rating 4.290076|Kids (1995)
ID 2291|Rating 4.252073|Edward Scissorhands (1990)
ID 3930|Rating 4.174327|Creature From the Black Lagoon, The (1954)
ID 3927|Rating 4.008096|Fantastic Voyage (1966)
ID 1580|Rating 3.940378|Men in Black (1997)
ID 3926|Rating 3.710525|Voyage to the Bottom of the Sea (1961)
===== ===== ===== ===== =====

User 1565, age 45, M,  has rated 37 movies in training set

ID 2396|Rating 4.019731|Shakespeare in Love (1998)
ID 2997|Rating 3.972599|Being John Malkovich (1999)
ID 34  |Rating 3.844083|Babe (1995)
ID 2599|Rating 3.788674|Election (1999)
ID 357 |Rating 3.756473|Four Weddings and a Funeral (1994)
ID 1732|Rating 3.732164|Big Lebowski, The (1998)
ID 348 |Rating 3.689384|Bullets Over Broadway (1994)
ID 3129|Rating 3.605549|Sweet and Lowdown (1999)
ID 224 |Rating 3.497573|Don Juan DeMarco (1995)
ID 619 |Rating 2.260920|Ed (1996)
===== ===== ===== ===== =====

User 1961, age 50, F,  has rated 71 movies in training set

ID 318 |Rating 4.274407|Shawshank Redemption, The (1994)
ID 1198|Rating 4.134890|Raiders of the Lost Ark (1981)
ID 858 |Rating 4.117238|Godfather, The (1972)
ID 2501|Rating 4.052022|October Sky (1999)
ID 1196|Rating 3.969681|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 457 |Rating 3.879099|Fugitive, The (1993)
ID 3578|Rating 3.860467|Gladiator (2000)
ID 1036|Rating 3.832394|Die Hard (1988)
ID 2951|Rating 3.640119|Fistful of Dollars, A (1964)
ID 296 |Rating 3.620668|Pulp Fiction (1994)
===== ===== ===== ===== =====

User 2088, age 50, M,  has rated 355 movies in training set

ID 1304|Rating 4.821420|Butch Cassidy and the Sundance Kid (1969)
ID 1222|Rating 4.818082|Full Metal Jacket (1987)
ID 589 |Rating 4.795949|Terminator 2: Judgment Day (1991)
ID 480 |Rating 4.750465|Jurassic Park (1993)
ID 541 |Rating 4.738379|Blade Runner (1982)
ID 3737|Rating 4.733196|Lonely Are the Brave (1962)
ID 702 |Rating 4.729947|Faces (1968)
ID 3469|Rating 4.723322|Inherit the Wind (1960)
ID 599 |Rating 4.681270|Wild Bunch, The (1969)
ID 1301|Rating 4.678211|Forbidden Planet (1956)
===== ===== ===== ===== =====

User 5020, age 56, F,  has rated 84 movies in training set

ID 938 |Rating 4.417951|Gigi (1958)
ID 3751|Rating 4.282625|Chicken Run (2000)
ID 3792|Rating 4.174982|Duel in the Sun (1946)
ID 3723|Rating 4.134624|Hamlet (1990)
ID 2872|Rating 4.092530|Excalibur (1981)
ID 1210|Rating 4.057086|Star Wars: Episode VI - Return of the Jedi (1983)
ID 3171|Rating 4.043768|Room at the Top (1959)
```

```
ID 2143|Rating 4.015172|Legend (1985)
ID 329 |Rating 3.970602|Star Trek: Generations (1994)
ID 592 |Rating 3.850215|Batman (1989)
===== ===== ===== ===== =====

User 5583, age 56, M,  has rated 95 movies in training set

ID 260 |Rating 4.686344|Star Wars: Episode IV - A New Hope (1977)
ID 2396|Rating 4.616703|Shakespeare in Love (1998)
ID 1198|Rating 4.601224|Raiders of the Lost Ark (1981)
ID 920 |Rating 4.486513|Gone with the Wind (1939)
ID 1291|Rating 4.467784|Indiana Jones and the Last Crusade (1989)
ID 2858|Rating 4.353408|American Beauty (1999)
ID 3406|Rating 4.259543|Captain Horatio Hornblower (1951)
ID 2571|Rating 4.236502|Matrix, The (1999)
ID 457 |Rating 4.211889|Fugitive, The (1993)
ID 1266|Rating 4.115752|Unforgiven (1992)
===== ===== ===== ===== =====
```

In [ ]:

In [ ]:

3. No gender, with age

In [59]:

```python
def get_result_df(datasets, model, best_weights, bs, save_path=None):
    '''
    model_parameter1_best.weights
    '''
    minmax = (1.0, 5.0)
    device = torch.device('cpu')
    model.load_state_dict(best_weights)
    groud_truth, predictions = [], []

    val_size = len(datasets['val'][0])
    # print("Total val size: {}".format(val_size))

    with torch.no_grad():
        for batch in batches(*datasets['val'], shuffle=False, bs=bs):
            x_batch, y_batch = [b.to(device) for b in batch]
            outputs = model(x_batch[:, 0], x_batch[:, 1], x_batch[:, 3], minmax)
            groud_truth.extend(y_batch.tolist())
            predictions.extend(outputs.tolist())

        last_num = val_size % bs
        # print("Last num: {}".format(last_num))
        dataset_last = (datasets['val'][0][-last_num:], datasets['val'][1][-last_num:])
        # print("Last dataset: {}".format(len(dataset_last[0])))
        for batch in batches(*dataset_last, shuffle=False, bs=1):
            x_batch, y_batch = [b.to(device) for b in batch]
            outputs = model(x_batch[:, 0], x_batch[:, 1], x_batch[:, 3], minmax)
            groud_truth.extend(y_batch.tolist())
            predictions.extend(outputs.tolist())

    groud_truth = np.asarray(groud_truth).ravel()
    predictions = np.asarray(predictions).ravel()

    assert(predictions.shape[0] == val_size)

    final_loss = np.sqrt(np.mean((predictions - groud_truth)**2))
    # print("RMSE: {:.4f}".format(final_loss))

    df_final = pd.DataFrame(datasets['val'][0])[['user_id','movie_id']]
    df_final['truth'] = datasets['val'][1]
    df_final['pred'] = predictions

    if save_path is not None:
        print("Save weight to:{}".format(save_path))
        with open(save_path, 'wb') as file:
            pickle.dump(best_weights, file)

    return df_final # note that here the sex and age is not included
```

In [60]:
```
n_factors = 150
hidden = [1000] * 3
a_factor = 15

embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
# training
lr = 1e-3
wd = 1e-5
bs =2000
n_epochs = 100
patience = 10

model = EmbeddingNetAge(n_users, n_movies, n_factors=n_factors, hidden=hidden, dropouts=dropouts, a_factor=a_
factor)
best_weights = pickle.load(open("saved_models/noSex_withAge_para12.weights",'rb'))
```

In [61]:
```
df_final = get_result_df(datasets, model, best_weights, bs, save_path=None)
for user_id in user_id_list:
    get_train_pred_top10(df, df_final, datasets, user_id=user_id, n=10)
    print("===== ===== ===== ===== =====\n")
```

```
User 1, age 1, F,  has rated 41 movies in training set

ID 1207|Rating 4.489090|To Kill a Mockingbird (1962)
ID 1961|Rating 4.385052|Rain Man (1988)
ID 595 |Rating 4.291056|Beauty and the Beast (1991)
ID 1287|Rating 4.273516|Ben-Hur (1959)
ID 260 |Rating 4.243756|Star Wars: Episode IV - A New Hope (1977)
ID 1   |Rating 4.169435|Toy Story (1995)
ID 2804|Rating 4.160616|Christmas Story, A (1983)
ID 608 |Rating 4.145622|Fargo (1996)
ID 1022|Rating 4.098650|Cinderella (1950)
ID 2791|Rating 4.021626|Airplane! (1980)
===== ===== ===== ===== =====

User 19, age 1, M,  has rated 189 movies in training set

ID 2571|Rating 4.618315|Matrix, The (1999)
ID 2858|Rating 4.527251|American Beauty (1999)
ID 2028|Rating 4.471524|Saving Private Ryan (1998)
ID 110 |Rating 4.467456|Braveheart (1995)
ID 1210|Rating 4.332788|Star Wars: Episode VI - Return of the Jedi (1983)
ID 1148|Rating 4.304241|Wrong Trousers, The (1993)
ID 1198|Rating 4.231340|Raiders of the Lost Ark (1981)
ID 356 |Rating 4.198155|Forrest Gump (1994)
ID 2502|Rating 4.141893|Office Space (1999)
ID 919 |Rating 4.115625|Wizard of Oz, The (1939)
===== ===== ===== ===== =====

User 176, age 18, F,  has rated 58 movies in training set

ID 2795|Rating 4.319016|Vacation (1983)
ID 3039|Rating 4.305584|Trading Places (1983)
ID 1196|Rating 4.250363|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 1197|Rating 4.174350|Princess Bride, The (1987)
ID 3552|Rating 4.136608|Caddyshack (1980)
ID 1544|Rating 4.105245|Lost World: Jurassic Park, The (1997)
ID 1307|Rating 3.962842|When Harry Met Sally... (1989)
ID 1968|Rating 3.958197|Breakfast Club, The (1985)
ID 2011|Rating 3.934531|Back to the Future Part II (1989)
ID 2378|Rating 3.920125|Police Academy (1984)
===== ===== ===== ===== =====

User 181, age 18, M,  has rated 242 movies in training set

ID 1213|Rating 4.887247|GoodFellas (1990)
ID 1136|Rating 4.665027|Monty Python and the Holy Grail (1974)
ID 1732|Rating 4.521598|Big Lebowski, The (1998)
ID 778 |Rating 4.462507|Trainspotting (1996)
ID 1291|Rating 4.346444|Indiana Jones and the Last Crusade (1989)
ID 3424|Rating 4.343379|Do the Right Thing (1989)
ID 2028|Rating 4.336349|Saving Private Ryan (1998)
ID 1704|Rating 4.287665|Good Will Hunting (1997)
ID 36  |Rating 4.188827|Dead Man Walking (1995)
ID 353 |Rating 4.113652|Crow, The (1994)
===== ===== ===== ===== =====

User 253, age 25, F,  has rated 41 movies in training set

ID 246 |Rating 4.271537|Hoop Dreams (1994)
ID 3083|Rating 4.237897|All About My Mother (Todo Sobre Mi Madre) (1999)
ID 3408|Rating 4.178385|Erin Brockovich (2000)
ID 2599|Rating 4.157721|Election (1999)
ID 3751|Rating 4.137932|Chicken Run (2000)
ID 446 |Rating 4.059258|Farewell My Concubine (1993)
ID 296 |Rating 3.992813|Pulp Fiction (1994)
ID 1089|Rating 3.963421|Reservoir Dogs (1992)
ID 1533|Rating 3.776638|Promise, The (La Promesse) (1996)
ID 1127|Rating 3.741367|Abyss, The (1989)
===== ===== ===== ===== =====

User 261, age 25, M,  has rated 297 movies in training set

ID 1221|Rating 4.741171|Godfather: Part II, The (1974)
ID 2692|Rating 4.545064|Run Lola Run (Lola rennt) (1998)
ID 3741|Rating 4.393841|Badlands (1973)
ID 1732|Rating 4.355713|Big Lebowski, The (1998)
ID 1283|Rating 4.322998|High Noon (1952)
ID 3424|Rating 4.259820|Do the Right Thing (1989)
ID 1225|Rating 4.243003|Amadeus (1984)
ID 1307|Rating 4.227256|When Harry Met Sally... (1989)
ID 2542|Rating 4.175666|Lock, Stock & Two Smoking Barrels (1998)
ID 1265|Rating 4.101122|Groundhog Day (1993)
===== ===== ===== ===== =====

User 704, age 35, F,  has rated 61 movies in training set

ID 2959|Rating 3.980768|Fight Club (1999)
ID 3384|Rating 3.857615|Taking of Pelham One Two Three, The (1974)
```

```
ID 2761|Rating 3.845630|Iron Giant, The (1999)
ID 3911|Rating 3.842279|Best in Show (2000)
ID 2390|Rating 3.429466|Little Voice (1998)
ID 2490|Rating 3.384069|Payback (1999)
ID 3298|Rating 3.332826|Boiler Room (2000)
ID 3893|Rating 3.308586|Nurse Betty (2000)
ID 3185|Rating 3.174899|Snow Falling on Cedars (1999)
ID 3896|Rating 2.829831|Way of the Gun, The (2000)
===== ===== ===== ===== =====

User 749, age 35, M,  has rated 375 movies in training set

ID 2762|Rating 4.453091|Sixth Sense, The (1999)
ID 1196|Rating 4.043712|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 1210|Rating 3.932474|Star Wars: Episode VI - Return of the Jedi (1983)
ID 1610|Rating 3.897503|Hunt for Red October, The (1990)
ID 912 |Rating 3.849423|Casablanca (1942)
ID 293 |Rating 3.824950|Professional, The (a.k.a. Leon: The Professional) (1994)
ID 1090|Rating 3.820425|Platoon (1986)
ID 1954|Rating 3.820174|Rocky (1976)
ID 3098|Rating 3.787492|Natural, The (1984)
ID 3114|Rating 3.786118|Toy Story 2 (1999)
===== ===== ===== ===== =====

User 1428, age 45, F,  has rated 41 movies in training set

ID 908 |Rating 5.156441|North by Northwest (1959)
ID 1206|Rating 5.059811|Clockwork Orange, A (1971)
ID 1282|Rating 4.960043|Fantasia (1940)
ID 2227|Rating 4.823973|Lodger, The (1926)
ID 175 |Rating 4.449563|Kids (1995)
ID 2291|Rating 4.441682|Edward Scissorhands (1990)
ID 1580|Rating 4.028945|Men in Black (1997)
ID 3930|Rating 4.006862|Creature From the Black Lagoon, The (1954)
ID 3926|Rating 4.002057|Voyage to the Bottom of the Sea (1961)
ID 3927|Rating 3.862918|Fantastic Voyage (1966)
===== ===== ===== ===== =====

User 1565, age 45, M,  has rated 37 movies in training set

ID 2396|Rating 4.527424|Shakespeare in Love (1998)
ID 34  |Rating 4.309412|Babe (1995)
ID 2997|Rating 4.110594|Being John Malkovich (1999)
ID 2599|Rating 4.068440|Election (1999)
ID 357 |Rating 4.006423|Four Weddings and a Funeral (1994)
ID 224 |Rating 3.816521|Don Juan DeMarco (1995)
ID 348 |Rating 3.769893|Bullets Over Broadway (1994)
ID 3129|Rating 3.648651|Sweet and Lowdown (1999)
ID 1732|Rating 3.618274|Big Lebowski, The (1998)
ID 619 |Rating 2.897219|Ed (1996)
===== ===== ===== ===== =====

User 1961, age 50, F,  has rated 71 movies in training set

ID 1198|Rating 4.401946|Raiders of the Lost Ark (1981)
ID 318 |Rating 4.127230|Shawshank Redemption, The (1994)
ID 2501|Rating 4.069268|October Sky (1999)
ID 858 |Rating 4.011808|Godfather, The (1972)
ID 457 |Rating 3.899220|Fugitive, The (1993)
ID 1036|Rating 3.809142|Die Hard (1988)
ID 1196|Rating 3.797349|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 296 |Rating 3.667643|Pulp Fiction (1994)
ID 608 |Rating 3.621137|Fargo (1996)
ID 3578|Rating 3.554532|Gladiator (2000)
===== ===== ===== ===== =====

User 2088, age 50, M,  has rated 355 movies in training set

ID 1207|Rating 4.869573|To Kill a Mockingbird (1962)
ID 1304|Rating 4.786379|Butch Cassidy and the Sundance Kid (1969)
ID 1303|Rating 4.725040|Man Who Would Be King, The (1975)
ID 2396|Rating 4.690739|Shakespeare in Love (1998)
ID 3469|Rating 4.678252|Inherit the Wind (1960)
ID 1197|Rating 4.664592|Princess Bride, The (1987)
ID 909 |Rating 4.644711|Apartment, The (1960)
ID 1028|Rating 4.613358|Mary Poppins (1964)
ID 599 |Rating 4.559749|Wild Bunch, The (1969)
ID 3062|Rating 4.546902|Longest Day, The (1962)
===== ===== ===== ===== =====

User 5020, age 56, F,  has rated 84 movies in training set

ID 3751|Rating 4.254084|Chicken Run (2000)
ID 3171|Rating 4.187455|Room at the Top (1959)
ID 3723|Rating 4.175472|Hamlet (1990)
ID 2872|Rating 4.106682|Excalibur (1981)
ID 592 |Rating 3.972631|Batman (1989)
ID 1210|Rating 3.960043|Star Wars: Episode VI - Return of the Jedi (1983)
ID 3792|Rating 3.937494|Duel in the Sun (1946)
```

```
ID 329 |Rating 3.890255|Star Trek: Generations (1994)
ID 938 |Rating 3.861818|Gigi (1958)
ID 2470|Rating 3.749801|Crocodile Dundee (1986)
===== ===== ===== ===== =====

User 5583, age 56, M,  has rated 95 movies in training set

ID 2396|Rating 4.670719|Shakespeare in Love (1998)
ID 920 |Rating 4.581555|Gone with the Wind (1939)
ID 1198|Rating 4.536726|Raiders of the Lost Ark (1981)
ID 1266|Rating 4.425125|Unforgiven (1992)
ID 457 |Rating 4.403832|Fugitive, The (1993)
ID 2858|Rating 4.361459|American Beauty (1999)
ID 260 |Rating 4.331812|Star Wars: Episode IV - A New Hope (1977)
ID 1291|Rating 4.310675|Indiana Jones and the Last Crusade (1989)
ID 3406|Rating 4.271886|Captain Horatio Hornblower (1951)
ID 3203|Rating 4.100273|Dead Calm (1989)
===== ===== ===== ===== =====
```

4. With gender, with age

In [62]:

```
def get_result_df(datasets, model, best_weights, bs, save_path=None):
    '''
    model_parameter1_best.weights
    '''
    minmax = (1.0, 5.0)
    device = torch.device('cpu')
    model.load_state_dict(best_weights)
    groud_truth, predictions = [], []

    val_size = len(datasets['val'][0])
    # print("Total val size: {}".format(val_size))

    with torch.no_grad():
        for batch in batches(*datasets['val'], shuffle=False, bs=bs):
            x_batch, y_batch = [b.to(device) for b in batch]
            outputs = model(x_batch[:, 0], x_batch[:, 1], x_batch[:, 2], x_batch[:, 3],minmax)
            groud_truth.extend(y_batch.tolist())
            predictions.extend(outputs.tolist())

        last_num = val_size % bs
        # print("Last num: {}".format(last_num))
        dataset_last = (datasets['val'][0][-last_num:], datasets['val'][1][-last_num:])
        # print("Last dataset: {}".format(len(dataset_last[0])))
        for batch in batches(*dataset_last, shuffle=False, bs=1):
            x_batch, y_batch = [b.to(device) for b in batch]
            outputs = model(x_batch[:, 0], x_batch[:, 1], x_batch[:, 2], x_batch[:, 3],minmax)
            groud_truth.extend(y_batch.tolist())
            predictions.extend(outputs.tolist())

    groud_truth = np.asarray(groud_truth).ravel()
    predictions = np.asarray(predictions).ravel()

    assert(predictions.shape[0] == val_size)

    final_loss = np.sqrt(np.mean((predictions - groud_truth)**2))
    # print("RMSE: {:.4f}".format(final_loss))

    df_final = pd.DataFrame(datasets['val'][0])[['user_id','movie_id']]
    df_final['truth'] = datasets['val'][1]
    df_final['pred'] = predictions

    if save_path is not None:
        print("Save weight to:{}".format(save_path))
        with open(save_path, 'wb') as file:
            pickle.dump(best_weights, file)

    return df_final # note that here the sex and age is not included
```

In [63]:
```
n_factors = 200
hidden = [1000] * 3
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
g_factor = 10
a_factor = 10

# training
lr = 1e-3
wd = 1e-5
bs =2000
n_epochs = 100
patience = 10


model = EmbeddingNetGenderAge(n_users, n_movies,
                        n_factors=n_factors, hidden=hidden, dropouts=dropouts, g_factor=g_factor, a_factor=a_f
actor)
best_weights = pickle.load(open("saved_models/both_para5.weights",'rb'))
```

In [64]:
```
df_final = get_result_df(datasets, model, best_weights, bs, save_path=None)
for user_id in user_id_list:
    get_train_pred_top10(df, df_final, datasets, user_id=user_id, n=10)
    print("===== ===== ===== ===== =====\n")
```

```
User 1, age 1, F,  has rated 41 movies in training set

ID 1207|Rating 4.742472|To Kill a Mockingbird (1962)
ID 608 |Rating 4.574499|Fargo (1996)
ID 260 |Rating 4.477774|Star Wars: Episode IV - A New Hope (1977)
ID 2804|Rating 4.427573|Christmas Story, A (1983)
ID 1   |Rating 4.336008|Toy Story (1995)
ID 595 |Rating 4.332149|Beauty and the Beast (1991)
ID 1287|Rating 4.243710|Ben-Hur (1959)
ID 1961|Rating 4.211578|Rain Man (1988)
ID 2791|Rating 4.161878|Airplane! (1980)
ID 2687|Rating 4.089414|Tarzan (1999)
===== ===== ===== ===== =====

User 19, age 1, M,  has rated 189 movies in training set

ID 2858|Rating 4.862138|American Beauty (1999)
ID 2571|Rating 4.595146|Matrix, The (1999)
ID 1198|Rating 4.542353|Raiders of the Lost Ark (1981)
ID 2028|Rating 4.518578|Saving Private Ryan (1998)
ID 1148|Rating 4.458857|Wrong Trousers, The (1993)
ID 1210|Rating 4.328912|Star Wars: Episode VI - Return of the Jedi (1983)
ID 356 |Rating 4.303192|Forrest Gump (1994)
ID 110 |Rating 4.270610|Braveheart (1995)
ID 2949|Rating 4.199471|Dr. No (1962)
ID 1278|Rating 4.189931|Young Frankenstein (1974)
===== ===== ===== ===== =====

User 176, age 18, F,  has rated 58 movies in training set

ID 1196|Rating 4.566271|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 1307|Rating 4.487173|When Harry Met Sally... (1989)
ID 1197|Rating 4.370369|Princess Bride, The (1987)
ID 2795|Rating 4.265700|Vacation (1983)
ID 3039|Rating 4.134203|Trading Places (1983)
ID 3552|Rating 4.088376|Caddyshack (1980)
ID 1968|Rating 4.066108|Breakfast Club, The (1985)
ID 3715|Rating 3.939243|Burglar (1987)
ID 1544|Rating 3.935678|Lost World: Jurassic Park, The (1997)
ID 2378|Rating 3.927120|Police Academy (1984)
===== ===== ===== ===== =====

User 181, age 18, M,  has rated 242 movies in training set

ID 1213|Rating 5.075797|GoodFellas (1990)
ID 1136|Rating 4.939441|Monty Python and the Holy Grail (1974)
ID 2028|Rating 4.829145|Saving Private Ryan (1998)
ID 3424|Rating 4.555366|Do the Right Thing (1989)
ID 1291|Rating 4.510047|Indiana Jones and the Last Crusade (1989)
ID 1732|Rating 4.479739|Big Lebowski, The (1998)
ID 1704|Rating 4.405294|Good Will Hunting (1997)
ID 1240|Rating 4.388027|Terminator, The (1984)
ID 778 |Rating 4.248252|Trainspotting (1996)
ID 2706|Rating 4.207235|American Pie (1999)
===== ===== ===== ===== =====

User 253, age 25, F,  has rated 41 movies in training set

ID 446 |Rating 4.609703|Farewell My Concubine (1993)
ID 3083|Rating 4.484456|All About My Mother (Todo Sobre Mi Madre) (1999)
ID 246 |Rating 4.368542|Hoop Dreams (1994)
ID 3408|Rating 4.078164|Erin Brockovich (2000)
ID 3751|Rating 3.972930|Chicken Run (2000)
ID 2599|Rating 3.938196|Election (1999)
ID 1533|Rating 3.908612|Promise, The (La Promesse) (1996)
ID 1127|Rating 3.715311|Abyss, The (1989)
ID 296 |Rating 3.711310|Pulp Fiction (1994)
ID 1089|Rating 3.585794|Reservoir Dogs (1992)
===== ===== ===== ===== =====

User 261, age 25, M,  has rated 297 movies in training set

ID 1221|Rating 4.728684|Godfather: Part II, The (1974)
ID 1225|Rating 4.486713|Amadeus (1984)
ID 1283|Rating 4.431026|High Noon (1952)
ID 3424|Rating 4.383136|Do the Right Thing (1989)
ID 3741|Rating 4.287354|Badlands (1973)
ID 2692|Rating 4.268642|Run Lola Run (Lola rennt) (1998)
ID 1732|Rating 4.256298|Big Lebowski, The (1998)
ID 1307|Rating 4.192164|When Harry Met Sally... (1989)
ID 3738|Rating 4.117684|Sugarland Express, The (1974)
ID 2542|Rating 4.092818|Lock, Stock & Two Smoking Barrels (1998)
===== ===== ===== ===== =====

User 704, age 35, F,  has rated 61 movies in training set

ID 3911|Rating 4.029378|Best in Show (2000)
ID 2959|Rating 4.018937|Fight Club (1999)
```

```
ID 3384|Rating 3.996497|Taking of Pelham One Two Three, The (1974)
ID 2761|Rating 3.957870|Iron Giant, The (1999)
ID 2390|Rating 3.736249|Little Voice (1998)
ID 3298|Rating 3.426171|Boiler Room (2000)
ID 2490|Rating 3.385803|Payback (1999)
ID 3896|Rating 3.247116|Way of the Gun, The (2000)
ID 3893|Rating 3.208550|Nurse Betty (2000)
ID 3185|Rating 2.915368|Snow Falling on Cedars (1999)
===== ===== ===== ===== =====

User 749, age 35, M,  has rated 375 movies in training set

ID 912 |Rating 4.309187|Casablanca (1942)
ID 2762|Rating 4.179867|Sixth Sense, The (1999)
ID 1196|Rating 4.139338|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 908 |Rating 4.019060|North by Northwest (1959)
ID 3114|Rating 3.976644|Toy Story 2 (1999)
ID 1204|Rating 3.945578|Lawrence of Arabia (1962)
ID 1090|Rating 3.912392|Platoon (1986)
ID 2791|Rating 3.910148|Airplane! (1980)
ID 1210|Rating 3.895487|Star Wars: Episode VI - Return of the Jedi (1983)
ID 1276|Rating 3.884572|Cool Hand Luke (1967)
===== ===== ===== ===== =====

User 1428, age 45, F,  has rated 41 movies in training set

ID 1206|Rating 5.144929|Clockwork Orange, A (1971)
ID 908 |Rating 4.884665|North by Northwest (1959)
ID 1282|Rating 4.669325|Fantasia (1940)
ID 2291|Rating 4.464828|Edward Scissorhands (1990)
ID 2227|Rating 4.437533|Lodger, The (1926)
ID 175 |Rating 4.327851|Kids (1995)
ID 3927|Rating 3.956764|Fantastic Voyage (1966)
ID 1580|Rating 3.852091|Men in Black (1997)
ID 3450|Rating 3.774438|Grumpy Old Men (1993)
ID 3930|Rating 3.772157|Creature From the Black Lagoon, The (1954)
===== ===== ===== ===== =====

User 1565, age 45, M,  has rated 37 movies in training set

ID 2997|Rating 4.326965|Being John Malkovich (1999)
ID 2396|Rating 4.082603|Shakespeare in Love (1998)
ID 2599|Rating 4.032507|Election (1999)
ID 34  |Rating 3.887970|Babe (1995)
ID 1732|Rating 3.823928|Big Lebowski, The (1998)
ID 357 |Rating 3.812804|Four Weddings and a Funeral (1994)
ID 348 |Rating 3.777414|Bullets Over Broadway (1994)
ID 3129|Rating 3.565782|Sweet and Lowdown (1999)
ID 224 |Rating 3.546489|Don Juan DeMarco (1995)
ID 619 |Rating 3.473640|Ed (1996)
===== ===== ===== ===== =====

User 1961, age 50, F,  has rated 71 movies in training set

ID 858 |Rating 4.472063|Godfather, The (1972)
ID 1198|Rating 4.229097|Raiders of the Lost Ark (1981)
ID 2501|Rating 3.993985|October Sky (1999)
ID 318 |Rating 3.988014|Shawshank Redemption, The (1994)
ID 1196|Rating 3.975174|Star Wars: Episode V - The Empire Strikes Back (1980)
ID 457 |Rating 3.907841|Fugitive, The (1993)
ID 1036|Rating 3.817840|Die Hard (1988)
ID 608 |Rating 3.781312|Fargo (1996)
ID 2966|Rating 3.621969|Straight Story, The (1999)
ID 2951|Rating 3.599778|Fistful of Dollars, A (1964)
===== ===== ===== ===== =====

User 2088, age 50, M,  has rated 355 movies in training set

ID 1222|Rating 4.947099|Full Metal Jacket (1987)
ID 589 |Rating 4.920642|Terminator 2: Judgment Day (1991)
ID 1303|Rating 4.879650|Man Who Would Be King, The (1975)
ID 1207|Rating 4.868109|To Kill a Mockingbird (1962)
ID 2947|Rating 4.852085|Goldfinger (1964)
ID 3737|Rating 4.842061|Lonely Are the Brave (1962)
ID 1304|Rating 4.811199|Butch Cassidy and the Sundance Kid (1969)
ID 541 |Rating 4.784345|Blade Runner (1982)
ID 702 |Rating 4.783013|Faces (1968)
ID 3469|Rating 4.738050|Inherit the Wind (1960)
===== ===== ===== ===== =====

User 5020, age 56, F,  has rated 84 movies in training set

ID 3792|Rating 4.405981|Duel in the Sun (1946)
ID 938 |Rating 4.348991|Gigi (1958)
ID 3723|Rating 4.287194|Hamlet (1990)
ID 3171|Rating 4.264806|Room at the Top (1959)
ID 2872|Rating 4.224791|Excalibur (1981)
ID 3712|Rating 4.182477|Soapdish (1991)
ID 3751|Rating 4.051610|Chicken Run (2000)
```

```
ID 2143|Rating 3.944847|Legend (1985)
ID 1210|Rating 3.910752|Star Wars: Episode VI - Return of the Jedi (1983)
ID 2405|Rating 3.904504|Jewel of the Nile, The (1985)
===== ===== ===== ===== =====

User 5583, age 56, M,  has rated 95 movies in training set

ID 920 |Rating 4.687065|Gone with the Wind (1939)
ID 2396|Rating 4.656053|Shakespeare in Love (1998)
ID 1198|Rating 4.611466|Raiders of the Lost Ark (1981)
ID 1266|Rating 4.485502|Unforgiven (1992)
ID 2571|Rating 4.403478|Matrix, The (1999)
ID 260 |Rating 4.379815|Star Wars: Episode IV - A New Hope (1977)
ID 2858|Rating 4.375957|American Beauty (1999)
ID 1291|Rating 4.359865|Indiana Jones and the Last Crusade (1989)
ID 553 |Rating 4.277444|Tombstone (1993)
ID 3176|Rating 4.163035|Talented Mr. Ripley, The (1999)
===== ===== ===== ===== =====
```