

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Indonesia merupakan negara maritim dengan luas lautan yang lebih besar daripada luas daratan. Dikelilingi dengan daerah kepulauan yang luas sehingga sangat rawan terhadap bencana alam khususnya banjir. Banjir merupakan bencana alam yang dapat diprediksi datangnya dengan memperhatikan curah hujan dan aliran air dan pada umumnya terjadi karena curah hujan tinggi terhadap suatu daerah, namun banjir juga bisa terjadi karena kondisi lingkungan seperti berkurangnya lahan terbuka hijau. Ancaman banjir sering terjadi di beberapa provinsi di Indonesia, khususnya Kalimantan Timur. Di Ibu Kota Kalimantan Timur sendiri yakni Samarinda, banjir menjadi fenomena rutin yang kerap kali melanda setiap tahun. Berdasarkan data pada website Badan Pusat Statistik Kota Samarinda (BPS, n.d.) terjadi peningkatan jumlah desa / kelurahan yang terkena dampak banjir pada tahun 2014 sebanyak 24 kali, pada tahun 2018 sebanyak 18 kali, pada tahun 2020 sebanyak 33 kali dan pada tahun 2021 sebanyak 32 kali. Dengan beberapa daerah yang terdampak antara lain Palaran, Samarinda Ilir, Samarinda Kota, Sambutan, Samarinda Seberang, Loa Janan Ilir, Sungai Kunjang, Samarinda Ulu, Samarinda Utara, dan Sungai Pinang.

Banjir di klasifikasikan menurut proses penyebabnya untuk membantu dalam meningkatkan akurasi estimasi frekuensi terjadinya banjir dan besaran banjir (Tarasova et al., 2019). Klasifikasi menjadi salah satu metode *Data mining* yang paling umum untuk digunakan. *Data mining* dalam pemanfaatannya juga dapat dihubungkan dengan berbagai aspek termasuk bencana alam, khususnya banjir. *Data mining* mempunyai peran penting dalam menyatukan teknologi dan penelitian. Dapat mengidentifikasi aturan asosiasi dengan klasifikasi serta pengenalan bekerja dengan kategorisasi yang mendapatkan beberapa hasil buruk, rata-rata dan baik (Mian & Ghabban, 2022).

Support Vector Machine (SVM) adalah salah satu algoritma yang paling sering digunakan pada prediksi/klasifikasi dalam *data mining*. SVM adalah salah satu algoritma yang digunakan untuk klasifikasi data menggunakan *hyperplane* (Fitriana & Sibaroni, 2020). Dalam penelitian yang melakukan klasifikasi risiko tsunami, algoritma SVM menghasilkan nilai akurasi yang tinggi, dimana semakin tinggi nilai maka semakin tepat prediksi dan klasifikasi yang dilakukan (Wahani & Prasetyo, 2022). Namun pada penelitian lain yang dilakukan oleh (Abdullah & Utami, 2018) algoritma SVM mempunyai tingkat akurasi yang rendah. Oleh karena itu akan diterapkan algoritma optimasi untuk membuat performa yang dihasilkan dapat lebih baik dengan menggunakan *feature selection*.

Berdasarkan penelitian yang sudah dilakukan (Pratama et al., 2022) menunjukkan bahwa optimasi SVM dalam klasifikasi curah hujan dapat meningkatkan akurasi sebesar 2%. Pada penelitian tersebut fitur-fitur yang biasa digunakan untuk mengetahui apa yang mempengaruhi frekuensi terjadinya banjir dan besaran banjir yaitu temperatur, kelembapan, curah hujan, lama penyinaran matahari, kecepatan angin dan arah angin. Penelitian tersebut juga menyarankan bahwa peneliti selanjutnya dapat meningkatkan hasil akurasi klasifikasi menggunakan metode seleksi fitur lain seperti *Wrapper based methods*. *Feature selection* tipe *wrapper* ini melakukan pemilihan bersamaan dengan pelaksanaan pemodelan. *Selection* tipe ini dapat dilakukan dengan memanfaatkan metode *Genetic Algorithm*.

Genetic Algorithm adalah salah satu algoritma optimasi yang kuat dan dapat digunakan pada berbagai studi kasus dengan penggunaan prinsip teori evolusi. Algoritma ini sering digunakan untuk menemukan solusi optimal pada kasus yang sederhana sampai yang rumit (Elva, 2019). Beberapa penelitian lain terkait peningkatan akurasi klasifikasi dengan *Feature Selection Wrapper Based Methods* sudah pernah dilakukan sebelumnya seperti pada penelitian (Novichasari, 2019) yang menggunakan algoritma SVM dengan GA pada data kinerja akademik mahasiswa yang mendapat akurasi *Support Vector Machine – Genetic Algorithm (SVM-GA)* tertinggi.

Dari beberapa uraian diatas maka penelitian ini menggunakan metode *Genetic Algorithm* untuk menyeleksi fitur mana yang paling berpengaruh pada dataset banjir kota Samarinda. Penelitian ini bertujuan untuk mencari nilai perubahan yang terjadi terhadap model klasifikasi banjir ketika menggunakan fitur dataset keseluruhan dan ketika menggunakan fitur yang telah diseleksi oleh GA. Algoritma SVM digunakan untuk mencari nilai akurasi dari kedua model klasifikasi tersebut.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang sudah ada, maka rumusan masalah pada penelitian ini yaitu:

- a) Pada dataset banjir di Kota Samarinda fitur manakah yang memiliki pengaruh besar ketika menggunakan *Support Vector Machine* dengan penggunaan optimasi *Genetic Algorithm*?
- b) Pada metode optimasi *Genetic Algorithm* terhadap klasifikasi kinerja Algoritma SVM apakah terdapat peningkatan akurasi pada model tersebut?

1.3. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

- a) Mengetahui fitur mana yang paling berpengaruh di dalam dataset banjir Kota Samarinda.
- b) Mengetahui perubahan akurasi yang akan dihasilkan oleh model klasifikasi banjir di kota Samarinda dengan optimasi *Genetic Algorithm* pada algoritma *Support Vector Machine*.

1.4. Batasan Masalah

Agar lingkup permasalahan yang dibahas tidak meluas, maka penulis membatasi penelitian sebagai berikut:

- a) Data yang digunakan adalah dataset Banjir Kota Samarinda yang diambil dari BPBD dan BMKG Kota Samarinda pada tahun 2019 – 2023.

- b) Metode algoritma data mining klasifikasi yang digunakan adalah Algoritma SVM dengan seleksi fiturnya yaitu *Genetic Algorithm*.
- c) Fitur yang akan digunakan pada penelitian ini yaitu:
- Temperatur minimum suhu
 - Temperature maksimum suhu
 - Temperature rata-rata
 - Kelembapan rata-rata
 - Curah hujan
 - Lamanya penyinaran matahari
 - Kecepatan angin maksimum
 - Arah angin saat kecepatan maksimum
 - Kecepatan angin rata-rata
 - Arah angin terbanyak.

1.5. Manfaat Penelitian

Penelitian ini diharapkan dapat memberi manfaat kepada beberap pihak, khususnya:

1. Penulis

Mengembangkan wawasan dan berkontribusi pada bidang ilmu pengetahuan, terutama dalam pemodelan klasifikasi banjir kota Samarinda menggunakan algoritma SVM dan optimasinya yaitu GA.

2. Mahasiswa

Penelitian ini dapat memberikan manfaat kepada mahasiswa untuk memperoleh pemahaman yang lebih baik tentang teknik klasifikasi banjir dan penggunaan algoritma SVM-GA pada kinerja model dan menjadi refrensi khususnya bagi mahasiswa dibidang Informatika yang dapat membantu proses pembelajaran.