ABSTRACT

By
Amira Salsabila Hidayat
16/395993/PA/17334

This thesis research is inspired by existing annotation schemes for emotion in text and will optimise machine learning techniques, Naive Bayes Classifiers and Support Vector Machines, for classification. The emotion model schemed to annotate the data in this research defines basic emotions into anger, happiness, sadness, positive surprise and negative surprise. This research will contribute knowledge towards a computational model of emotions that is most suitable for use in the process of analyzing feedback. Moreover, this thesis examines the emotion contents of Twitter posts on and joins these contents in products' feedback analysis strategies.

The datasets produced in the research are the iPhone 11 Camera dataset consisting of 4,046 annotated tweets, the Amazon Alexa dataset consisting of 747 annotated tweets and the Kylie Lip Kit dataset consisting of 293 annotated tweets. All datasets were implemented on two Naive Bayes Classifiers and three kernels of Support Vector Machines.

This research concluded that for emotion classification of products' feedback, SVM's linear and RBF kernel performed best, with accuracies of 95% and 96% on the Alexa Dataset respectively and of 91% and 92% on the iPhone 11 Camera dataset. As for the Naive Bayes Classifiers, Complement Naive performed best, even though it was still not up to par with SVM's RBF and Linear kernels, it outperformed SVM's polynomial kernels.

Keywords: emotions, Naive Bayes, SVM, feedbacks, tweets