

ABSTRAK

Kualitas udara di Indonesia yang setiap harinya memburuk dengan keadaan sekarang. Jaman sekarang yang sudah semakin modern dengan adanya perindustrian kian maju serta transportasi yang semakin lama semakin maju ini sangat berbanding terbalik terhadap kualitas udara. Seperti polutan Konsentrasi partikulat, Karbon Monoksida, Ozon, Sulfur Dioksida dan Nitrogen dioksida ini yang nanti penulis monitorkan dengan keinginan mampu mengklasifikasikan polusi udara yang nanti ada di kota kota besar seperti halnya pada ramalan cuaca.

Dengan klasifikasi berbasis *machine learning* dengan metode *Adaboost* yang nantinya akan mampu mengklasifikasi data dari polusi udara di setiap waktunya. Sebelum data masuk ke dalam model *machine learning*, data dilatih terlebih dahulu menggunakan ISPU (Indeks Standar Pencemaran Udara). Dari situ kita bisa membuat data Training dataset yang nantinya akan bisa menjadi data yang bersih akan *Missing value* dan *Imbalance Dataset*. Dengan data sampel dari ISPU yaitu Konsentrasi partikulat (PM10), Karbon monoksida (CO), Ozon (O3), Sulfur dioksida (SO2) dan Nitrogen dioksida (NO2).

Dari sini penulis memakai beberapa metode untuk bisa mengklasifikasikan dataset yang nantinya akan di masukan kedalam sebuah model *machine learning*. Dengan melakukan beberapa test terhadap training dataset, lalu setelah itu baru dicari fill missing valuenya menggunakan *MICE*, dan setelah itu baru muncul imbalance data yang nanti akan dilatih menggunakan *Random oversampling*. Untuk klasifikasinya sendiri penulis memakai *Adaboost* sebagai metode yang digunakan dengan akurasi sekitar 90% menurut program yang sudah di jalan kan di dalam model *machine learningnya*.

Kata Kunci : *Machine Learning, Adaboost, RandomOversampling, MICE*