

ABSTRACT

Cloud computing on cloud data warehousing is a new technology that is being hotly discussed by information technology experts. Cloud data warehousing technology is here to answer the challenge of the need for more efficient cloud-based big data technology. Many things can be implemented with cloud data warehousing technology such as data storage services, creating cloud-based applications, and many more. In this Final Project, cloud data warehousing technology is implemented using snowflake. All data that is already available on the snowflake is then analyzed and tested using the SVM algorithm. Testing big data on the snowflake with the implementation of the SVM algorithm After 9 tests were carried out with various test parameters, it was found that the 4th test with parameters $C = 8$, $\text{Gamma} = 0.0078125$, with the rbf kernel which has a precision value of 0.96, with a recall of 0, 96, f1 score 0.960847701 (0.96) has the highest level of accuracy with an accuracy value of 96% being the highest value compared to testing with other test parameters.

Keywords: *Big Data, Cloud Data Warehousing, Snowflake, Support Vector Machine, t-SNE, Python, and Grid Search web.*

ABSTRAK

Cloud computing pada cloud data warehousing adalah teknologi baru yang sedang hangat dibahas oleh pakar teknologi informasi. Teknologi cloud data warehousing hadir untuk menjawab tantangan akan kebutuhan teknologi big data berbasis cloud yang semakin efisien. Banyak hal yang dapat diimplementasikan dengan teknologi cloud data warehousing seperti layanan penyimpanan data, pembuatan aplikasi berbasis cloud, dan banyak lagi. Dalam Tugas Akhir ini teknologi cloud data warehousing yang diimplementasikan menggunakan snowflake. Semua data yang sudah tersedia pada snowflake kemudian dianalisis dan diuji dengan menggunakan algoritma SVM. Pengujian big data pada snowflake dengan implementasi algoritma SVM Setelah dilakukan 9 pengujian dengan berbagai parameter uji, didapatkan bahwa pengujian ke-4 dengan parameter $C = 8$, $\text{Gamma} = 0.0078125$, dengan kernel rbf yang memiliki nilai precision senilai 0,96, dengan recall 0,96, f1 score 0.960847701 (0,96) memiliki tingkat akurasi tertinggi dengan nilai akurasi sebesar 96% menjadi nilai tertinggi dibanding pengujian dengan parameter uji lainnya.

Kata Kunci : *Big Data, Cloud Data Warehousing, Snowflake, Support Vector Machine, t-SNE, Python, dan Grid Search.*