

## 1. Pendahuluan

### Latar Belakang

Melihat UHI (Urban Heat Island) dari tahun 2008 hingga 2018 dapat dilihat bahwa kota Jakarta menunjukkan daerah yang hangat, hal tersebut sejalan dengan meningkatnya tingkat pertumbuhan jumlah penduduk dan juga urbanisasi yang terjadi di Jakarta serta penggunaan lahan di Jakarta[1]. Penggunaan lahan di kota dapat berupa perumahan, komersial dan jasa, industri dan pergudangan, dan fasilitas transportasi. Menurut pendekatan neoclassical khususnya untuk penentuan harga suatu lahan di perkotaan ditentukan oleh perusahaan dan preferensi suatu pihak. Dari pendekatan tersebut didapatkan bawah faktor yang mempengaruhi nilai suatu lahan biasanya ditentukan berdasarkan lokasi, jarak kedekatan dengan berbagai layanan perkotaan, ketersediaan fasilitas, kualitas lingkungan, faktor social, dan transportasi beserta beberapa atribut lainnya[2]. Menurut *Rumah.com Indonesia Property Market Index* (RIPMI) hasil analisis dari 700.000 listing properti dijual dan disewakan di seluruh Indonesia mendapatkan lebih dari 17 juta kunjungan dan diakses oleh lebih dari 5,5 juta pencari properti setiap bulannya[3]. Kurangnya informasi mengenai klasifikasi lahan membuat para pencari lahan kesulitan dalam menentukan lahan mana yang akan dibeli maka diperlukan informasi terkait dengan harga lahan pada tiap kelasnya beserta persebaran harga lahan di Jakarta.

Pada penelitian ini, klasifikasi dilakukan untuk mengetahui kelas pada data harga lahan. Klasifikasi yang dilakukan menggunakan *machine learning* dengan metode *Random Forest*. *Random Forest* adalah salah satu teknik data mining yang digunakan untuk mengatasi masalah klasifikasi dan regresi[4]. *Random Forests* menarik karena mereka dapat dilatih dan diprediksi dengan cukup cepat, hanya bergantung pada satu atau dua parameter tuning, memiliki estimasi kesalahan generalisasi yang dibangun di dalamnya, dan mudah untuk dibangun secara paralel[5]. Dengan *Random Forest* memungkinkan untuk melakukan klasifikasi dengan hasil yang baik sehingga prediksi kelas dengan harga lahan akan menunjukkan akurasi yang memuaskan. Dibantu dengan melakukan analisis spasial dengan metode *Ordinary Kriging* menggunakan semivariogram informasi mengenai harga lahan dan persebarannya akan lebih baik. *Ordinary Kriging* umumnya berkinerja baik bahkan ketika pengukuran secara spasial tidak terdistribusi secara seragam dan ketika densitas pengukuran sangat rendah[6]. Dengan *Ordinary Kriging* persebaran harga lahan di Jakarta akan dihasilkan dimana metode tersebut dapat memprediksi harga lahan pada suatu titik dengan melihat harga lahan di titik dekatnya.

### Topik dan Batasannya

Topik yang dibahas pada tugas akhir ini adalah klasifikasi harga lahan dengan menggunakan *Random Forest* dan *mapping* dengan analisis spasial dengan *Ordinary Kriging* menggunakan *Semivariogram*. Batasan masalah dari tugas akhir ini yaitu menggunakan data harga lahan di Jakarta.

### Tujuan

Tujuan tugas akhir ini yaitu mengimplementasikan *Random Forest* untuk klasifikasi harga lahan di Jakarta dan menghitung akurasi dari *Random Forest* untuk klasifikasi serta melakukan analisis spasial menggunakan *Random Forest*.

### Organisasi Tulisan

Organisasi tulisan pada tugas akhir ini setelah pendahuluan yaitu studi terkait mengenai klasifikasi menggunakan *Random Forest*, *Semivariogram*, dan *Ordinary Kriging*, ADASYN, SMOTE