

---

## 1. Pendahuluan

### Latar Belakang

Cuaca dan iklim merupakan fenomena alam yang keberadaannya penting dalam menentukan aktivitas manusia. Misalnya, informasi cuaca dan iklim digunakan sebagai acuan di bidang pertanian untuk menentukan masa tanam dan transportasi untuk jadwal keberangkatan pesawat. Cuaca merupakan fenomena alam yang cenderung tidak terkendali[1]. Salah satu bentuk cuaca dan iklim adalah curah hujan. Curah hujan yang tinggi biasanya terjadi pada bulan Desember-Maret setiap tahun. Salah satu dampak dari curah hujan yang tinggi adalah banjir. Salah satu daerah yang mengalami banjir di Jakarta. Pada tahun 2007 Jakarta dilanda banjir besar[2]. Dilaporkan hampir 70% wilayah Jakarta tergenang air.

Informasi cuaca dan iklim meningkat seiring dengan meningkatnya fenomena alam yang tidak biasa[3]. Dampak curah hujan yang tinggi dapat diminimalisir dengan memberikan informasi prediksi terjadinya curah hujan yang tidak biasa di masa mendatang. Salah satu Instansi Pemerintah yang menyediakan data prediksi curah hujan adalah BMKG. Menurut BMKG, prediksi 2022-2023 akan lebih akurat karena perubahan data zona musiman yang digunakan. Sebelumnya menggunakan zona musim 1981-2010, kini diperbarui menjadi 1991-2020. Namun, ada kemungkinan bahwa prediksi yang dibuat tidak benar.

Besar kecilnya intensitas curah hujan tidak dapat diatur dengan campur tangan manusia tetapi dapat diprediksi. Prediksi curah hujan yang akurat diperlukan untuk mengantisipasi dampak negatif curah hujan yang tinggi. Dengan demikian, dimungkinkan untuk melakukan penelitian tentang prediksi curah hujan. Penelitian[4] adalah suhu udara, angin, waktu pemaparan, dan kelembaban relatif. menggunakan atribut suhu udara, angin, waktu pemaparan, dan kelembaban relatif. Prediksi dilakukan dengan klasifikasi menggunakan machine learning pada data yang disediakan. Berdasarkan penelitian ini masih terdapat kekurangan yaitu atribut yang digunakan masih sedikit dan hasil yang diperoleh tidak menampilkan pemetaan hanya nilai prediktif.

Salah satu metode machine learning yang terkenal untuk proses klasifikasi adalah Naive Bayes. Naive Bayes ditemukan pada abad ke-18 oleh Thomas Bayes menggunakan rumus Bayes[5]. Penelitian yang dilakukan oleh Slamet Triyanto dkk pada tahun 2021 memprediksi klasifikasi bencana banjir berdasarkan curah hujan menggunakan naive Bayes dan memperoleh akurasi sebesar 79,16% [6]. Pada tahun 2018 juga dilakukan penelitian oleh Nabila dan Siti untuk memprediksi curah hujan di kota Subang[7]. Penelitian membandingkan beberapa metode klasifikasi dengan evaluasi menggunakan RMSE dan untuk model nave Bayes nilai RMSE adalah 0,37 yang merupakan nilai yang cukup baik. Namun pada penelitian ini hanya 5 parameter yang digunakan. Pada tahun 2018, A. K. Sharma dkk membandingkan 4 metode untuk prediksi curah hujan di India[8]. Salah satu metode yang digunakan adalah nave Bayes yang memperoleh akurasi di atas 96% untuk setiap area yang berbeda. Penelitian[9] yang dilakukan oleh A.U.Uzmi dkk pada tahun 2021 melakukan klasifikasi curah hujan menggunakan Naive Bayes dan diperoleh akurasi sebesar 96%. Berdasarkan beberapa penelitian sebelumnya, metode nave bayes merupakan metode yang cukup baik untuk mengklasifikasikan curah hujan.

Ketersediaan peta prediksi curah hujan akan sangat membantu masyarakat khususnya petani untuk lebih mudah memahami hasil prediksi curah hujan. Namun pada beberapa penelitian sebelumnya, belum ada penelitian yang menghasilkan output berupa peta prediksi. S.S. Prasetyowati dkk melakukan penelitian pencemaran udara tahun 2020 di wilayah Bandung dan menghasilkan output berupa peta prediksi pencemaran udara beberapa tahun ke depan dengan menggunakan metode Simple Kriging[10]. Pada tahun yang sama, penelitian M. Hassim et al. membandingkan beberapa metode interpolasi untuk memetakan curah hujan di wilayah sungai Langat, Malaysia, dan hasil akhir yang diperoleh adalah metode simple kriging merupakan metode yang paling optimal karena memiliki nilai RMSE terkecil[11]. Pada tahun 2021, [12] penelitian akan menginterpolasi curah hujan di Pakistan. Salah satu teknik interpolasi yang digunakan adalah Simple Kriging. Nilai RMSE rata-rata adalah 0,365 untuk setiap model semivariogram yang digunakan dan termasuk RMSE terkecil dibandingkan teknik interpolasi lainnya. Pada tahun yang sama juga dilakukan studi interpolasi menggunakan simple kriging dan nilai RMSE adalah 0,93[13].

Pada beberapa studi prediksi curah hujan di atas, atribut yang digunakan masih sedikit dan belum ada output berupa pemetaan. Pada penelitian interpolasi, output yang dihasilkan masih belum menghasilkan pemetaan sesuai dengan bentuk peta yang ada. Maka dari beberapa faktor tersebut, membuka peluang untuk melakukan penelitian tentang prediksi klasifikasi curah hujan menggunakan Naive Bayes dan pemetaan dengan interpolasi kriging sederhana. Penambahan atribut dan output yang menyesuaikan bentuk peta akan ditambahkan pada penelitian ini. Hasil dari penelitian ini adalah prediksi untuk beberapa hari dan bulan ke depan sesuai dengan data yang digunakan. Diharapkan penelitian ini dapat bermanfaat bagi masyarakat luas untuk lebih mengantisipasi dampak curah hujan.