

BAB V PENUTUP

5.1 Kesimpulan

Penelitian ini menggunakan metode resampling oversampling SMOTE dan ADASYN dengan melakukan perbandingan dari kedua metode tersebut untuk meangani distribusi kelas yang tidak seimbang pada dataset. Evaluasi model yang digunakan dalam perbandingan kedua metode tersebut adalah *balance accuracy*, *True Negative Rate (TNR)*, *True Positive Rate (TPR)*, *geometric mean*. Berdasarkan penelitian yang dilakukan, dapat ditarik kesimpulan sebagai berikut.

1. Hasil pengujian evaluasi *geometric - mean* pada dataset *wine_quality* dengan menggunakan algoritma klasifikasi K – Nearest Neighbor dan Naïve Bayes menunjukkan bahwa metode ADASYN dengan menggunakan algoritma klasifikasi Naïve Bayes menunjukkan nilai lebih tinggi dengan rata-rata *geometric mean* = 0,749367 dibandingkan dengan Metode SMOTE menggunakan algoritma klasifikasi yang sama yaitu = 0,734976, dengan demikian terdapat peningkatan akurasi sebesar 0,014391 pada metode ADASYN.
2. Hasil pengujian evaluasi *geometric - mean* pada dataset *us_crime* dengan menggunakan algoritma klasifikasi K – Nearest Neighbor dan Naïve Bayes menunjukkan bahwa metode SMOTE dengan menggunakan algoritma klasifikasi Naïve Bayes menunjukkan nilai lebih tinggi dengan rata-rata *geometric - mean* = 0,899661 dibandingkan dengan metode ADASYN menggunakan algoritma klasifikasi yang sama yaitu = 0,888171, dengan demikian terdapat peningkatan akurasi sebesar 0,01149 pada metode SMOTE.
3. Hasil pengujian evaluasi *balance accuracy* pada dataset *wine_quality* dengan menggunakan algoritma klasifikasi K – Nearest Neighbor dan Naïve Bayes pada metode ADASYN dengan algoritma Naïve Bayes menunjukkan nilai yang optimal yaitu 0,751718 dibandingkan dengan menggunakan algoritma K – Nearest Neighbor yaitu 0,631102. Sehingga dengan ketiga

kondisi tersebut diperoleh nilai yang paling optimal adalah metode ADASYN dengan menggunakan algoritma klasifikasi Naïve Bayes.

4. Hasil pengujian evaluasi balance accuracy pada dataset *us_crime* dengan menggunakan algoritma klasifikasi K – Nearest Neighbor dan Naïve Bayes pada metode SMOTE dengan algoritma Naïve Bayes menunjukkan nilai yang optimal yaitu 0,901907 dibandingkan dengan menggunakan algoritma K – Nearest Neighbor yaitu 0,813352. Sehingga dengan ketiga kondisi tersebut diperoleh nilai yang paling optimal adalah metode SMOTE dengan menggunakan algoritma klasifikasi Naïve Bayes.

5.2 Saran

Berdasarkan penelitian yang sudah dilakukan terdapat saran untuk penelitian selanjutnya terkait masalah dataset imbalance :

1. Penelitian selanjutnya dapat melakukan pengujian dengan menggunakan kumpulan data *multiclass*.
2. Dataset yang digunakan dapat diuji untuk metode *undersampling* maupun kombinasi (*oversampling-undersampling*).
3. Dapat ditambahkan jumlah dataset maupun algoritma serta evaluasi, sehingga lebih memiliki hasil akurasi yang beragam.
4. Penelitian selanjutnya dapat dilakukan kombinasi resampling menggunakan algoritma *ensemble* maupun yang lainnya.