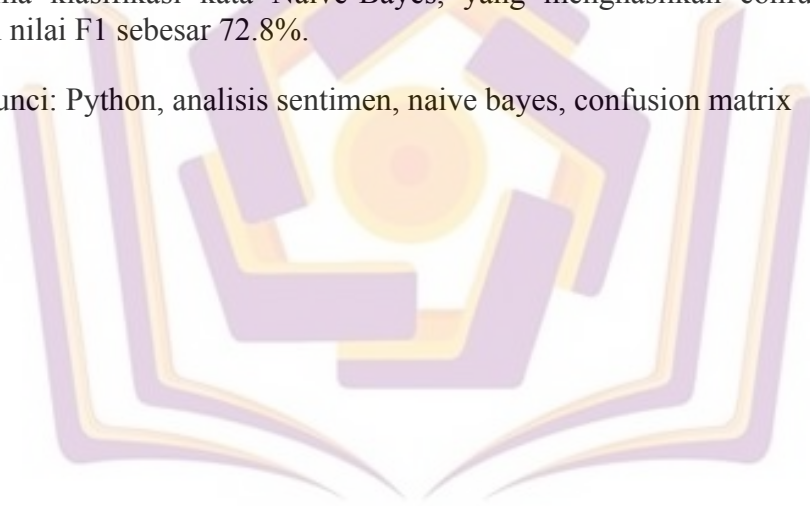


INTISARI

Internet telah memberikan platform untuk berbagai orang maupun kelompok untuk memberikan opini secara massa, tidak jarang opini-opini ini akan mempengaruhi marketability oleh suatu produk/figur/perusahaan, namun sangat sulit bagi suatu entitas untuk mengetahui informasi tersebut, mengingat cepatnya transfer informasi di internet. Tujuan dari penelitian ini adalah membangun sistem yang mampu mengelompokkan emosi dari berbagai sumber data dan dengan demikian gambaran besar opini massa. Penelitian akan dilakukan dengan menggunakan data-data yang diambil dari berbagai situs web dengan menggunakan web scraper yang ditulis menggunakan Python. Proses pengolahan data juga akan dilakukan dengan menggunakan Python, yang dipilih karena bahasa yang mudah dipelajari dan banyaknya tools yang dibutuhkan untuk melakukan studi seperti penelitian yang diajukan yang natif pada bahasa pemrograman, sehingga meminimalisir proses persiapan yang harus dilakukan. Penelitian ini menggunakan algoritma klasifikasi kata Naive-Bayes, yang menghasilkan confusion matrix dengan nilai F1 sebesar 72.8%.

Kata kunci: Python, analisis sentimen, naive bayes, confusion matrix



ABSTRACT

The Internet has given the masses a platform to share opinions, quite often these opinions affect some entity/product's marketability, however things are not as easy on the other end of this interaction, considering the speed in which these informations moves and changes in the internet. And so, this study's goal is decided. This study will document the process of the development of a system capable of discerning a text's sentiment, and therefore that of the masses. This study will be conducted using various data taken from various websites using an in-house web crawler/scrapper written in Python. The data processing itself will be done using the Python programming language, chosen for its relatively low learning curve and the available native tools are mature enough, should ease the weight of study. This study uses the Naive-Bayes Classifier as its algorithm of choice, the model produced outputs a confusion matrix with an F1 score of 72.8%

Keywords: Python, sentiment analysis, naive bayes, confusion matrix

