

ABSTRAK

Pada era ini, teknologi sangat berkembang salah satunya adalah teknologi data mining memberikan kontribusi signifikan bagi para pengambil keputusan dalam menentukan solusi yang tepat terhadap berbagai permasalahan. Data Mining sendiri merupakan metode analisis data yang memadukan unsur statistika, matematika, serta pembelajaran mesin (machine learning). Dalam pendekatan Hybrid Data Mining, proses pembentukan model umum dipengaruhi oleh pembuatan classifier yang bervariasi. Penelitian ini mengajukan pendekatan dengan menggabungkan dua teknik Data Mining, yaitu *clustering* dan *classification*. Penggabungan kedua teknik ini dikenal dengan istilah Hybrid Data Mining.. Dengan memanfaatkan teknik pengelompokan data berdasarkan persamaan dan perbedaan, kesalahan data dapat diminimalkan sehingga meningkatkan akurasi model klasifikasi. Dengan menggunakan metode *Hybrid Data Mining*, dimana data di *cluster* menggunakan model *k-medoid* kemudian diklasifikasi, menunjukkan bahwa 3 dari 5 model dapat diterapkan secara efektif. Selanjutnya untuk model naive bayes mendapatkan nilai akurasi sebesar 86% kemudian dilakukan hybrid data mining menjadi 92%, dari hasil penggabungan model naive bayes mendapat kenaikan 6% model naive bayes cocok dalam konsep hybrid data mining. Kemudian pada model random forest nilai akurasi sebelum dilakukan hybrid data mining sebesar 95% namun saat menggunakan hybrid data mining nilai akurasi menurun menjadi 91%, untuk model random forest tidak cocok dilakukan hybrid data mining. Dan terakhir dengan model K-Nearest Neighbours (KNN) nilai akurasi sebesar 75% dan setelah dilakukan hybrid data mining menjadi 84% dari hasil tersebut dapat dikatakan bahwa model K-Nearest Neighbours (KNN) cocok dalam konsep hybrid data mining.

Kata kunci: *Hybrid Data Mining, clustering, classification*