

# Prediksi Banjir Lahar Dingin Pada Lereng Merapi Menggunakan Data Curah Hujan Dari Satelit

Rosa Andrie Asmara<sup>1</sup>, Arief Prasetyo<sup>2</sup>, Siska Stevani<sup>3</sup>, Ratih Indri Hapsari<sup>4</sup>

<sup>1,2,3,4</sup>Program Studi Teknik Informatika, Jurusan Teknologi Informasi, Politeknik Negeri Malang  
<sup>1</sup>rosa.andrie@polinema.ac.id, <sup>2</sup>arp.ariefprasetyo@gmail.com, <sup>3</sup>siska.stevani@gmail.com, <sup>4</sup>ratihindri@gmail.com

---

## Abstrak

Banjir lahar dingin merupakan sekumpulan lahar yang dimuntahkan oleh gunung berapi dan sampai ke permukaan yang lebih rendah dengan bantuan atau dorongan dari air hujan yang mengakibatkan material-material vulkanik dari lahar akan menerjang lahan yang berada di bawahnya ataupun pemukiman penduduk. Faktor yang menyebabkan banjir lahar adalah curah hujan (mm/jam) dan akumulasi hujan (mm/7hari). Terjadinya banjir lahar dapat dideteksi oleh beberapa alat salah satunya adalah Geofon. Alat sering rusak dan hanyut karena dipasang atau ditempatkan pada permukaan tanah disetiap stasiun sungai, dan pada saat terjadinya banjir lahar hingga sampai ke permukaan yang lebih rendah maka alat tersebut tidak dapat mengirimkan informasi getaran. Oleh karena itu pada penelitian ini di buat sebuah sistem untuk membantu sensor Geofon dalam memprediksi banjir lahar pada kawasan Lereng Merapi. Sistem akan mengeluarkan status getaran yang terdiri dari 4 kelas yaitu banjir rendah, banjir sedang, banjir tinggi dan tidak terjadi banjir lahar dengan memperhitungkan atribut curah hujan dan akumulasi hujan dari satelit menggunakan metode K-NN (*K-Nearest Neighbor*). Metode *10-Fold Cross Validation* dan Uji Akurasi digunakan untuk mengetahui nilai K Optimal pada tiap lokasi penelitian yaitu Gendol, Putih 1 dan Putih 2. Berdasarkan hasil pengujian yang didapat adalah pada lokasi Gendol dan Putih 1 menggunakan 3-NN dengan akurasi rata-rata 72.307% dan 81.429%, lokasi Putih 2 menggunakan 1-NN dengan akurasi rata-rata 86.955%. Data pengujian pada lokasi Gendol menggunakan data *1-Fold Cross Validation* dengan akurasi 3-NN 92.31%, Putih 1 data *8-Fold Cross Validation* dengan akurasi 3-NN 95.24%, dan Putih 2 data *10-Fold Cross Validation* dengan akurasi 1-NN 91.3%.

**Kata kunci:** banjir lahar, *k-nearest neighbor*, *10-fold cross validation*

---

## 1. Pendahuluan

Banjir lahar dingin merupakan sekumpulan lahar yang dimuntahkan oleh gunung berapi dan sampai ke permukaan yang lebih rendah dengan bantuan atau dorongan dari air hujan. Lahar yang berada di sekitar gunung berapi tersebut akan terbawa turun melalui lereng gunung ketika hujan turun dengan deras. Akibatnya, air hujan yang membawa serta material-material vulkanik dari lahar ini akan menerjang lahan yang berada di bawahnya ataupun pemukiman penduduk. Hal ini akan berakibat banyaknya kerusakan ataupun dampak-dampak lain yang akan dihasilkan oleh banjir lahar dingin ini [1]. Menurut Badan Penanggulangan Bencana Daerah (BPBD) Kabupaten Sleman, Daerah Istimewa Yogyakarta tetap mewaspadaai potensi terjadinya banjir lahar hujan atau banjir lahar dingin di sejumlah sungai berhulu gunung Merapi saat puncak musim hujan seperti Sungai Gendol, Kali Putih, Boyong, dan lain-lain yang memiliki potensi diterjang lahar dingin. Ancaman banjir lahar dingin yang masih tinggi di kawasan Lereng Merapi mendorong Pemkab Sleman untuk mengingatkan warganya dan para wisatawan agar terus waspada, terutama jika hujan deras turun [2]. Faktor yang

menyebabkan banjir lahar adalah intensitas atau curah hujan (mm/jam) dan akumulasi hujan (mm). Terjadinya banjir lahar dapat dideteksi oleh beberapa alat salah satunya adalah Geofon. Alat sering rusak dan hanyut karena dipasang atau ditempatkan pada permukaan tanah disetiap stasiun sungai, dan pada saat terjadinya banjir lahar hingga sampai ke permukaan yang lebih rendah maka alat tersebut tidak dapat mengirimkan informasi getaran.

Berdasarkan permasalahan diatas maka penulis membangun aplikasi Prediksi Banjir Lahar Dingin pada Lereng Merapi yang diharapkan dapat membantu sensor geofon dalam memprediksi dan memberikan informasi getaran dalam bentuk klasifikasi status getaran yaitu terjadi banjir lahar rendah, sedang tinggi atau tidak terjadi banjir lahar dengan menggunakan parameter curah hujan dan akumulasi hujan pada kawasan lereng merapi khususnya di Kali Putih 1, Putih 2 dan Gendol yang rawan terjadinya banjir lahar dan juga dekat dengan permukiman penduduk. Pendekatan yang digunakan untuk memprediksi banjir lahar adalah dengan menerapkan metode K-NN (*K-Nearest Neighbor*), karena memiliki akurasi yang tinggi dengan rasio kesalahan kecil. Algoritma K-NN adalah sebuah

metode klasifikasi terhadap sekumpulan data berdasarkan pembelajaran data yang sudah terklasifikasi sebelumnya dan termasuk dalam *supervised learning*, dimana hasil dari sampel uji yang baru diklasifikasikan berdasarkan kedekatan mayoritas dari kategori pada k-NN untuk dijadikan prediksi dari sample uji [3].

## 2. Landasan Teori

### 2.1 Penelitian Terdahulu

Penelitian yang dilakukan oleh Bima Ahida Indaka Sugna mahasiswa Politeknik Negeri Malang pada tahun 2019 dengan judul “Sistem Penentuan Status Bencana Banjir Pada Daerah Lereng Merapi” dan berkolaborasi dengan penelitian pada *IOP Conference Series: Earth and Environmental Science* dengan judul “*Predictability of Naïve Bayes classifier for lahar hazard mapping by weather radar*”. Penelitian yang dilakukan berfokus pada pemetaan pada daerah lereng merapi dan menggunakan metode *Naïve Bayes* untuk kerentanan bahaya lahar sebagai sistem peringatan dengan parameter curah hujan dari radar cuaca, kemiringan tanah, kelembaban tanah dari SMAP dan jenis tanah. Keluaran dari klasifikasi sistem ini untuk menghasilkan 2 kelas yaitu aman dan rawan bencana banjir yang ditampilkan pada peta dengan tingkat akurasi 84,6%, dari hasil pengambilan 10% dari data latih [4][5][6].

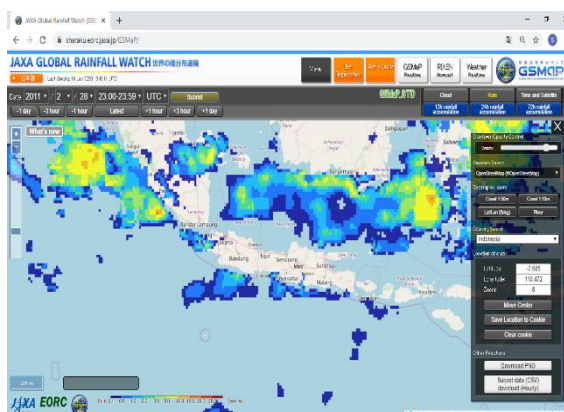
Penelitian yang dilakukan oleh M. Reza Noviansyah mahasiswa Universitas Tanjungpura pada tahun 2018 dengan judul “Penerapan Data Mining Menggunakan Metode K-Nearest Neighbor Untuk Klasifikasi Indeks Cuaca Kebakaran Berdasarkan Data Aws (*Automatic Weather Station*) (Studi Kasus: Kabupaten Kubu Raya)”. Penelitian ini menerapkan *Data Mining* dengan algoritma klasifikasi K-NN dalam mengklasifikasi tingkatan bahaya kebakaran hutan berdasarkan kelas *Fire Weather Index (FWI)* dengan menggunakan atribut data berupa suhu, kelembapan, curah hujan dan kecepatan angin untuk diklasifikasi pada kelas FWI. Berdasarkan hasil pada pengujian dengan menggunakan 252 data uji dihasilkan persentase keberhasilan 80.16% dengan nilai K=5 [7].

Selanjutnya penelitian yang dilakukan oleh Firman Tempola mahasiswa Universitas Khairun Ternate pada tahun 2018 dengan judul “Perbandingan Klasifikasi Antara K-NN Dan *Naive Bayes* Pada Penentuan Status Gunung Berapi Dengan K-Fold Cross Validation”. Sedangkan untuk validasi data menggunakan *k-fold cross validation*. Ada 5 kriteria yang digunakan dalam melakukan klasifikasi yaitu empat faktor kegempaan diantaranya gempa vulkanik dangkal, gempa tektonik jauh, gempa vulkanik dalam, gempa hembusan dan ditambah satu kriteria yaitu status sebelumnya. Ada 3 status yang di yang

diklasifikasi yaitu normal, waspada dan siaga. Hasil penelitian yang dibagi kedalam *3fold* disetiap metode klasifikasi diperoleh perbandingan akurasi sistem rata-rata tertinggi pada k-nn 63.68 % dengan standar deviasi 7.47 %. Sedangkan dengan menggunakan *naive bayes* diperoleh rata-rata akurasi sebesar 79.71 % dengan standar deviasi 3.55 % [8].

### 2.2 Data Satelit

Salah satu pengamatan curah hujan dari penginderaan jauh adalah GSMaP / *The Global Satellite Mapping of Precipitation* melalui JAXA *Global Rainfall Watch System*. Tujuan dari Proyek GSMaP adalah pengembangan algoritma gelombang mikro gelombang mikro canggih berdasarkan algoritma pengambilan hujan deterministik dari dan produksi peta curah hujan global resolusi tinggi yang tepat. GPM-GSMaP Ver.6 adalah algoritma terbaru yang dikembangkan oleh Proyek Pemetaan Satelit Global Presipitasi (GSMaP), menghasilkan peta curah hujan global dengan resolusi tinggi dan presisi tinggi menggunakan data satelit. Fitur utama dari algoritma GSMaP adalah pemanfaatan berbagai atribut yang berasal dari radar presipitasi ruang angkasa, TRMM / PR dan GPM / DPR. Satuan tingkat curah hujan adalah mm / jam dalam format kisi. Resolusi spasial adalah 0.1 derajat lintang / bujur dan resolusi temporal adalah 1 jam [9]. Gambar 1 dibawah ini adalah GSMaP melalui JAXA *Global Rainfall Watch System*.

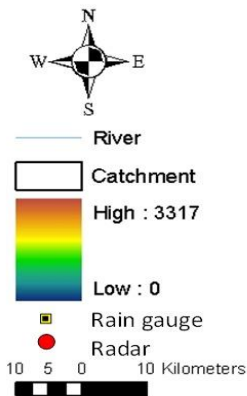
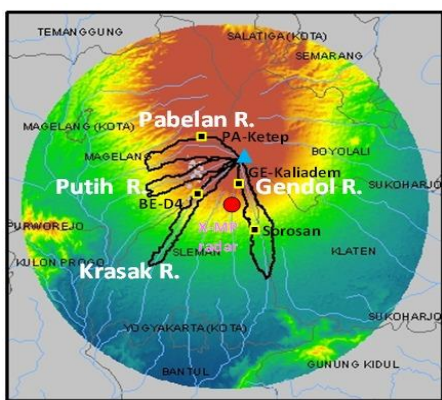


Gambar 1. Global Satelit Rainfall

### 2.3 Sensor Geofon

Salah satu instrumentasi yang digunakan untuk memonitoring getaran akibat aktivitas vulkanik adalah *Real-time Siesmic Amplitude Measurement (RSAM)*. RSAM merupakan sistem yang menyediakan pengukuran berkelanjutan mengenai amplitudo seismik rata-rata absolut dari jumlah stasiun *seismic*. RSAM menempati peranan strategis dalam monitoring aktivitas seismik gunung api terutama pada saat krisis menjelang erupsi. Geofon mengubah energi seismik, atau vibrasi, menjadi tegangan listrik yang dapat diukur secara akurat.

Sistem pengukuran amplitudo seismik *real-time* terdiri dari komputer laptop model 100 Tandy (*Radio Shack*), unit akuisisi data *Elexor*, dan pengkondisi sinyal / papan multipleksor yang dirancang sendiri. Di Gunung Merapi, RSAM yang digunakan adalah sensor geofon L4C. RSAM yang ada saat ini, melakukan sampling dengan frekuensi 60 Hz untuk tiap stasiun, atau 60 sampel per detik. Untuk masing-masing stasiun amplitudo dirata-ratakan untuk periode 10 menit dan dikirim ke memori komputer [10]. Gambar 2 dibawah ini adalah gambar letak geofon yang berada di 3 titik, yaitu di Stasiun Gendol (-7.685500°, 110.471389°, 319 m AMSL). Stasiun Putih1 (-7.577083°, 110.372055°, 730 m AMSL), dan stasiun Putih2 (-7.584623°, 110.354679°, 612 m AMSL).



Gambar 2. Lokasi Sensor Geofon

2.4 K-Nearest Neighbor

Algoritma K-NN (K-Nearest Neighbor) adalah algoritma untuk melakukan klasifikasi terhadap objek dimana hasil dari *instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori K-tetangga terdekat atau berdasarkan data pembelajaran yang jaraknya paling dekat dengan obyek tersebut. Tujuan dari algoritma ini adalah mengklasifikasi objek baru berdasarkan atribut dan sampel latih. Pengklasifikasian tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik uji, maka akan ditemukan sejumlah K objek (titik *training*) yang paling dekat

dengan titik uji [11][12]. Langkah-langkah untuk menghitung metode *k-Nearest Neighbor* antara lain:

1. Menentukan parameter K
2. Menghitung jarak *Euclidean Distancen*: Menghitung jarak antara data yang akan dievaluasi dengan semua pelatihan menggunakan perhitungan jarak *Euclidean* untuk menentukan jumlah kemiripan atau dekat atau jauhnya tetangga. Persamaan dari Jarak Euclidean adalah:

$$d(x, y) = (\sqrt{\sum_{i=1}^n (x_i - y_i)^2}) \quad (1)$$

Keterangan:

- $x_i$  : Sampel data
- $y_i$  : Data uji
- $i$  : Variabel data
- $d(x, y)$  : *Dissimilarity*/jarak
- $n$  : Dimensi data

3. Mengurutkan jarak yang terbentuk dari yang terkecil sampai terbesar.
4. Menentukan jarak terdekat sampai urutan K atau mengumpulkan kategori Y (Klasifikasi *nearest neighbor* berdasarkan nilai k).
5. Dengan menggunakan kategori *nearest neighbor* yang paling mayoritas maka dapat diprediksikan kategori objek.

2.5 K-Fold Cross Validation

*K-Fold Cross Validation* merupakan teknik validasi yang membagi keseluruhan data menjadi dua subset yaitu data latih dan data uji. Seluruh data diacak dan dibagi menjadi K bagian atau subset dan juga dilakukan percobaan sebanyak K. Secara umum pengujian nilai K dilakukan sebanyak 10 kali untuk memperkirakan akurasi estimasi. Dalam penelitian ini juga nilai K yang digunakan berjumlah 10 atau *10-fold Cross Validation* [13]. *10-Fold Cross Validation*, data dibagi menjadi 10 *fold* berukuran kira-kira sama, sehingga kita memiliki 10 subset data untuk mengevaluasi kinerja model atau algoritma. Untuk masing-masing dari 10 subset data tersebut, maka akan menggunakan *9fold* untuk pelatihan dan *1fold* untuk pengujian. Setiap data untuk pengujian akan dilakukan prediksi dengan menggunakan metode K-NN dengan K yang bervariasi.

$$b = \frac{n}{k} \quad (2)$$

Keterangan:

- $b$  : *10-Fold Cross Validation*
- $n$  : Jumlah data
- $k$  : Nilai *k-Fold Cross Validation*

Akurasi model akan diuji menggunakan data uji pada setiap *fold*, dan berlanjut ke *fold* selanjutnya sampai selesai. Akurasi akan dirata-rata setiap K pada

K-NN dan tingkat akurasi yang tinggi itulah yang terpilih menjadi nilai K terbaik atau K optimal pada K-NN. Tingkat akurasi dapat dihitung dengan jumlah data benar dibagi dengan seluruh data uji dikali 100.

### 3 Metodologi

#### 3.1 Pengolahan Data

Data yang digunakan adalah data dari primer dari metode pengambilan data yaitu data satelit untuk curah hujan dan data sensor geofon atau getaran pada masing-masing daerah yang dijadikan daerah penelitian. Langkah pertama adalah data curah hujan di *download* berupa .csv melalui situs dari JAXA *Global Rainfall Watch System* dengan memasukan tanggal-bulan-tahun, Negara beserta dengan *Latitude* dan *Longitude* [14]. Setelah didapat data curah hujan, langkah kedua adalah mencari nilai curah hujan akumulasi dengan cara menghitung nilai curah hujan seminggu sebelumnya. Karena data curah hujan tersebut waktunya perjam sedangkan pada waktu data getarannya per menit, maka langkah ketiga adalah menggabungkan data curah hujan dengan data getaran dengan cara menyesuaikan waktu getaran sesuai dengan waktu curah hujan yaitu perjam. Maka hasil dari data gabungan tersebut terdiri dari waktu, curah hujan, curah hujan akumulasi, dan nilai getaran. Langkah keempat adalah mengklasifikasikan nilai getaran menjadi 4 kelas seperti Tabel 1 dibawah ini yaitu banjir rendah, sedang, tinggi dan tidak terjadi banjir lahar pada tiap lokasi.

Table 1 Klasifikasi Banjir Lahar

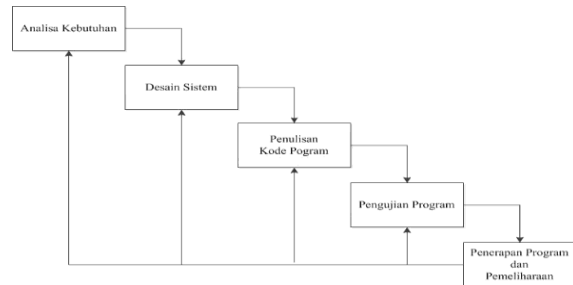
Status	Gendol	Putih 1	Putih 2
Tidak Terjadi Banjir Lahar	0	0	0
Banjir Rendah	1 – 2.500	1 – 2.000	1 – 1.650
Banjir Sedang	2.501 – 5000	2.001 – 4.000	1.651 – 3.300
Banjir Tinggi	5.001 – 7.500	4.001 – 6000	3.301 – 5000

Dari data gabungan beserta klasifikasi banjir lahar akan digunakan model *10-Fold Cross Validation* untuk perhitungan K-NN agar mendapatkan keluaran yang diharapkan yaitu klasifikasi status getaran banjir lahar berdasarkan dari *inputan* berupa curah hujan dan akumulasi hujan [15].

#### 3.2 Pengembangan Perangkat Lunak

Sistem prediksi banjir lahar dingin pada daerah Lereng Merapi menggunakan proses pengembangan perangkat lunak yaitu metode *waterfall*. Metode *waterfall* merupakan metode yang sering digunakan oleh penganalisa sistem pada umumnya. Inti dari metode *waterfall* adalah pengerjaan dari suatu sistem dilakukan secara berurutan atau secara

linear. Jadi jika langkah ke-1 belum dikerjakan, maka langkah 2 tidak dapat dikerjakan. Jika langkah ke-2 belum dikerjakan maka langkah ke-3 juga tidak dapat dikerjakan, begitu seterusnya. Tahapan pada metode *waterfall* dapat dilihat pada gambar 3 dibawah ini.



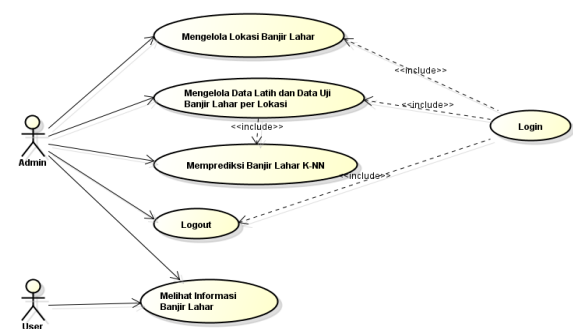
Gambar 3. Metode Waterfall

##### a. Analisis Kebutuhan

Pada analisa kebutuhan sistem prediksi banjir lahar dingin pada Lereng Merapi adalah untuk memprediksi getaran banjir lahar pada daerah Lereng Merapi khususnya di 3 titik daerah studi. Kebutuhan data yang diperlukan adalah curah hujan dan akumulasi hujan dari data satelit dan getaran dari sensor geofon.

##### b. Desain Sistem

Desain sistem yang digunakan pada penelitian ini adalah desain sistem dalam bentuk *use case* diagram untuk mengetahui fungsi apa saja yang ada di dalam sebuah sistem dan siapa saja yang berhak menggunakan fungsi-fungsi tersebut. Pada *use case* sistem prediksi banjir lahar dingin dapat dilihat pada gambar 4 dibawah ini.



Gambar 4. Use Case Diagram

Gambar 4 terdapat 2 aktor yaitu Admin dan User. Fungsi Admin diawali dengan *login* terlebih dahulu, kemudian setelah Admin *login* maka Admin dapat mengelola data lokasi, data latih dan data uji yang digunakan untuk memprediksi status getaran banjir lahar dingin pada lereng merapi menggunakan metode K-NN. Sedangkan fungsi pada User adalah dapat melihat informasi status getaran banjir lahar pada daerah Lereng Merapi.

c. Penulisan Kode Program

Pada penulisan kode program dilakukan dengan menggunakan bahasa pemrograman yang telah ditentukan yaitu PHP (*Hypertext Preprocessor*) dengan *framework* atau kerangka kerja *Codeigniter* menggunakan metode pengembangan berbasis MVC (*Model-View-Controller*) dan juga dengan proses metode K-NN.

d. Pengujian Program

Tahapan pengujian digunakan untuk memastikan sistem yang dibangun memiliki kelayakan dan sesuai dengan kebutuhan dan memastikan bahwa semua bagian sudah diuji baik secara fungsional yang disebut dengan pengujian *blackbox*. Pengujian program juga dilakukan dengan mencari nilai K optimal K-NN dari rata-rata akurasi tertinggi dengan metode *10-Fold Cross Validation*. Pengujian dilakukan untuk meminimalisir kesalahan (*error*) dan memastikan keluaran yang dihasilkan sesuai dengan yang diinginkan yaitu status getaran banjir lahar.

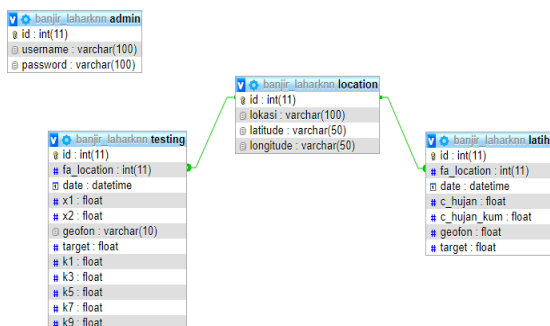
e. Penerapan Program dan Pemeliharaan

Tahap ini adalah akhir dalam model *waterfall*. Jika semua tahapan sudah di selesaikan dan sudah menjadi sistem yang diharapkan, maka akan dijalankan serta dilakukan pemeliharaan. Pemeliharaan bertujuan dalam memperbaiki kesalahan yang tidak ditemukan pada langkah testing sebelumnya.

4 Implementasi

4.1 Implementasi Database

Implementasi basis data yang dilakukan sesuai dengan perancangan basis data. Basis data yang dibuat diberi nama “banjir\_laharknn” menggunakan *MySQL*. Implementasi basis data adalah seperti pada gambar 5 dibawah ini.



Gambar 5. Implementasi Database

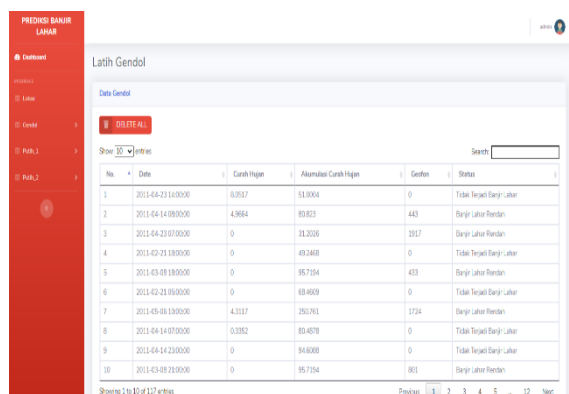
Gambar 5 terdapat beberapa table yaitu tabel admin, latihan, location, dan testing.

4.2 Implementasi Halaman

Implementasi antarmuka (*interface*) dari perangkat lunak dilakukan berdasarkan rancangan yang telah dilakukan. Implementasi ditampilkan dari screenshot dari halaman website yang digunakan sebagai bahan penelitian yang telah dirincikan pada bab sebelumnya.

a. Halaman Data Latih

Gambar 6 dibawah ini merupakan implementasi halaman pada data latih.

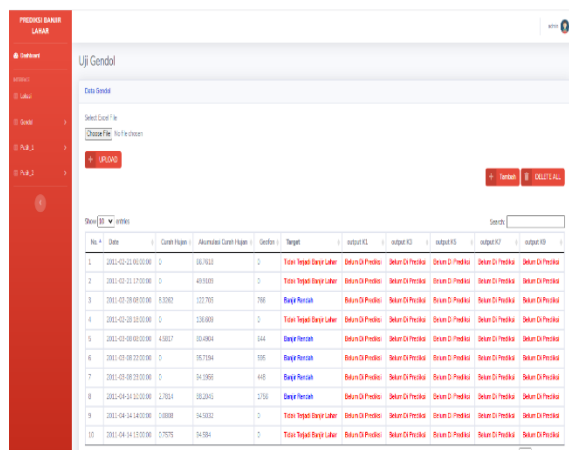


Gambar 6. Implementasi Halaman Data Latih

Gambar 6 adalah implementasi Admin untuk melakukan pengelolaan data latih tiap lokasi dengan *upload excel* ke dalam sistem..

b. Halaman Data Uji

Gambar 7 dibawah ini merupakan implementasi halaman pada data uji.



Gambar 7. Implementasi Halaman Data Uji

Gambar 7 adalah implementasi Admin untuk melakukan pengelolaan data uji tiap lokasi dengan *upload excel* ke dalam sistem. “Tambah” digunakan untuk menambah data uji untuk diprediksi. “K-Nearest Neighbor” digunakan untuk melakukan proses perhitungan algoritma K-NN pada semua data latih dan data uji tiap lokasi.

c. Halaman *K-Nearest Neighbor*

Gambar 8 dibawah ini merupakan implementasi hasil perhitungan menggunakan metode K-NN.

No.	Target	K1	K3	K5	K7	K9
1	Tidak Terjadi Banjir Lahar	TRUE	TRUE	TRUE	TRUE	TRUE
2	Tidak Terjadi Banjir Lahar	TRUE	TRUE	TRUE	TRUE	TRUE
3	Banjir	TRUE	TRUE	FALSE	FALSE	FALSE
4	Tidak Terjadi Banjir Lahar	TRUE	TRUE	TRUE	TRUE	TRUE
5	Banjir	TRUE	TRUE	TRUE	FALSE	FALSE
6	Banjir	TRUE	TRUE	TRUE	TRUE	TRUE
7	Banjir	FALSE	FALSE	FALSE	FALSE	FALSE
8	Banjir	TRUE	TRUE	TRUE	TRUE	TRUE
9	Tidak Terjadi Banjir Lahar	TRUE	TRUE	TRUE	TRUE	TRUE
10	Tidak Terjadi Banjir Lahar	TRUE	TRUE	TRUE	TRUE	TRUE
11	Banjir	FALSE	TRUE	FALSE	TRUE	TRUE
12	Tidak Terjadi Banjir Lahar	TRUE	TRUE	TRUE	TRUE	TRUE
13	Tidak Terjadi Banjir Lahar	TRUE	TRUE	TRUE	TRUE	TRUE

Gambar 8. Implementasi Halaman *K-Nearest Neighbor*

Gambar 8 adalah implementasi Admin dalam melakukan proses algoritma K-NN untuk memprediksi status getaran banjir lahar beserta dengan akurasi nya.

d. Halaman *User*

Gambar 9 dibawah ini merupakan implementasi hasil perhitungan menggunakan metode K-NN pada *User*.

No.	Date	Cuan Nyan	Alameda Cuan Nyan	Status
1	2020-10-02 08:00:00	2.234	10.000	Tidak Terjadi Banjir Lahar
2	2020-10-02 08:00:00	2.234	10.000	Banjir
3	2020-10-02 08:00:00	2.234	10.000	Banjir
4	2020-10-02 08:00:00	2	10.000	Tidak Terjadi Banjir Lahar
5	2020-10-02 08:00:00	2.234	10.000	Tidak Terjadi Banjir Lahar
6	2020-10-02 08:00:00	2	10.000	Tidak Terjadi Banjir Lahar
7	2020-10-02 08:00:00	2	10.000	Tidak Terjadi Banjir Lahar
8	2020-10-02 08:00:00	2.234	10.000	Banjir
9	2020-10-02 08:00:00	2.234	10.000	Banjir
10	2020-10-02 08:00:00	2	10.000	Tidak Terjadi Banjir Lahar
11	2020-10-02 08:00:00	2	10.000	Tidak Terjadi Banjir Lahar
12	2020-10-02 08:00:00	2	10.000	Banjir
13	2020-10-02 08:00:00	2.234	10.000	Tidak Terjadi Banjir Lahar

Gambar 9. Implementasi Halaman *User*

Gambar 9 adalah implementasi halaman *user* untuk melihat hasil status getaran banjir lahar pada tiap lokasi.

5 Pengujian Pembahasan

5.1 10-Fold Cross Validation

Pengujian 10-Fold Cross Validation adalah teknik validasi dengan membagi data secara acak kedalam 10 bagian dan masing-masing bagian akan dilakukan percobaan sebanyak 10 kali dari data latih dan data uji dan dengan nilai K pada K-NN secara bervariasi untuk mengetahui K-optimal dari akurasi tertinggi di tiap lokasi. Klasifikasi banjir lahar pada penelitian ini terdiri dari 4 klasifikasi yaitu terjadi

banjir lahar dalam tingkat rendah, sedang, tinggi, dan tidak terjadi banjir. Pengujian penentuan nilai K optimal menggunakan bilangan ganjil dikarenakan ada 4 klasifikasi banjir. Terdapat 5 nilai K yang digunakan pada setiap bagian adalah 1,3,5,7 dan 9 agar menghasilkan nilai yang valid atau mendapatkan klasifikasi paling mayoritas, sedangkan jika nilai k menggunakan bilangan genap hasil yang didapatkan akan memungkinkan untuk menghasilkan klasifikasi yang sama sehingga tidak didapatkan klasifikasi yg paling mayoritas. Di bawah ini adalah pengujian 10-Fold Cross Validation pada lokasi Gendol, Putih 1 dan Putih 2.

a. Gendol

Lokasi Gendol mempunyai dataset berjumlah 130 data yang dibagi menjadi 117 data latih dan 13 data uji dan akan dilakukan proses 10-Fold Cross Validation untuk mencari nilai K-optimal. Tabel 2 adalah hasil dari 10-Fold Cross Validation.

Tabel 2 10-Fold Cross Validation Gendol

10-Fold Cross Validation					
K	K1 (%)	K3 (%)	K5 (%)	K7 (%)	K9 (%)
1	84.62	92.31	76.92	76.92	76.92
2	76.92	76.92	76.92	84.62	69.23
3	53.85	38.46	30.77	30.77	38.46
4	69.23	69.23	69.23	69.23	69.23
5	84.62	84.62	76.92	69.23	61.54
6	61.54	69.23	61.54	76.92	69.23
7	69.23	69.23	76.92	76.92	76.92
8	53.85	76.92	76.92	46.15	53.85
9	69.23	69.23	76.92	76.92	76.92
10	69.23	76.92	76.92	69.23	69.23
<b>Rata - Rata</b>	<b>69.23</b>	<b>72.30</b>	<b>69.99</b>	<b>67.69</b>	<b>66.15</b>
	<b>2</b>	<b>7</b>	<b>8</b>	<b>1</b>	<b>3</b>

Tabel 2 adalah hasil pengujian 10-Fold Cross Validation dengan tiap nilai K pada K-NN untuk mencari nilai K optimal. Berdasarkan tabel diatas, menunjukkan bahwa K tertinggi adalah K=3 dengan akurasi rata-rata 72.307 %.

b. Putih 1

Lokasi Putih 1 mempunyai dataset berjumlah 210 data yang dibagi menjadi 189 data latih dan 21 data uji. Tabel 3 adalah hasil dari 10-Fold Cross Validation.

Tabel 3 10-Fold Cross Validation Putih 1

10-Fold Cross Validation					
K	K1 (%)	K3 (%)	K5 (%)	K7 (%)	K9 (%)
1	76.19	80.95	80.95	71.43	61.9
2	76.19	76.19	71.43	76.19	76.19
3	80.95	90.48	95.24	95.24	90.48
4	85.71	90.48	90.48	90.48	80.95
5	80.95	85.71	80.95	71.43	71.43
6	80.95	85.71	76.19	80.95	85.71
7	61.9	76.19	71.43	76.19	76.19
8	95.24	95.24	80.95	80.95	85.71
9	66.67	66.67	57.14	47.62	47.62
10	76.19	66.67	66.67	61.9	61.9
<b>Rata - Rata</b>	<b>78.094</b>	<b>81.429</b>	<b>77.143</b>	<b>75.238</b>	<b>73.808</b>

Tabel 3 adalah hasil pengujian 10-Fold Cross Validation dengan tiap nilai K pada K-NN untuk mencari nilai K optimal. Berdasarkan tabel diatas, menunjukkan bahwa K tertinggi adalah K=3 dengan akurasi rata-rata 81.429 %.

c. Putih 2

Lokasi Putih 2 mempunyai dataset berjumlah 230 data yang dibagi menjadi 207 data latih dan 23 data uji. Tabel 4 adalah hasil dari 10- Fold Cross Validation.

Tabel 4 10-Fold Cross Validation Putih 2

10-Fold Cross Validation					
K	K1 (%)	K3 (%)	K5 (%)	K7 (%)	K9 (%)
1	91.3	91.3	82.61	82.61	82.61
2	86.96	78.26	69.57	69.57	65.22
3	73.91	69.57	69.57	69.57	69.57
4	91.3	82.61	78.26	82.61	82.61
5	91.3	86.96	82.61	82.61	82.61
6	78.26	82.61	86.96	82.61	78.26
7	86.96	86.96	82.61	86.96	82.61
8	95.65	100	100	100	100
9	82.61	91.3	86.96	86.96	86.96
10	91.3	86.96	91.3	91.3	91.3
<b>Rata-Rata</b>	<b>86.955</b>	<b>85.653</b>	<b>83.045</b>	<b>83.488</b>	<b>82.175</b>

Tabel 4 adalah hasil pengujian 10-Fold Cross Validation dengan tiap nilai K pada K-NN untuk

mencari nilai K optimal. Berdasarkan tabel diatas, menunjukkan bahwa K tertinggi adalah K=1 dengan akurasi rata-rata 86.955 %.

Pada pengujian penelitian di atas terlihat bahwa hasil pengujian K-3 pada beberapa lokasi menunjukkan hasil yang lebih baik, hal tersebut dikarenakan pada pengujian K-3 lebih sering ditemukan hasil yang paling mayoritas dan sesuai dengan hasil output data latih. Misal K-3 dibandingkan dengan K-5, pada K-3 terdapat hasil klasifikasi banjir tingkat sedang sebanyak 1 dan rendah sebanyak 2 maka pada K-3 hasil yang mayoritas adalah banjir tingkat rendah. Sedangkan pada K-5 bisa terdapat hasil klasifikasi sedang sebanyak 1, rendah sebanyak 2, dan tinggi sebanyak 2 maka pada hasil ini tidak bisa ditentukan data yang mayoritas dikarenakan hasil klasifikasi rendah dan tinggi berjumlah sama. Maka dari itu perlu adanya 10-Fold Cross Validation, untuk validasi data dan dilakukan pengujian sebanyak 10 kali, dengan beberapa nilai K untuk bisa diminimalisir dan dari hasil klasifikasi tersebut dicari tingkat akurasi dan kemudian dicari hasil rata-rata akurasi, akurasi tertinggi itu lah yang nantinya akan menjadi K optimal yang digunakan untuk data testing. Begitu juga hal tersebut dilakukan untuk nilai K lainnya di setiap lokasi.

## 6 Kesimpulan dan Saran

### 6.1 Kesimpulan

Dari penelitian yang telah dilakukan sampai dengan tahap pengujian maka dapat diambil kesimpulan bahwa penelitian ini berhasil dalam membangun dan merancang sebuah sistem prediksi banjir lahar pada Lereng Merapi yang dapat digunakan untuk membantu dalam memprediksi dan memberikan informasi getaran dengan menampilkan status terjadi banjir lahar dan tidak terjadi banjir lahar.

Algoritma K-Nearest Neighbor dapat diterapkan dalam menentukan klasifikasi status getaran banjir lahar pada tiap lokasinya dengan hubungan dari parameter curah hujan dan akumulasi hujan khususnya di Kali Putih 1, Putih 2 dan Gendol yang rawan terjadinya banjir lahar dan juga dekat dengan permukiman penduduk. Nilai K optimal pada algoritma K-Nearest Neighbor yang digunakan untuk prediksi banjir lahar pada tiap lokasinya berdasarkan dari akurasi rata-rata dengan menggunakan 10-Fold Cross Validation adalah Gendol dan Putih 1 menggunakan 3-NN dengan akurasi rata-rata 72.307% (Gendol) dan 81.429% (Putih 1), dan Putih 2 menggunakan 1-NN dengan akurasi rata-rata 86.955%.

## 6.2 Saran

Penulis mempunyai saran untuk pengembangan selanjutnya antara lain adalah diharapkan dalam penelitian lebih lanjut bisa menggunakan lebih banyak data latih, data uji dan parameter yang lebih bervariasi agar mendapatkan tingkat akurasi yang lebih tinggi. Keluaran pada sistem selain klasifikasi status getaran namun juga bisa dalam bentuk nilai getaran. Penelitian selanjutnya juga pada model algoritma *K-Nearest Neighbor* dapat dimodifikasi, dibandingkan maupun menggunakan algoritma klasifikasi dan peramalan lainnya.

### Daftar Pustaka:

- Sari, M. (2016): *Banjir Lahar, Pengertian, Ciri-ciri, Penyebab dan Dampaknya* [online]. Available:<https://ilmugeografi.com/bencana-alam/banjir-lahar>
- Aini, N. (2019): *Potensi Banjir Lahar Dingin Gunung Merapi Ancam Yogya* [online]. Available:<https://www.republika.co.id/berita-nasional/daerah/19/01/18/pli811382-potensi-banjir-lahar-dingin-gunung-merapi-ancam-yogya>
- Ismail, A. M. (2018): *Cara Kerja Algoritma k-Nearest Neighbor (k-NN)* [online]. Available:<https://medium.com/bee-solution-partners/cara-kerja-algoritma-k-nearest-neighbor-k-nn-389297de543e>
- Sugna, B. A. (2019): *Sistem Penentuan Status Bencana Banjir Pada Daerah Lereng Merapi. Malang: Politeknik Negeri Malang.*
- Hapsari, R. I., Sugan, B. A., Novianto, D., Asmara, R. A., & Oishi, S. (2019): *Predictability of Naïve Bayes classifier for lahar hazard. Retrieved from Predictability of Naïve Bayes classifier for lahar hazard:* Available:<https://iopscience.iop.org/article/10.1088/1755-1315/437/1/012049>
- Hapsari, R. I., Sugna, B. A., Asmara, R. A., & Oishi, S. (2020): *Naïve Bayes Classifier for Debris Flow Disaster Mitigation in Mount Merapi Volcanic Rivers, Indonesia, Using X-band Polarimetric Radar* [online]. Available:<https://link.springer.com/article/10.1007/s13753-020-00321-7>
- Noviansyah, M. R. (2018): *Penerapan Data Mining menggunakan Metode K-nearset Neighbor untuk Klasifikasi Indeks Cuaca Kebakaran Berdasarkan Data AWS (AUTOMATIC WEATHER STATION) (STUDI KASUS: KABUPATEN KUBU RAYA). Jurnal Coding, Sistem Komputer Untan, 48-56.*
- Tempola, F. (2018): *PERBANDINGAN KLASIFIKASI ANTARA KNN DAN NAIVE BAYES PADA. Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK), 577-584.*
- Bowden, G. B. (2003): *Calibration of Geophone Microseismic Sensors.* Stanford University, Stanford: Stanford Linear Accelerator Center.
- Elliot TEndo, T. M. (1991): *Real-time Seismic Amplitude Measurement (RSAM). a volcano monitoring and prediction tool*, hal. 533-545.
- Ketutrare. (2018): *Algoritma K-Nearest Neighbor dan Contoh Soal* [online]. Available:<https://www.ketutrare.com/2018/11/algorithm-k-nearest-neighbor-dan-contoh-soal.html>
- Advernesia. (2017): *Pengertian dan Cara Kerja Algoritma K-Nearest Neighbors (KNN)* [online]. Available:<https://www.advernesia.com/blog/d-ata-science/pengertian-dan-cara-kerja-algoritma-k-nearest-neighbours-knn/>.
- University, B. (2017): *10 FOLD-CROSS VALIDATION* [online]. Available:<https://mti.binus.ac.id/2017/11/24/10-fold-cross-validation/>.
- GSMAP. (2011): *JAXA GLOBAL RAINFALL WATCH* [online]. Available:<https://sharaku.eorc.jaxa.jp/GSMaP>
- Siska.stevani@gmail.com. (2020): *Data Banjir Lahar* [online]. Available:<https://drive.google.com/drive/folders/1akQfBMFTTN5bVo2Ke4z3SxCrDmx4GgVX>