

Implementasi *Linear Regression* Untuk Estimasi Realisasi Pendapatan Asli Daerah Pemerintah Kota

Tina Tri Wulansari^{1*}, Dody Novandi²

Sistem Informasi, Universitas Mulia Kampus Samarinda, Samarinda, Indonesia
 *e-mail *Corresponding Author*: tina@universitasmulia.ac.id

Abstract

*The Government of Samarinda City every year always sets a target for local revenue, but in setting the target there is always no data regarding the estimated realization of the target that has been set. One method that can be implemented to obtain information on the estimation of the realization of local revenue is data mining using the Linear Regression algorithm. The research carried out applies the standardization of a process, namely the Cross-Industry Standard Process Model for Data Mining (CRISP-DM), using the Samarinda City Government's local revenue realization dataset for the 2017 to 2021 fiscal year as much as 178 data. The ideal ratio scenario for the percentage of training and testing data for the dataset used is 95:5 and applies the Range Transformation normalization method, so that the resulting model to calculate the realization estimate is by the formula: $1.133 * TARGET + 0.003$.*

Keywords: Data Mining; *Linear Regression*; CRISP-DM

Abstrak

Pemerintah Kota Samarinda setiap tahunnya selalu menetapkan target pendapatan asli daerah, namun dalam penetapan target selalu tidak tersaji data terkait estimasi realisasi dari target yang telah ditetapkan. Salah satu metode yang bisa diimplementasikan untuk memperoleh informasi estimasi realisasi pendapatan asli daerah adalah data *mining* dengan menggunakan algoritme *Linear Regression*. Penelitian yang dilakukan menerapkan standarisasi dari sebuah proses yaitu *Cross-Industry Standard Process Model for Data Mining (CRISP-DM)*, dengan menggunakan dataset realisasi pendapatan asli daerah Pemerintah Kota Samarinda tahun anggaran 2017 sampai dengan 2021 sebanyak 178 data. Skenario ratio prosentase data *training* dan *testing* yang ideal untuk dataset yang digunakan yaitu 95:5 serta menerapkan metode normalisasi *Range Transformation*, sehingga model yang dihasilkan untuk melakukan perhitungan estimasi realisasi yaitu dengan formula: $1,133 * TARGET + 0,003$.

Kata kunci: Data Mining; *Linear Regression*; CRISP-DM

1. Pendahuluan

Berdasarkan Undang-undang Nomor 33 Tahun 2004 tentang Perimbangan Keuangan antara Pemerintah dan Pemerintahan Daerah, Pendapatan Asli Daerah (PAD) merupakan sumber penerimaan daerah yang digali untuk digunakan sebagai modal dasar pemda dalam membiayai pembangunan dan usaha-usaha daerah untuk memperkecil ketergantungan dana dari pemerintah pusat [1]. PAD merupakan salah komponen yang selalu diproyeksi dalam penyusunan anggaran pemerintah Kabupaten/Kota. Dalam setiap penyusunan anggaran pendapatan dan belanja daerah (APBD), PAD terdiri dari beberapa komponen yaitu: Pos Penerimaan Pajak yang terdiri atas Pajak Daerah dan Retribusi Daerah, Pos Penerimaan Non Pajak berupa penerimaan hasil Perusahaan Milik Daerah, serta pos penerimaan investasi serta pengelolaan sumber daya alam [2].

Pemerintah Kabupaten/Kota selalu menetapkan target penyampaian PAD. Kontribusi yang berhasil didapatkan dari PAD memiliki peranan untuk mewujudkan daerah/kota supaya lebih tumbuh dan dapat menumbuhkan kesejahteraan masyarakat. Dalam penyusunan anggaran pendapatan dan belanja daerah (APBD), Pemerintah Kota Samarinda setiap tahunnya selalu menetapkan target PAD dimana terdapat beberapa komponen yaitu pajak daerah, retribusi daerah, hasil pengelolaan kekayaan daerah, dan lain-lain pendapatan daerah

yang sah secara bersama-sama. Dalam penetapan target PAD selalu tidak tersaji informasi terkait proyeksi realisasi penerimaan dari target yang telah ditetapkan sehingga kondisi ini dapat membuat proses penyusunan APBD pada tahun berikutnya menjadi terhambat. Untuk dapat melakukan proyeksi realisasi penerimaan PAD, maka diperlukan sebuah rumus yang bisa melakukan estimasi secara cepat dan mudah.

Salah satu metode yang dapat digunakan sebagai suatu solusi untuk memperoleh informasi tentang estimasi realisasi PAD adalah dengan menggunakan teknik *data mining*. *Data mining* dapat membantu sebuah organisasi atau individual untuk melakukan eksplorasi data yang tersembunyi dan analisis data serta bisa mendukung dalam pembuatan kebijakan yang tepat berdasarkan pengetahuan yang didapatkan [3]. Tujuan akhir dari *data mining* adalah untuk menemukan pengetahuan dari data, dengan memanfaatkan teknologi dan algoritme yang tepat maka akan didapatkan wawasan yang dicapai [4]. Salah satu algoritme yang dapat digunakan dalam melakukan estimasi dengan pendekatan *data mining* adalah algoritme *Linear Regression*.

Penelitian mengenai peramalan dengan memanfaatkan algoritme *Regresi Linear* dengan menggunakan data aktual selama 5 tahun dapat meramalkan jumlah jasa penerjemahan di Azzam Translator dengan hasil RMSE yang dihasilkan dengan nilai 0,832 (predikat bagus) [5]. Hasil penelitian yang dilakukan peneliti lainnya terkait prediksi nilai tukar mata uang rupiah terhadap US dollar menggunakan algoritme *linear regression*, menunjukkan bahwa algoritme *linear regression* cukup baik diimplementasikan pada kasus prediksi nilai tukar mata uang menurut tanggal spesifik. Dengan menggunakan sebanyak 100 record untuk data uji dihasilkan kemampuan akurasi sebesar 95% dengan nilai threshold sebesar 30 dan nilai RMSE sebesar 14,951 [6].

Selanjutnya penelitian tentang prediksi harga minyak kelapa sawit dengan menggunakan algoritme *Linear Regression* dan *Random Forest* dengan menggunakan dataset minyak kelapa sawit sejak tanggal 1 Januari 2018 sampai dengan 30 November 2021 dengan skema prosentase data latih dan data uji yaitu 90:10, 80:20 dan 70:30 memberikan kesimpulan bahwa algoritme *Linear Regression* lebih bagus daripada *Random Forest* yang dibuktikan dengan nilai *root mean square error* (RMSE) untuk algoritme *Linear Regression* sebesar 30.227 sedangkan *Random Forest* adalah 32,924 dengan skenario data latih dan uji 70:30 [7].

Algoritme yang digunakan dalam penelitian ini untuk melakukan estimasi realisasi PAD terhadap target yang sudah ditentukan adalah *Linear Regression*. Sebuah penelitian yang pernah dilakukan sebelumnya dalam melakukan prediksi penjualan properti pada PT. XYZ menyimpulkan bahwa pemanfaatan algoritme *Linear Regression* bisa melakukan peramalan penjualan properti dengan cukup baik [8].

Penelitian ini bertujuan menguji algoritme *Linear Regression* untuk menghasilkan sebuah model berupa persamaan regresi linear yang dapat melakukan perhitungan estimasi realisasi penerimaan PAD terhadap target PAD yang telah ditetapkan sebelumnya.

2. Tinjauan Pustaka

Pada penelitian ini dideskripsikan beberapa penelitian terdahulu yang berhubungan dengan implementasi algoritme *Linear Regression* untuk melakukan estimasi. Sebuah penelitian melakukan peramalan pendapatan reksa dana dalam rentang waktu 1 tahun dengan menerapkan algoritme *regresi linier* sederhana sehingga mampu melihat trend masa depan grafik pendapatan yang didapatkan dari investasi itu sendiri. Data yang digunakan dalam penelitian merupakan data pendapatan dalam setahun sebelumnya. Setelah dilakukan peramalan, didapatkan hasil perkiraan pendapatan pada bulan selanjutnya pada tahun tersebut. Hasil peramalan digabungkan dengan data pendapatan realisasi hingga terbentuk satu garis pada grafik. Untuk garis hasil peramalan dan garis pendapatan realisasi diberikan warna yang berbeda. Berdasarkan hasil pengujian, penelitian mengambil kesimpulan bahwa metode regresi linier sederhana cocok untuk meramalkan pendapatan Reksa Dana dalam setahun dengan syarat parameter yang digunakan hanya bulan dan pendapatan [9].

Selanjutnya penelitian tentang perencanaan anggaran pinjaman dengan prediksi regresi linier sederhana dan optimasi menggunakan metode *fuzzy tsukamoto* sehingga bisa melakukan prediksi perencanaan anggaran sehingga pinjaman dapat dimaksimalkan untuk menentukan target koperasi dalam usaha simpan pinjaman dan menjaga kas koperasi tetap stabil. Dataset yang digunakan dalam penelitian yaitu data modal dan realisasi pinjaman dari Januari 2018 sampai Maret 2019. Setelah dilakukan prediksi berdasarkan rumus persamaan

regresi linier sederhana yang ditetapkan, maka didapatkan hasil perkiraan realisasi anggaran pinjaman untuk bulan April 2019, dengan syarat parameter yang digunakan hanya modal dan realisasi pinjaman. Berdasarkan hasil pengujian, penelitian mengambil kesimpulan bahwa metode regresi linier sederhana cocok untuk memprediksi anggaran pinjaman bulan April 2019 dengan jumlah modal Rp. 500.000.000 dimana prediksi realisasi anggaran pinjaman sebesar Rp. 441.690.000, setelah optimasi menggunakan Metode *Fuzzy Tsukamoto* adalah Rp. 396.960.000 lebih hemat Rp. 44.730.000 atau 10.12% dengan batas maksimal jumlah anggaran yang bisa dioptimasi adalah Rp. 535.000.000 [10].

Persamaan dari kedua penelitian tersebut di atas dengan penelitian yang akan dilakukan yaitu menggunakan algoritme *Linear Regression* sedangkan perbedaannya terletak penggunaan metode normalisasi dan *Cross-Industry Standard Process Model for Data Mining (CRISP-DM)*.

3. Metodologi

Penelitian yang dilakukan menggunakan standarisasi dari sebuah proses yaitu *Cross-Industry Standard Process Model for Data Mining (CRISP-DM)* [11], dimana diawali dari *Business Understanding* sampai dengan tahapan *Deployment* [12]. Dengan menggunakan prosedur dari *CRISP-DM*, maka tahapan penelitian yang dijalankan dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

Berikut ini beberapa tahapan penelitian yang diimplementasikan dalam penelitian ini adalah sebagai berikut:

a. *Business Understanding*

Wawasan yang didapatkan adalah sebuah model untuk melakukan estimasi realisasi pendapatan pajak daerah terhadap target pendapatan asli daerah (PAD) yang sudah ditetapkan. Model yang dihasilkan berupa rumus untuk melakukan perhitungan estimasi realisasi dimana output tersebut didapatkan berdasarkan dataset yang telah dilatih.

b. *Data Understanding*

Dataset yang didapatkan dalam penelitian ini berupa data realisasi pendapatan asli daerah (PAD) Pemerintah Kota Samarinda tahun anggaran 2017 sampai dengan 2021. Ada beberapa atribut yang terdapat pada *dataset* yang didapatkan dari Bapenda Kota Samarinda yaitu: Kode Rekening, Uraian, Target Perubahan, Realisasi Bulan Ini (Desember), Realisasi (%), dan Realisasi s/d Bulan Lalu.

Setiap atribut-atribut pada *dataset* yang digunakan mempunyai nilai sebagai berikut:

- 1) Kode Rekening merupakan kolom untuk kode rekening.
- 2) Uraian merupakan kolom untuk jenis pendapatan.
- 3) Target Perubahan merupakan kolom untuk jumlah target yang akan dicapai.
- 4) Realisasi hingga Bulan Lalu merupakan kolom untuk jumlah realisasi pendapatan pada periode sebelumnya.

- 5) Realisasi Bulan Ini (Desember) merupakan kolom untuk jumlah realisasi pendapatan pada periode bulan Desember.
- 6) Realisasi merupakan kolom untuk jumlah realisasi pendapatan dalam bentuk prosentase.
- 7) Realisasi hingga Bulan Lalu merupakan kolom untuk jumlah keseluruhan realisasi pendapatan pada periode 1 tahun berjalan.

c. *Data Preparation*

Pada tahapan ini akan dilakukan *pre-processing* pada *dataset* yang dengan beberapa tahapan yang dilakukan adalah sebagai berikut:

- 1) Pembersihan data
Setelah memperhatikan data yang didapatkan, maka dilakukan proses pembersihan *dataset* yang digunakan dalam penelitian.
- 2) Melakukan seleksi data
Dalam penelitian ini diambil beberapa atribut dari *dataset* yang didapatkan untuk digunakan dalam proses *data mining*. Adapun atribut yang digunakan yaitu: Uraian, Target Perubahan dan Realisasi s/d Bulan Lalu. Jumlah data yang digunakan sebanyak 178 data.

Tabel 1 Dataset Realisasi PAD Samarinda Tahun 2017- 2021

Uraian	Target Perubahan	Realisasi s/d Bulan Lalu
Hotel Bintang Lima 2017	8.000.000.000	5.151.229.848
Hotel Bintang Empat 2017	8.472.000.000	9.726.236.277
Hotel Bintang Tiga 2017	3.000.000.000	5.979.531.474
Hotel Bintang Dua 2017	3.400.000.000	2.629.028.343
Hotel Bintang Satu 2017	148.000.000	71.376.300
Hotel Melati Satu 2017	1.500.000.000	1.549.513.564
Losmen/Rumah		
Penginapan/Pesanggrahan/Rumah Kos 2017	980.000.000	650.085.567
Restoran 2017	25.500.000.000	23.042.888.917
.....		
Pajak Mineral Bukan Logam dan Batuan 2021	1.348.910.000	1.351.754.000

- 3) Melakukan transformasi data
Dilakukan proses perubahan identitas atribut-atribut pada *dataset* yang digunakan dalam penelitian ini. Target Perubahan dirubah menjadi Target, atribut Realisasi s/d Bulan Lalu menjadi atribut Realisasi.

Tabel 2 Perubahan Atribut Dataset Realisasi PAD Samarinda Tahun 2017- 2021

Uraian	Target	Realisasi
Hotel Bintang Lima 2017	8.000.000.000	5.151.229.848
Hotel Bintang Empat 2017	8.472.000.000	9.726.236.277
Hotel Bintang Tiga 2017	3.000.000.000	5.979.531.474
Hotel Bintang Dua 2017	3.400.000.000	2.629.028.343
Hotel Bintang Satu 2017	148.000.000	71.376.300
Hotel Melati Satu 2017	1.500.000.000	1.549.513.564
Losmen/Rumah		
Penginapan/Pesanggrahan/Rumah Kos 2017	980.000.000	650.085.567
Restoran 2017	25.500.000.000	23.042.888.917
.....		
Pajak Mineral Bukan Logam dan Batuan 2021	1.348.910.000	1.351.754.000

d. *Modeling*

Penelitian ini menerapkan sebuah salah satu metode dalam data *mining* yaitu *estimasi*. Untuk mendukung semua kegiatan penambangan data, penelitian ini

menggunakan algoritme *linear regression*, normalisasi data, aplikasi *Rapid Miner* dan *Microsoft Excel*. Bentuk persamaan yang diterapkan adalah regresi linear sederhana dengan model persamaannya adalah: $y = a + bX$, dimana y = variabel dependent, x = variabel independent, a = konstanta dan b = koefisien regresi [13].

Skenario prosentase pembagian data *training* dan *testing* yaitu 60:40, 65:35, 70:30, 75:25, 80:20, 85:15, 90:10, dan 95:5 sedangkan metode normalisasi yang digunakan yaitu *Z-Transformation* dan *Range Transformation* [14]. Proses normalisasi ini dilakukan karena dataset yang digunakan memiliki rentang nilai yang tidak seimbang pada atribut target dan realisasi [15].

e. *Evaluation*

Pada tahapan ini dilakukan interpretasi terhadap hasil data *mining* pada dataset yang sudah didapatkan. Evaluasi ini dilakukan agar proses pada *modeling phase* sesuai dengan *business understanding*.

Setelah dilakukan beberapa pengujian data, dilakukan evaluasi akhir dengan menggunakan parameter *root mean square error (RMSE)*. Pengujian dan evaluasi dilakukan agar bisa menentukan pilihan ratio prosentase dan metode normalisasi data yang ideal digunakan dengan membandingkan nilai *root mean square error (RMSE)*. Nilai RMSE yang paling rendah menunjukkan proses estimasi yang baik [16]. Rumus persamaan yang digunakan untuk menghitung RMSE adalah sebagai berikut[17]:

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n}} \dots\dots\dots (1)$$

keterangan:

$\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ = nilai yang diprediksi

y_1, y_2, \dots, y_n = nilai yang diamati

n = jumlah data yang diamati

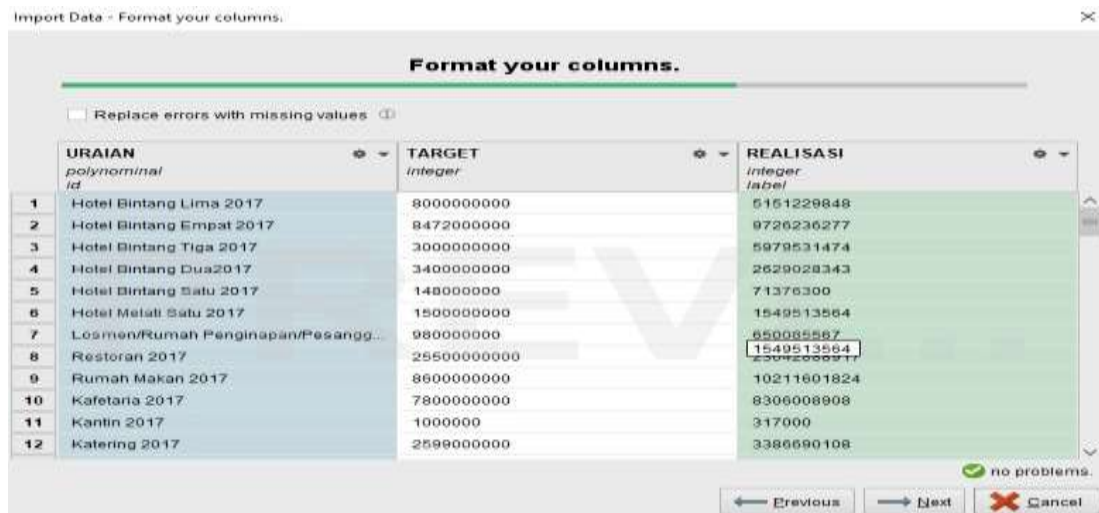
f. *Deployment*

Pada tahapan ini dilakukan penyampaian presentasi tentang proses pembentukan model untuk perhitungan aksiran realisasi pendapatan asli daerah (PAD) Pemerintah Kota Samarinda.

4. Hasil dan Pembahasan

Berdasarkan algoritme *Linear Regression*, tahapan dalam proses pembentukan model untuk fungsi estimasi yang dilakukan adalah sebagai berikut:

- a. Mengimport *dataset* kedalam *Rapid Miner*
- b. Menentukan role atribut sesuai kebutuhan, dimana atribut Uraian sebagai *id* dan atribut Realisasi sebagai *label*.



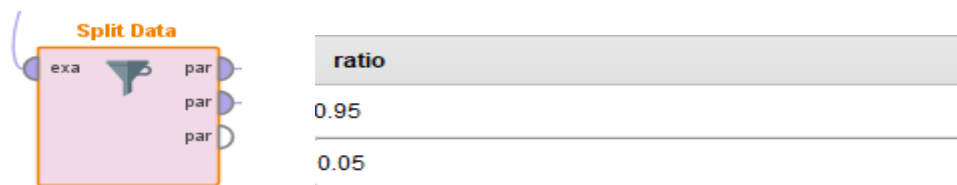
Gambar 2. Penentuan Role Atribut

- c. Melakukan normalisasi data pada atribut target dan realisasi dengan menggunakan metode tertentu. Proses ini dilakukan karena *dataset* yang digunakan memiliki rentang nilai yang tidak seimbang pada atribut target dan realisasi.



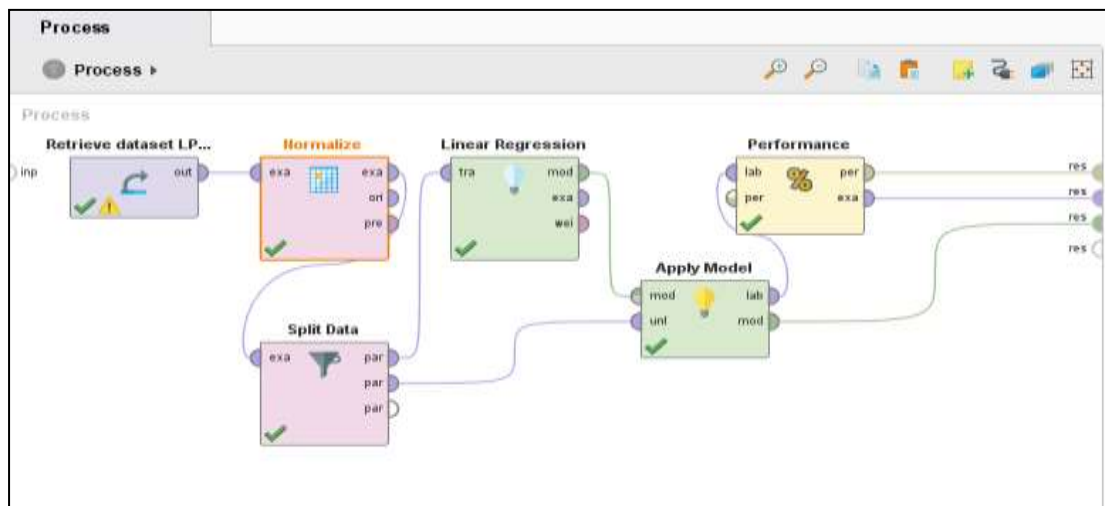
Gambar 3. Normalisasi Data

- d. Melakukan split data terhadap *dataset* yang telah di *import*. Proses ini dilakukan untuk menentukan ratio prosentase jumlah data untuk *training* dan *testing* terhadap keseluruhan *dataset*.



Gambar 4. Proses Penentuan Ratio Data *Training* dan *Testing*

- e. Memasukan operator *Linear Regression*.
- f. Melakukan pengujian metode dan evaluasi hasil dengan memilih operator *apply model* dan *performance* untuk *Linear Regression*. Pengujian dan evaluasi dilakukan agar bisa menentukan pilihan ratio prosentase dan metode normalisasi data yang ideal digunakan dengan membandingkan nilai *root mean square error* (RMSE) yang dihasilkan.



Gambar 5 Proses *Linear Regression* di Rapid Miner

Setelah melakukan pengujian dengan menggunakan Rapid Miner, maka didapatkan beberapa nilai *root mean square error* (RMSE) yang dapat dilihat pada tabel 2.

Tabel 2 Nilai RMSE

Ratio Split Data		Metode Normalisasi Rapid Miner	
Data Training	Data Testing	Z-Transformation	Range Transformation
60	40	0,219	0,043
65	35	0,171	0,034
70	30	0,185	0,036

Ratio Split Data		Metode Normalisasi Rapid Miner	
Data Training	Data Testing	Z-Transformation	Range Transformation
75	25	0,161	0,032
80	20	0,177	0,035
85	15	0,203	0,040
90	10	0,245	0,048
95	5	0,086	0,017

Dengan melihat hasil ujicoba, maka nilai RMSE yang paling rendah dihasilkan adalah 0.017, dengan skenario ratio prosentase data *training* dan *testing* yang ideal adalah dengan perbandingan 95:5 serta menggunakan metode normalisasi *Range Transformation*. Berdasarkan hasil pengujian metode dengan algoritme *Linear Regression*, maka model untuk melakukan perhitungan estimasi realisasi terhadap target pendapatan asli daerah (PAD) dapat menggunakan formulasi: $1,133 * TARGET + 0,003$. Hasil penelitian yang telah dihasilkan selaras dengan peneliti yang telah melakukan penelitian tentang prediksi perencanaan anggaran dengan menggunakan model persamaan regresi linear yang telah dihasilkan [10]. Terkait tentang parameter nilai RMSE, penelitian ini selaras dengan peneliti yang memprediksi jumlah jasa penerjemahan di Azzam Translator yang menggunakan nilai RMSE sebagai acuan menilai sebuah prediksi yang baik [5]. Temuan ini menjadi menarik untuk dilakukan penelitian lebih lanjut, karena belum dilakukannya ujicoba untuk melakukan penilaian tingkat akurasi terhadap formulasi yang dihasilkan.

5. Simpulan

Dengan telah dilakukannya percobaan sebanyak 8 skenario skenario ratio prosentase data *training* dan *testing* yaitu: 60:40, 65:35, 70:30, 75:25, 80:20, 85:15, 90:10, 95:5, maka akurasi yang ideal untuk menghasilkan model yang baik adalah perbandingan 95:5 serta menerapkan metode normalisasi *Range Transformation*. Dari hasil model *Linear Regression* yang berhasil disajikan, maka dataset pada kegiatan penelitian ini bisa dijadikan sebagai acuan untuk pembuatan sebuah sistem dalam melakukan estimasi realisasi penerimaan PAD. Rekomendasi untuk keberlanjutan penelitian ini perlunya dilakukan pengukuran tingkat akurasi kinerja dari model yang telah dihasilkan.

Daftar Referensi

- [1] D. Saraswati, Y. S. Rioni, "Pengaruh Pendapatan Asli Daerah, Ukuran Pemerintah Daerah, Leverage terhadap Kinerja Keuangan Pemerintah Daerah," *J. Akunt. Bisnis Publik*, vol. 9, no. 2, pp. 110–120, 2019.
- [2] M. S. Nasir, "Analisis Sumber-Sumber Pendapatan Asli Daerah Setelah Satu Dekadeotonomi Daerah," *J. Din. Ekon. Pembang.*, vol. 2, no. 1, pp. 30-45., 2019, doi: 10.14710/jdep.2.1.
- [3] V. Fitriyani, "Analisis Clustering Provinsi Indonesia Berdasarkan Persebaran Virus Corona (Covid-19) Menggunakan Algoritme K-Means," *Pros. Pendidik. Mat. dan Mat.*, vol. 3, no. 2721, pp. 1-8, 2021.
- [4] R. T. Vuldari, *Data Mining - Teori dan Aplikasi Rapidminer*. Yogyakarta: Gava Media, 2017.
- [5] W. Warsino, "Metode Peramalan Permintaan Jasa Penerjemahan Bahasa Asing Dengan Algorithma Linear Regression, Menggunakan Rapidminer. Studi Kasus: Azzam Translator Bekasi," *Santika J. Ilm. Sains dan Teknol.*, vol. 7, no. 2, pp. 621–628, 2017, [Online]. Available: <https://www.jurnal.ummi.ac.id/index.php/santika/article/view/657/301>
- [6] T. Rachman, "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier," *Angew. Chemie Int. Ed.* 6(11), 951–952., pp. 10–27, 2018.
- [7] H. s Toluhula, "Jurnal Ilmiah Wahana Pendidikan," *J. Ilm. Wahana Pendidik.* <https://jurnal.unibrah.ac.id/index.php/JIWP>, vol. 8, no. 3, pp. 178–183, 2022, doi: 10.5281/zenodo.6559603.
- [8] G. N. Ayuni and D. Fitriyah, "Penerapan metode Regresi Linear untuk prediksi penjualan properti pada PT XYZ," *J. Telemat.*, vol. 14, no. 2, pp. 79–86, 2019, [Online]. Available: <https://journal.ithb.ac.id/telematika/article/view/321>
- [9] H. Jayanti, "Peramalan Pendapatan Reksa Dana Dalam Setahun Menggunakan Metode

- Regresi Linier Sederhana,” *J. FIKI*, vol. VIII, no. 2, pp. 2087–2372, 2018, [Online]. Available: <http://jurnal.unnur.ac.id/index.php/jurnalfiki>
- [10] Y. Yanitasari, D. Dedih, and U. Mustofa, “Perencanaan Anggaran Pinjaman Dengan Prediksi Regresi Linier Sederhana Dan Optimasi Menggunakan Metode Fuzzy Tsukamoto,” *Ilk. J. Ilm.*, vol. 11, no. 3, pp. 206–213, 2019.
- [11] Y. Suhanda, I. Kurniati, and S. Norma, “Penerapan Metode Crisp-DM Dengan Algoritme K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik,” *J. Teknol. Inform. dan Komput.*, vol. 6, no. 2, pp. 12–20, 2020, doi: 10.37012/jtik.v6i2.299.
- [12] D. Astuti, “Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritme K-Means Clustering,” *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 1, no. 2, pp. 60–72, 2019, doi: 10.20895/inista.v1i2.71.
- [13] Petrus Katemba and R. Koro, “Menggunakan Regresi Linear,” *J. Ilm. Flash*, no. 3, pp. 42–51, 2015.
- [14] RAPIDMINER, “Rapidminer Documentation :Normalize,” 2022. <https://docs.rapidminer.com/9.10/tutorial/operators/cleansing/normalization/normalize.html> (accessed Aug. 11, 2022).
- [15] D. A. Nasution, H. H. Khotimah, and N. Chamidah, “Perbandingan Normalisasi Data untuk Klasifikasi Wine Menggunakan Algoritme K-NN,” *Comput. Eng. Sci. Syst. J.*, vol. 4, no. 1, pp. 78-82, 2019, doi: 10.24114/cess.v4i1.11458.
- [16] D. I. Mulyana, “Optimasi Prediksi Harga Udang Vaname dengan Metode RMSE dan MAE Dalam Algoritme Regresi Linier,” Vol. 13, no. 01, pp. 50–58.
- [17] A. S. B. Karno *et al.*, “Optimalisasi Data Terbatas Prediksi Jangka Panjang Covid-19 Dengan Kombinasi Lstm Dan GRU,” *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 4, no. 1, pp. 453–457, 2020, [Online]. Available: <https://ejournal.stitpn.ac.id/index.php/bintang%0Aejournal.citrabakti.ac.id/jipcb/article/download>