

Sentiment Classification for Film Reviews by Reducing Additional Introduced Sentiment Bias

Fery Ardiansyah Effendi¹, Dr. Yuliant Sibaroni, S.Si., M.T.²

^{1,2}Fakultas Informatika, Universitas Telkom, Bandung

¹feryardiansyaheff@students.telkomuniversity.ac.id, ²yuliant@telkomuniversity.ac.id,

Abstract

Individual reviews are crucial aspects of the film business as it grows more prominent, with mixed reviews help film sales exponentially. IMDb is a credible source of film reviews posted through a public medium. The problem arises with IMDb site reviews being unstructured and bias-heavy that bias processing classification methods are required to determine the site review's polarity. This paper evaluates sentiment classification methods by reducing additional introduced sentiment bias to create a balanced classification with a lower polarity bias. Elimination of additional sentiment bias would improve the accuracy in determining the review sentiment. This research limits the dataset by 50.000 rows of randomly extracted reviews from the IMDb website using dataset preparation methods such as Preprocessing, POS-Tagging, and Word Embeddings. Then preprocessed data is used in classification methods such as Artificial Neural Network, and SentiWordNet, SO-Cal. This paper also used bias processing methods such as Hyperparameter Tuning and Additional Introduced Bias Processing Method in SentiWordNet, with outputs evaluated using Accuracy and Polarity Bias Rate metrics. This research yields 77.39 % for ANN, 66.32% for BPM, 75.6% for SO-Cal, and 76.26% for Hybrid classification. Best PBR resulted in two lexicon-based methods on 0.0009 for BPM, and 0.00006 for SO-Cal.

Keywords: *Sentiment Classification, Machine Learning, Artificial Neural Network, Lexicon-based method, Bias Aware Thresholding, Semantic Orientation Calculator.*
