ML-1M DL +gender +age

In [8]:
```
# import io
# import os
import math
import copy
import pickle
# import zipfile
# from textwrap import wrap
from pathlib import Path
from itertools import zip_longest
from collections import defaultdict
# from urllib.error import URLError
# from urllib.request import urlopen

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

%matplotlib inline
```

In [9]:
```
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)
```

In [10]:
```
# load preprocessed df
df = pd.read_csv("ml-1m_dl.csv")
print(df.shape)
df.head()
```

```
(1000209, 9)
```

Out[10]:

	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 [11]:
```
# load dataset
datasets = pickle.load(open('ml-1m_dl.pkl','rb'))
datasets['val'][1]
```

Out[11]:
```
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
```

In [12]:
```
n_users = 6040
n_movies = 3706
dataset_sizes = {'train': 800167, 'val': 200042}
```

Define the network

In [13]:
```
# consider both gender and age
class EmbeddingNetGenderAge(nn.Module):
    def __init__(self, n_users, n_movies, n_factors=50, embedding_dropout=0.02, hidden=10, dropouts=0.2, g_fa
ctor=25, a_factor=25):
        super().__init__()
        hidden = get_list(hidden)
        dropouts = get_list(dropouts)
        n_last = hidden[-1]
        def gen_layers(n_in):
            """
            A generator that yields a sequence of hidden layers and
            their activations/dropouts.

            Note that the function captures `hidden` and `dropouts`
            values from the outer scope.
            """
            nonlocal hidden, dropouts
            assert len(dropouts) <= len(hidden)
            for n_out, rate in zip_longest(hidden, dropouts):
                yield nn.Linear(n_in, n_out)
                yield nn.ReLU()
                if rate is not None and rate > 0.:
                    yield nn.Dropout(rate)
                n_in = n_out


        self.u = nn.Embedding(n_users+1 , n_factors) # hard code
        self.m = nn.Embedding(4000, n_factors)  # hardcode

        self.g_factor = g_factor
        self.g = nn.Embedding(2, self.g_factor)
        self.a_factor = a_factor
        self.a = nn.Embedding(7, self.a_factor)


        self.drop = nn.Dropout(embedding_dropout)
        self.hidden = nn.Sequential(*list(gen_layers(n_factors * 2 + self.g_factor + self.a_factor)))
        self.fc = nn.Linear(n_last, 1)
        self._init()

    def forward(self, users, movies, genders, ages, minmax=None):
        uu = self.u(users)
        gg = self.g(genders)
        aa = self.a(ages)
        mm = self.m(movies)
        features = torch.cat([uu, gg, aa, mm], dim=1)
        x = self.drop(features)
        x = self.hidden(x)
        out = torch.sigmoid(self.fc(x))
        if minmax is not None:
            min_rating, max_rating = minmax
            out = out*(max_rating - min_rating + 1) + min_rating - 0.5
        return out

    def _init(self):
        def init(m):
            if type(m) == nn.Linear:
                torch.nn.init.xavier_uniform_(m.weight)
                m.bias.data.fill_(0.01)
        self.u.weight.data.uniform_(-0.05, 0.05)
        self.m.weight.data.uniform_(-0.05, 0.05)
        self.hidden.apply(init)
        init(self.fc)

def get_list(n):
    if isinstance(n, (int, float)):
        return [n]
    elif hasattr(n, '__iter__'):
        return list(n)
    raise TypeError('layers configuraiton should be a single number or a list of numbers')
```

In [14]:

```
# test
testnet = EmbeddingNetGenderAge(n_users, n_movies, n_factors=150, hidden=100, dropouts=0.5, g_factor=25, a_fa
ctor=25)
print(testnet)
testnet = EmbeddingNetGenderAge(n_users, n_movies, n_factors=150, hidden=[100, 200, 300], dropouts=[0.25, 0.5
], g_factor=25, a_factor=25)
print(testnet)
```

```
EmbeddingNetGenderAge(
  (u): Embedding(6041, 150)
  (m): Embedding(4000, 150)
  (g): Embedding(2, 25)
  (a): Embedding(7, 25)
  (drop): Dropout(p=0.02, inplace=False)
  (hidden): Sequential(
    (0): Linear(in_features=350, out_features=100, bias=True)
    (1): ReLU()
    (2): Dropout(p=0.5, inplace=False)
  )
  (fc): Linear(in_features=100, out_features=1, bias=True)
)
EmbeddingNetGenderAge(
  (u): Embedding(6041, 150)
  (m): Embedding(4000, 150)
  (g): Embedding(2, 25)
  (a): Embedding(7, 25)
  (drop): Dropout(p=0.02, inplace=False)
  (hidden): Sequential(
    (0): Linear(in_features=350, out_features=100, bias=True)
    (1): ReLU()
    (2): Dropout(p=0.25, inplace=False)
    (3): Linear(in_features=100, out_features=200, bias=True)
    (4): ReLU()
    (5): Dropout(p=0.5, inplace=False)
    (6): Linear(in_features=200, out_features=300, bias=True)
    (7): ReLU()
  )
  (fc): Linear(in_features=300, out_features=1, bias=True)
)
```

In [15]:

```
# 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 [16]:
```
class CyclicLR(_LRScheduler):
    def __init__(self, optimizer, schedule, last_epoch=-1):
        assert callable(schedule)
        self.schedule = schedule
        super().__init__(optimizer, last_epoch)
    def get_lr(self):
        return [self.schedule(self.last_epoch, lr) for lr in self.base_lrs]

def triangular(step_size, max_lr, method='triangular', gamma=0.99):

    def scheduler(epoch, base_lr):
        period = 2 * step_size
        cycle = math.floor(1 + epoch/period)
        x = abs(epoch/step_size - 2*cycle + 1)
        delta = (max_lr - base_lr)*max(0, (1 - x))

        if method == 'triangular':
            pass  # we've already done
        elif method == 'triangular2':
            delta /= float(2 ** (cycle - 1))
        elif method == 'exp_range':
            delta *= (gamma**epoch)
        else:
            raise ValueError('unexpected method: %s' % method)

        return base_lr + delta

    return scheduler

def cosine(t_max, eta_min=0):

    def scheduler(epoch, base_lr):
        t = epoch % t_max
        return eta_min + (base_lr - eta_min)*(1 + math.cos(math.pi*t/t_max))/2

    return scheduler

def plot_lr(schedule, label):
    ts = list(range(1000))
    y = [schedule(t, 0.001) for t in ts]
    plt.plot(ts, y, label=label)
```

In [17]:
```
def train_model(datasets, model, lr, wd, bs, n_epochs, patience):
    minmax = (1.0, 5.0)
    device = torch.device('cpu')

    # Training
    no_improvements = 0
    best_loss = np.inf
    best_weights = None
    history = []
    lr_history = []

    start_time = time()
    model.to(device)
    criterion = nn.MSELoss(reduction='sum')
    optimizer = optim.Adam(model.parameters(), lr=lr, weight_decay=wd)
    iterations_per_epoch = int(math.ceil(dataset_sizes['train'] // bs))
    scheduler = CyclicLR(optimizer, cosine(t_max=iterations_per_epoch * 2, eta_min=lr/10))

    start_time = time()

    for epoch in range(n_epochs):
        stats = {'epoch': epoch + 1, 'total': n_epochs}

        for phase in ('train', 'val'):
            training = phase == 'train'
            running_loss = 0.0
            n_batches = 0

            for batch in batches(*datasets[phase], shuffle=training, bs=bs):
                x_batch, y_batch = [b.to(device) for b in batch] # [2000, 4], [2000, 1]
                optimizer.zero_grad()
                # compute gradients only during 'train' phase
                with torch.set_grad_enabled(training):
                    outputs = model(x_batch[:, 0], x_batch[:, 1], x_batch[:, 2], x_batch[:, 3], minmax)
                    loss = criterion(outputs, y_batch)
                    # don't update weights and rates when in 'val' phase
                    if training:
                        loss.backward()
                        optimizer.step()
                        scheduler.step()
                        lr_history.extend(scheduler.get_lr())
                running_loss += loss.item()

            epoch_loss = running_loss / dataset_sizes[phase]
            stats[phase] = epoch_loss

            # early stopping: save weights of the best model so far
            if phase == 'val':
                if epoch_loss < best_loss:
                    print('loss improvement on epoch: %d' % (epoch + 1))
                    best_loss = epoch_loss
                    best_weights = copy.deepcopy(model.state_dict())
                    no_improvements = 0
                else:
                    no_improvements += 1

        history.append(stats)
        cost_time = (time() - start_time) / 60.
        print('[{:03d}/{:03d}]|train {:.4f}|val {:.4f}|Time {:.2f}mins'.format(
                                                                        stats['epoch'], stats['total'],
                                                                        stats['train'], stats['val'], cost_ti
me))
        if no_improvements >= patience:
            print('early stopping after epoch {:03d}'.format(stats['epoch']))
            break
    return best_weights
```

In [18]:
```
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 [19]:
```
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

# i = 0
# top_n = get_top_n(df_final, n=10)
# for user_id, pred_ratings in top_n.items():
#     print("User id: {}".format(user_id))
#     for (movie_id, rating) in pred_ratings:
#         print("Movie {:<5d}|Rating {:.2f}".format(movie_id, rating))
#     print("----- -----")
#     i += 1
#     if i > 0:
#         break

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)
```

Now let's begin

In [20]:

```
# para 1
n_factors = 200
hidden = [1000] * 3
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
g_factor = 15
a_factor = 15

# 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 = train_model(datasets, model, lr, wd, bs, n_epochs, patience)
df_final = get_result_df(datasets, model, best_weights, bs, save_path='both_para1.weights')
get_precision_recall(df_final, k=10, threshold=3.5)
```

```
loss improvement on epoch: 1
[001/100]|train 0.9325|val 0.8179|Time 0.90mins
loss improvement on epoch: 2
[002/100]|train 0.7942|val 0.7996|Time 1.72mins
loss improvement on epoch: 3
[003/100]|train 0.7953|val 0.7905|Time 2.69mins
loss improvement on epoch: 4
[004/100]|train 0.7543|val 0.7768|Time 3.56mins
loss improvement on epoch: 5
[005/100]|train 0.7583|val 0.7662|Time 4.43mins
loss improvement on epoch: 6
[006/100]|train 0.7156|val 0.7589|Time 5.31mins
[007/100]|train 0.7277|val 0.7600|Time 6.18mins
loss improvement on epoch: 8
[008/100]|train 0.6902|val 0.7553|Time 7.05mins
[009/100]|train 0.7066|val 0.7567|Time 7.93mins
[010/100]|train 0.6717|val 0.7577|Time 8.84mins
[011/100]|train 0.6897|val 0.7568|Time 9.71mins
[012/100]|train 0.6549|val 0.7586|Time 10.55mins
[013/100]|train 0.6744|val 0.7600|Time 11.43mins
[014/100]|train 0.6391|val 0.7627|Time 12.34mins
[015/100]|train 0.6582|val 0.7616|Time 13.28mins
[016/100]|train 0.6217|val 0.7679|Time 14.21mins
[017/100]|train 0.6424|val 0.7659|Time 15.12mins
[018/100]|train 0.6036|val 0.7762|Time 16.03mins
early stopping after epoch 018
RMSE: 0.8691
Save weight to:both_para1.weights
Prec10 0.8023|Rec10 0.5653
```

In [21]:
```
# para 2
n_factors = 200
hidden = [1000] * 3
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
g_factor = 15
a_factor = 25

# 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 = train_model(datasets, model, lr, wd, bs, n_epochs, patience)
df_final = get_result_df(datasets, model, best_weights, bs, save_path='both_para2.weights')
get_precision_recall(df_final, k=10, threshold=3.5)
```

```
loss improvement on epoch: 1
[001/100]|train 0.9517|val 0.8221|Time 0.89mins
loss improvement on epoch: 2
[002/100]|train 0.7968|val 0.8014|Time 1.77mins
loss improvement on epoch: 3
[003/100]|train 0.7984|val 0.7946|Time 2.66mins
loss improvement on epoch: 4
[004/100]|train 0.7611|val 0.7827|Time 3.54mins
loss improvement on epoch: 5
[005/100]|train 0.7645|val 0.7713|Time 4.44mins
loss improvement on epoch: 6
[006/100]|train 0.7227|val 0.7626|Time 5.28mins
loss improvement on epoch: 7
[007/100]|train 0.7336|val 0.7594|Time 6.13mins
loss improvement on epoch: 8
[008/100]|train 0.6969|val 0.7560|Time 6.97mins
loss improvement on epoch: 9
[009/100]|train 0.7125|val 0.7555|Time 7.80mins
[010/100]|train 0.6786|val 0.7557|Time 8.67mins
[011/100]|train 0.6965|val 0.7569|Time 9.56mins
[012/100]|train 0.6636|val 0.7592|Time 10.40mins
[013/100]|train 0.6823|val 0.7564|Time 11.24mins
[014/100]|train 0.6493|val 0.7642|Time 12.06mins
[015/100]|train 0.6686|val 0.7641|Time 12.90mins
[016/100]|train 0.6351|val 0.7674|Time 13.77mins
[017/100]|train 0.6544|val 0.7650|Time 14.61mins
[018/100]|train 0.6194|val 0.7744|Time 15.44mins
[019/100]|train 0.6399|val 0.7667|Time 16.29mins
early stopping after epoch 019
RMSE: 0.8688
Save weight to:both_para2.weights
Prec10 0.8019|Rec10 0.5682
```

In [22]:
```
# para 3
n_factors = 200
hidden = [1000] * 3
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
g_factor = 25
a_factor = 15

# 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 = train_model(datasets, model, lr, wd, bs, n_epochs, patience)
df_final = get_result_df(datasets, model, best_weights, bs, save_path='both_para3.weights')
get_precision_recall(df_final, k=10, threshold=3.5)
```

```
loss improvement on epoch: 1
[001/100]|train 0.9439|val 0.8194|Time 0.82mins
loss improvement on epoch: 2
[002/100]|train 0.7955|val 0.8011|Time 1.71mins
loss improvement on epoch: 3
[003/100]|train 0.7975|val 0.7943|Time 2.58mins
loss improvement on epoch: 4
[004/100]|train 0.7576|val 0.7805|Time 3.45mins
loss improvement on epoch: 5
[005/100]|train 0.7627|val 0.7708|Time 4.30mins
loss improvement on epoch: 6
[006/100]|train 0.7223|val 0.7607|Time 5.20mins
loss improvement on epoch: 7
[007/100]|train 0.7332|val 0.7580|Time 6.07mins
loss improvement on epoch: 8
[008/100]|train 0.6973|val 0.7566|Time 6.94mins
loss improvement on epoch: 9
[009/100]|train 0.7130|val 0.7559|Time 7.80mins
loss improvement on epoch: 10
[010/100]|train 0.6792|val 0.7539|Time 8.66mins
[011/100]|train 0.6967|val 0.7593|Time 9.56mins
[012/100]|train 0.6647|val 0.7575|Time 10.43mins
[013/100]|train 0.6830|val 0.7568|Time 11.32mins
[014/100]|train 0.6495|val 0.7607|Time 12.22mins
[015/100]|train 0.6686|val 0.7609|Time 13.07mins
[016/100]|train 0.6346|val 0.7672|Time 13.89mins
[017/100]|train 0.6538|val 0.7679|Time 14.76mins
[018/100]|train 0.6185|val 0.7725|Time 15.66mins
[019/100]|train 0.6388|val 0.7661|Time 16.59mins
[020/100]|train 0.6017|val 0.7776|Time 17.50mins
early stopping after epoch 020
RMSE: 0.8685
Save weight to:both_para3.weights
Prec10 0.8015|Rec10 0.5648
```

In [23]:
```
# para 4
n_factors = 200
hidden = [1000] * 3
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
g_factor = 25
a_factor = 25

# 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 = train_model(datasets, model, lr, wd, bs, n_epochs, patience)
df_final = get_result_df(datasets, model, best_weights, bs, save_path='both_para4.weights')
get_precision_recall(df_final, k=10, threshold=3.5)
```

```
loss improvement on epoch: 1
[001/100]|train 0.9572|val 0.8214|Time 0.80mins
loss improvement on epoch: 2
[002/100]|train 0.7969|val 0.8012|Time 1.60mins
loss improvement on epoch: 3
[003/100]|train 0.7987|val 0.7962|Time 2.45mins
loss improvement on epoch: 4
[004/100]|train 0.7630|val 0.7844|Time 3.24mins
loss improvement on epoch: 5
[005/100]|train 0.7684|val 0.7751|Time 4.11mins
loss improvement on epoch: 6
[006/100]|train 0.7274|val 0.7643|Time 4.90mins
loss improvement on epoch: 7
[007/100]|train 0.7375|val 0.7620|Time 5.69mins
loss improvement on epoch: 8
[008/100]|train 0.7011|val 0.7573|Time 6.57mins
loss improvement on epoch: 9
[009/100]|train 0.7158|val 0.7566|Time 7.34mins
loss improvement on epoch: 10
[010/100]|train 0.6819|val 0.7565|Time 8.25mins
[011/100]|train 0.6987|val 0.7568|Time 9.13mins
[012/100]|train 0.6660|val 0.7604|Time 9.98mins
[013/100]|train 0.6838|val 0.7586|Time 10.82mins
[014/100]|train 0.6508|val 0.7625|Time 11.65mins
[015/100]|train 0.6697|val 0.7597|Time 12.49mins
[016/100]|train 0.6369|val 0.7668|Time 13.35mins
[017/100]|train 0.6561|val 0.7620|Time 14.23mins
[018/100]|train 0.6224|val 0.7741|Time 15.14mins
[019/100]|train 0.6422|val 0.7703|Time 16.00mins
[020/100]|train 0.6069|val 0.7782|Time 16.85mins
early stopping after epoch 020
RMSE: 0.8699
Save weight to:both_para4.weights
Prec10 0.8009|Rec10 0.5640
```

In [24]:
```
# para 5
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 = train_model(datasets, model, lr, wd, bs, n_epochs, patience)
df_final = get_result_df(datasets, model, best_weights, bs, save_path='both_para5.weights')
get_precision_recall(df_final, k=10, threshold=3.5)
```

```
loss improvement on epoch: 1
[001/100]|train 0.9196|val 0.8178|Time 0.97mins
loss improvement on epoch: 2
[002/100]|train 0.7953|val 0.7998|Time 1.88mins
loss improvement on epoch: 3
[003/100]|train 0.7963|val 0.7923|Time 2.58mins
loss improvement on epoch: 4
[004/100]|train 0.7543|val 0.7776|Time 3.32mins
loss improvement on epoch: 5
[005/100]|train 0.7571|val 0.7654|Time 4.12mins
loss improvement on epoch: 6
[006/100]|train 0.7120|val 0.7581|Time 4.90mins
loss improvement on epoch: 7
[007/100]|train 0.7245|val 0.7580|Time 5.63mins
loss improvement on epoch: 8
[008/100]|train 0.6855|val 0.7562|Time 6.34mins
[009/100]|train 0.7022|val 0.7583|Time 7.11mins
[010/100]|train 0.6650|val 0.7586|Time 7.88mins
loss improvement on epoch: 11
[011/100]|train 0.6835|val 0.7549|Time 8.66mins
[012/100]|train 0.6458|val 0.7643|Time 9.44mins
[013/100]|train 0.6662|val 0.7571|Time 10.21mins
[014/100]|train 0.6269|val 0.7699|Time 10.94mins
[015/100]|train 0.6470|val 0.7617|Time 11.72mins
[016/100]|train 0.6066|val 0.7758|Time 12.46mins
[017/100]|train 0.6271|val 0.7670|Time 13.22mins
[018/100]|train 0.5842|val 0.7843|Time 13.92mins
[019/100]|train 0.6070|val 0.7830|Time 14.71mins
[020/100]|train 0.5634|val 0.7929|Time 15.45mins
[021/100]|train 0.5863|val 0.7877|Time 16.19mins
early stopping after epoch 021
RMSE: 0.8688
Save weight to:both_para5.weights
Prec10 0.8007|Rec10 0.5698
```

In [25]:
```
# para 6
n_factors = 200
hidden = [1000] * 3
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
g_factor = 10
a_factor = 15

# 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 = train_model(datasets, model, lr, wd, bs, n_epochs, patience)
df_final = get_result_df(datasets, model, best_weights, bs, save_path='both_para6.weights')
get_precision_recall(df_final, k=10, threshold=3.5)
```

```
loss improvement on epoch: 1
[001/100]|train 0.9304|val 0.8192|Time 0.87mins
loss improvement on epoch: 2
[002/100]|train 0.7947|val 0.7991|Time 1.63mins
loss improvement on epoch: 3
[003/100]|train 0.7956|val 0.7928|Time 2.35mins
loss improvement on epoch: 4
[004/100]|train 0.7565|val 0.7793|Time 3.06mins
loss improvement on epoch: 5
[005/100]|train 0.7595|val 0.7687|Time 3.78mins
loss improvement on epoch: 6
[006/100]|train 0.7169|val 0.7585|Time 4.55mins
[007/100]|train 0.7279|val 0.7596|Time 5.30mins
loss improvement on epoch: 8
[008/100]|train 0.6898|val 0.7558|Time 6.03mins
[009/100]|train 0.7068|val 0.7575|Time 6.83mins
loss improvement on epoch: 10
[010/100]|train 0.6706|val 0.7551|Time 7.58mins
[011/100]|train 0.6894|val 0.7592|Time 8.34mins
[012/100]|train 0.6547|val 0.7612|Time 9.07mins
[013/100]|train 0.6736|val 0.7583|Time 9.83mins
[014/100]|train 0.6375|val 0.7629|Time 10.64mins
[015/100]|train 0.6577|val 0.7605|Time 11.40mins
[016/100]|train 0.6198|val 0.7721|Time 12.10mins
[017/100]|train 0.6401|val 0.7632|Time 12.84mins
[018/100]|train 0.6014|val 0.7751|Time 13.56mins
[019/100]|train 0.6229|val 0.7680|Time 14.28mins
[020/100]|train 0.5830|val 0.7849|Time 15.02mins
early stopping after epoch 020
RMSE: 0.8690
Save weight to:both_para6.weights
Prec10 0.8017|Rec10 0.5666
```

In [26]:

```
# para 7
n_factors = 200
hidden = [1000] * 3
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
g_factor = 15
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 = train_model(datasets, model, lr, wd, bs, n_epochs, patience)
df_final = get_result_df(datasets, model, best_weights, bs, save_path='both_para7.weights')
get_precision_recall(df_final, k=10, threshold=3.5)
```

```
loss improvement on epoch: 1
[001/100]|train 0.9257|val 0.8178|Time 0.74mins
loss improvement on epoch: 2
[002/100]|train 0.7935|val 0.7992|Time 1.52mins
loss improvement on epoch: 3
[003/100]|train 0.7953|val 0.7939|Time 2.27mins
loss improvement on epoch: 4
[004/100]|train 0.7539|val 0.7790|Time 3.02mins
loss improvement on epoch: 5
[005/100]|train 0.7583|val 0.7680|Time 3.81mins
loss improvement on epoch: 6
[006/100]|train 0.7145|val 0.7602|Time 4.57mins
loss improvement on epoch: 7
[007/100]|train 0.7272|val 0.7595|Time 5.34mins
loss improvement on epoch: 8
[008/100]|train 0.6888|val 0.7569|Time 6.08mins
loss improvement on epoch: 9
[009/100]|train 0.7055|val 0.7547|Time 6.80mins
[010/100]|train 0.6690|val 0.7580|Time 7.52mins
[011/100]|train 0.6867|val 0.7584|Time 8.28mins
[012/100]|train 0.6511|val 0.7608|Time 9.05mins
[013/100]|train 0.6710|val 0.7614|Time 9.85mins
[014/100]|train 0.6341|val 0.7655|Time 10.64mins
[015/100]|train 0.6547|val 0.7635|Time 11.41mins
[016/100]|train 0.6161|val 0.7735|Time 12.19mins
[017/100]|train 0.6374|val 0.7637|Time 12.94mins
[018/100]|train 0.5972|val 0.7807|Time 13.69mins
[019/100]|train 0.6192|val 0.7733|Time 14.43mins
early stopping after epoch 019
RMSE: 0.8684
Save weight to:both_para7.weights
Prec10 0.8040|Rec10 0.5676
```

In [27]:
```
# para 8
n_factors = 200
hidden = [1000] * 3
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
g_factor = 10
a_factor = 25

# 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 = train_model(datasets, model, lr, wd, bs, n_epochs, patience)
df_final = get_result_df(datasets, model, best_weights, bs, save_path='both_para8.weights')
get_precision_recall(df_final, k=10, threshold=3.5)
```

```
loss improvement on epoch: 1
[001/100]|train 0.9441|val 0.8188|Time 0.87mins
loss improvement on epoch: 2
[002/100]|train 0.7955|val 0.8007|Time 1.69mins
loss improvement on epoch: 3
[003/100]|train 0.7981|val 0.7977|Time 2.58mins
loss improvement on epoch: 4
[004/100]|train 0.7606|val 0.7836|Time 3.46mins
loss improvement on epoch: 5
[005/100]|train 0.7653|val 0.7731|Time 4.22mins
loss improvement on epoch: 6
[006/100]|train 0.7251|val 0.7641|Time 5.04mins
loss improvement on epoch: 7
[007/100]|train 0.7354|val 0.7635|Time 5.89mins
loss improvement on epoch: 8
[008/100]|train 0.6976|val 0.7562|Time 6.68mins
[009/100]|train 0.7129|val 0.7573|Time 7.51mins
[010/100]|train 0.6783|val 0.7576|Time 8.42mins
loss improvement on epoch: 11
[011/100]|train 0.6958|val 0.7546|Time 9.28mins
[012/100]|train 0.6623|val 0.7587|Time 10.04mins
[013/100]|train 0.6812|val 0.7594|Time 10.77mins
[014/100]|train 0.6483|val 0.7614|Time 11.51mins
[015/100]|train 0.6672|val 0.7613|Time 12.34mins
[016/100]|train 0.6340|val 0.7668|Time 13.14mins
[017/100]|train 0.6534|val 0.7623|Time 13.92mins
[018/100]|train 0.6179|val 0.7698|Time 14.68mins
[019/100]|train 0.6383|val 0.7674|Time 15.41mins
[020/100]|train 0.6020|val 0.7782|Time 16.15mins
[021/100]|train 0.6233|val 0.7778|Time 16.93mins
early stopping after epoch 021
RMSE: 0.8682
Save weight to:both_para8.weights
Prec10 0.8027|Rec10 0.5667
```

In [28]:

```
# para 9
n_factors = 200
hidden = [1000] * 3
embedding_dropout = 0.05
dropouts = [0.5,0.5,0.25]
g_factor = 25
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 = train_model(datasets, model, lr, wd, bs, n_epochs, patience)
df_final = get_result_df(datasets, model, best_weights, bs, save_path='both_para9.weights')
get_precision_recall(df_final, k=10, threshold=3.5)
```

```
loss improvement on epoch: 1
[001/100]|train 0.9375|val 0.8201|Time 0.75mins
loss improvement on epoch: 2
[002/100]|train 0.7961|val 0.8024|Time 1.51mins
loss improvement on epoch: 3
[003/100]|train 0.7987|val 0.7950|Time 2.31mins
loss improvement on epoch: 4
[004/100]|train 0.7595|val 0.7805|Time 3.07mins
loss improvement on epoch: 5
[005/100]|train 0.7628|val 0.7697|Time 3.84mins
loss improvement on epoch: 6
[006/100]|train 0.7207|val 0.7602|Time 4.63mins
loss improvement on epoch: 7
[007/100]|train 0.7320|val 0.7575|Time 5.42mins
loss improvement on epoch: 8
[008/100]|train 0.6955|val 0.7551|Time 6.19mins
[009/100]|train 0.7116|val 0.7552|Time 6.99mins
[010/100]|train 0.6773|val 0.7573|Time 7.79mins
loss improvement on epoch: 11
[011/100]|train 0.6955|val 0.7549|Time 8.55mins
[012/100]|train 0.6616|val 0.7600|Time 9.31mins
[013/100]|train 0.6803|val 0.7573|Time 10.07mins
[014/100]|train 0.6458|val 0.7601|Time 10.87mins
[015/100]|train 0.6651|val 0.7609|Time 11.63mins
[016/100]|train 0.6298|val 0.7646|Time 12.39mins
[017/100]|train 0.6498|val 0.7614|Time 13.21mins
[018/100]|train 0.6127|val 0.7731|Time 14.02mins
[019/100]|train 0.6332|val 0.7678|Time 14.82mins
[020/100]|train 0.5951|val 0.7823|Time 15.63mins
[021/100]|train 0.6164|val 0.7803|Time 16.41mins
early stopping after epoch 021
RMSE: 0.8689
Save weight to:both_para9.weights
Prec10 0.8006|Rec10 0.5695
```

In [ ]: