

ABSTRAK

Implementasi *K Nearest Neighbor* (KNN) dalam Klasifikasi Status Kerja Lulusan Sekolah Menengah Kejuruan (SMK) dengan *Oversampling Synthetic Minority Oversampling Technique* (SMOTE) dan *Adaptive Synthetic* (ADASYN)

Halima Tusyakdiah

Program Studi Statistika, Fakultas MIPA

Universitas Islam Indonesia

Permasalahan yang sering dihadapi suatu negara yaitu masalah pengangguran. Jika dilihat berdasarkan pendidikan terakhir yang ditamatkan, data dari Badan Pusat Statistik (BPS) menunjukkan bahwa selama tiga tahun terakhir Tingkat Pengangguran Terbuka (TPT) tertinggi yaitu pada lulusan Sekolah Menengah Kejuruan (SMK). Padahal harapan dibentuknya SMK agar lulusannya langsung menjadi tenaga kerja yang siap pakai di dunia kerja. Pada pelaporan BPS periode Februari 2020 Provinsi Banten menjadi penyumbang terbanyak pengangguran di Indonesia (8,01%) dan juga penyumbang terbanyak untuk pengangguran lulusan SMK (13,13%). Hal ini cukup kontradiktif karena Banten dikenal dengan wilayah strategis industri dan wilayah yang terkena spasial dari Provinsi DKI Jakarta. Oleh karena itu penelitian ini bertujuan untuk melakukan analisis klasifikasi status kerja penduduk lulusan SMK di Provinsi Banten menggunakan metode *K Nearest Neighbor* (KNN) dengan perhitungan jarak *Heterogeneous Euclidean-Overlap Metric* (HEOM). Dari data yang diperoleh, kelas pengangguran memiliki jumlah jauh lebih sedikit dibandingkan kelas bukan pengangguran, yang mengindikasikan terdapat ketidakseimbangan kelas data. Hal ini dapat membuat model klasifikasi sangat buruk dalam memprediksi objek kelas pengangguran (minoritas), sehingga dilakukan penanganan ketidakseimbangan data dengan menggunakan teknik *Synthetic Minority Oversampling Technique* (SMOTE) dan *Adaptive Synthetic* (ADASYN). Nilai evaluasi klasifikasi KNN dengan *oversampling* SMOTE memberikan performa model terbaik dengan nilai sensitivitas, *G-mean*, dan AUC berturut-turut sebesar 0,69, 0,768, dan 0,773. Berdasarkan hasil ini maka dapat dikatakan bahwa model klasifikasi sudah cukup baik dalam mengklasifikasi status kerja terutama pada kelas pengangguran (minoritas).

Kata Kunci: ADASYN, HEOM, ketidakseimbangan data, KNN, pengangguran, SMOTE