

1. Pendahuluan

1.1 Latar Belakang

Curah hujan merupakan informasi yang sangat penting karena dapat berdampak pada banyak elemen kehidupan manusia, baik secara langsung maupun tidak langsung. Banjir dan tanah longsor adalah dua contoh bencana alam yang dapat diakibatkan oleh curah hujan. Atau dampak terhadap kehidupan perkotaan, termasuk dampak pada sistem saluran pembuangan, dampak lalu lintas, dll [1]. Letak geografis curah hujan, seperti lokasi dengan iklim tropis, khususnya di Indonesia yang cenderung memiliki iklim yang beragam, serta bentuk dan arah pulau-pulau, semuanya dapat mempengaruhi hal tersebut [2, 15]. Daerah iklim tropis biasanya memiliki suhu panas, dengan suhu berfungsi sebagai kriteria utama. Namun pada kenyataannya, perubahan iklim tidak sepenuhnya bergantung pada suhu. Sebaliknya, curah hujan—terutama selama musim hujan—memainkan peran penting dalam perubahan iklim [3]. Perubahan volume curah hujan dapat berdampak pada kegiatan ekonomi, produksi pangan, bahkan bencana alam.

Untuk mengatasi hal tersebut perlu adanya pendekatan prediksi curah hujan yang bertujuan untuk mengetahui kapan dan jenis curah hujan yang akan terjadi. Metode machine learning *Random Forest* merupakan metode yang dapat digunakan untuk memprediksi curah hujan. Suhaila Zainudin, dkk dalam penelitiannya menunjukkan hasil perbandingan lima teknik klasifikasi (Naïve Bayes, Decision Tree, Support Vector Machine, Neural Network dan *Random Forest*) yang dilakukan untuk prediksi curah hujan di Malaysia, dimana hasilnya adalah Pohon Keputusan dan *Random Forest* dilakukan dengan baik untuk prediksi. curah hujan karena kemampuannya untuk melatih lebih sedikit data dan memprediksi jumlah data yang lebih besar dengan ukuran F yang lebih tinggi, dan *Random Forest* mengidentifikasi 1043 kejadian dengan data pelatihan kecil (10%) dari 1581 kejadian [4]. Sedangkan Jie Dou, dkk dalam penelitiannya membandingkan kinerja dua model machine learning Decision Tree dan *Random Forest* untuk memodelkan kejadian longsor besar yang dipicu oleh curah hujan. Hasil penelitian menunjukkan bahwa model Decision Tree dan *Random Forest* menghasilkan peta kerentanan yang hampir akurat (AUC N 0.9). *Random Forest* (AUC = 0,956) memiliki efisiensi keseluruhan yang jauh lebih tinggi daripada Decision Tree (AUC = 0,928) [5].

Namun penelitian [4] dan [5] hanya berfokus pada hasil prediksi dari model machine learning yang dibangun. telah mengkaji hasil prediksi menggunakan pola hamburan dengan memprediksi distribusi curah hujan di Taiwan bagian tengah menggunakan metode Inverse Distance Weighting (IDW) dari data curah hujan antara tahun 1981 dan 2010 yang diambil dari 46 stasiun pemantau curah hujan. di mana 12 stasiun pemantau curah hujan milik Asosiasi Irigasi Taichung (TIA) digunakan untuk validasi silang. Pengaruh nilai radius, dan parameter kontrol-a dipilih untuk menghitung nilai RMSE guna mendapatkan data interpolasi curah hujan yang optimal. Dari hasil penelitian, parameter IDW yang memiliki nilai optimal untuk data interpolasi curah hujan pada kebanyakan kasus berada pada radius hingga 10–30 km. Dan dari hasil penerapannya, metode IDW memiliki akurasi prediksi yang lebih baik pada musim kemarau (Oktober hingga April) dibandingkan pada musim banjir (Mei–September).

Berdasarkan permasalahan dan penelitian tersebut, maka dibuatlah suatu sistem peta prediksi klasifikasi curah hujan dengan menggunakan metode *Random Forest* untuk mengetahui identifikasi terhadap intensitas curah hujan pada suatu daerah tertentu dengan mengidentifikasi hasil dari evaluasi metode yang dihasilkan. Dan menambahkan metode analisis spasial yaitu *Inverse Distance Weighted* (IDW) untuk menggambarkan pola penyebaran curah hujan berdasarkan ragam karakteristik iklim yang ada.