

Bài toán:

Sentiment analysis of movie reviews ([overview](#))

Input:

Review về một bộ phim từ nhà phê bình/ người xem

Output:

Phân loại reviews đó mang tính tiêu cực hay tích cực

Metrics:

Precision

Recall

F1-score

Ứng dụng:

Đánh giá bộ phim có đáng xem hay không ứng dụng trong các mô hình Recommendation

Dataset:

Tên: Movie_Review

Nguồn: được thu thập từ việc crawl reviews từ IMDb và Rotten Tomatoes

Feature engineering (quy trình xử lý dữ liệu - preprocessing):

Crawl data → Label data → Preprocessing data → Train data → Evaluation

Models:

Logistic Regression

SVM

Naive Bayes