PREDIKSI TINGKAT KEKERASAN PADA PEREMPUAN DAN ANAK DI KABUPATEN GORONTALO MENGGUNAKAN METODE MACHINE LEARNING

Citra Ayu[1], Alter Lasarudin[2], Wahyudin Hasyim[3]*
[1,2,3]Universitas Muhammadiyah Gorontalo
citraayumhn@gmail.com*
Universitas Muhammadiyah Gorontalo, Jl. Prof. Mansoer Pateda, Kab. Gorontalo, Indonesia

ABSTRACT

Violence against women and children remained a social issue that required special attention, especially in Gorontalo Regency, which had the highest case numbers in Gorontalo Province. This study aimed to predict the levels of violence against women and children using machi ne learning methods with the Naïve Bayes and Multiple Linear Regression algorithms. The results indicated that Naïve Bayes produced an RMSE of 0.174 for cases involving women and 0.254 for cases involving children. Meanwhile, multiple linear regression achieved a very low RMSE of 0.000 for both women's and children's cases. These results concluded that the Multiple Linear Regression algorithm was more accurate in predicting violence levels. It was hoped that this research could assist the government and related agencies in taking preventive measures and faster actions, thus reducing cases of violence against women and children in Gorontalo Regency.

Keywords*: Prediction, Violence, Women and Children, Machine Learning, Naïve Bayes, Multiple Linear Regression.*

ABSTRAK

Kekerasan terhadap perempuan dan anak merupakan isu sosial yang masih perlu mendapatkan perhatian khusus, terutama di Kabupaten Gorontalo yang memiliki angka kasus tertinggi di Provinsi Gorontalo. Penelitian ini bertujuan untuk memprediksi tingkat kekerasan terhadap perempuan dan anak menggunakan metode Machine Learning dengan algoritma Naïve Bayes dan Regresi Linier Berganda. Hasil penelitian menunjukkan bahwa Naïve Bayes menghasilkan RMSE sebesar 0.174 untuk kasus perem puan dan 0.254 untuk kasus anak. Sementara itu, Regresi Linier Berganda menghasilkan RMSE yang sangat kecil yaitu 0.000, baik untuk kasus perempuan maupun kasus anak. Dari hasil tersebut, dapat disimpulkan bahwa algoritma Regresi Linier Berganda lebih akurat dalam memprediksi tingkat kekerasan. Dengan hasil penelitian ini, diharapkan dapat membantu pemerintah dan lembaga terkait dalam mengambil langkah pencegahan serta tindakan yang lebih cepat, sehingga kasus kekerasan terhadap perempuan dan anak di Kabupaten Gorontalo dapat dicegah

Kata Kunci: Prediksi, Kekerasan, Perempuan dan Anak, Machine learning, Naïve Bayes, Regresi Linier Berganda.

1. PENDAHULUAN

Kekerasan adalah suatu tindakan tercela yang dilakukan dengan sengaja oleh seseorang atau sekelompok orang yang bisa menyebabkan kerugian, baik dalam bentuk cedera fisik maupun gangguan psikologis. Meskipun semua orang berisiko mengalami kekerasan, fakta menunjukkan bahwa perempuan dan anak-anak sering menjadi kelompok yang paling rentan menjadi korban kekerasan.

Di Indonesia, kekerasan masih menjadi isu sosial yang masih memerlukan perhatian besar. Adapun kekerasan yang sering terjadi umumnya terbagi dalam beberapa bentuk, seperti kekerasan fisik, psikis, seksual, eksploitasi dan lain sebagainya. Tingkat kekerasan dapat dipengaruhi oleh berbagai faktor seperti psikologis yang terganggu, lingkungan keluarga, lingkungan sosial, kemiskinan, serta norma-norma budaya yang patriakis menjadi penyebab utama tingginya angkat kekerasan. Bentuk pelanggaran yang satu ini tidak hanya terjadi di wilayah perkotaan, kasus-kasus serupa juga banyak ditemukan di daerah-daerah kecil, termasuk di Kabupaten Gorontalo.

Berdasarkan informasi yang tercatat dalam Sistem Informasi Online Perlindungan Perempuan dan Anak (SIMFONI PPA) pada tahun 2023, Kabupaten Gorontalo mencatat jumlah kasus kekerasan tertinggi di Provinsi Gorontalo, yaitu sebanyak 101 kasus. Angka ini lebih tinggi dibandingkan kabupaten atau kota lain di provinsi tersebut, seperti Kabupaten Bone Bolango dengan 83 kasus, Kabupaten Boalemo 69 kasus, Kabupaten Pohuwato 64 kasus, Kota Gorontalo 42 kasus. dan Kabupaten Gorontalo Utara menjadi wilayah dengan jumlah kasus paling sedikit, yakni sebanyak 28 kasus.

Sementara itu data dari Dinas Pemberdayaan Perempuan dan Anak Kabupaten Gorontalo menunjukkan bahwa pada tahun 2024 dari bulan januari hingga bulan juni sudah tercatat sebanyak 27 kasus kekerasan yang terjadi di wilayah ini. Dengan kasus kekerasan yang terlapor yaitu kasus seksual, kasus fisik, kasus psikis, dan kasus penelantaran. Angka kelerasan yang terjadi diwilayah ini menunjukkan tingkat kekerasan yang mengkhawatirkan dan memerlukan perhatian yang mendalam. Untuk mengatasi hal tersebut diperlukan langkah-langkah pencegahan yang lebih efektif. Salah satu

cara yang efektif untuk menangani masalah ini adalah dengan memprediksi terjadinya kekerasan, sehingga langkah pencegahan bisa diambil sebelum hal buruk terjadi.

Machine learning adalah cabang dari kecerdasan buatan yang didefinisikan sebagai metode yang belajar dari pengalaman dan menyesuaikan diri dengan informasi baru. Dengan cara ini, machine learning dapat mengambil keputusan secara otomatis dan tepat. Teknologi ini mendukung manusia dalam proses pengambilan keputusan dengan menganalisis data, mengenali pola, dan memberikan rekomendasi yang akurat (Prasetyo & Dewayanto, 2024).

Salah satu algoritma machine learning yang digunakan untuk memprediksi tingkat kekerasan pada penelitian ini yaitu algoritma naïve bayes. Naïve bayes merupakan teknik sederhana untuk memprediksi peluang dengan menerapkan aturan Teorema Bayes yaitu menghitung kemungkinan suatu kejadian berdasarkan data yang ada. Keunggulan menggunakan Naïve Bayes adalah kemampuannya dalam menangani data yang tidak seimbang atau kecil.

Selain itu ada algoritma Regresi Linier Berganda yang merupakan metode regresi yang melibatkan banyak variabel independen. Salah satu keunggulan utama dari pendekatan ini adalah kemampuannya untuk memprediksi kondisi masa depan dengan menganalisis hubungan antara beberapa variabel independen (X) dan variabel dependen (Y). Teknik ini sangat sesuai untuk digunakan dalam penelitian ini, karena melibatkan beberapa variabel independen dalam menganalisis tingkat kekerasan di Kabupaten Gorontalo.

Berdasarkan penjelasan latar belakang yang telah disampaikan, maka penelitian ini akan membandingan dua algoritma Machine learning yaitu Naïve Bayes dan Regresi Linier Berganda, untuk memprediksi tingkat kekerasan yang dialami oleh perempuan dan anak di Kabupaten Gorontalo. Untuk itu judul dari penelitian ini adalah : "Prediksi Tingkat Kekerasan pada Perempuan dan Anak di Kabupaten Gorontalo Menggunakan Metode Machine learning".

2. TINJAUAN PUSTAKA

Machine Learning

Machine learning pertama kali disebukan oleh Arthur Samuel pada tahun 1959. Secara istilah, machine learning atau pembelajaran mesin adalah model komputasi statistik yang fokus pada prediksi menggunakan sebuah komputer (Ramadhandi et al., 2020). Machine learning dapat diartikan sebagai bidang ilmu yang memanfaatkan komputer dan algoritma matematika untuk mempelajari pola atau informasi dari data yang tersedia. Tujuan dari proses ini adalah menghasilkan prediksi atau keputusan yang lebih tepat di masa depan berdasarkan data yang telah diproses. Pembelajaran dalam machine learning terdiri dari dua langkah utama: pelatihan (training) dan pengujian (testing)(Roihan et al., 2020).

Dalam jurnal (Sidik & Ansawarman, 2022) secara umum, machine learning dikelompokkan ke dalam tiga kategori: Supervised Learning, Unsupervised Learning, dan Reinforcement Learning.

a. Supervised Learning adalah algoritma machine learning yang proses pembelajarannya dilakukan dengan adanya pengawasan. Contoh metode dalam supervised learning meliputi klasifikasi dan regresi.

b. Unsupervised Learning adalah algoritma machine learning yang proses pembelajarannya tidak memerlukan pengawasan. Contoh metode dalam unsupervised learning termasuk pengelompokan (clustering) dan pengurangan dimensi (dimensionality reduction).

c. Reinforcement Learning adalah algoritma machine learning yang memungkinkan agen perangkat lunak untuk bekerja secara mandiri dan menentukan perilaku ideal guna memaksimalkan kinerja algoritma. Contoh penerapannya meliputi pengambilan keputusan real-time, navigasi robot, learning tasks, skills acquisition dam game AI.

Data Mining

Data mining merupakan salah satu bagian dari machine learning yang berfungsi untuk menemukan pola atau informasi penting dalam data menggunakan teknik atau metode tertentu (Raharjo & Windarto, 2021).

Data mining adalah metode yang digunakan untuk menggali informasi penting dari database. Dalam prosesnya, teknik memanfaatkan berbagai pendekatan, seperti kecerdasan buatan, pembelajaran mesin, dan statistik, untuk mengekstraksi dan mengenali data yang dibutuhkan dari database yang besar (Dodi Nofri Yoliad, 2023).

Prediksi

Secara etimologi, prediksi merupakan kegiatan yang bertujuan untuk meramalkan atau memperkirakan sesuatu. Menurut Kamus Besar Bahasa Indonesia (KBBI), prediksi diartikan sebagai hasil dari proses memperkirakan, meramal, atau memprediksi sesuatu. Sedangkan dalam konteks ilmiah, prediksi adalah suatu metode sistematis yang digunakan untuk memperoleh pengetahuan berdasarkan bukti fisik. Perapan prediksi banyak ditemukan dalam berbagai bidang, seperti peramalan penjualan, nilai tukar mata uang, dan aliran air sungai (Lasarudin & Maku, 2022).

Algortima Naïve Bayes

Naive Bayes adalah teknik prediksi yang menggunakan konsep probabilitas sederhana untuk membuat keputusan, berdasarkan pada aturan Bayes dengan asumsi bahwa variabel-variabel yang ada tidak saling bergantung. Selain itu, Naive Bayes juga dapat menentukan variabel mana yang paling berpengaruh dengan melihat peluang (Lasarudin et al., 2023). Dalam pengujian, algoritma Naive Bayes digunakan untuk memprediksi kelas target dari data yang belum diketahui dengan mempertimbangkan probabilitas yang dihitung selama proses pelatihan. Naive Bayes mengasumsikan bahwa pengaruh suatu nilai variabel terhadap kelas yang ditentukan tidak bergantung pada nilai-nilai variabel lainnya. Algoritma ini memungkinkan pembuatan model

yang cepat dengan kemampuan untuk melakukan prediksi (Cahyo, 2023).

Algoritma Regresi Linier Berganda

Regresi Linier Berganda adalah teknik prediksi yang melibatkan dua atau lebih variabel independen yang mempengaruhi satu variabel dependen. Teknik ini digunakan untuk menganalisis dan menjelaskan hubungan antara variabel-variabel tersebut, serta bagaimana perubahan pada variabel independen memengaruhi variabel dependen.(Maharadja et al., 2021).

Penelitian Terkait

Penelitian yang dilakukan oleh (Murni Pratiwi et al., 2024) bertujuan untuk membangun model prediksi dalam proses rekrutmen karyawan menggunakan algoritma Naïve Bayes. Studi ini menggunakan data rekrutmen dari PT Atma Darma Apta dengan 372 data pelamar dan 8 atribut. Evaluasi model dilakukan dengan membandingkan akurasi serta nilai Mean Squared Error (MSE), Root Mean Squared Error (RMSE), dan R^2 Score, di mana hasil terbaik diperoleh pada pembagian data 90:10 dengan akurasi 97,14%, MSE 2,86, RMSE 16,90, dan R^2 1,00.

Penelitian yang dilakukan oleh (Kurniawan et al., 2022) bertujuan untuk memprediksi jumlah penduduk di Jakarta Selatan berdasarkan data historis dari tahun 2010 hingga 2018. Dengan dua variabel independen, yaitu jumlah penduduk laki-laki dan perempuan, model ini mampu memperkirakan pertumbuhan penduduk dengan hasil prediksi kenaikan sebesar 4.195 jiwa pada tahun 2019, sehingga total populasi diperkirakan mencapai 2.298.974 jiwa. Evaluasi model menggunakan Root Mean Square Error (RMSE) menghasilkan nilai 0.43, yang menunjukkan tingkat akurasi yang baik.

3. METODE PENELITIAN

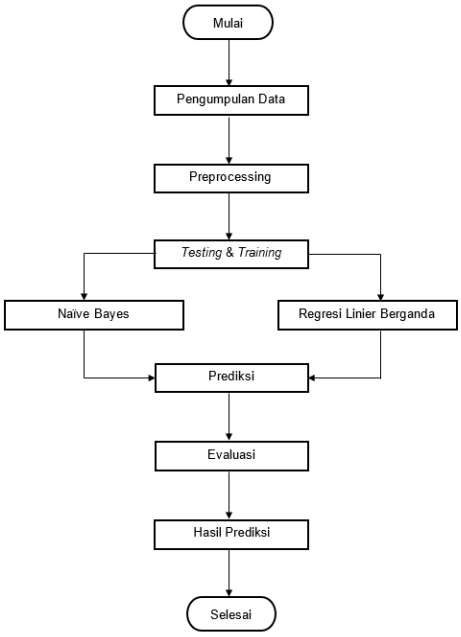
Objek Penelitian

Objek penelitian ini berfokus pada data perbulan kasus kekerasan terhadap perempuan dan anak yang tercatat dari Januari 2018 hingga Desember 2024. Dengan lokasi penelitian yaitu Kantor Dinas Pemberdayaan Perempuan dan Anak (DPP-PA) Kabupaten Gorontalo.

Pengumpulan Data

Pada tahap pengumpulan data, peneliti menggunakan metode observasi dan wawancara. Observasi dilakukan dengan mengunjungi langsung Kantor Dinas Pemberdayaan Perempuan dan Anak (DPP-PA) Kabupaten Gorontalo. Sementara itu, wawancara dilakukan dengan pihak terkait di bidang penanganan kasus kekerasan untuk mendapatkan informasi yang relevan dengan kebutuhan penelitian.

Tahapan Penelitian



4. HASIL DAN PEMBAHASAN

Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data bulanan yang diperoleh dari Dinas Pemberdayaan Perempuan dan Anak Kabupaten Gorontalo, Data tersebut mencatat kejadian kekerasan terhadap perempuan dan anak dari Januari 2018 hingga Desamber 2024.

Preprocessing

Pada tahap preprocessing, data mentah yang sudah dikumpulkan akan disiapkan menjadi data yang bisa digunakan dengan baik pada saat pengolahan data. Salah satu langkah yang dilakukan adalah mengubah kolom tahun dan bulan menjadi format tanggal yang mencakup keduanya. Selain itu, beberapa kolom atribut yang tidak relevan untuk prediksi, seperti Kasus Penelantaran Anak, Kasus Psikis Anak, Kasus Penelantaran Perempuan, dan Kasus Psikis Perempuan, dihapus karena memiliki frekuensi sangat rendah dan banyak berisi angka nol yang dapat mengganggu model prediksi. Dengan menghapus kolom-kolom tersebut, diharapkan kemampuan prediksi dapat berjalan dengan baik.

Tabel 1. Data Setelah Preprocessing

Date	Kasus Seksual Anak	Kasus Fisik Anak	Jumlah Kasus Anak	Kasus Seksual Perempuan	Kasus Fisik Perempuan	Jumlah Kasus Perempuan
2018-01	3	2	5	0	1	1
2018-02	1	1	2	0	2	2
2018-03	1	0	1	0	2	2
2018-04	1	1	2	0	0	0
2018-05	1	1	2	0	2	2
2018-06	0	1	1	0	0	0
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
2024-08	3	0	3	0	0	0
2024-09	5	0	5	0	0	0
2024-10	3	1	4	1	0	1
2024-11	5	1	6	4	0	4
2024-12	5	0	5	0	1	1

Training dan Testing

Setelah proses preprocessing, langkah berikutnya adalah membagi dataset menjadi dua bagian yaitu data training dan data testing. Dalam penelitian ini, data training yang digunakan 80% dan data testing sebanyak 20% dari 84 baris data.:

Hasil Evaluasi Menggunakan RapidMiner

a) Algoritma Naïve Bayes

Hasil evaluasi yang di dapatkan dalam pengujian 10 kali pengulangan menggunakan data kasus perempuan dan anak adalah sebagai berikut.

Tabel 2. Hasil Cross Validation Naive Bayes

X - Validation	Kasus Perempuan	Kasus Anak
2	0.412	0.337
3	0.329	0.254
4	0.183	0.407
5	0.277	0.307
6	0.174	0.360
7	0.265	0.402
8	0.265	0.378
9	0.268	0.324
10	0.212	0.311

Hasil pengujian menggunakan X-Validation menunjukkan bahwa untuk atribut kasus perempuan, nilai RMSE terbaik diperoleh pada parameter ke-6, dengan hasil RMSE sebesar 0.174. Sementara itu, untuk atribut kasus anak, hasil cross-validation terbaik ditemukan pada parameter ke-3, dengan nilai RMSE 0.254. Ini menunjukkan bahwa pada parameter ini, model memberikan prediksi yang lebih akurat dengan kesalahan yang lebih kecil.

b) Algoritma Regresi Linier Berganda

Selanjutnya hasil yang diperoleh dari proses cross-validation menggunakan algoritma Regresi Linier Berganda dapat dilihat pada tabel berikut.

Tabel 3. Hasil Cross Validation Regresi

Kasus Perempuan	Kasus Anak
0.000	0.000

Tabel di atas menunjukkan bahwa hasil evaluasi dari kasus perempuan dan anak mendapatkan hasil RMSE yang sangat kecil yaitu 0.000.

Hasil Evaluasi Menggunakan Colab

a) Algortima Naïve Bayes

Saat dilakukan pengujian menggunakan algoritma Naïve Bayes di Google Colab, diperoleh nilai RMSE sebagai berikut:

Tabel 4. Hasil Evaluasi Naive Bayes di Colab

Kasus Perempuan	Kasus Anak
0.766	0.342

Berdasarkan tabel 4, hasil evaluasi model Naïve Bayes di Google Colab menunjukkan bahwa model memiliki tingkat akurasi yang berbeda untuk dua jenis kasus. Untuk kasus perempuan, model mencapai skor evaluasi sebesar 0.766, sedangkan untuk kasus anak, skornya lebih rendah, yaitu 0.342.

b) Algoritma Regresi Linier Berganda

Hasil evaluasi algoritma Regresi Linier Berganda setelah dilakukan pengujian menggunakan Colab disajikan pada tabel berikut.

Tabel 5. Hasil Evaluasi Regresi di Colab

Kasus Perempuan	Kasus Anak
0.000	0.000

Berdasarkan tabel hasil evaluasi model Regresi di Google Colab menunjukkan nilai RMSE sebesar 0.000 untuk kasus perempuan dan kasus anak. Hasil ini mengindikasikan bahwa model memiliki tingkat kesalahan yang sangat rendah.

Hasil Prediksi

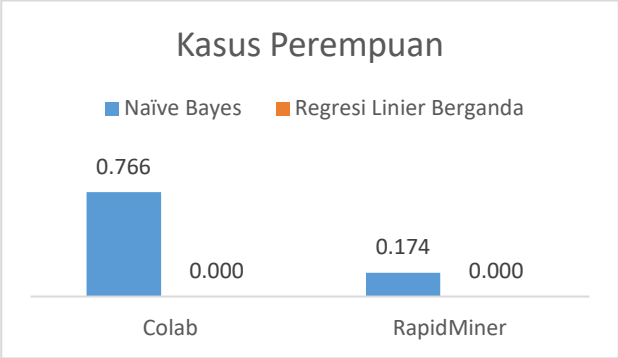
Pada penelitian ini, hasil prediksi diperoleh dengan menggunakan Microsoft Excel dan RapidMiner sebagai tools utama dalam pengolahan data. Selain itu, Google Colab digunakan sebagai tools tambahan untuk memeriksa apakah hasil prediksi yang diperoleh dari RapidMiner tetap konsisten. Berikut ini disajikan tabel yang menunjukkan hasil prediksi kasus perempuan yang diperoleh dari tools tersebut.

Tabel 6. Hasil Prediksi Kasus Perempuan

	Colab	RapidMiner
Naïve Bayes	0.766	0.174
Regresi Linier Berganda	0.000	0.000

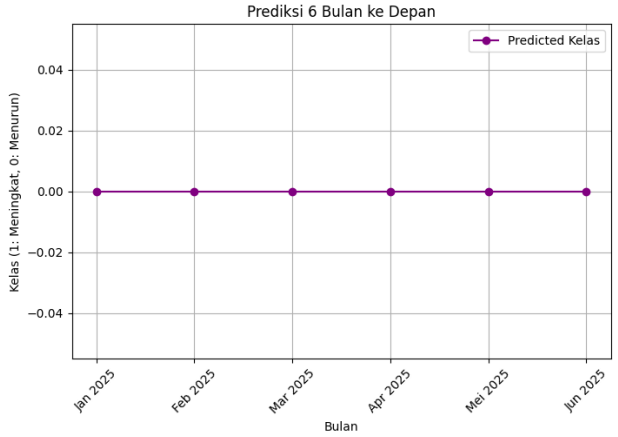
Tabel menunjukkan bahwa model Naive Bayes di Google Colab menghasilkan nilai RMSE 0.766, sedangkan di RapidMiner, RMSE-nya lebih rendah, yaitu 0.174. Ini berarti model Naive Bayes di RapidMiner lebih akurat dibandingkan dengan yang dihasilkan di Google Colab. Sementara itu, model Regresi Linier menunjukkan nilai RMSE yang sangat rendah di kedua tools. Di Google Colab, RMSE untuk regresi linier adalah 0.000, begitu juga di RapidMiner, yang menghasilkan nilai RMSE 0.000. Hal ini menunjukkan bahwa model regresi linier sangat akurat

di kedua tools. Berikut adalah diagram mengenai gambaran perbandingan kedua model prediksi.



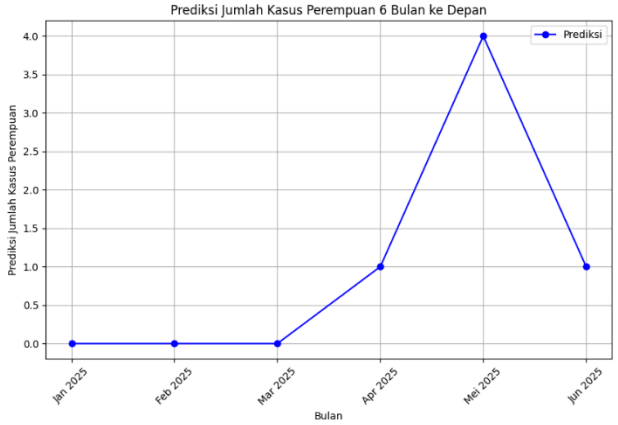
Gambar 1. Diagram Hasil Prediksi Kasus Perempuan

Untuk mendapatkan gambaran lebih jelas mengenai bagaimana perubahan tingkat kekerasan terhadap perempuan dalam enam bulan ke depan, hasil prediksi divisualisasikan dalam bentuk grafik berikut ini.



Gambar 2. Grafik Hasil Prediksi Ke Depan Kasus Perempuan Menggunakan Naive Bayes

Grafik untuk kasus perempuan di atas menunjukkan bahwa selama enam bulan ke depan, kelas prediksi tetap berada pada nilai 0, yang berarti model memperkirakan bahwa kasus kekerasan terhadap perempuan tidak mengalami peningkatan.



Gambar 3. Grafik Hasil Prediksi Ke Depan Kasus Perempuan Menggunakan Regresi Linier

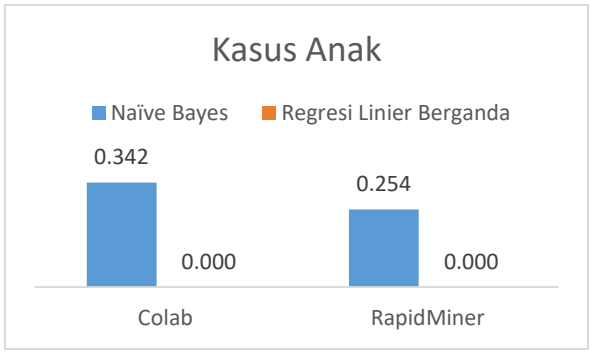
Kemudian hasil prediksi menggunakan algoritma Regresi Linier Berganda menunjukkan bahwa jumlah kasus kekerasan terhadap perempuan diperkirakan tetap nol dari Januari hingga Maret 2025. Namun, mulai April 2025, kasus mulai meningkat, mencapai puncaknya pada Mei 2025 dengan empat kasus, sebelum kembali menurun pada Juni 2025.

Selanjutnya adalah tabel hasil prediksi kasus anak, yang membandingkan kinerja model Naive Bayes dan Regresi Linear di Google Colab dan RapidMiner.

Tabel 7. Hasil Prediksi Kasus Anak

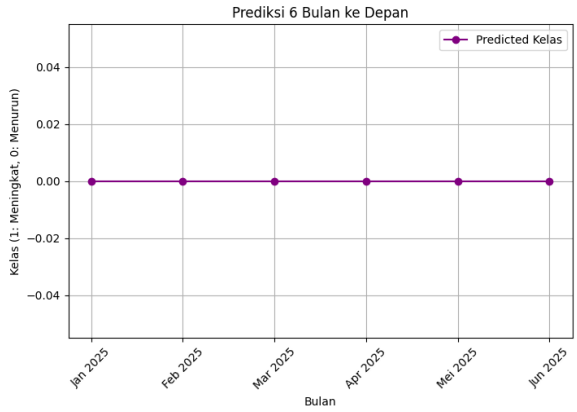
	Colab	RapidMiner
Naïve Bayes	0.342	0.254
Regresi Linier Berganda	0.000	0.000

Tabel di atas menunjukkan hasil prediksi kasus anak menggunakan model Naive Bayes dan Regresi Linier Berganda di Google Colab dan RapidMiner. Naive Bayes di Google Colab memiliki nilai RMSE 0.342, sedangkan di RapidMiner lebih rendah, yaitu 0.254, yang menunjukkan prediksi lebih akurat di RapidMiner. Untuk Regresi Linier Berganda, baik di Google Colab maupun RapidMiner, nilai RMSE yang dihasilkan adalah 0.000, hal ini menandakan hasil yang sangat akurat di kedua tools.



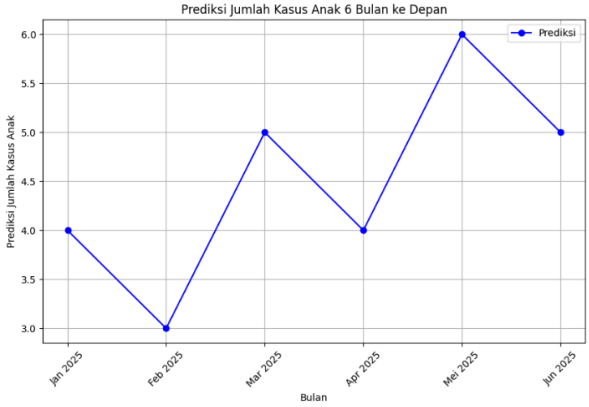
Gambar 4. Diagram Hasil Prediksi Kasus Anak

Diagram di atas merupakan gambaran hasil prediksi kasus anak menggunakan algoritma *Naive Bayes* dan Regresi Linear pada dua tools. *Naive Bayes* menunjukkan akurasi yang lebih baik saat dijalankan di *RapidMiner* dibandingkan di Colab. Sementara itu, Regresi Linear memberikan hasil prediksi yang sempurna di kedua tools tanpa adanya kesalahan. Hal ini menunjukkan bahwa regresi linear lebih unggul dalam memprediksi kasus kekerasan, sementara *Naive Bayes* tetap memberikan performa yang cukup baik, terutama di Rapid. Untuk melihat lebih jelas bagaimana perubahan tingkat kekerasan terhadap anak dalam enam bulan ke depan, hasil prediksi ditampilkan dalam grafik berikut ini.



Gambar 5. Grafik Hasil Prediksi Ke Depan Kasus Anak Menggunakan Naive Bayes

Grafik untuk kasus anak di atas menunjukkan bahwa selama enam bulan ke depan, kelas prediksi tetap berada pada nilai 0, yang berarti model memperkirakan bahwa tingkat kekerasan terhadap anak tidak mengalami peningkatan.



Gambar 6. Grafik Hasil Prediksi Ke Depan Kasus Anak Menggunakan Regresi Linier

Sementara itu grafik di atas, menunjukkan jumlah kasus anak dalam enam bulan ke depan yang diprediksi mengalami naik turun. Pada Januari 2025, kasus diperkirakan sebanyak 4 kasus, lalu menurun menjadi 3 kasus pada Februari, kemudian jumlah kasus kembali meningkat menjadi 5 kasus di Maret, namun kembali turun ke 4 kasus pada April.

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan, kesimpulan dapat dirumuskan sebagai berikut:

1. Dari hasil prediksi metode *Machine Learning* menggunakan algoritma *Naïve Bayes* dan Regresi Linier Berganda untuk memprediksi tingkat kekerasan di Kabupaten Gorontalo, ditemukan bahwa algoritma Regresi Linier Berganda menghasilkan performa prediksi yang lebih baik dibandingkan dengan *Naïve Bayes*. Hal ini tercermin dari nilai *Root Mean Square Error* (RMSE) yang sangat kecil, yang telah diuji menggunakan tools *RapidMiner* dan Google Colab.
2. Berdasarkan hasil penerapan menggunakan metode *Machine Learning*, Regresi Linier Berganda menunjukkan nilai *Root Mean Square Error* (RMSE) yang sangat kecil, yaitu 0.000, baik untuk kasus perempuan maupun kasus anak. Sementara itu, *Naïve Bayes* menghasilkan nilai RMSE sebesar 0.174 untuk kasus perempuan dan 0.254 untuk kasus anak. Hal ini menunjukkan performa prediksi yang dihasilkan *Naïve Bayes* cukup baik namun masih kurang akurat dibandingkan dengan Regresi Linier Berganda. Dengan demikian, Regresi Linier Berganda lebih efektif dan dapat diandalkan untuk memprediksi tingkat kekerasan pada perempuan dan anak di Kabupaten Gorontalo.

Dari hasil penelitian, saran yang dapat diberikan untuk penelitian berikutnya adalah disarankan untuk memperluas cakupan data, baik dengan memperbesar wilayah penelitian maupun menggunakan dataset yang berbeda. Hal ini bertujuan untuk mengevaluasi konsistensi dan meningkatkan akurasi prediksi dengan variasi data yang lebih beragam, sehingga model yang dihasilkan dapat memberikan hasil yang lebih akurat dan dapat diandalkan.

DAFTAR PUSTAKA

[1] Alfarizi, A. D., & Andri, A. (2021). Pemanfaatan Data Mining Dalam Memprediksi Produksi Pada PT Pupuk Sriwidjaja Palembang Menggunakan Algoritma Regresi Liniear Berganda. Jurnal Nasional Ilmu Komputer, 2(1), 51–63. https://doi.org/10.47747/jurnalnik.v2i1.522

[2] Asohi, Y., & Andri, A. (2020). Impelementasi Algoritma Regresi Linier Berganda Untuk Prediksi Penjualan. Jurnal Nasional Ilmu Komputer, 1(3), 149–158. https://doi.org/10.47747/jurnalnik.v1i3.161

[3] Baharuddin, F., & Tjahyanto, A. (2022). Peningkatan Performa Klasifikasi Machine learning Melalui Perbandingan Metode Machine learning dan Peningkatan Dataset. Jurnal Sisfokom (Sistem Informasi Dan Komputer), 11(1), 25–31. https://doi.org/10.32736/sisfokom.v11i1.1337

[4] Cahyo, A. D. (2023). Metode Naïve Bayes Untuk Klasifikasi Masa Studi Sarjana | Junal Teknologi Pintar. Junal Teknologi Pintar, 3(4). http://teknologipintar.org/index.php/teknologipintar/article/view/385

[5] Dodi Nofri Yoliad. (2023). Data mining Dalam Analisis Tingkat Penjualan Barang Elektronik Menggunakan Algoritma K-means. Insearch (Information System Research) Journal, 3(1).

[6] Fricco, A., & Setiawan, K. (2023). Implementasi Algoritma Naive Bayes Menggunakan Particle Swarm Optimization Untuk Prediksi Penjualan Produk Air Mineral. Jurnal Sains Dan Teknologi, 5(1), 193–197. https://doi.org/10.55338/saintek.v5i1.1430

[7] Hasmaulina, B. (2022). Penerapan Data Mining Untuk Membentuk Kelompok Belajar Menggunakan Metode Clustering Di SMK Negeri 3 Seluma. JUKOMIKA (Jurnal Ilmu Komputer Dan Informatika), 4(2), 57–71. https://doi.org/10.54650/jukomika.v4i2.368

[8] Kurniawan, P., Rossa, H., Permana, A., Ramadan, W. A., Aji, A. B. W., Hidayatulloh, S., Iksan, N., & Arief, U. M. (2022). Prediksi Jumlah Penduduk Jakarta Selatan Menggunakan Metode Regresi Linear Berganda. Jurnal Sistem Dan Teknologi Informasi (JustIN), 10(4), 518. https://doi.org/10.26418/justin.v10i4.48331

[9] Lasarudin, A., Hilmansyah Gani, & Misran Tomayahu.

(2023). Perbandingan Metode Naïve Bayes dan C4.5 Klasifikasi Status Gizi Bayi Balita. SPECTA Journal of Technology, 6(3), 273–283. https://doi.org/10.35718/specta.v6i3.789

[10] Lasarudin, A., & Maku, R. (2022). Prediksi Pertumbuhan Jumlah Penduduk Menggunakan Algoritma Neural Network. Jurnal Ilmu Komputer (JUIK), 2(2), 37. https://doi.org/10.31314/juik.v2i2.1715

[11] Lois, J., Kurnia, Y., Lasut, D., & Fenriana, I. (2022). Aplikasi Pengolahan Data Mining Berbasis Web Menggunakan Algoritma Apriori Untuk Menganalisis Data Penjualan Toko Lumbini Mart. Jurnal Algor, 3(2), 12–24.

[12] Maharadja, A. N., Maulana, I., & Dermawan, B. A. (2021). Penerapan Metode Regresi Linear Berganda untuk Prediksi Kerugian Negara Berdasarkan Kasus Tindak Pidana Korupsi. Journal of Applied Informatics and Computing, 5(1), 95–102. https://doi.org/10.30871/jaic.v5i1.3184

[13] Maku, R. (2022). Neural Network Algorithm for Budget Expenditure Prediction in LPP RRI Gorontalo. Journal La Multiapp, 3(2), 61–67. https://doi.org/10.37899/journallamultiapp.v3i2.596

[14] Muhadi, A., & Octaviano, A. (2023). Penerapan Data Mining Untuk Prediksi Hasil Keuntungan Lelang Mesin X-Ray Tahun 2020 Dengan Metode K-Nearest Neighbor (Studi Kasus : PT.Ramadika Mandiri). Jurnal Informatika MULTI, 1(2), 2985–8860. https://jurnal.publikasitecno.id/index.php/jim126

[15] Muharromah, O., Suarna, N., & Prihartono, W. (2024). Implementasi Algoritma Regresi Linear Berganda Untuk Prediksi Produksi Padi Di Kabupaten Cirebon. JATI (Jurnal Mahasiswa Teknik Informatika), 7(6), 3815–3820. https://doi.org/10.36040/jati.v7i6.8206

[16] Murni Pratiwi, I., Fauzi, A., Arum Puspita Lestari, S., Cahyana, Y., Ilmu Komputer, F., & Buana Perjuangan Karawang, U. (2024). Penerapan Algoritma Naïve Bayes Untuk Prediksi Penerimaan Karyawan. Jurnal TEKINKOM, 7(1), 236–243. https://doi.org/10.37600/tekinkom.v7i1.1282

[17] Pambudi, H. K., Kusuma, P. G. A., Yulianti, F., & Julian, K. A. (2020). Prediksi Status Pengiriman Barang Menggunakan Metode Machine learning. Jurnal Ilmiah Teknologi Infomasi Terapan, 6(2), 100–109. https://doi.org/10.33197/jitter.vol6.iss2.2020.396

[18] Prasetyo, S., & Dewayanto, T. (2024). Penerapan Machine learning, Deep Learning, Dan Data Mining Dalam Deteksi Kecurangan Laporan Keuangan-a Systematic Literature Review. Diponegoro Journal of Accounting, 13(3), 1–12. http://ejournal-s1.undip.ac.id/index.php/accounting

[19] Raharjo, M. R., & Windarto, A. P. (2021). Penerapan Machine learning dengan Konsep Data Mining Rough Set (Prediksi Tingkat Pemahaman Mahasiswa terhadap Matakuliah). Jurnal Media Informatika Budidarma, 5(1), 317. https://doi.org/10.30865/mib.v5i1.2745

[20] Ramadhandi, R., Santoso, Rani Megasari, Hambali, & Yudi Ahmad. (2020). Implementasi Metode Machine learning Menggunakan Algoritma Evolving Artificial Neural Network Pada Kasus Prediksi Diagnosis Diabetes Implementation of Machine learning Method Using Evolving Artificial Neural Network Algorithm in Prediction of Diabetes Dia. Jatikom, 3(2), 9–20. https://ejournal.upi.edu/index.php/JATIKOM

[21] Roihan, A., Sunarya, P. A., & Rafika, A. S. (2020). Pemanfaatan Machine learning dalam Berbagai Bidang: Review paper. IJCIT (Indonesian Journal on Computer and Information Technology), 5(1), 75–82. https://doi.org/10.31294/ijcit.v5i1.7951

[22] Sidik, A. D., & Ansawarman, A. (2022). Prediksi Jumlah Kendaraan Bermotor Menggunakan Machine learning. Formosa Journal of Multidisciplinary Research, 1(3), 559–568. https://doi.org/10.55927/fjmr.v1i3.745

[23] Sitompul, A. M., Suhada, & Saifulah. (2021). Teknik Data Mining Dalam Prediksi Jumlah Siswa Baru Dengan Algoritma Naive Bayes. KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen), 2(2), 108–117.

[24] Zahra, F., Ridla, M. A., & Azise, N. (2024). Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang (Studi Kasus : Toko Sinar Harahap). JUSTIFY : Jurnal Sistem Informasi Ibrahimy, 3(1), 55–65. https://doi.org/10.35316/justify.v3i1.5335