

# **Implementasi Machine Learning Pada Sistem Pemetaan Daerah Rawan Banjir Di Desa Pahlawan Kabupaten Batu Bara**

**Zulham Sitorus, Eko Hariyanto, Fahmi Kurniawan\***

Fakultas Sain Dan Teknologi, Program Studi Sistem Komputer, Universitas Pembangunan Panca Budi, Medan, Indonesia  
Email: <sup>1</sup>zulhamsitorus@dosen.pancabudi.ac.id, <sup>2</sup>ekohariyanto@dosen.pancabudi.ac.id, <sup>3,\*</sup>fahmikurniawan@dosen.pancabudi.ac.id  
Email Penulis Korespondensi: fahmikurniawan@dosen.pancabudi.ac.id

**Abstrak**—Salah satu daerah di Kabupaten Batu Bara Desa Pahlawan Kecamatan Tanjung Tiram memiliki luas 173,79 km<sup>2</sup> dan terletak di daerah dataran rendah dengan ketinggian 0,-4.5M yang berbatasan langsung dengan Selat Malaka disebelah Timur. Dimana hampir separuh daerahnya terkena pasang surut air laut, Desa pahlawan beriklim tropis dengan dua musim yaitu musim hujan dan musim kemarau, Masyarakat yang tinggal di Desa Pahlawan Kecamatan Tanjung Tiram. Begitu banyak kendala yang dihadapi oleh masyarakat Desa Pahlawan antara lain permasalahan banjir yang mengakibatkan dampak Kesehatan dan perekonomian bagi masyarakat. Kurangnya penyuluhan dan pengetahuan, serta kesadaran masyarakat akan terjadinya banjir pada saat pasang air besar, dan apabila musim penghujan akan meningkatkan debit air pada permukaan laut akan naik, sehingga bisa menyebabkan terjadinya banjir. Dalam penelitian ini Implementasi Machine Learning digunakan sebagai Sistem Pemetaan Daerah Rawan Banjir Di Desa Pahlawan Kabupaten Batu Bara, dengan analisis data yang digunakan adalah data primer dan skunder yang bersifat kualitatif dan kuantitatif. Akibat terjadinya banjir dampak kerugiannya yang mempengaruhi material dan non-material, maka sangat penting untuk memetakan daerah rawan banjir untuk perencanaan pembangunan daerah. Identifikasi potensi banjir melibatkan machine learning menggunakan metode Random Forest, dengan mempertimbangkan faktor pemicu banjir. Metode Random Forest juga memberikan parameter sensitivitas menggunakan kurva Receiver Operating Characteristic (ROC) yang menunjukkan area rentan banjir misalnya Desa Pahlawan, Kecamatan Tanjung Tiram.

**Kata Kunci:** Machine Learning; Banjir; Desa Pahlawan

**Abstract**—One of the areas in Batu Bara Regency, Pahlawan Village, Tanjung Tiram District, has an area of 173.79 km<sup>2</sup> and is located in a lowland area with an altitude of 0.-4.5m which is directly adjacent to the Malacca Strait to the east. Where almost half of the area is affected by sea tides, Hero Village has a tropical climate with two seasons namely the rainy season and the dry season. The people who live in Pahlawan Village, Tanjung Tiram District. There are so many obstacles faced by the people of Pahlawan Village, including the problem of flooding which has an impact on the health and the economy of the community. Lack of counseling and knowledge, as well as public awareness of the occurrence of flooding during high tides, and when the rainy season will increase the water discharge at sea level will rise so that it can cause flooding. In this study, the implementation of machine learning was used as a mapping system for flood-prone areas in Pahlawan Village, Batu Bara District, with data analysis used using primary and secondary data, both qualitative and quantitative. Due to the occurrence of floods, and the impact of losses that affect material and non-material, it is very important to map flood-prone areas for regional development planning. Identification of potential flooding involves machine learning using the Random Forest method, taking into account the triggering factors for flooding. The Random Forest method also provides sensitivity parameters using a Receiver Operating Characteristic (ROC) curve which indicates areas prone to flooding, for example, Pahlawan Village, Tanjung Tiram District.

**Keywords:** Machine Learning; Flood; Hero's Village

## **1. PENDAHULUAN**

Banjir dan genangan merupakan masalah tahunan dan memberikan pengaruh besar terhadap kondisi masyarakat baik secara sosial, ekonomi maupun lingkungan. Banjir bukan merupakan masalah pribadi yang diteliti berdasarkan disiplin ilmu, tetapi banjir diakibatkan sistem lingkungan yang rusak dan mata rantai fisis lingkungan yang terganggu, sehingga untuk mengatasi masalah banjir perlu dikaji secara terpadu. Banjir di Desa Pahlawan, Kecamatan Tanjung Tiram Kabupaten Batu Bara atau daerah lainnya disebabkan pada badan aluran air terjadi pendangkalan dan dekat dengan garis pantai, serta penumpukan sampah pada aliran air yang ada di Desa Pahlawan.

Peningkatan debit aliran pada air laut dan curah hujan yang turun sebagai akibat tidak/kurang meresap ke dalam tanah, sehingga curah hujan menjadi aliran permukaan. Aliran permukaan yang bergerak di permukaan tanah mengikis tanah dan membawa ke badan sungai, karena itu aliran sungai bukan saja debit meningkat juga ditambah material hasil erosi. Fokus besar dalam penelitian kali ini adalah machine learning yang dapat menerapkan bagaimana mendeteksi secara berkala pada pola kompleks dan membuat keputusan cerdas berdasarkan data yang tersedia pada tahun-tahun sebelumnya. Machine learning dapat mempelajari pola dari data historis yang terdapat pada database tahun-tahun sebelumnya untuk pemetaan banjir dari akibat curah hujan dan air pasang dari laut, hal tersebut dapat dilakukan untuk mengetahui seberapa besar fungsi dari hasil pemetaan banjir di daerah Desa Pahlawan.

Machine learning memiliki makna berupa metode yang digunakan untuk membuat program yang bisa belajar dari data yang tersedia dari beberapa sumber yang dapat di percaya untuk di gunakan data tersebut. Berbeda dengan program komputer biasa dilakukan secara statis karena sudah ada ketentuan apa yang harus di lakukan terlebih dahulu, namun berbeda dengan program machine learning merupakan sebuah program yang dirancang untuk mampu belajar sendiri dengan data yang telah tersedia. Hasil penelitian yang dilakukan nantinya dapat membantu masyarakat Desa Pahlawan mengetahui titik-titik rawan banjir pada daerah yang akan digunakan oleh masyarakat tersebut dan juga untuk meningkatkan kualitas sumbu daya manusia. Sistem yang dirancang juga dapat memberikan pengetahuan lebih luas untuk mengembangkan konsep machine learning dalam ilmu pengetahuan.

## 2. METODOLOGI PENELITIAN

### 2.1 Materi Penelitian

Data yang digunakan untuk menganalisa meningkatkan kemampuan minat baca yaitu data dari hasil FGD dan hasil survei lapangan di Desa Pahlawan. Pada bab ini akan diuraikan metodologi penelitian dan kerangka kerja penelitian yang digunakan dalam penyelesaian penelitian ini. Kerangka kerja ini merupakan langkah-langkah yang akan dilakukan dalam rangka penyelesaian masalah yang akan dibahas. Adapun kerangka kerja dalam penelitian ini digambarkan pada gambar dibawah.



Gambar 1. Kerangka Penelitian

Metode yang dipakai dalam analisis data adalah metode Analisis Penelitian Kuantitatif, dimana data-data yang diperoleh berdasarkan dari hasil statistik, pengukuran dan pengujian yang dilakukan dilapangan.

Metodologi penelitian berhubungan erat pengumpulan informasi dari suatu data yang ingin di angkat pada judul ini dan diambil pada beberapa website resmi, seperti BMKG, BPS, dan BNPD kabupaten Batu Bara. pada penelitian ini berupa data bencana banjir yang ada di daerah kabupaten Batu Bara, Selain itu penggunaan teknik machine learning atau pembelajaran mesin dapat digunakan untuk meningkatkan ketangguhan dan kesiapsiagaan data yang didapatkan terhadap bencana banjir.

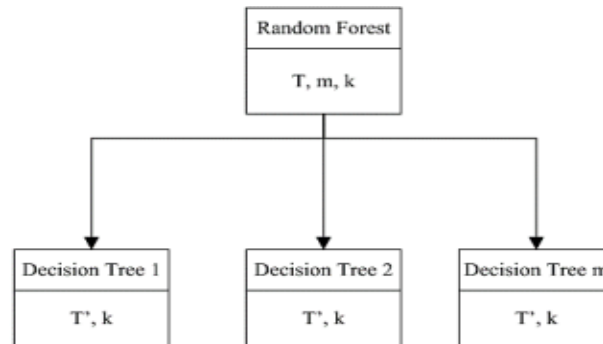


Gambar 2. Peta Daerah Kabupaten Batu Bara

Penggunaan model pembelajaran mesin untuk memprediksi banjir secara spasial merupakan bidang riset yang disarankan untuk penelitian selanjutnya. Metode machine learning telah diterapkan dalam penilaian resiko, dampak banjir, dan dampak iklim yang khususnya dalam pengembangan langkah-langkah mitigasi, persiapan tanggap darurat dan perencanaan pemulihan banjir. Penelusuran dilakukan dari berbagai sumber seperti buku, jurnal, website tertentu, khususnya penelusuran melalui internet, Tujuannya adalah didapakkannya gambaran terkini penelitian dan penerapan upaya mitigasi bencana banjir. Penelitian ini bertujuan untuk mengetahui hasil nilai accuracy yang di tampilkan jika penelitian tersebut menggunakan Metode Random Forest yang juga memberikan parameter sensitivitas menggunakan kurva Receiver Operating Characteristic dalam pemetaan daerah rawan banjir.

Pemanfaatan machine learning dengan mengabungkan sistem random forest merupakan metode bagging yaitu metode yang membangkitkan sejumlah tree dari data sample dimana pembuatan satu tree pada saat training tidak

bergantung pada tree sebelumnya kemudian keputusan diambil berdasarkan voting terbanyak (Wibowo, Saikhu, & Soelaiman, 2016). Dua konsep yang menjadi dasar dari random forest adalah membangun ensemble dari tree via bagging dengan replacement dan penyeleksian fitur secara acak untuk tiap tree yang dibangun. Pertama, setiap sample yang diambil dari dataset untuk training tree bisa dipakai lagi untuk training tree yang lain. Kedua, fitur yang digunakan pada saat training untuk tiap tree merupakan subset dari fitur yang dimiliki oleh dataset (Wibowo et al., 2016). Random forest mempunyai dua parameter utama, yaitu: m jumlah tree yang akan dipakai dan k yaitu maksimal banyaknya fitur yang dipertimbangkan ketikan proses percabangan. Semakin banyak nilai m maka semakin bagus hasil klasifikasi, sedangkan untuk nilai k direkomendasikan sebesar akar kuadrat atau logaritma dari jumlah total fitur (Wibowo et al., 2016).



Gambar 3. Ilustrasi Random Forest

Gambar diatas menunjukkan proses training untuk random forest menggunakan dataset T dengan sejumlah m tree sebagai basic leaner dan k fitur yang dipilih secara acak dari total fitur yang ada untuk percabangan pada setiap tree. Proses training pada setiap tree menggunakan dataset T' yang merupakan hasil dari bootstrap dari dataset yang dijadikan parameter untuk random forest. Bootstrap merupakan proses memilih sample dari dataset yang akan digunakan proses training tree. Metode ensemble, bootstrap merupakan proses sampling dengan replacement, sehingga sample yang diambil untuk proses training tree yang satu masih bisa dipakai lagi untuk proses training tree yang lainnya (Wibowo et al., 2016). Nilai akurasi GCP ditunjukkan oleh nilai *Root Mean Square Error* (RMSError). RMS-error menyatakan nilai kesalahan dari proses koreksi geometrik. Akurasi yang baik ditunjukkan oleh nilai RMS-error yang sangat kecil mendekati nol. "*Root Mean Square Error (RMSE) merupakan besarnya tingkat kesalahan hasil prediksi, dimana semakin kecil (mendekati 0) nilai RMSE maka hasil prediksi akan semakin akurat.*" Perhitungan RMS-error dengan menggunakan persamaan.

$$RMSError = \sqrt{(X' - X_{orig}) + (Y' - Y_{orig})} \quad (1)$$

Keterangan:

X' dan Y' = Koordinat citra keluaran (output) Xorig dan Yorig= koordinat citra asli (input)

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil

Penelitian ini dilakukan di sejumlah wilayah yang rawan banjir di Desa Pahlawan Kabupaten Batu Bara Sumatera Utara Tahun 2022 pada saat tingginya intensitas hujan yang dapat menyebabkan banjir yang meliputi, 12 dusun merupakan dusun yang berada di wilayah Desa Pahlawan tergolong dalam zona rawan banjir yang tersebar 12 Dusun yang ada pada Desa Pahlawan.

Tabel 1. Daftar Dusun Desa Pahlawan

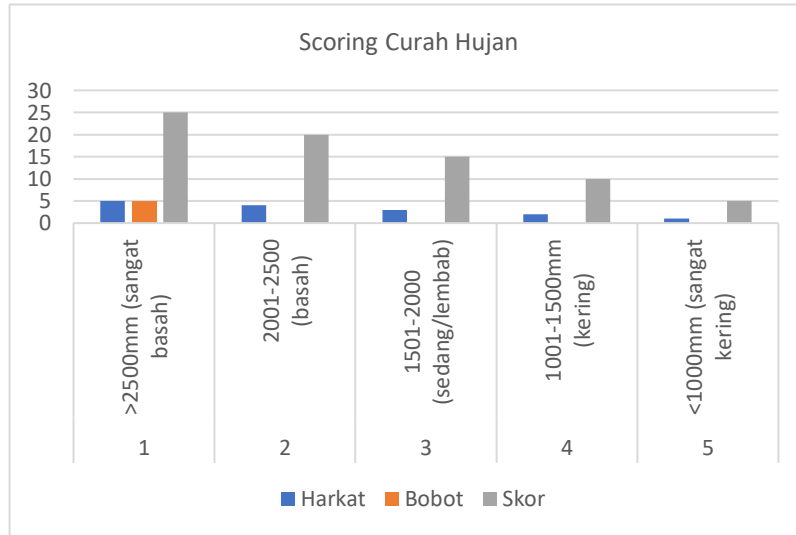
No.	Nama Dusun
1.	Dusun Adi Daya
2.	Dusun Amanah
3.	Dusun Bogak
4.	Dusun Pabrik
5.	Dusun Sejahtera
6.	Dusun Sejarah
7.	Dusun Bandar
8.	Dusun Bunga Jumpa
9.	Dusun LobaiAbbas
10.	Dusun Nelayan
11.	Dusun Nilam
12.	Dusun Wan Ahad

**a. Pengolahan Data Scoring Curah Hujan**

Peta curah hujan didapatkan dari hasil interpolasi data curah hujan bulanan yang kemudian dilakukan skoring pada tiap klasifikasi yang digunakan.

**Tabel 2.** Scoring Curah Hujan

No.	Kelas	Harkat	Bobot	Skor
1.	>2500mm (sangat basah)	5		25
2.	2001-2500 (basah)	4		20
3.	1501-2000 (sedang/lembab)	3	5	15
4.	1001-1500mm (kering)	2		10
5.	<1000mm (sangat kering)	1		5



**Gambar 4.** Ilustrasi Random Forest

Koreksi Radiometrik dilakukan dengan dua tahapan yaitu kalibrasi radiometric dan koreksi atmosfer. Kalibrasi radiometric digunakan untuk mengubah nilai digital number menjadi reflectance, sedangkan koreksi atmosfer digunakan untuk menghilangkan bias atmosferic yang ada pada citra.

Klasifikasi citra bertujuan untuk mengelompokkan dan melakukan segmentasi terhadap kenampakan-kenampakan yang homogen menggunakan klasifikasi citra terbimbing. Klasifikasi citra pada daerah penelitian akan dibagi menjadi beberapa kelas, yakni Tanah terbuka, sawah, Pertanian, pemukiman, lahan kering, Semak, Perkebunan, Hutan.

**b. Pengolahan Data Scoring Jenis Tanah**

Peta jenis tanah didapatkan dari digitasi peta jenis tanah Provinsi Sumatera utara, khususnya wilayah Batu Bara, lalu dilakukan clip pada daerah penelitian kemudian diberikan nilai skor sesuai dengan acuan yang digunakan. Nilai Skoring jenis tanah.

**Tabel 3.** Scoring Jenis Tanah

No.	Jenis Tanah	Kelas	Harkat	Bobot	Skor
1.	Aluvial, Planosol, Hidromorf Kelabu, Laterik, Air Tanah	Sangat Halus	4	2	18
2.	Latosol	Halus	3		14
3.	Tanah Hutan Coklat, Tanah Mediteranian	Sedang	2		10
4.	Andosol, Laterik, Grumosol, podsol, Podsolik	Kasar	2		8
5.	Regosol, Litosol, Organosol, Renzina	Sangat Kasar	1		2

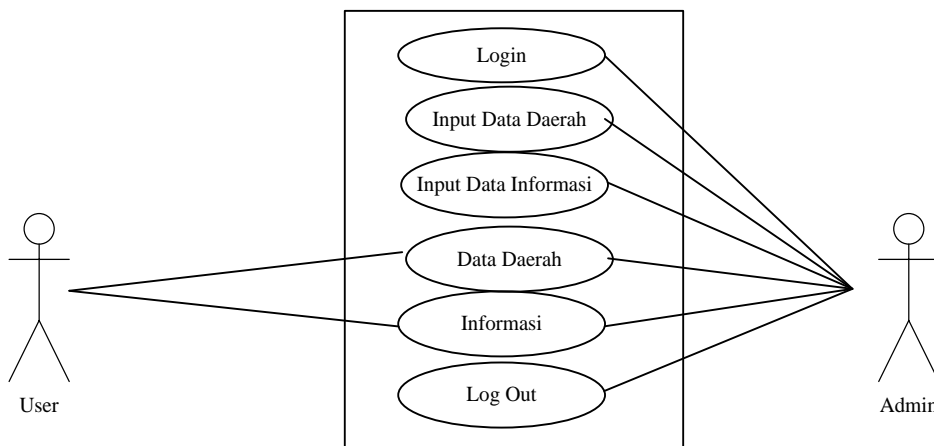
Dari proses pemetaan data scoring tanah dengan melakukan pemotongan citra untuk dapat memetakan area yang sesuai dengan kajian sehingga dapat mempermudah dalam proses klasifikasi citra dengan konsep machine learning. Koreksi Geometrik dilakukan pada citra dengan mengidentifikasi *Ground Control Points* (GCP) atau titik-titik ikat yang mudah ditentukan seperti percabangan sungai atau perpotongan jalan. Nilai akurasi GCP ditunjukkan oleh nilai *Root Mean Square Error* (RMSError). RMS-error menyatakan nilai kesalahan dari proses koreksi geometrik. Akurasi yang baik ditunjukkan oleh nilai RMS-error yang sangat kecil mendekati nol. "*Root Mean Square Error (RMSE) merupakan besarnya tingkat kesalahan hasil prediksi, dimana semakin kecil (mendekati 0) nilai RMSE maka hasil prediksi akan semakin akurat.*"

Untuk mengetahui hasil dan pembahasan pada penelitian pada kegiatan dalam meningkatkan minat baca di Rumah Baca ini merupakan program dari pemerintahan Kabupaten Karo, agar meningkatnya minat baca di masyarakat,

khususnya di kalangan remaja. Karena remaja akhir-akhir ini lebih banyak bermain Sosmed dan online games daripada membaca buku, apalagi untuk membaca buku di Perpustakaan atau Rumah Baca.

### 3.2 Use Case Diagram

Use case atau diagram use case merupakan pemodelan untuk kelakuan (*behavior*) sistem informasi yang akan dibuat. Use case mendeskripsikan sebuah intraksi antara satu atau lebih aktor dengan sistem informasi yang akan dibuat. Secara kasar, use case digunakan untuk mengetahui fungsi apa saja yang ada didalam sebuah sistem informasi dan siapa saja yang berhak menggunakan fungsi-fungsi itu.

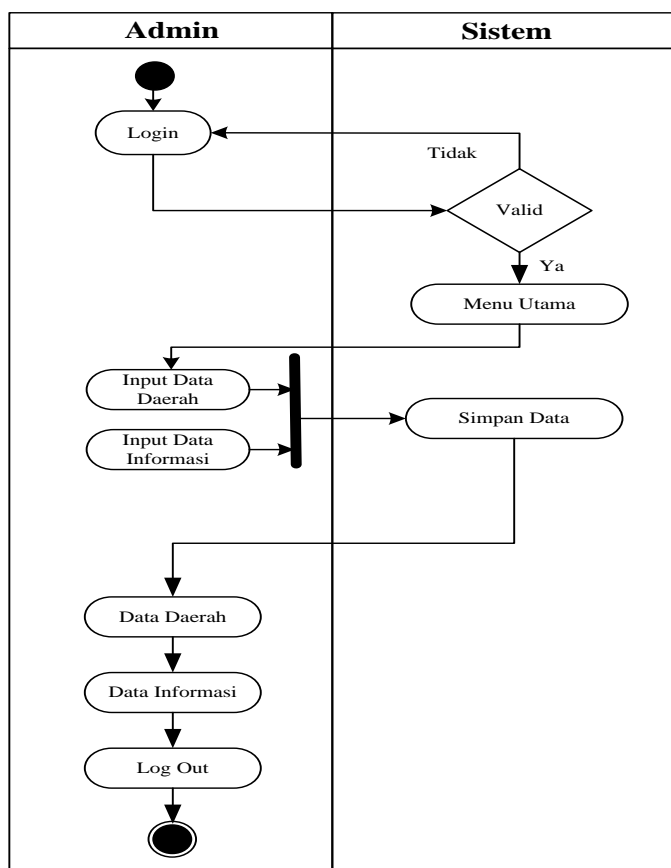


Gambar 2. Use Case Sistem

Penjelasan tentang gambar use case diagram untuk perancangan aplikasi ini dapat dilihat pada tabel dibawah ini.

### 3.3 Activity Diagram

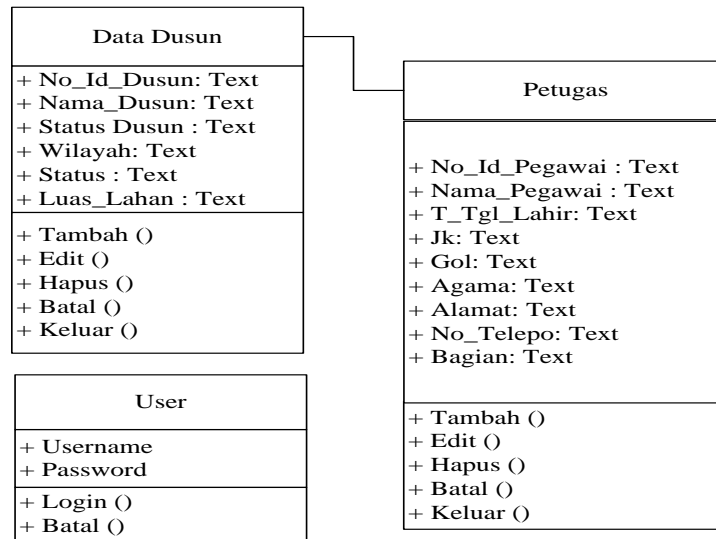
Diagram aktivitas atau Activity diagram menggambarkan c (aliran kerja) atau aktivitas dari sebuah sistem atau proses bisnis atau menu yang ada pada perangkat lunak. Yang perlu diperhatikan disini adalah bahwa diagram aktivitas menggambarkan aktivitas sistem bukan apa yang dilakukan aktor, jadi aktivitas yang dapat dilakukan oleh sistem.



Gambar 3. Activity Diagram

### 3.4 Class Diagram

*Class Diagram* merupakan pemodelan untuk melakukan sistem informasi yang dibuat, dan gambaran proses dari sistem yang berhak menggunakan sistem tersebut.



Gambar 4. Class Diagram Sistem

## 4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan maka dapat disimpulkan bahwa, dalam melakukan pemetaan daerah rawan banjir di Desa Pahlawan Kecamatan Tanjung Tiram Kabupaten Batubara, digunakan untuk proses digitalisasi pengolahan data yang dilakukan ini telah berhasil mengelompokkan jumlah data daerah rawan bencana banjir di 12 dusun yang terdapat pada Desa Pahlawan Perbandingan dilakukan dengan membandingkan data bencana alam yang sejenis pada kejadian actual yaitu tahun 2021 – 2022 yang dihasilkan lebih tinggi dibandingkan kejadian prediksi yaitu tahun 2022-2023. Machine Learning adalah sebuah proses melakukan pengolahan data historis kejadian kejadian terdahulu dan dipergunakan sebagai dasar untuk membangun sebuah pengetahuan. Untuk penelitian selanjutnya pada dataset sebaiknya dilengkapi dengan beberapa indikator seperti curah hujan ataupun tingkat kelembaban yang telah dikelompokkan dalam kelas tertentu, sehingga hasil prediksi data bencana lebih akurat.

## REFERENCES.

- [1] Anglin, dkk. (2011). Analysis Effect Of Mother Tongue Using In Mathematics Learning Toward Conceptual Understanding Ability In Elementary Students. Dipresentasikan Dalam International Conference (ICREAM 5). Bandung.
- [2] Ahmad, Abu. (2017). Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning.
- [3] Ahmad, Adeel. (2017, November 13). An overview of activation functions used in neural networks. Retrieved May 3, 2020, from adl1995.github.io website: <https://adl1995.github.io/an-overview-of-activation-functions-used-inneural-networks.html>
- [4] Ali, Z. (2019, January 7). A simple Word2vec tutorial. Retrieved May 1, 2020, from Medium website: <https://medium.com/@zafaralibagh6/a-simpleword2vec-tutorial-61e64e38a6a1>
- [5] Arni, D. (2018a, October 25). Apa Itu Text Mining? Retrieved April 24, 2020, from Garuda Cyber Indonesia website: <https://garudacyber.co.id/artikel/1254-apa-itu-text-mining>.
- [6] Ali, Z. (2019, January 7). A simple Word2vec tutorial. Retrieved May 1, 2020, from Medium website: <https://medium.com/@zafaralibagh6/a-simpleword2vec-tutorial-61e64e38a6a1>
- [7] Britz, D. (2016, July 8). Recurrent Neural Networks Tutorial, Part 1 – Introduction to RNNs. Retrieved from WildML website: <http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part1-introduction-to-rnns/>.
- [8] Brownlee, J. (2017, October 3). How to Use Word Embedding Layers for Deep Learning with Keras. Retrieved from Machine Learning Mastery website: <https://machinelearningmastery.com/use-word-embedding-layers-deeplearning-keras/>.
- [9] Ganesan, K. (2019). All you need to know about text preprocessing for NLP and Machine Learning. Retrieved from Kdnuggets.com website: <https://www.kdnuggets.com/2019/04/text-preprocessing-nlp-machinelearning.html>
- [9] Nugroho, K. S. (2020, February 8). Dasar Text Preprocessing dengan Python. Retrieved from Medium website: <https://medium.com/@ksnugroho/dasartext-preprocessing-dengan-python-a4fa52608ffe>
- [11] Prijono, B. (2018a, April 4). Pengenalan Recurrent Neural Network (RNN) – Bagian 1. Retrieved from Belajar Pembelajaran Mesin Indonesia website: <https://indoml.com/2018/04/04/pengenalan-rnn-bag-1/>
- [12] Putra, A. (2019, November 1). PROSES DATA TESTING DAN TRAINING PADA MACHINE LEARNING. Retrieved May 3, 2020, from Medium website: <https://medium.com/@anrelputranew/proses-data-testing-dantraining-pada-machine-learning-cdf8fcd93355>.
- [13] Thomas, A. (2017, August 30). A Word2Vec Keras tutorial. Retrieved May 1, 2020, from Adventures in Machine Learning website: <https://adventuresinmachinelearning.com/word2vec-keras-tutorial/>.