Topic Modelling of high and low ratings

In this section we use topic modelling to find common threads between reviews of score 5 and reviews of score 1, with the aim of discerning any patterns in the terms that contribute to the topic.

We explored several models for topic modelling namely NMF, SVD and LDA.

NMF

Non-negative matrix factorization

The NMF model decomposes its input matrix into two smaller approximate product matrices that only contain nonnegative values these are iteratively adjusted until they more closely result into the input matrix due to this process the features are clustered as the error value is reduce during each iteration.

Singular value decomposition

Like NMF the input matrix is decomposed but in SVD three product matrices are approximated.