

IMPLEMENTASI ALGORITMA REGRESI LINEAR UNTUK MEMPREDIKSI PENDAPATAN PT PLN BERDASARKAN PENGUNAAN PER KELOMPOK PELANGGAN

Taufan Maulana, Rini Astuti, Fadhil Muhammad Basysyar

Teknik Informatika, Sekolah Tinggi Manajemen dan Informatika IKMI Cirebon
Jalan Perjuangan No.10B Karyamulya Kec.Kesambi Kota Cirebon Jawa barat 45135, Indonesia
taufanmaulana201201@gmail.com

ABSTRAK

Perusahaan Listrik Negara (PLN) merupakan perusahaan milik negara yang bertanggung jawab atas penyediaan listrik di Indonesia. Metode yang digunakan pada penelitian kali ini adalah menggunakan data mining yaitu regresi linear berganda untuk memprediksi pendapatan PT PLN. Penelitian ini data didapatkan dengan melakukan pencarian di website resmi PT PLN yaitu <https://web.pln.co.id/stakeholder/laporan-statistik>. data yang akan diteliti pada penelitian ini merupakan data penggunaan perkelompok pelanggan yang meliputi data jumlah pelanggan per jenis pelanggan, daya tersambung per kelompok pelanggan (MVA), energi terjual per kelompok pelanggan (GWh), pendapatan per kelompok pelanggan (juta Rp), energi terjual rata-rata per jenis pelanggan (kWh), tarif listrik rata-rata per kelompok pelanggan (Rp/kWh), dan tahun. Data tersebut terkumpul dari tahun 2014 sampai tahun 2022. Kemudian akan dilakukan analisa dan *preprocessing* sebelum akhirnya siap dilakukan analisa prediksi menggunakan algoritma regresi linear berganda. Dari penelitian yang telah dilakukan dan setelah dilakukan pemodelan didapatkan penggunaan per kelompok pelanggan sangat mempengaruhi tingkat pendapatan PT PLN, dengan didapati akurasi dari hasil model menggunakan *r_score* adalah 0.9924 dari hasil pengujian *matric* dengan menggunakan *Mean Absolute error* (MAE) didapatkan 583656.0650, *Mean Squared Error* (MSE) didapatkan 894668323229.3309, dan *Root Mean Squared Error* (RMSE) didapatkan 945869.0835. Dengan demikian pemodelan tersebut dikatakan efektif untuk melakukan prediksi menggunakan algoritma regresi linear berganda.

Kata kunci : Regresi Linear, Data Mining, PT PLN (persero), Listrik

1. PENDAHULUAN

Perusahaan Listrik Negara (PLN) ialah Badan Usaha Milik Negara (BUMN) yang mempunyai tanggung jawab terhadap penyediaan tenaga listrik di Indonesia. Sebagai entitas utilitas terbesar di Indonesia, PLN memegang peran yang amat vital dalam memberikan pasokan tenaga listrik untuk warga, industri, dan sektor usaha lainnya. Sejak pendiriannya pada tahun 1945, PLN telah mengalami pertumbuhan yang cepat dalam memberikan layanan tenaga listrik untuk komunitas Indonesia. Misi inti perusahaan ini adalah menjamin ketersediaan, keandalan, dan keselamatan tenaga listrik bagi semua pelanggan di segala wilayah Indonesia.

Dalam menjalankan operasionalnya, PT PLN (Persero) tidak hanya bertanggung jawab dalam menyediakan listrik bagi masyarakat, tetapi juga perlu mengelola aspek keuangan perusahaan. Salah satu aspek penting dalam pengelolaan keuangan adalah memprediksi pendapatan yang akan diperoleh oleh perusahaan. Penghasilan PT PLN dapat dipengaruhi oleh berbagai faktor salah satunya adalah penggunaan perkelompok, yang merujuk pada pembagian penggunaan listrik dalam kategori-kategori tertentu, seperti rumah tangga, industri, dan komersial. Setiap kelompok pengguna memiliki kebiasaan dan karakteristik penggunaan listrik yang berbeda. Memprediksi penghasilan berdasarkan penggunaan perkelompok dapat memberikan manfaat

signifikan bagi PT PLN. Dengan memiliki perkiraan yang akurat tentang penghasilan di masa mendatang, PT PLN dapat melakukan perencanaan keuangan yang lebih baik, termasuk alokasi sumber daya dan pengambilan keputusan investasi yang tepat. Salah satu pendekatan yang dapat digunakan untuk memprediksi penghasilan adalah algoritma regresi linear. Regresi linear adalah metode statistik yang digunakan untuk memodelkan hubungan linier antara variabel independen (penggunaan perkelompok) dan variabel dependen (pendapatan). Dengan menganalisis data historis penggunaan perkelompok dan penghasilan PT PLN, algoritma regresi linear dapat mengidentifikasi pola dan hubungan yang ada di antara keduanya, serta memprediksi penghasilan di masa depan berdasarkan penggunaan perkelompok.

2. TINJAUAN PUSTAKA

Dalam penelitian sebelumnya, peneliti mendapati setidaknya tujuh jurnal yang dianalisis sebagai acuan dalam penelitian ini. perbandingan sebelumnya akan menunjukkan kejelasan dan justifikasi yang kuat untuk relevansi dan kontribusi penelitian ini.

Salah satunya ialah Paper [1] dilakukan pengembangan sistem prediksi penggunaan listrik jangka panjang melalui platform web dengan menggunakan regresi linear. Tujuannya adalah meramalkan konsumsi listrik di masa depan dengan

mempertimbangkan variabel seperti kapasitas daya terhubung, populasi, Produk Domestik Regional Bruto (PDRB), jumlah pelanggan listrik, beban faktor, kerugian distribusi, dan volume penjualan tenaga listrik. Metodologi pengembangan yang digunakan adalah pendekatan *Agile*, melibatkan tahap perencanaan, desain, pengembangan, pengujian, peluncuran, dan perbaikan berkelanjutan. Hasil uji coba menggunakan metode *blackbox* menunjukkan bahwa sistem prediksi ini beroperasi efektif dan memberikan proyeksi yang akurat, dengan kapasitas daya terhubung pada tahun 2021 sekitar 3.848.954.690. Dibandingkan dengan penggunaan *Microsoft Excel*, terdapat selisih sebesar 0,11%, disebabkan oleh pembatasan dan keterbatasan komputasi. Uji fungsionalitas halaman login, pengisian data baru, dan pengubahan data menegaskan bahwa sistem ini berfungsi sesuai harapan. Sistem ini juga dilengkapi dengan dashboard yang menampilkan grafik dan data perkiraan konsumsi listrik. Diskusi lebih lanjut mencakup proses pengembangan sistem, spesifikasi perangkat keras dan lunak, serta analisis mendalam tentang penerapan regresi linear. Secara keseluruhan, penelitian ini berhasil menghasilkan sistem prediksi konsumsi listrik jangka panjang dengan akurasi tinggi untuk meramalkan kebutuhan listrik di masa depan.

Selanjutnya berdasarkan Paper [2] bertujuan untuk meramalkan penggunaan energi listrik di PT.PLN Medan Selatan menggunakan metode Data Mining, khususnya Regresi Linier Berganda. Metode ini diterapkan untuk mengidentifikasi hubungan antara variabel yang tergantung (*regressand*) dengan satu atau lebih variabel yang mempengaruhi (*regressor*). Hasil penelitian menunjukkan bahwa penerapan Regresi Linier Berganda efektif dalam menghasilkan perkiraan penggunaan energi listrik di PT.PLN Medan Selatan. Prediksi untuk bulan Januari tahun 2020 mencapai 36,204,961.91 kWh. Penelitian ini juga melibatkan formulir seperti Form Login untuk meningkatkan keamanan sistem dan Form Menu Utama sebagai antarmuka utama. Dengan memanfaatkan metode Data Mining, terutama Regresi Linier Berganda, penelitian ini memberikan kontribusi berharga dalam meramalkan kebutuhan energi listrik di PT.PLN Medan Selatan. Kontribusi ini diharapkan dapat mendukung upaya perencanaan dan peningkatan layanan yang dilakukan oleh PT.PLN Medan Selatan.

Paper [3] jurnal tersebut membahas proyeksi pertumbuhan permintaan energi listrik di Provinsi Jambi dari 2020 hingga 2025 dengan menggunakan data dari PT. PLN Persero (2009-2014). Studi ini menerapkan analisis regresi linear untuk mengkaji konsumsi energi listrik di sektor rumah tangga, industri, bisnis, dan sosial. Hasilnya menunjukkan peningkatan konsumsi di setiap sektor, terbesar pada sektor rumah tangga (23,89%), diikuti sektor industri (21,03%), bisnis (18,29%), dan sosial (7,06%). Penelitian ini menggunakan metode ekonometrik

untuk memprediksi konsumsi energi listrik di masa mendatang, melibatkan penetapan fungsi utilitas dan analisis regresi. Proyeksi kebutuhan energi listrik di Provinsi Jambi antara 2020-2025 diperoleh melalui pendekatan ini. Data yang digunakan termasuk konsumsi energi rumah tangga, pendapatan rumah tangga, jumlah pelanggan rumah tangga, dan harga energi listrik bagi rumah tangga. Semua data ini menjadi variabel independen dalam analisis regresi untuk meramalkan kebutuhan energi listrik di masa depan.

Paper [4] Penelitian ini bertujuan untuk meramalkan dan menganalisis permintaan energi listrik di Provinsi Sumatera Barat sampai tahun 2024. Metode yang dipakai dalam penelitian ini ialah analisis regresi linear berganda. Temuan dari penelitian ini mengindikasikan bahwa beban listrik di Provinsi Sumatera Barat terus meningkat seiring berjalannya waktu. Pada tahun 2020, diperhitungkan bahwa beban listrik akan mencapai 31.100.794 kWh dan pada tahun 2024 diperkirakan mencapai 35.212.420 kWh. Ini menggambarkan bahwa kebutuhan energi listrik di wilayah tersebut terus bertambah sejalan dengan pertumbuhan penduduk dan kemajuan ekonomi.

Paper [5] Dalam jurnal ini, dilakukan evaluasi terhadap Ranting Kayu Aro, Kabupaten Solok, selama periode 2004 hingga 2011, fokus pada jumlah pelanggan, kapasitas terhubung, dan penggunaan listrik. Analisis menunjukkan peningkatan tahunan sekitar 4,33% dalam penggunaan listrik. Uji regresi dilakukan untuk menilai kecocokan model regresi linear dalam meramalkan jumlah pelanggan. Proyeksi pertumbuhan antara tahun 2012 dan 2020 menunjukkan pertumbuhan rata-rata pelanggan sebesar 2,57%, kapasitas terhubung sebesar 3,6%, dan penggunaan listrik sebesar 4,47%. Kesimpulannya, terdapat peningkatan signifikan dalam jumlah pelanggan, kapasitas terhubung, dan penggunaan listrik di Ranting Kayu Aro, Kabupaten Solok. Fenomena ini mencerminkan peningkatan kebutuhan energi listrik di daerah tersebut, menekankan pentingnya perencanaan yang akurat untuk menyesuaikan kapasitas produksi listrik dengan tuntutan masyarakat.

Paper [6] Penelitian ini bertujuan menguji dampak rasio profitabilitas terhadap perubahan laba perusahaan dengan fokus pada anak perusahaan PT PLN (Persero). Variabel yang diuji meliputi *Gross Profit Margin (GPM)*, *Net Profit Margin (NPM)*, *Return on Assets (ROA)*, dan *Return on Equity (ROE)*. Hasil analisis menunjukkan secara keseluruhan, GPM, NPM, ROA, dan ROE secara signifikan mempengaruhi perubahan laba. Namun, ketika dianalisis terpisah, GPM, NPM, dan ROA memiliki pengaruh yang signifikan terhadap perubahan laba, sementara ROE tidak berpengaruh. Ini menunjukkan bahwa GPM, NPM, dan ROA lebih relevan dalam meramalkan perubahan laba dibandingkan dengan ROE. Persamaan regresi linear

berganda menggunakan variabel tersebut sebagai faktor prediksi perubahan laba perusahaan. Hasil statistik deskriptif menunjukkan rata-rata laba perusahaan sekitar 1,516,942.27 juta rupiah, memberikan gambaran kinerja profitabilitas selama periode penelitian.

Paper [7] Dalam studi tersebut, dilakukan analisis data menggunakan metode tren linear untuk memprediksi pertumbuhan populasi, Produk Domestik Regional Bruto (PDRB), dan beban listrik hingga tahun 2026. Hasil prediksi menunjukkan peningkatan beban listrik dari 1541,146 MW pada tahun 2022 menjadi 1970,241 MW pada tahun 2026. PDRB dan jumlah penduduk juga menunjukkan peningkatan konsisten setiap tahun. Selain itu, analisis masa pakai transformator distribusi dilakukan, melibatkan berbagai parameter seperti persentase beban daya, tegangan output, arus beban, kesalahan persentase masa pakai, degradasi, dan sisa masa pakai transformator distribusi. Berdasarkan analisis regresi, disimpulkan bahwa nilai PDRB dan jumlah penduduk berpengaruh pada penggunaan listrik di penyulang Area Ambon. Metode regresi linier berganda digunakan untuk memprediksi beban di penyulang tersebut. Hasil penelitian ini dapat digunakan sebagai panduan untuk merencanakan dan mengelola sistem listrik di wilayah penyulang Area Ambon, memungkinkan perencanaan kebutuhan listrik yang lebih efisien dengan pemahaman terhadap prediksi beban dan faktor-faktor yang memengaruhi.

3. METODE PENELITIAN

Penelitian ini menggunakan *Knowledge Discovery in Database (KDD)* yang melibatkan *data selection*, *data preprocessing*, *data transformation*, *data mining*, dan *evaluation*.

3.1. Data Selection

Pada tahap ini peneliti memilih subset data yang telah dikumpulkan dari website resmi pln untuk kemudian akan diteliti. data tersebut adalah data penggunaan perkelompok pelanggan yang meliputi jumlah pelanggan per jenis pelanggan, daya tersambung per kelompok pelanggan (MVA), energi terjual per kelompok pelanggan (GWh), pendapatan per kelompok pelanggan (juta Rp), energi terjual rata-rata per jenis pelanggan (kWh), tarif listrik rata-rata per kelompok pelanggan (Rp/kWh) dan tahun.

3.2. Data Preprocessing

Langkah selanjutnya adalah membersihkan, mentransformasi, dan mempersiapkan data tersebut agar lebih sesuai dan siap untuk digunakan dalam analisis data atau dalam pemodelan mesin. Tujuan dari data *preprocessing* adalah untuk menghilangkan kesalahan, *noise*, atau ketidaksesuaian dalam data, sehingga hasil analisis atau model yang dibangun lebih akurat dan dapat diandalkan.

3.3. Data Transformation

Setelah dilakukan *preprocessing* selanjutnya data tersebut dilakukan penyesuaian terhadap format data yang sesuai agar dapat dilakukan proses pemodelan.

3.4. Data Mining

Setelah data siap dan sesuai dengan format yang dibutuhkan untuk pemodelan. Selanjutnya implementasi data mining menggunakan algoritma regresi linear dengan variable y adalah pendapatan per kelompok pelanggan (juta Rp) dan variable x adalah jumlah pelanggan per jenis pelanggan, daya tersambung per kelompok pelanggan (MVA), energi terjual per kelompok pelanggan (GWh), energi terjual rata-rata per jenis pelanggan (kWh), tarif listrik rata-rata per kelompok pelanggan (Rp/kWh) dan tahun.

3.5. Evaluation

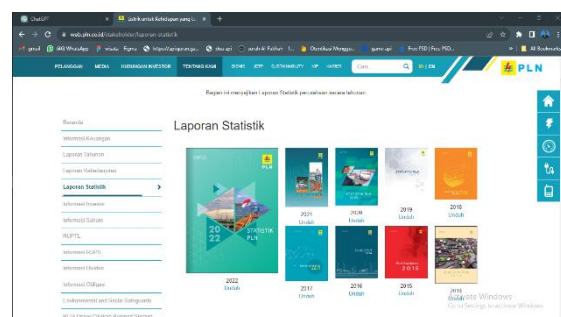
Setelah selesai beberapa tahapan diatas dan di dapatkan pemodelan yang sesuai selanjutnya dilakukan evaluasi terhadap model yang telah di buat apakah model tersebut sesuai dan layak untuk selanjutnya sebagai acuan untuk melakukan prediksi terhadap pendapatan PT PLN.

4. HASIL DAN PEMBAHASAN

Pada bab ini akan diberikan pemaparan mengenai hasil penelitian dan pembahasan yang telah dilakukan oleh peneliti saat melakukan penelitian di PT PLN. Penjelasan yang diberikan berupa penjelasan tahap yang dilakukan dalam melakukan penelitian.

4.1. Data Selection

Tahap pertama dalam melakukan penelitian ialah pengumpulan data. Pengumpulan data pada penelitian ini dilakukan melalui website resmi PT PLN (Persero), yang dapat diakses melalui tautan <https://web.pln.co.id/stakeholder/laporan-statistik>.



rata-rata per kelompok pelanggan (dalam Rupiah per kWh). Keseluruhan data ini bersifat periodik dan telah terkumpul selama sembilan tahun, mulai dari tahun 2014 hingga tahun 2022. Hasil pengumpulan data ini terdokumentasi pada gambar 1.2 hingga gambar 1.7 berikut.

No	Satuan PLN/Provinsi	Tahun	Rumah Tangga	Industri	Busnis	Sosial	Gdg. Kantor Pemerintahan	Penerangan Jalan Umum	Jumlah	(%)
1	Wilayah Aceh	2014	1.067.703	1.441	76.605	33.473	6.470	1.207	1.186.899	2,06
2	Wilayah Sumatera Utara	2014	2.851.592	1.499	106.876	55.721	7.413	14.409	3.051.822	5,31
3	Wilayah Sumatera Barat	2014	1.019.075	130	62.401	28.469	5.200	2.471	1.118.778	2,02
4	Wilayah Riau	2014	1.160.085	301	109.383	21.481	5.205	2.578	1.299.634	2,26
5	Riau	2014	992.411	223	88.842	18.217	3.840	2.237	1.105.800	1,92
6	Kepulauan Riau	2014	168.273	78	20.541	3.264	1.357	361	193.834	0,34
7	Wilayah Sumat. Jambi, dan Bengkulu	2014	2.351.107	247	97.812	41.083	9.233	5.971	2.505.953	4,36
8	Sumatera Selatan	2014	1.630.885	547	65.108	27.772	5.640	4.891	1.729.134	3,01
9	Jambi	2014	338.765	148	23.961	8.273	1.787	748	371.128	0,65
10	Bengkulu	2014	381.409	32	14.540	7.686	1.789	748	405.649	0,71
11	Wilayah Bangka Belitung	2014	314.399	215	16.584	5.197	1.859	811	339.005	0,59
12	Distrik Lampung	2014	1.564.817	347	15.522	12.467	3.060	1.406	1.617.624	2,85
13	Wilayah Kalimantan Barat	2014	799.788	17	17.201	17.221	1.411	1.170	829.179	1,31
14	Wilayah Kalimantan dan Kalteng	2014	1.254.465	641	84.966	32.711	7.579	3.389	1.383.045	2,41
15	Kalimantan Selatan	2014	875.818	304	41.265	22.105	4.093	2.380	946.013	1,63
...	...	...	...	...	...	...	...	...	...	...
423	Jawa	2022	47.009.082	146.192	2.464.141	1.180.371	85.843	243.520	51.329.549	99,9

Gambar 1. 2 Data Pelanggan Perjenis Pelanggan

No	Satuan PLN/Provinsi	Tahun	Rumah Tangga	Industri	Busnis	Sosial	Gdg. Kantor Pemerintahan	Penerangan Jalan Umum	Jumlah	(%)
1	Wilayah Aceh	2014	761.74	44	213.04	76.13	54.34	26.57	1.171.90	1,17
2	Wilayah Sumatera Utara	2014	2.326.08	817.26	717.65	158.81	64.68	88.06	4.202.54	4,2
3	Wilayah Sumatera Barat	2014	890.76	148	293.1	61.79	36.77	15.43	1.404.96	1,41
4	Wilayah Riau	2014	1.270.05	377	338.36	86.28	79.41	37.62	2.064.97	2,05
5	Riau	2014	1.060.53	82,97	412,64	70,18	56,17	33,35	1.778,81	1,74
6	Kepulauan Riau	2014	209.52	14,61	65,52	15,96	18,44	4,3	359,16	0,36
7	Wilayah Sumat. Jambi, dan Bengkulu	2014	2.471.23	313,34	805,32	171,22	91,18	91,1	3.444,28	3,44
8	Sumatera Selatan	2014	1.552,23	276,68	415,64	86,12	66,17	26,42	2.423,58	2,42
9	Jambi	2014	375,41	64,31	180,96	20,17	18,17	7,21	596,23	0,6
10	Bengkulu	2014	318,56	16,1	16,54	14,68	14,68	6,16	348,15	0,32
11	Wilayah Bangka Belitung	2014	347,24	27,75	96,04	18,52	17,47	4,38	513,79	0,51
12	Distrik Lampung	2014	1.371,19	138,13	241,49	65,18	30,14	17,1	2.061,83	2,06
13	Wilayah Kalimantan Barat	2014	869,59	49,91	241,76	45,04	27,13	13,18	1.094,63	1,06
14	Wilayah Kalimantan dan Kalteng	2014	992,12	101,99	302,96	61,26	37,22	26,5	1.406,78	1,41
15	Kalimantan Selatan	2014	671,82	87,72	235,57	40,78	22,11	13,38	1.069,40	1,06
...	...	...	...	...	...	...	...	...	...	...
423	Jawa	2022	47.531,88	28.111,08	22.772,88	1.180.371	85.843	243.520	51.329.549	99,9

Gambar 1. 3 Daya Tersambung Perkelompok Pelanggan

No	Satuan PLN/Provinsi	Tahun	Rumah Tangga	Industri	Busnis	Sosial	Gdg. Kantor Pemerintahan	Penerangan Jalan Umum	Jumlah	(%)
1	Wilayah Aceh	2014	1.206.78	130	72.61	31.53	74	161	1.462,39	2,46
2	Wilayah Sumatera Utara	2014	4.377,31	2.084,13	1.251,63	255,27	96,54	796,13	8.271,01	8,18
3	Wilayah Sumatera Barat	2014	1.497,11	844,41	412,08	108,1	51,21	75,51	3.005,26	3,1
4	Wilayah Riau	2014	2.485,58	614,18	276,46	141,48	102,57	158,7	3.572,48	3,57
5	Riau	2014	2.020,52	155,92	816,19	120,1	78,11	117,61	3.194,48	3,19
6	Kepulauan Riau	2014	376,41	28,28	162,29	21,13	24,68	21,03	633,72	0,62
7	Wilayah Sumat. Jambi, dan Bengkulu	2014	2.864,76	88,76	1.066,58	184,08	128,79	145,28	4.399,53	4,39
8	Sumatera Selatan	2014	2.689,72	690,07	796,1	131,96	81,23	108,47	4.431,99	4,43
9	Jambi	2014	630,43	101,72	227,55	36,17	23,43	23,95	1.037,44	0,52
10	Bengkulu	2014	544,62	27,17	102,94	22,34	20,13	12,44	729,44	0,37
11	Wilayah Bangka Belitung	2014	574,58	12,62	120,45	21,51	21,78	8,25	748,41	0,41
12	Distrik Lampung	2014	2.060,38	109,5	399,1	91,79	45,7	36,97	3.092,43	3,72
13	Wilayah Kalimantan Barat	2014	1.212,87	15,77	391,58	67,85	63,12	34,38	1.862,44	1,84
14	Wilayah Kalimantan dan Kalteng	2014	2.036,75	213,67	376,43	88,76	89,27	76,54	3.065,18	3,04
15	Kalimantan Selatan	2014	1.394,28	187,44	380,13	58,94	48,31	31,3	2.091,03	2,03
...	...	...	...	...	...	...	...	...	...	...
423	Jawa	2022	72.982,48	72.252,34	33.831,13	6.369,38	2.899,67	1.833,63	188.948,23	99,9

Gambar 1. 4 Energi Terjual Perkelompok Pelanggan

No	Satuan PLN/Provinsi	Tahun	Rumah Tangga	Industri	Busnis	Sosial	Gdg. Kantor Pemerintahan	Penerangan Jalan Umum	Jumlah	(%)
1	Wilayah Aceh	2014	426.230	78.474	334.428,81	75.894,97	86.512,81	111.689,38	1.323.055,87	9,71
2	Wilayah Sumatera Utara	2014	2.441.109,84	1.024.843,01	1.008.794,98	377.773,92	122.129,68	498.850,81	4.779.851,23	9,82
3	Wilayah Sumatera Barat	2014	823.717,69	784.860,21	489.399,98	75.789,77	86.934,14	101.498,52	3.308.245,12	3,42
4	Wilayah Riau	2014	1.532.448,82	181.218,81	1.149.168,47	111.020,18	154.877,17	174.517,82	3.269.879,16	3,27
5	Riau	2014	1.205.761,28	172,67	841.778,68	91.963,08	102.175,76	151.138,18	2.465.106,12	2,42
6	Kepulauan Riau	2014	351.185,34	28.848,08	204.718,47	17.075,73	32.177,17	23.889,04	607.772,42	0,63
7	Wilayah Sumat. Jambi, dan Bengkulu	2014	2.111.888,88	84,1	1.102.888,84	154.488,88	102.129,68	111.129,68	4.431,99	4,43
8	Sumatera Selatan	2014	1.603.575,82	875.521,68	802.888,51	105.988,15	115.424,54	159.862,13	3.495.051,81	3,49
9	Jambi	2014	541.621,00	105.888,01	280.022,08	24.247,08	32.121,08	38.789,08	1.032.658,08	0,93
10	Bengkulu	2014	369.831,28	16,1	14.650,44	26.460,44	26.460,44	26.460,44	70.962,16	0,16
11	Wilayah Bangka Belitung	2014	255.987,89	49.152,89	144.458,89	14.020,22	28.900,34	10.166,72	496.645,16	0,72
12	Distrik Lampung	2014	1.059.983,28	688.818,22	488.337,38	63.039,38	58.505,08	86.833,18	2.423.528,22	2,42
13	Wilayah Kalimantan Barat	2014	624.428,77	64.494,68	417.624,68	64.002,13	72.852,47	89.002,13	1.012.617,78	1,01
14	Wilayah Kalimantan dan Kalteng	2014	1.025.987,20	216.146,81	607.008,78	60.370,12	118.525,78	87.185,18	2.345.626,93	2,34
15	Kalimantan Selatan	2014	569.338,78	188.268,68	429.115,78	38.474,18	64.015,49	70.287,72	1.339.705,85	1,33
...	...	...	...	...	...	...	...	...	...	...
423	Jawa	2022	80.534.069,48	79.209.946,74	41.760.701,50	5.947.421,48	4.120.831,24	2.888.203,80	213.896.176,08	88,7

Gambar 1. 5 Pendapatan Perkelompok Pelanggan

No	Satuan PLN/Provinsi	Tahun	Rumah Tangga	Industri	Busnis	Sosial	Gdg. Kantor Pemerintahan	Penerangan Jalan Umum	Jumlah	(%)
1	Wilayah Aceh	2014	1.179,93	55.988,48	4.137,13	1.458,13	11.536,13	86.128,93	1.058,04	1,05
2	Wilayah Sumatera Utara	2014	1.458,77	566.896,28	11.710,89	4.581,13	32.980,57	27.824,19	2.779,18	2,78
3	Wilayah Sumatera Barat	2014	1.460,81	2.412.608,92	4.974,95	5.637,17	9.843,38	38.653,99	2.591,49	2,59
4	Wilayah Riau	2014	2.077,83	611.909,28	8.948,52	8.593,13	28.783,28	81.577,48	1.056,28	1,05
5	Riau	2014	2.040,08	699.211,07	1.189,17	6.807,69	20.250,53	61.500,79	3.918,82	3,91
6	Kepulauan Riau	2014	2.218,93	162.290,60	7.898,72	6.512,13	18.174,32	61.450,12	3.209,43	3,20
7	Wilayah Sumat. Jambi, dan Bengkulu	2014	1.643,83	1.082.945,12	10.950,42	4.495,17	13.951,17	24.116,24	2.473,72	2,47
8	Sumatera Selatan	2014	1.649,24	1.342.871,68	10.210,03	773,19	33.098,77	39.199,79	2.840,10	2,84
9	Jambi	2014	1.860,97	687.297,79	9.866,12	4.840,83	13.111,39	21.082,85	2.799,03	2,79
10	Bengkulu	2014	1.427,72	322.319,12	7.089,18	3.174,83	11.218,08	16.713,37	1.798,70	1,79
11	Wilayah Bangka Belitung	2014	1.819,21	209.411,92	8.966,18	4.118,29	12.220,19	11.810,01	2.379,44	2,37
12	Distrik Lampung	2014	1.322,44	1.297.086,44	11.203,48	2.837,08	44.909,12	54.742,98	2.071,02	2,07
13	Wilayah Kalimantan Barat	2014	1.165,19	241.186,28	4.819,28	3.038,48	14.544,13	16.917,78	2.110,03	2,11
14	Wilayah Kalimantan dan Kalteng	2014	1.497,66	180.810,63	7.786,77	2.702,29	11.738,29	20.878,17	2.211,88	2,21
15	Kalimantan Selatan	2014	1.544,63	171.811,69	9.211,13	2.837,18	11.949,67	27.891,47	2.212,28	2,21
...	...	...	...	...	...	...	...	...	...	...
423	Jawa	2022	1.511,21	494.226,10	12.522,11	5.489,26	33.778,76	7.514,84	3.681,08	3,68

Gambar 1. 6 Energi Terjual Rata-Rata

No	Satuan PLN/Provinsi	Tahun	Rumah Tangga	Industri	Busnis	Sosial	Gdg. Kantor Pemerintahan	Penerangan Jalan Umum	Jumlah	(%)
1	Wilayah Aceh	2014	497,09	975,22	1.011,18	455,35	1.487,05	1.074,18	671,09	671,09
2	Wilayah Sumatera Utara	2014	584,85	966,52	1.205,44	779,78	1.249,23	1.327,69	824,99	824,99
3	Wilayah Sumatera Barat	2014	151,57	926,47	1.090,44	722,11	1.197,16	1.197,16	1.017,07	1.017,07
4	Wilayah Riau	2014	645,25	983,91	1.171,23	783,89	1.306,83	1.090,69	826,77	826,77
5	Riau	2014	381,63	978,51	1.110,22	779,78	1.054,89	1.097,68	791,44	791,44
6	Kepulauan Riau	2014	626,64	1.266,77	1.112,77	807,45	1.112,77	1.112,77	1.027,65	1.027,65
7	Wilayah Sumat. Jambi, dan Bengkulu	2014	651,54	980,74	1.221,94	765,82	1.116,14	1.112,17	821,83	821,83
8	Sumatera Selatan	2014	586,19	990,52	1.190,43	762,53	1.295,58	1.111,97	797,97	797,97
9	Jambi	2014	855,48	987,20	1.120,74	798,48	1.157,92	1.117,28	857,85	857,85
10	Bengkulu	2014	872,23	990,26	1.211,42	742,91	1.159,54	1.112,28	794,79	794,79
11	Bengkulu Bengkulu	2014	444,31	985,88	1.104,22	742,91	1.159,54	1.112,28	794,79	794,79
12	Distribusi Lampung	2014	686,47	975,86	1.086,48	686,77	1.260,19	1.104,44	713,73	713,73
13	Wilayah Kalimantan Barat	2014	522,26	1.026,76	1.068,17	564,83	1.211,31	1.099,70	706,72	706,72
14	Wilayah Kalimantan Tengah	2014	588,72	1.014,54	1.040,23	564,83	1.211,31	1.099,70	706,72	706,72
15	Kalimantan Selatan	2014	420,41	1.000,48	1.094,29	652,77	1.200,48	1.100,60	649,49	649,49
423	total	2022	1.118,80	1.066,90	1.543,54	810,53	1.477,86	1.575,43	1.132,03	1.132,03

No	Tahun	JumlahPelanggan	DayaTersambung	EnergiTerjual	EnergiTerjualRataRata	HargaJualListrik	Pendapatan
1	2014	1186899	1173	1965	1656	673	1323000
2	2014	3051822	4202	8271	2710	814	6735801
3	2014	1158772	1429	3005	2593	768	2308240
4	2014	1105800	1735	3338	3018	791	2642106
5	2014	1105800	359	633	3269	1037	657772
6	2014	1729134	2423	4431	2563	787	3492061
7	2014	371170	596	1037	2795	985	1022656
8	2014	405649	424	729	1798	794	579962
9	2014	339065	511	805	2375	616	496640
10	2014	1637824	2063	3392	2071	713	2421526
...	...	...	...	...	...	...	...
300	2022	1630528	4213	5470	3355	1306	7147173
301	2022	1782714	2096	2359	1323	1010	2383894
302	2022	1104170	1453	1206	1093	1209	1460257
303	2022	357673	1700	2943	8238	1420	4180901
304	2022	13275133	23767	40546	3054	1085	44002627
305	2022	11458105	16291	27564	2409	1038	28621809
306	2022	1413515	2313	3326	2953	1142	3802130
307	2022	16310301	30196	56226	3447	1118	62900363
308	2022	3828077	12702	26705	6976	1118	29870292
309	2022	5064418	20889	34578	6827	1292	44698953

Gambar 1. 8 Hasil Transformasi Data

4.2.3. Penanganan Missing Value

Pada tahap ini dilakukan pengecekan terhadap data apakah terdapat *missing value* pada data set dengan melakukan pengecekan menggunakan *tools python*. Tahap ini dapat dilihat pada gambar 1.9 berikut.

```
[ ] # melihat apakah terdapat null value pada data
df.isnull().sum()

Tahun                0
JumlahPelanggan      0
DayaTersambung       0
EnergiTerjual        0
EnergiTerjualRataRata 0
HargaJualListrik     0
Pendapatan           0
dtype: int64
```

Gambar 1. 9 Pengecekan Missing value

Dari gambar 1.9 dapat dilihat bahwa dalam dataset tersebut tidak terdapat *missing value* sehingga tidak terdapat masalah yang menghambat proses pemodelan berikutnya.

4.2.4. Data Training dan Data Testing

Langkah selanjutnya adalah menentukan data *training* data data *testing* dengan melakukan pembagian dari keseluruhan dataset yang telah dilakukan beberapa proses sebelumnya. Sebelum melakukan pembagian data terlebih dahulu ditentukan *variable x* sebagai *variable* yang akan dicari pengaruh terhadap *variable y*, dan *variable y* yang merupakan *variable* yang menjadi acuan prediksi. *Variable x* pada penelitian kali ini adalah Tahun, Jumlah Pelanggan, Daya Tersambung, Energi Terjual, Energi Terjual Rata-Rata, Harga Jual Listrik. Sedangkan variabel *y* adalah Pendapatan. Perbandingan antara data *training* dan data *testing* adalah 70% dan 30% [9], sehingga didapati data *training* sebanyak 216 data dan data *testing* sebanyak 93 data. Pengolahan data dilakukan menggunakan *python*, untuk lebih jelas dapat dilihat pada gambar 1.10 berikut.

```
[ ] # melakukan split data menjadi training dan testing dengan perbandingan 70 30
x_train,x_test,y_train,y_test = train_test_split(x,y, test_size=0.3, random_state=0)
```

Gambar 1. 10 Pembagian Data Training Dan Testing Hasil dari proses pembagian data training

Dan data testing diatas diatas dapat dilihat pada gambar 1.11 – gambar 1.13 dan tabel 1.2 dan 1.3 berikut.

No	Tahun	JumlahPelanggan	DayaTersambung	EnergiTerjual	EnergiTerjualRataRata	HargaJualListrik
262	2021	404578	498	637	1574.0	1113
217	2020	1270916	1797	2715	2136.0	1039
168	2018	10011388	12908	23558	2353.0	820
184	2019	1245177	1612	2819	2264.0	1120
298	2022	270809	519	621	2294.0	1298
...	...	...	...	...	...	...
251	2021	1351906	1936	2912	2154.0	1064
192	2019	277883	339	380	631.0	1150
117	2017	556745	685	1135	527.0	1091
47	2015	471124	555	1049	2226.0	881
172	2018	4395066	18563	32779	7458.0	1057

Gambar 1. 11 Data X Training

No	Tahun	JumlahPelanggan	DayaTersambung	EnergiTerjual	EnergiTerjualRataRata	HargaJualListrik
63	2015	45379	110	207	4551.0	1382
231	2020	1528854	3845	4946	3235.0	1256
167	2018	11469970	19484	35817	3122.0	904
159	2018	341998	429	597	1746.0	805
189	2019	699040	796	1146	1652.0	1078
...	...	...	...	...	...	...
158	2018	255036	304	345	1353.0	1163
267	2021	1018487	1332	1160	1139.0	1106
37	2015	1220930	1511	3063	2509.0	808
132	2017	308124	1333	2067	6709.0	1249
176	2019	1735225	2850	4646	17176.0	1228

Gambar 1. 12 Data X Testing

Tabel 1. 2 Data Y Train

No	Pendapatan
262	709519
217	2822749
168	19335553
184	3157847
298	807040
...	...
251	3102227
192	437293
117	1238072
47	2292112
172	34647774

Tabel 1. 3 Tabel Y Test

No	Pendapatan
63	285355
231	6214049
167	32402294
159	481093
189	1236311
...	...
158	401758
267	1283192
37	2474928
132	2581151
176	5710642

4.3. Data Mining

Tahap utama dalam penelitian kali ini adalah pemodelan data mining menggunakan algoritma regresi linear, *tools* yang digunakan dalam pemodelan adalah *google collaboratory* dengan Bahasa pemrograman *python*. Untuk dapat melakukan pemodelan regresi linear pada *python* dibutuhkan *library sklearn* dengan mengimport modul *LinearRegression* dari *sklearn model*. Proses pemodelan dilakukan dengan melakukan *fit* terhadap data *training x* dan data *training y* yang telah dibuat

sebelumnya, untuk lebih jelas proses pemodelan dapat dilihat pada gambar 1.13 berikut.

```
# membuat model regresi linear
reg = LinearRegression()
# melakukan pemodelan dari data train
reg.fit(x_train, y_train)
```

LinearRegression
LinearRegression()

Gambar 1. 13 Pemodelan Regresi Linear

Dari hasil pemodelan tersebut kita dapat mengetahui nilai coefficient (b) dan intercept (a) dari hasil pemodelan yang telah dilakukan. untuk melihat nilai *coefficient* (b) dan *intercept* (a) digunakan perintah *reg.coef_* dan *reg.intercept_*, untuk lebih jelas dalam mengetahui nilai *coefficient* dan *intercept* dapat dilihat pada gambar 1.14 berikut.

```
[ ] # melihat nilai coefficients(b) dan intercept(a) dari hasil pemodelan
b = reg.coef_
a = reg.intercept_
```

Gambar 1. 14 Coefficient dan Intercept

Hasil *coefficient* yang didapatkan dari proses pemodelan dapat menentukan variabel mana yang lebih mempengaruhi dalam melakukan prediksi terhadap pendapatan PT PLN (Persero). Hasil dari *coefficient* dapat dilihat pada tabel 1.4 berikut.

Tabel 1. 4 Coefficient Hasil Pemodelan

No	features	coef_value
1	Tahun	112.701.729.828
2	JumlahPelanggan	-0.666624
3	DayaTersambung	1.457.991.746
4	EnergiTerjual	420.319.625
5	EnergiTerjualRataRata	50.814.261
6	HargaJualListrik	664.143.193

Dari tabel 1.4 tersebut dapat diketahui bahwa hampir semua variabel x yang dilakukan pengujian memiliki pengaruh yang sangat signifikan terhadap prediksi pendapatan, dengan demikian variabel tahun, Daya tersambung, Energi Terjual, Energi Terjual Rata-Rata, dan Harga Jual Listrik lebih mempengaruhi terhadap prediksi pendapatan dibandingkan dengan variabel Jumlah Pelanggan. Dari penjelasan hasil pemodelan tersebut tentunya membuktikan bahwa pemodelan terhadap data perkelompok pelanggan untuk memprediksi pendapatan PT PLN sangat baik.

4.4. Evaluation

Setelah dilakukan tahap pemodelan langkah selanjutnya adalah melakukan pengujian dan evaluasi model dengan tujuan mengetahui sejauh keberhasilan dari pemodelan yang telah dilakukan sebelumnya. Untuk pengujian, akan dilakukan prediksi menggunakan model yang telah dibuat dengan data uji x. Selanjutnya, hasil prediksi akan dibandingkan dengan data uji y dan untuk mengevaluasi model regresi linear berganda secara umum menggunakan 3

matrik yaitu *Mean Absolute Error (MAE)*, *Mean Squared Error (MSE)*, dan *Root Mean Squared Error (RMSE)* [10] sedangkan untuk mengetahui akurasi dari model dilakukan evaluasi menggunakan matrik *r_score*. Untuk lebih jelas dari tahapan tersebut dapat dilihat pada sub bab berikut.

4.4.1. Pengujian Model

Seperti yang sudah dijelaskan diatas pengujian dilakukan dengan menguji model dengan data uji x dan membandingkan hasil prediksi dengan data uji y. Dengan melihat selisih dari hasil prediksi, akan terlihat akurasi dari pemodelan yang dibuat apakah model tersebut layak digunakan sebagai acuan prediksi pendapatan PT PLN. Untuk implementasi model terhadap data uji x dapat dilihat dari gambar 1.15 kemudian untuk hasil dan selisih dari pengujian tersebut dapat dilihat pada tabel 1.5 berikut.

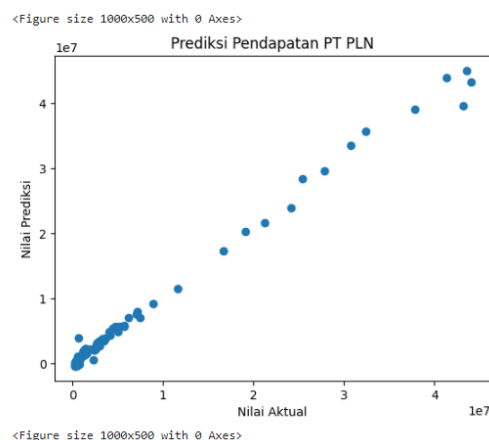
```
[ ] # melakukan pengujian data testing terhadap model
prediction = reg.predict(x_test)
# print("hasil prediksi Pendapatan ",prediction)
pd.DataFrame(prediction)
```

Gambar 1. 15 Pengujian Model

Tabel 1. 5 Hasil dan Selisih Prediksi

No	nilai aktual	nilai prediksi	selisih
63	285355	1,05E+11	-1,80E+11
231	6214049	6,97E+12	7,52E+11
167	32402294	3,57E+13	3,25E+12
159	481093	3,49E+11	-1,33E+11
189	1236311	1,17E+12	-7,07E+10
...	...	...	...
158	401758	3,36E+11	-6,56E+10
267	1283192	1,96E+12	6,75E+11
37	2474928	2,08E+12	-3,96E+11
132	2581151	2,74E+12	1,60E+11
176	5710642	5,83E+12	1,18E+11

Untuk melihat visualisasi diagram *scatter* dari hasil pengujian terhadap data *training* dapat dilihat pada gambar 1.16 berikut



Gambar 1. 16 Grafik Scatter Pengujian Model

4.4.2. Evaluasi Model

Evaluasi model dilakukan untuk mengetahui berapa tingkat *error* dan akurasi yang didapatkan dari hasil pemodelan yang dilakukan. Untuk melihat nilai *error* yang dihasilkan dari model dilakukan dengan 3 matrik yaitu *Mean Absolute Error (MAE)*, *Mean Squared Error (MSE)*, dan *Root Mean Squared Error (RMSE)*. Matrik tersebut dipanggil dengan menggunakan *library sklearn* dengan mengimport masing-masing matrik dari *sklearn.metrics*. untuk lebih memahami dan mengetahui hasil dari nilai *error* model dapat dilihat pada gambar 1.17 berikut.

```
[ ] from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
from sklearn.metrics import mean_squared_error
import math
# pengujian error menggunakan MAE MSE RMSE
print("Mean Absolute Error (MAE)", mean_absolute_error(prediction, y_test))
print("Mean Squared Error (MSE)", mean_squared_error(prediction, y_test))
print("Root Mean Squared Error (RMSE)", math.sqrt(mean_squared_error(prediction, y_test)))

Mean Absolute Error (MAE) 587819.4573281321
Mean Squared Error (MSE) 900306740153.5876
Root Mean Squared Error (RMSE) 948844.9505338517
```

Gambar 1. 17 Pengujian Error

Dari gambar 1.17 diketahui *Mean Absolute Error (MAE)* bernilai 587819.4573, *Mean Squared Error (MSE)* bernilai 900306740153.5876, dan *Root Mean Squared Error (RMSE)* bernilai 94884.9505. selanjutnya untuk melihat akurasi dari model peneliti menggunakan matrik *r2 score* yang di *import* dari *library sklearn.metrics* dengan hasil dari akurasi yang didapatkan adalah 0.9923. untuk lebih jelas hasil dari akurasi pemodelan dapat dilihat pada gambar 1.18 berikut.

```
[ ] from sklearn.metrics import r2_score
# melakukan pengujian dari model dengan r2_score yang telah dibuat diatas
r2_score(y_test, prediction)

0.9923863967546535
```

Gambar 1. 18 Pengujian Akurasi

5. KESIMPULAN DAN SARAN

Kesimpulan yang dapat diambil dalam penelitian ini adalah dapat diketahui dari hasil pemodelan yang dilakukan untuk memprediksi pendapatan PT PLN (Persero) dengan menggunakan algoritma regresi linear berganda, bahwa data per kelompok pelanggan hampir secara keseluruhan dapat memengaruhi tingkat pendapatan PT PLN. Variabel yang secara signifikan mempengaruhi pendapatan PT PLN antara lain adalah Tahun, Daya tersambung, Energi Terjual, Energi Terjual Rata-Rata, dan Harga Jual Listrik. Hal tersebut membuktikan bahwa data historis dapat digunakan dalam melakukan analisis prediksi terhadap pendapatan PT PLN (Persero) dengan tingkat akurasi 0.9923, yang berarti hasil pemodelan yang dilakukan merupakan pemodelan yang sesuai dengan data per kelompok pelanggan. penelitian selanjutnya juga disarankan untuk melakukan Analisa terhadap pendapatan PT PLN menggunakan *tools* dan metode yang berbeda untuk menguji tingkat efektifitas model terhadap prediksi yang dilakukan.

DAFTAR PUSTAKA

- [1] A. Hartawan, "Rancang Bangun Sistem Prediksi

Penggunaan Listrik Jangka Panjang Menggunakan Metode Regresi Linier Berbasis WEB (Studi Kasus PT. PLN (Persero) Unit Induk, Distribusi Bali)," *J. Petik*, vol. 8, no. 1, pp. 35–41, 2022, doi: 10.31980/jpetik.v8i1.1337.

- [2] N. Nazeriandy, Y. Syahra, and M. Syaifudin, "Penerapan Data Mining Untuk Memprediksi Penggunaan Daya Listrik Pada PT.PLN (Persero) Rayon Medan Selatan Dengan Menggunakan Metode Regresi Linier Berganda," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 20, no. 1, p. 20, 2021, doi: 10.53513/jis.v20i1.2431.
- [3] T. I. Janwardi, "Prediksi Peningkatan Kebutuhan Tenaga Listrik Provinsi Jambi Tahun 2020-2025," *J. Elektron. List. dan Teknol. Inf. Terap.*, vol. 2, no. 1, p. 21, 2020, doi: 10.37338/e.v2i1.125.
- [4] P. H. Hutabarat, M. F. Zambak, and S. Suwarno, "Prediksi Kebutuhan Energi Listrik Wilayah Pln Kota Parapat Simalungun Sampai Tahun 2024," *J. Electr. Syst. Control Eng.*, vol. 5, no. 2, pp. 52–58, 2022, doi: 10.31289/jesce.v5i2.5757.
- [5] S. Bandri, "Prediksi Perkembangan Kebutuhan Energi Listrik di Unit PLN Kayu Aro," *Menara Ilmu*, vol. XIII, no. 6, pp. 187–205, 2019.
- [6] S. Sinaga, "MENILAI TINGKAT PERUBAHAN LABA PERUSAHAAN BERDASARKAN PERUBAHAN NILAI RATIO PROFITABILITAS," *J. Sci. Soc. Res.*, vol. 4307, no. 1, pp. 66–71, 2023.
- [7] A. S. Sukri, J. J. Rikumahu, and E. T. Mbitu, "Peramalan Beban Penyulang Passo Pt. Pln (Persero) Area Ambon Menggunakan Metode Regresi Linier Berganda," *J. ELKO (Elektrikal dan Komputer)*, vol. 2, no. 2, pp. 152–164, 2023, doi: 10.54463/je.v2i2.50.
- [8] D. B. Lomet, "Bulletin of the Technical Committee on Data Engineering," *Bull. Tech. Comm. Data Eng.*, vol. 24, no. 4, pp. 1–56, 2001, [Online]. Available: papers2://publication/uuid/30073F7F-1B7C-4496-ADA4-94FF4E6EE8F7
- [9] Baiq Nurul Azmi, Arief Hermawan, and Donny Avianto, "Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma Decision Tree untuk Klasifikasi Penderita Penyakit Liver," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 4, pp. 281–290, 2023, doi: 10.35746/jtim.v4i4.298.
- [10] I. D. Kurniawan, A. Mhm, and S. J. Legowo, "Comparative Study of Error Estimation on Linear Model and Instantaneous Model in Estimating Travel Time Based on Instantaneous Speed (Study Location: North Ring Road Surakarta)," *Civ. Eng. Matrix J.*, vol. 4, no. 4, pp. 1264–1272, 2016, [Online]. Available: https://jurnal.uns.ac.id/matriks/article/view/3706