

Project definition

MALIS

Torstein Mellingen Langan

November 2019

Context and problem

Methodology

Data

The dataset used for this project are the Netflix prize data set[1]. It consists of `movie_titles.txt` which lists all movies on the format given below.

```
[MovieID],[YearOfRelease],[Title]
1,2003,Dinosaur Planet
2,2004,Isle of Man TT 2004 Review
3,1997,Character
.
.
```

In addition, there are a .txt file for every movie in the `training_set` folder that contains information about what rating a given user has given the movie, and on what date the movie was rated. It is formatted as following

```
[MovieID]
[CustomerID],[Rating],[Date]
1:
1488844,3,2005-09-06
```

References

- [1] Netflix. Netflix Prize Data Set. <http://archive.ics.uci.edu/ml/datasets/Netflix+Prize>, 2009. [Online; accessed 29-October-2019].