

ABSTRAK

Nabila Affah Mumtazah, 2021, Perbandingan hasil metode *Support Vector Machine* (SVM) dengan *Ensemble Smote Bagging* dan *Smote Boosting* pada data kelulusan mahasiswa UNIMUS Skripsi, Program Studi Statistika, Universitas Muhammadiyah Semarang. Pembimbing I. Tiani Wahyu Utami, M.Si, II. Dr. Rochdi Wasono, M.Si.

Imbalanced data merupakan suatu kondisi dimana jumlah dari data yang mempresentasikan satu kelas (minoritas) sangatlah kecil dibandingkan kelas lain (mayoritas). Data yang digunakan dalam penelitian ini adalah data kelulusan mahasiswa unimus 2010-2020 yang berukuran besar yaitu 87 % data yang berada di kategori tepat waktu, sedangkan kategori tidak tepat waktu hanya 13% sehingga dikategorikan *imbalanced data*. Metode klasifikasi SVM tidak bisa mengatasi masalah tersebut karena akan bias terhadap kelas mayor dan memiliki kinerja rendah pada kelas minor. Sehingga membutuhkan metode baru untuk mengatasi masalah *imbalanced data* maka di perkenalkan metode *smote bagging* dan *smote boosting*. Tujuan menggunakan metode *smote bagging* dan *smote boosting* adalah untuk mengatasi permasalahan *imbalanced* dari hasil klasifikasi SVM yang diperoleh. Selain itu pada penelitian ini membagi *data training* dan *testing* dengan menggunakan *10-fold cross validation*. Hasil dari klasifikasi dapat memberikan masukan kepada pihak Universitas Muhammadiyah Semarang dalam perumusan kebijakan dalam rangka meminimalisir lama studi kelulusan mahasiswa berdasarkan faktor-faktor yang mempengaruhinya. Hasil penelitian menunjukkan bahwa model *smote bagging* terpilih memiliki nilai *g-mean* sebesar 0.87721 yang artinya model sudah baik dalam menjelaskan ketepatan kelulusan mahasiswa dibandingkan *smote boosting* dengan *g-mean* sebesar 0.86779.

Kata Kunci : Bagging, Boosting, SMOTE, SVM.