

HbaseMovieRatings

December 29, 2017

1 Import Connection from starbase and define the connection

This is the IP and port of our HBase Rest server.

```
In [3]: from starbase import Connection
        c = Connection("127.0.0.1", "8000")
```

2 Define the ratings table

```
In [6]: ratings = c.table('ratings')
        if(ratings.exists()):
            print("Dropping existing ratings from table\n")
            ratings.drop()

        ratings.create('rating')
```

Dropping existing ratings from table

```
Out[6]: 201
```

3 Parse the data file

```
In [8]: print("Parsing the ml-100k ratings data...\n")
        ratingFile = open("ml-100k/u.data", "r")

        batch = ratings.batch()
        for line in ratingFile:
            (userID, movieID, rating, timestamp) = line.split()
            batch.update(userID, {'rating': {movieID: rating}})

        ratingFile.close()
```

Parsing the ml-100k ratings data...

4 Commit the ratings to HBase through the connected server

```
In [9]: print("Committing ratings data to HBase via REST service\n")
        batch.commit(finalize=True)
```

Committing ratings data to HBase via REST service

```
Out[9]: {'method': 'POST', 'response': [200], 'url': 'ratings/MTk2'}
```

5 Retrieve the data we wrote

```
In [ ]: print("Get back ratings for some users...\n")
        print("Ratings for user ID 1:\n")
        print(ratings.fetch("1"))
        print()
        print("Ratings for user ID 12:\n")
        print(ratings.fetch("12"))
```

Get back ratings for some users...

Ratings for user ID 1:

```
{'rating': {'1': '5', '10': '3', '100': '5', '101': '2', '102': '2', '103': '1', '104': '1', '105': '2', '106': '4',
'107': '4', '108': '5', '109': '5', '11': '2', '110': '1', '111': '5', '112': '1', '113': '5', '114': '5', '115': '5',
'116': '3', '117': '3', '118': '3', '119': '5', '12': '5', '120': '1', '121': '4', '122': '3', '123': '4', '124': '5',
'125': '3', '126': '2', '127': '5', '128': '4', '129': '5', '13': '5', '130': '3', '131': '1', '132': '4', '133': '4',
'134': '4', '135': '4', '136': '3', '137': '5', '138': '1', '139': '3', '14': '5', '140': '1', '141': '3', '142': '2',
'143': '1', '144': '4', '145': '2', '146': '4', '147': '3', '148': '2', '149': '2', '15': '5', '150': '5', '151': '4',
'152': '5', '153': '3', '154': '5', '155': '2', '156': '4', '157': '4', '158': '3', '159': '3', '16': '5', '160': '4',
'161': '4', '162': '4', '163': '4', '164': '3', '165': '5', '166': '5', '167': '2', '168': '5', '169': '5', '17': '3',
'170': '5', '171': '5', '172': '5', '173': '5', '174': '5', '175': '5', '176': '5', '177': '5', '178': '5', '179': '3',
'18': '4', '180': '3', '181': '5', '182': '4', '183': '5', '184': '4', '185': '4', '186': '4', '187': '4', '188': '3',
'189': '3', '19': '5', '190': '5', '191': '5', '192': '4', '193': '4', '194': '4', '195': '5', '196': '5', '197': '5',
'198': '5', '199': '4', '2': '3', '20': '4', '200': '3', '201': '3', '202': '5', '203': '4', '204': '5', '205': '3', '206':
'4', '207': '5', '208': '5', '209': '4', '21': '1', '210': '4', '211': '3', '212': '4', '213': '2', '214': '4', '215': '3',
'216': '5', '217': '3', '218': '3', '219': '1', '22': '4', '220': '3', '221': '5', '222': '4', '223': '5', '224': '5',
'225': '2', '226': '3', '227': '4', '228': '5', '229': '4', '23': '4', '230': '4', '231': '1', '232': '3', '233': '2',
'234': '4', '235': '5', '236': '4', '237': '2', '238': '4', '239': '4', '24': '3', '240': '3', '241': '4', '242': '5',
'243': '1', '244': '2', '245': '2', '246': '5', '247': '1', '248': '4', '249': '4', '25': '4', '250': '4', '251': '4',
'252': '2', '253': '5', '254': '1', '255': '2', '256': '4', '257': '4', '258': '5', '259': '1', '26': '3', '260': '1',
'261': '1', '262': '3', '263': '1', '264': '2', '265': '4', '266': '1', '267': '4', '268': '5', '269': '5', '27': '2',
'270': '5', '271': '2', '272': '3', '28': '4', '29': '1', '3': '4', '30': '3', '31': '3', '32': '5', '33': '4', '34': '2', '35':
'1', '36': '2', '37': '2', '38': '3', '39': '4', '4': '3', '40': '3', '41': '2', '42': '5', '43': '4', '44': '5', '45': '5',
'46': '4', '47': '4', '48': '5', '49': '3', '5': '3', '50': '5', '51': '4', '52': '4', '53': '3', '54': '3', '55': '5', '56':
'4', '57': '5', '58': '4', '59': '5', '6': '5', '60': '5', '61': '4', '62': '3', '63': '2', '64': '5', '65': '4', '66': '4',
'67': '3', '68': '4', '69': '3', '7': '4', '70': '3', '71': '3', '72': '4', '73': '3', '74': '1', '75': '4', '76': '4', '77':
'4', '78': '1', '79': '4', '8': '1', '80': '4', '81': '5', '82': '5', '83': '3', '84': '4', '85': '3', '86': '5', '87': '5',
```

'88': '4', '89': '5', '9': '5', '90': '4', '91': '5', '92': '3', '93': '5', '94': '2', '95': '4', '96': '5', '97': '3', '98': '4', '99': '3'}}

Ratings for user ID 12:

{'rating': {'127': '4', '132': '5', '133': '4', '143': '5', '15': '5', '157': '5', '159': '4', '161': '5', '168': '4', '170': '4', '172': '4', '174': '5', '191': '5', '195': '4', '196': '5', '200': '1', '202': '4', '203': '3', '204': '5', '215': '4', '216': '5', '228': '4', '238': '5', '242': '5', '276': '4', '28': '5', '282': '5', '300': '4', '318': '5', '328': '4', '381': '4', '392': '4', '4': '5', '402': '5', '416': '3', '471': '5', '480': '4', '50': '4', '591': '5', '684': '5', '69': '5', '708': '3', '71': '4', '735': '5', '753': '5', '754': '4', '82': '4', '88': '5', '96': '4', '97': '5', '98': '5'}}