

## **ABSTRACT**

*Value Added Tax (VAT) is a tax imposed by individuals or groups on the consumption of goods and services. This activity takes place continuously with the aim of improving the welfare of the community. The government's policy of increasing VAT has led to various public comments, especially those on social media. Sentiment analysis is the process of determining sentiment and grouping data based on sentiment categories. Analysis is also often referred to as opinion mining based on the information gathered. SVM is a machine learning algorithm used in classifying. SVM has a basic concept, namely by separating the various classes of data by determining the best hyperplane.*

*In this final project, the application is built using the python programming language using data modeling used, namely Data Flow Diagrams. The application that is built can retrieve data from social media Twitter and Instagram, the data that has been collected will be cleaned by preprocessing the data. After that, the data will be tested based on the type of data and the type of test to get the classification results and the accuracy value of the tests carried out. The highest test results obtained from several tests are in the combined data test with a data partition of 80%-20% with an accuracy value of 83,6% with an execution time of 1440 seconds.*

*Keywords: Text Mining, Sentiment Analysis, Support Vector Machine, Value Added Tax, Social Media.*

## ABSTRAK

Pajak Pertambahan Nilai merupakan pajak yang dibebani oleh individu atau kelompok atas konsumsi barang maupun jasa. Kegiatan ini berlangsung secara terus-menerus yang memiliki tujuan untuk meningkatkan kesejahteraan masyarakat yang baik. Kebijakan pemerintah menaikan PPN memunculkan berbagai komentar masyarakat terlebih khusus yang ada pada media sosial. Analisis sentimen adalah proses menentukan sentimen dan pengelompokkan data berdasarkan kategori sentimen. Analisis sentimen sering disebut sebagai penggalian opini berdasarkan informasi yang dikumpulkan. SVM adalah algoritma *machine learning* yang digunakan dalam melakukan klasifikasi. SVM memiliki konsep dasar yaitu dengan memisahkan beberapa kelas data dengan menentukan *hyperplane* terbaik.

Pada tugas akhir ini aplikasi dibangun dengan menggunakan bahasa pemrograman python dengan menggunakan pemodelan data yang digunakan yaitu *Data Flow Diagram*. Aplikasi yang dibangun dapat mengambil data dari media sosial *Twitter* dan *Instagram*, data yang telah dikumpulkan akan dilakukan pembersihan data dengan dilakukan *preprocessing* data. Setelah itu data akan dilakukan pengujian berdasarkan jenis data dan jenis pengujian untuk mendapatkan hasil klasifikasi dan nilai akurasi dari pengujian yang dilakukan. Hasil pengujian tertinggi yang didapatkan dari beberapa kali pengujian terdapat pada pengujian data gabungan dengan partisi data 80%-20% dengan nilai akurasi 83,6% dengan waktu eksekusi 1440 detik.

Kata kunci : *Text Mining*, Analisis Sentimen, *Support Vector Machine*, Pajak Pertambahan Nilai, Media Sosial.