
ANALISA DATA MINING UNTUK MENGESTIMASI POTENSI CURAH HUJAN DENGAN MENGGUNAKAN METODE REGRESI LINEAR BERGANDA

Anzelina Marbun. *, Dicky Nofriansyah. **, Elfitriani. ***

* Program Studi Sistem Informasi, STMIK Triguna Dharma

** Program Studi Sistem Informasi, STMIK Triguna Dharma

*** Program Studi Sistem Informasi, STMIK Triguna Dharma

Article Info

Article history:

Received Feb 12th, 2021

Revised Feb 20th, 2021

Accepted Feb 26th, 2021

Keyword:

Curah Hujan

Data Mining

Regresi Linear Berganda

ABSTRACT

Prakiraan curah hujan telah menjadi salah satu masalah yang paling ilmiah dan teknologi menantang seluruh dunia pada abad terakhir. Ini disebabkan oleh dua faktor utama yaitu digunakan untuk melakukan berbagai kegiatan manusia dan kemajuan komputasi. Curah Hujan merupakan salah satu elemen pada iklim. Curah hujan juga merupakan faktor yang berpengaruh langsung terhadap perubahan cuaca, baik perubahan cuaca yang baik maupun perubahan cuaca yang buruk. Terdapat beberapa faktor yang utama yang dapat mempengaruhi terjadinya curah hujan antara lain suhu udara, kelembaban udara, kecepatan angin, dan lama penyinaran. Berdasarkan hal tersebut

Permasalahan pada perusahaan tersebut, maka penerapan dalam prediksi potensi curah hujan menggunakan keilmuan Data Mining dengan menggunakan Regresi Linear Berganda. Algoritma Regresi Linear sangat sesuai dan dapat sebagai alat bantu untuk memprediksi potensi curah hujan konsep keilmuan Data Mining.

Hasil yang didapatkan, merupakan perhitungan potensi curah hujan yang menggunakan Metode Regresi Linear Berganda yang dapat menyelesaikan masalah dalam memprediksi potensi curah hujan yang akan datang di Kota Medan.

Kata Kunci: Curah Hujan, Data Mining, Regresi Linear Berganda

Copyright © 2021 STMIK Triguna Dharma.

All rights reserved.

Corresponding Author:

Nama : Anzelina Marbun

Program Studi : Sistem Informasi

STMIK Triguna Dharma

Email : anzelinamarbun12@gmail.com

1. PENDAHULUAN

Sanitasi atau kesehatan lingkungan sangatlah penting untuk tetap menjaga kesehatan suatu lingkungan. Iklim merupakan unsur geografi yang penting, karena iklim memiliki pengaruh yang besar terhadap aktivitas manusia di lingkungan, seperti bidang transportasi, pertanian, perkebunan, dan lain-lain. Terdapat beberapa macam iklim di dunia. Iklim merupakan cuaca yang terjadi pada rentang waktu yang panjang dan wilayah yang luas. Cuaca merupakan kondisi udara pada wilayah tertentu dan dalam waktu yang relatif pendek. Cuaca dapat berubah dalam waktu yang singkat ditandai dengan adanya siang dan malam. Salah satu faktor yang paling penting dalam perubahan cuaca adalah curah hujan.

Prakiraan curah hujan telah menjadi salah satu masalah yang paling ilmiah dan teknologi menantang seluruh dunia pada abad terakhir. Ini disebabkan oleh dua faktor utama yaitu digunakan untuk melakukan berbagai kegiatan manusia dan kemajuan komputasi. Untuk membuat prediksi yang akurat adalah salah satu tantangan utama yang dihadapi ahli Meteorologi seluruh dunia. Sejak zaman kuno, prediksi cuaca maupun curah hujan telah menjadi salah satu yang paling menarik. Indonesia memiliki iklim tropis, dengan iklim tersebut akibat penguapan air ke udara menjadi sangat besar, sehingga intensitas curah hujan menjadi tidak stabil. Intensitas curah hujan menjadi tidak stabil. Intensitas curah hujan yang tinggi atau yang sering disebut dengan hujan ekstrem dapat mengakibatkan terjadinya banjir. Hujan ekstrem ini terjadi jika curah hujan mencapai >50 mm/hari (BMKG).

Curah Hujan merupakan salah satu elemen pada iklim. Curah hujan juga merupakan faktor yang berpengaruh langsung terhadap perubahan cuaca, baik perubahan cuaca yang baik maupun perubahan cuaca yang buruk. Terdapat beberapa faktor yang utama yang dapat mempengaruhi terjadinya curah hujan antara lain suhu udara, kelembaban udara, kecepatan angin, dan lama penyinaran. Berdasarkan hal tersebut penerapan dalam prediksi potensi curah hujan menggunakan keilmuan *Data Mining*.

Penerapan *Data Mining* telah banyak digunakan dalam pengolahan data untuk menghasilkan pengetahuan, salah satunya pada [1] yang menggunakan *Data Mining* untuk penentuan prediksi potensi curah hujan yang akan dikeluarkan. Selanjutnya algoritma yang akan digunakan untuk pengelolaan *Data Mining* pada kasus memprediksi potensi curah hujan adalah *Regresi Linear berganda*. *Regresi Linear* yang merupakan salah satu metode statistik yang dipergunakan dalam produksi untuk melakukan peramalan ataupun prediksi tentang karakteristik kualitas maupun kuantitas. *Simple Regresi Linear* terdiri dari satu buah variabel bebas (x) dengan satu buah variabel terikat (y) [2]. Selain itu metode *Regresi Linear Berganda* dapat digunakan untuk memprediksi sesuatu di masa depan dengan menggunakan lebih dari 3 faktor atau variabel yang memiliki pengaruh terhadap apa yang akan diprediksi [3]. Sehingga Algoritma *Regresi Linear* sangat sesuai dan dapat sebagai alat bantu untuk memprediksi potensi curah hujan konsep keilmuan *Data Mining*.

Dalam hal ini Balai Wilayah Sungai Sumatera II mengharapkan adanya perhitungan potensi curah hujan yang menggunakan *Metode Regresi Linear Berganda* yang dapat menyelesaikan masalah dalam memprediksi potensi curah hujan yang akan datang di Kota Medan. Penelitian ini juga diharapkan dapat memberikakan mamfaat agar bisa mengantisipasi Bencana Banjir akibat dari curah hujan yang tinggi.

2. METODE PENELITIAN

Untuk menyelesaikan permasalahan curah hujan pada Balai Wilayah Sungai Sumatera II. *Regresi linier Berganda* merupakan suatu algoritma yang memungkinkan untuk menemukan informasi (hubungan) antara dua atau lebih variabel independen ($X_1, X_2, X_3, \dots, X_n$) dengan variabel dependen (Y). Analisis antara variabel-variabel independen dan dipenden untuk mengetahui arah hubungan variabel-variabel yang diuji, Suatu objek data termasuk dalam suatu normalisasi jika memiliki variabel tersebut.

Maka dilakukan tahapan-tahapan sebagai berikut:

1. Inisialisasi data kedalam variabel X1, X2, X3, X4 dan Y
2. Menormalisasikan Data
3. Menyederhanakan Persamaan *Regresi Linear Berganda*
4. Hasil Prediksi Potensi Curah Hujan

Algoritma sistem merupakan penjelasan langkah-langkah penyelesaian masalah dalam perancangan *Data Mining* dalam menyelesaikan permasalahan curah hujan pada Balai Wilayah Sungai Sumatera II berdasarkan variabel dengan menggunakan metode *Regresi Linear Berganda*. Hal ini dilakukan untuk meningkat kinerja pada Balai Wilayah Sungai Sumatera II.

X1, X2, X3 dan X4 merupakan variabel bebas, sedangkan Y merupakan terikat yang dipengaruhi oleh keberadaan variabel X1, X2, X3 dan X4 Karena estimasi memprediksi Curah Hujan 2020 dan Realisasi 2021 dianggap mempengaruhi memprediksi Curah Hujan,

maka diinisialisasikan :

- X1 = TAVG
 X2 = RH_avg
 X3 = Lamanya penyinaran matahari (jam)
 X4 = $W_x(\text{Not})$
 Y = Jumlah Prediksi Curah Hujan

Tabel 3.3 Inialisasi Variabel

No	Variabel				
	X1	X2	X3	X4	Y
1	30,5	74	6,7	4,2	85
2	29,7	71	4,64	2,67	97
3	30,2	64	3,2	3,8	100

4	17	36,7	5,48	6,85	98
5	29,6	68	6,11	6,49	112
6	30	69	6,42	5,86	106,7
7	141,6	58	5,12	4,89	158,5
8	29,8	70	5,31	4,99	96,5
9	31,6	54	2,3	5,91	160,1
10	30,4	72	3,6	3,65	126,5

Normalisasi data berfungsi untuk memudahkan proses perhitungan. Berikut adalah hasil normalisasi data dengan membagi variabel X1,X2,X3, X4 menjadi pengelompokkan data. Dari tabel diatas maka diketahui :

3.4 Tabel Data Hasil Sigma

Variabel	Sigma
Y	1140
X1	400
X2	637
X3	49
X4	49
YX1	50900
YX2	71853
YX3	5394
YX4	5684
X1^2	27651
X2^2	41709
X3^2	258
X4^2	260
X1X2	25199
X1X3	1971
X1X4	1945
Y2	136236
X2X3	3142
X2X4	3055
X3X4	246

Dari tabel 3.4 diatas maka dilakukan perhitungan proses *Regresi Linier Berganda* dengan menggunakan rumus persamaan sebagai berikut :

$$\sum Y = na + \sum b1\sum X1 + \sum b2\sum X2 + \sum b3\sum X3 + \sum b4\sum X4 \dots\dots\dots (1)$$

$$\sum YX1 = a\sum X1 + b1\sum X1^2 + b2\sum X1X2 + b3\sum X1X3 + b4\sum X1X4 \dots\dots\dots (2)$$

$$\sum YX2 = a\sum X2 + b1\sum X1X2 + b2\sum X2^2 + b3\sum X2X3 + b4\sum X2X4 \dots\dots\dots (3)$$

$$\sum YX3 = a\sum X3 + b1\sum X1X3 + b2\sum X2X3 + b3\sum X3^2 + b4\sum X3X4 \dots\dots\dots (4)$$

$$\sum YX4 = a\sum X4 + b1\sum X1X4 + b2\sum X2X4 + b3\sum X3X4 + \sum X4^2 \dots\dots\dots (5)$$

Dengan menggunakan rumus di atas maka didapatkan persamaan sebagai berikut :

$$1140 = 10b0 + 400b1 + 636,70b2 + 49b3 + 49 [1]$$

$$50900 = 400b0 + 27651b1 + 25199,20b2 + 1971b3 + 1945b4 [2]$$

$$71853 = 637b0 + 25199b1 + 41709b2 + 3142b3 + 3055b4 [3]$$

$$5394 = 49b0 + 1971b1 + 3141,68b2 + 258b3 + 246b4 [4]$$

$$5684 = 49b0 + 1945b1 + 3055,49b2 + 246b3 + 260b4 [5]$$

Kemudian melakukan proses eliminasi antara persamaan (1) dengan persamaan (2) adalah sebagai berikut :

$$1140 = 10b_0 + 400b_1 + 636,70b_2 + 49b_3 + 49 \quad [1]$$

$$50900 = 400b_0 + 27651b_1 + 25199,20b_2 + 1971b_3 + 1945b_4 \quad [2]$$

$$-52420,48 = 0 - 116188,44b_1 + 2942,68b_2 - 136,088b_3 + 290,174b_4 \quad [5]$$

Kemudian melakukan proses eliminasi antara persamaan (1) dengan persamaan (3) adalah sebagai berikut :

$$1140 = 10b_0 + 400b_1 + 636,70b_2 + 49b_3 + 49 \quad [1]$$

$$71853 = 637b_0 + 25199b_1 + 41709b_2 + 3142b_3 + 3055b_4 \quad [3]$$

$$1802,674 = 0 - 136,088b_1 - 294,864b_2 - 189,7356b_3 - 47,4132 \quad [6]$$

Kemudian setelah melakukan proses eliminasi antara persamaan (1) hingga persamaan (4), maka diperoleh persamaan baru yaitu sebagai berikut.

$$1140 = 10b_0 + 400b_1 + 636,70b_2 + 49b_3 + 49 \quad [1]$$

$$5394 = 49b_0 + 1971b_1 + 3141,68b_2 + 258b_3 + 246b_4 \quad [4]$$

$$-611,287 = 0 + 290,174b_1 + 840,827b_2 - 47,4132b_3 - 165,0629b_4 \quad [7]$$

Kemudian setelah melakukan proses eliminasi antara persamaan (2) hingga persamaan (3), maka diperoleh persamaan baru yaitu sebagai berikut:

$$50900 = 400b_0 + 27651b_1 + 25199,20b_2 + 1971b_3 + 1945b_4 \quad [2]$$

$$71853 = 637b_0 + 25199b_1 + 41709b_2 + 3142b_3 + 3055b_4 \quad [3]$$

$$328410,3732 = 0 + 562480,1312b_1 - 26190,1744b_2 - 6931,81528b_3 - 3316,79504b_4 \quad [8]$$

Kemudian setelah melakukan proses eliminasi antara persamaan (2) hingga persamaan (4), maka diperoleh persamaan baru yaitu sebagai berikut.

$$50900 = 400b_0 + 27651b_1 + 25199,20b_2 + 1971b_3 + 1945b_4 \quad [2]$$

$$5394 = 49b_0 + 1971b_1 + 3141,68b_2 + 258b_3 + 246b_4 \quad [4]$$

$$234009,4554 = 0 + 584543,7646b_1 - 19156,358b_2 - 1227,3746b_3$$

$$- 8039,96651b_4 \quad [9]$$

Kemudian setelah melakukan proses eliminasi antara persamaan (3) hingga persamaan (4), maka diperoleh persamaan baru yaitu sebagai berikut.

$$71853 = 637b_0 + 25199b_1 + 41709b_2 + 3142b_3 + 3055b_4 \quad [3]$$

$$5394 = 49b_0 + 1971b_1 + 3141,68b_2 + 258b_3 + 246b_4 \quad [4]$$

$$-75883,4686 = 0 + 3965,0235b_1 - 111238,0664b_2 - 1564,82406b_3 -$$

$$14655,67278b_4 \quad [10]$$

Dari persamaan yang di peroleh maka dapat nilai di eliminasi setiap persamaan dan mendapatkan nilai b_0, b_1, b_2, b_3, b_4 adalah sebagai berikut :

$$B_0 = 15,08$$

$$B_1 = 0,49$$

$$B_2 = 0,44$$

$$B_3 = 0,00$$

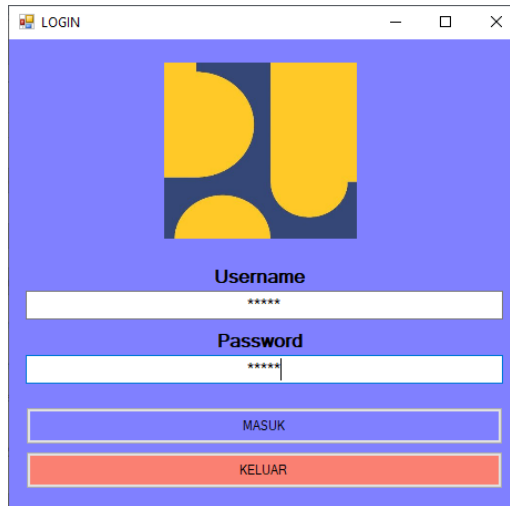
$$B_4 = 9,42$$

3. ANALISA DAN HASIL

Data mining ini dilengkapi dengan tampilan yang bertujuan untuk memudahkan penggunaannya. Fungsi dari *interface* (antarmuka) ini adalah untuk memberikan *input* dan menampilkan *output* dari aplikasi. pada aplikasi ini memiliki *interface* yang terdiri dari *menu login*, data curah hujan dan *menu* proses regresi linear berganda. Dalam halaman utama untuk menampilkan pada tampilan *menu* pada awal sistem yaitu *menu login* dan *menu* utama. Adapun *menu* halaman utama sebagai berikut.

1. Menu Login

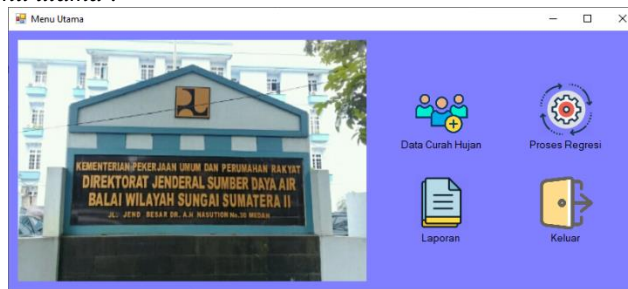
Menu login digunakan untuk mengamankan sistem dari *user-user* yang tidak bertanggung jawab sebelum masuk ke *menu* utama. Berikut adalah tampilan *menu login* :



Gambar 1 Menu Login

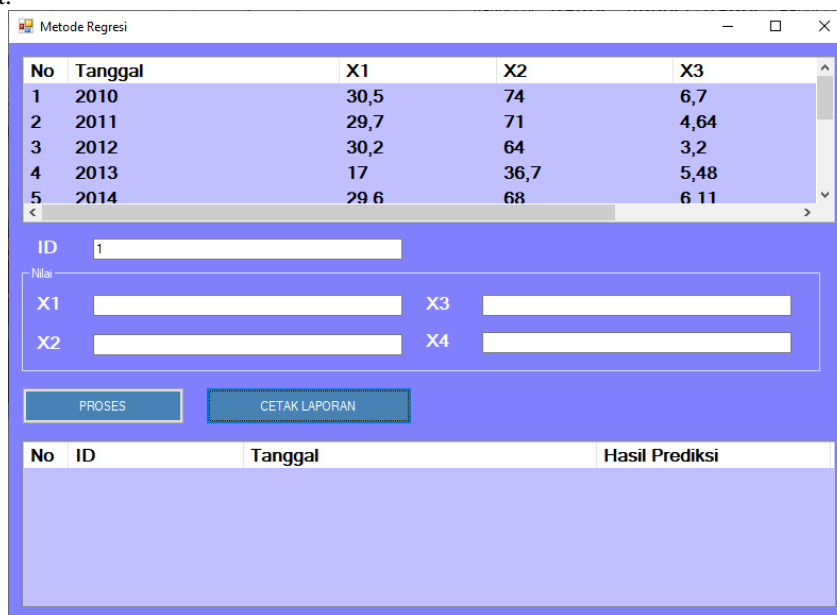
2. *Menu Utama*

Menu utama digunakan sebagai penghubung untuk *menu* data curah hujan, proses dan laporan. Berikut adalah tampilan *menu utama* :

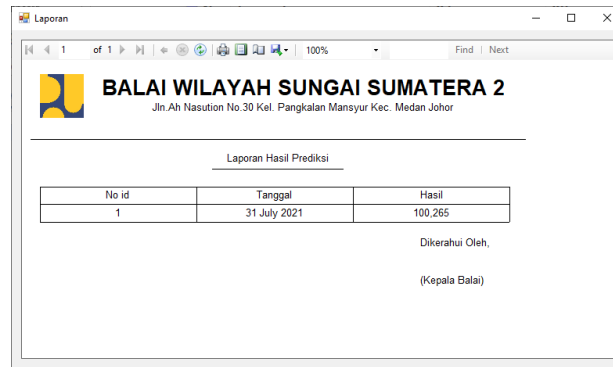


Gambar 2 Menu Utama

Pada bagian ini anda diminta untuk melakukan pengujian dengan sampling data baru dan pada bagian ini anda diminta untuk dapat menguji keakuratan sistem yang anda rancang dengan *tools-tools* yang sudah teruji dan terkalibrasi sebelumnya. Adapun hasil proses program dalam memprediksi curah hujan adalah sebagai berikut.



Gambar 3 Hasil Proses Metode Regresi Linear Berganda



BALAI WILAYAH SUNGAI SUMATERA 2
Jln.Ah Nasution No.30 Kel. Pangkalan Mansyur Kec. Medan Johor

Laporan Hasil Prediksi

No id	Tanggal	Hasil
1	31 July 2021	100,265

Diketahui Oleh,

(Kepala Balai)

Gambar 4 Laporan Hasil Prediksi

4. KESIMPULAN

Berdasarkan hasil analisa dari permasalahan yang terjadi dengan kasus yang di bahas tentang mengetahui prediksi curah hujan dengan menerapkan metode *Regresi* terhadap sistem yang dirancang dan dibangun maka dapat ditarik kesimpulan sebagai berikut :

1. Dalam menganalisa masalah prediksi curah hujan dilakukan pengumpulan data dan melakukan observasi ataupun wawancara untuk mengamati masalah dalam prediksi curah hujan dengan menggunakan metode *Regresi Linear Berganda*.
2. Dengan merancang aplikasi sistem data mining dengan melakukan pemodelan UML dan dengan membangun sistem berbasis *desktop* dalam mengestimasi curah hujan dengan menggunakan metode *Regresi Linear Berganda*.
3. Dengan menguji sistem dengan melakukan login sistem dan pengolahan data variabel dan proses dalam memprediksi curah hujan dengan menampilkan laporan hasil prediksi potensi curah hujan.




UCAPAN TERIMA KASIH

Terima Kasih diucapkan kepada kedua orang tua serta keluarga yang selalu memberi motivasi, Doa dan dukungan moral maupun materi, serta pihak-pihak yang telah mendukung dalam proses pembuatan jurnal ini yang tidak dapat disebutkan satu persatu. Kiranya jurnal ini bisa memberi manfaat bagi pembaca dan dapat meningkatkan kualitas jurnal selanjutnya.

REFERENSI

- [1] D. Sunia, K. and A. P. Jusia, "Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means," *STIKOM Dinamika Bangsa*, pp. 121-134, 2019.
- [2] P. Katemba and R. K. Djoh, "Prediksi Tingkat Produksi Kopi Menggunakan Regresi Linear," *Jurnal Ilmiah FLASH*, vol. III, 2017.
- [3] E. Triyanto, H. Sismoro and D. A. Laksito, "Implementasi Algoritma Regresi Linear Berganda Untuk Memprediksi Produksi Padi Di Kabupaten Bantul," *Jurnal Teknologi dan Sistem Informasi Univrab*, vol. IV, no. 2477-2062, pp. 73-86, 2019.
- [4] Alfannisa Annurullah Fajrin and Algifanri Maulana, " Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen Dengan Algoritma FP-growth Pada Data Transaksi Penjualan Spare Part Motor," *Kumpulan jurnal Ilmu Komputer (KLIK)*, vol. Volume 05, 2018.
- [5] M. R. Julianti, M. I. Dzulhaq and A. Subroto, "Sistem Informasi Pendataan Alat Tulis Kantor Berbasis Web pada PT Astari Niagara Internasional," *JURNAL SISFOTEK GLOBAL* , vol. IX, no. 2088 – 1762, pp. 92-97, 2019.
- [6] Y. Darmi, A. Setiawan, J. Bali, K. Kampung Bali, K. Teluk Segara and K. Bengkulu, "Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk," 2016.

BIBLIOGRAFI PENULIS

	<p>Nama Lengkap : Anzelina Marbun</p> <p>NIRM : 2017020508</p> <p>Tempat/Tgl.Lahir : Medan, 02 September 1996</p> <p>Jenis Kelamin : Perempuan</p> <p>Alamat : Jln Luku III No.23 LK II Medan</p> <p>No/Hp : 852-7045-7903</p> <p>Email : anzelinamarbun12@gmail.com</p> <p>Program Keahlian : Pemmograman Berbasis Desktop</p>
	<p>Nama Lengkap : Dr. Dicky Nofriansyah, S.Kom., M.Kom.</p> <p>NIDN : 0125048902</p> <p>Tempat/Tgl.Lahir : Medan, 31 Mei 1989</p> <p>Jenis Kelamin : Laki-Laki</p> <p>No/Hp : 0852-9666-8800</p> <p>Email : Nofriansyah@gmail.com</p> <p>Pendidikan : - S1 – STMIK Triguna Dharma - S2 – Universitas Putra Indonesia Yptk Padang - S3 – Universitas Buddhi Dharma</p> <p>Bidang Keahlian : Sistem Pendukung Keputusan, Data Mining, Kriptografi, Sistem Pakar, IT In Education, STEM, Sistem Informasi</p>
	<p>Nama Lengkap : Elfitriani, S.Pd, M.Si</p> <p>NIDN : 0124097301</p> <p>Tempat/Tgl.Lahir : Medan, 11 Mei 1988</p> <p>Jenis Kelamin : Perempuan</p> <p>No/Hp : 0813-7075-3204</p> <p>Email : Elfitrian@trigunadharma.ac.id</p> <p>Pendidikan : - S1 – Universitas Muslim Nusantara Al-wasliyah - S2 – Universitas Sumatera Utara</p> <p>Bidang Keahlian : Bahasa Inggris dan TOFEL</p>