Implementasi SMOTE Guna Penanganan Data Imbalance pada Citra Multispektral Lahan Sawah untuk Klasifikasi Kesuburan Lahan Sawah

ALFARIZ GILANG SEPTIAN

(LOGO IPB)

DEPARTEMEN ILMU KOMPUTER

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

INSTITUT PERTANIAN BOGOR

BOGOR

2023

I PENDAHULUAN

1.1 Latar Belakang

Populasi penduduk, lahan sawah, dan produksi pangan kemas kaitannya dengan ketahanan pangan yang akan menjerat ketahanan nasional, (Ariesa & Khairani, 2019). Menurut publikasi BPS tahun 2021, penambahan jumlah penduduk Indonesia rata-rata sebanyak 3.26 juta setiap tahun. Melihat rata-rata peningkatan jumlah penduduk Indonesia yang tinggi, mengakibatkan kebutuhan pangan Indonesia ikut meningkat. Laporan BPS pada tahun 2019 menunjukkan konsumsi beras nasional mencapai 28.69 juta ton, lebih tinggi dibandingkan konsumsi bahan pokok lainnya seperti daging sapi atau kerbau dengan jumlah konsumsi 768 ribu ton, daging ayam dengan jumlah konsumsi 4 juta ton, lalu diikuti dengan telur ayam atau bebek, ikan segar, dan jagung. Beras merupakan penyumbang sumber energi terbesar bagi pertumbuhan tubuh manusia, karena mengandung kalori dan protein utama. Sehingga dapat dikatakan bahwa beras menyumbangkan pengaruh yang signifikan terhadap kestabilan perekonomian nasional. Beras juga memiliki peran yang krusial pada ketahanan pangan, ketahanan ekonomi, dan stabilitas politik nasional, (BPS, 2019). Berdasarkan *Global Food Security Index* (GFSI), ketahanan pangan Indonesia pada tahun 2021 mengalami penurunan, pada tahun 2020, skor indeks ketahanan pangan Indonesia mencapai level 61.4, namun pada tahun 2021, skor indeks ketahanan pangan hanya mencapai level 59.2, sehingga menempatkan indonesia pada peringkat ke 69 dari 113 negara.

Berbagai teknik telah dikembangkan dan diterapkan sepanjang sejarah kehidupan manusia untuk meningkatkan produksi pertanian, terutama di daerah persawahan yang sumber dayanya terbatas. Optimalisasi sektor padi mutlak diperlukan, dalam mengoptimalkan produksi padi. Salah satu cara yang dapat diperlukan adalah meningkatkan kesuburan tanah lahan sawah. Kesuburan tanah merupakan kemampuan tanah dalam menyediakan unsur hara dengan takaran dan kesetimbangan tertentu yang dilakukan oleh tanah secara berkesinambungan dengan tujuan untuk menjaga pertumbuhan suatu tanaman tertentu dalam keadaan yang menguntungkan. Semakin subur tanah menandakan bahwa tanah tersebut mengandung unsur hara yang tinggi. Kandungan unsur hara tidak selalu sama di setiap tanah, hal itu bergantung pada jenis tanaman yang tertanam, musim, dan jenis pengolahan tanah yang diberikan, (Prabowo et al, 2018).

Salah satu teknologi yang dapat membantu meningkatkan produktivitas sektor pertanian adalah penggunaan citra multispektral yang diambil oleh *drone* UAV (*Unmanned Aerial Vehicle*). Citra multispektral atau multispektral imaging adalah salah satu cara atau metode menangkap data gambar pada frekuensi tertentu di seluruh spektrum elektromagnetik (Fitriyanti et al., 2021) Penggunaan Citra multispektral yang diambil oleh UAV salah satunya digunakan untuk klasifikasi, dan data citra yang ingin diklasifikasi lumrah terjadi imbalance data. Permasalahan imbalance data seringkali kita jumpai pada berbagai domain, dimana jumlah data pada tiap kelas tidak seimbang. Imbalance data terjadi ketika jumlah data dalam satu kelas jauh lebih tinggi (*majority class*) atau lebih rendah (*minority class*) dibandingkan kelas lainnya. Imbalanced data akan lebih sulit untuk dilakukan pengolahan dan analisis data seperti klasifikasi, pengklusteran, prediksi, dan sebagainya. Hal tersebut terjadi karena model analisis data tidak dirancang untuk mempertimbangkan distribusi kelas dalam meningkatkan akurasi dari model. Banyak penelitian yang melaporkan analisis data dengan imbalance data seringkali memberikan hasil yang keliru (Indrawati, 2021). Satu diantara banyak penanganan imbalance data adalah dengan menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE). SMOTE dikenalkan pertama kali oleh Chawla pada tahun 2002. Cara kerja SMOTE adalah dengan menambahkan data-data buatan pada kelas minority dengan melakukan interpolasi pada data-data asli, sehingga data buatan yang dihasilkan sangat bervariasi (Indrawati, 2021).

1.2 Rumusan Masalah

Berdasarkan latar belakang tersebut, maka penulis merumuskan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan Algoritme SMOTE untuk menangani data imbalance pada citra multispektral lahan sawah untuk klasifikasi kesuburan lahan sawah?

2. Bagaimana pengaruh Algoritme SMOTE untuk menangani data imbalance pada citra multispektral lahan sawah untuk klasifikasi kesuburan lahan sawah?

1.3 Tujuan

Penelitian ini bertujuan untuk menangani data imbalance pada citra multispektral lahan sawah untuk klasifikasi kesuburan lahan sawah dengan algoritme SMOTE dan pengaruh penggunaan SMOTE pada data imbalance pada citra multispektral lahan sawah

1.4 Manfaat

Manfaat yang diproyeksikan setelah penanganan data imbalance pada citra multispektral lahan sawah adalah mampu membantu pemerintah dan peneliti lain untuk menangani imbalance data pada citra multispektral persebaran kesuburan lahan sawah

1.5 Ruang Lingkup

Ruang lingkup pada penelitian ini adalah sebagai berikut:

1. Penelitian ini akan menggunakan data citra multispektral lahan sawah
2. Implementasi Algortime SMOTE akan menggunakan Bahasa Pemrograman Python