

LAMPIRAN

Lampiran 1

Data Rating.csv

userId,movieId,rating,timestamp
1,2,3,5,1112486027
3,1,29,3,5,1112486476
4,1,32,3,5,1112484819
5,1,47,3,5,1112484727
6,1,50,3,5,1112484580
7,1,112,3,5,1094785740
8,1,151,4,0,1094785734
9,1,223,4,0,1112485573
10,1,253,4,0,1112484940
11,1,260,4,0,1112484826
12,1,293,4,0,1112484703
13,1,296,4,0,1112484767
14,1,318,4,0,1112484798
15,1,337,3,5,1094785709
16,1,367,3,5,1112485980
17,1,541,4,0,1112484603
18,1,589,3,5,1112485557
19,1,593,3,5,1112484661
20,1,653,3,0,1094785691
21,1,919,3,5,1094785621
22,1,924,3,5,1094785598
23,1,1099,3,5,1112486013

Lampiran 2

Data Movies.csv

movieId,title,genres
1,movieId,title,genres
2,1,Toy Story (1995),Adventure Animation Children Comedy Fantasy
3,2,Jumanji (1995),Adventure Children Fantasy
4,3,Grumpier Old Men (1995),Comedy Romance
5,4,Waiting to Exhale (1995),Comedy Drama Romance
6,5,Father of the Bride Part II (1995),Comedy
7,6,Heat (1995),Action Crime Thriller
8,7,Sabrina (1995),Comedy Romance
9,8,Tom and Huck (1995),Adventure Children
10,9,Sudden Death (1995),Action
11,10,GoldenEye (1995),Action Adventure Thriller
12,11,"American President, The (1995)",Comedy Drama Romance
13,12,Dracula: Dead and Loving It (1995),Comedy Horror
14,13,Balto (1995),Adventure Animation Children
15,14,Nixon (1995),Drama
16,15,Cutthroat Island (1995),Action Adventure Romance
17,16,Casino (1995),Crime Drama
18,17,Sense and Sensibility (1995),Drama Romance
19,18,Four Rooms (1995),Comedy
20,19,Ace Ventura: When Nature Calls (1995),Comedy
21,20,Money Train (1995),Action Comedy Crime Drama Thriller
22,21,Get Shorty (1995),Comedy Crime Thriller
23,22,Copycat (1995),Crime Drama Horror Mystery Thriller

Lampiran 3

Data Judul Film yang mempunyai rating lebih besar dari 3



	movieID	title
1	1	Toy Story
2	2	Jumanji
3	7	Sabrina
4	29	City of Lost Children, The (CitÃ© des enfants perdus, La)
5	32	Twelve Monkeys (a.k.a. 12 Monkeys)
6	47	Seven (a.k.a. Se7en)
7	50	Usual Suspects, The
8	52	Mighty Aphrodite
9	110	Braveheart
10	112	Rumble in the Bronx (Hont faan kui)
11	116	Anne Frank Remembered
12	150	Apollo 13
13	151	Rob Roy
14	223	Clerks (1994)
15	253	Interview with the Vampire: The Vampire Chronicles
16	260	Star Wars: Episode IV - A New Hope
17	293	The Professional (a.k.a. The Professional)
18	296	Pulp Fiction
19	318	Shawshank Redemption, The (1994)
20	337	What's Eating Gilbert Grape
21	356	Forrest Gump
22	364	Lion King, The
23		

Lampiran 4

Source code implementasi algoritma Apriori

```
# APRIORI Alogrithm
# STEP 1 . Buat frequent itemsets
from collections import defaultdict

def find_frequent_itemsets(favorable_reviews_by_users, k_1_itemsets, min_support):
    counts = defaultdict(int)
    for user, reviews in favorable_reviews_by_users.items():
        for itemset in k_1_itemsets:
            if itemset.issubset(reviews):
                for other_reviewed_movie in reviews - itemset:
                    current_superset = itemset | frozenset((other_reviewed_movie,))
                    counts[current_superset] += 1
    return dict([(itemset, frequency) for itemset, frequency in counts.items() if frequency >= min_support])

import sys
frequent_itemsets = {} # itemsets diurutkan berdasarkan kombinasi item
min_support = 50

frequent_itemsets[1] = dict((frozenset((movie_id,)), row["favorable"])
                           for movie_id, row in num_favorable_by_movie.iterrows()
                           if row["favorable"] > min_support)

print("Terdapat {} film dengan lebih dari {} ulasan bagus".format(len(frequent_itemsets[1]), min_support))
sys.stdout.flush()
for k in range(2, 20):
    cur_frequent_itemsets = find_frequent_itemsets(favorable_reviews_by_users, frequent_itemsets[k-1],
                                                    min_support)
    if len(cur_frequent_itemsets) == 0:
        print("Tidak terdapat frequent itemsets dengan kombinasi {} item".format(k))
        sys.stdout.flush()
        break
    else:
        print("Terdapat {} frequent itemsets dengan kombinasi {} item".format(len(cur_frequent_itemsets), k))
        #print(cur_frequent_itemsets)
        sys.stdout.flush()
        frequent_itemsets[k] = cur_frequent_itemsets

# hapus frek itemset yang hanya terdiri dari 1 item
del frequent_itemsets[1]
```