Experiment Design
Predicting the suitability of movies for an inflight viewing context, Group 9

Pia Pachinger, 01408624
Sebastian Dolezel, 01126006
Elisabeth Harlander, 01106899

Vienna University of Technology
Faculty of Informatics

January 15, 2020

Strategies

- Visual Data: 2 rows for each movie we tried different aggregation methods (sum, mean, max, min)
- Try different seeds

- Use only the data from the folders Dev$_S etand Test_S et$

Difficulties

- For table 2: Which data set was used? Dev or Test?
- Why are the ratings not treated?
- Inconsistent names (goodforairplane / goodforairplanes)
- Audio Features: different dimensions in the single data sets
- Visual Features: no specification of how they handled the two rows for each movie keep both or use an aggregation method?

Difficulties

- No seed specified
- 10-fold cross validation on a training set of 95 instances
- No number of used features specified
- Actual implementation was not specified

Key findings, Visual descriptors

- KNN - Keep only columnwise sum Precision of KNN : 0.48 Recall of KNN : 0.5 F1 of KNN : 0.4897959183673469
- KNN - Keep only columnwise mean Precision of KNN : 0.48 Recall of KNN : 0.5 F1 of KNN : 0.4897959183673469
- KNN - Keep only columnwise max Precision of KNN : 0.5769230769230769 Recall of KNN : 0.625 F1 of KNN : 0.6
- KNN - Keep only columnwise min Precision of KNN : 0.52 Recall of KNN : 0.5416666666666666 F1 of KNN : 0.5306122448979592