

PREDIKSI PRESTASI SISWA SMAS KATOLIK SANTO YOSEPH DENPASAR BERDASARKAN KEDISIPLINAN DAN TINGKAT EKONOMI ORANG TUA MENGGUNAKAN METODE *KNOWLEDGE DISCOVERY IN DATABASE* DAN ALGORITMA REGRESI LINIER BERGANDA

Maria Atalya Angelus Leza, Nengah Widya Utami, Putri Anugrah Cahya Dewi

Sistem Informasi, Primakara University
Jalan Tukad Badung No. 135, Renon, Denpasar, Bali
mariaatalya1@gmail.com

ABSTRAK

Pendidikan merupakan suatu sistem yang didalamnya mengandung unsur penting yang saling berkaitan dan mempunyai tujuan untuk menghasilkan taraf SDM yang kompeten demi kemajuan bangsa. SMAS Katolik Santo Yoseph Denpasar merupakan salah satu lembaga pendidikan yang sangat memperhatikan pendidikan siswanya, tetapi mempunyai permasalahan dimana pihak sekolah kesulitan dalam mengidentifikasi kemampuan dan kebutuhan siswa sehingga siswa tidak mendapatkan perhatian dan bantuan yang tepat. Maka dari itu, perlu dilakukan prediksi prestasi siswa agar lembaga pendidikan dapat memberikan bimbingan dan pengajaran yang tepat dengan kualitas yang baik terhadap siswa sesuai dengan kemampuannya. Penelitian ini menggunakan metode KDD (*Knowledge Discovery in Database*) dan Algoritma Regresi Linier Berganda yang bertujuan untuk mengetahui hasil prediksi prestasi siswa SMAS Katolik Santo Yoseph Denpasar berdasarkan kedisiplinan dan tingkat ekonomi orang tua dan mengukur tingkat akurasi prediksi menggunakan MAE (*Mean Absolute Error*), MSE (*Mean Squared Error*) dan RMSE (*Root Mean Square Error*). KDD adalah teknik untuk mendapatkan informasi berupa data dari basis data yang sudah ada dengan beberapa tahapan, yaitu *data selection*, *pre-processing/cleaning*, *transformation*, *data mining* dan *interpretation/evaluation*. Algoritma Regresi Linier Berganda adalah teknik *data mining* yang digunakan untuk mengetahui apakah ada hubungan antara variabel yang ingin diprediksi (variabel dependen) dengan variabel lain (variabel independen) yang memiliki persamaan $Y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n$. Penelitian ini menghasilkan hubungan positif antara kedisiplinan siswa, tingkat ekonomi orang tua dan prestasi siswa dengan dataset sebanyak 2672 data dengan perbandingan 70% *data training* dan 30% *data testing*. Perbandingan 70% *data training* dan 30% *data testing* merupakan tingkat akurasi tertinggi dengan nilai MAE sebesar 0,026, nilai MSE sebesar 0,001 dan nilai RMSE sebesar 0,032.

Kata kunci: *Data Mining, Kedisiplinan, Ekonomi, Regresi, Prediksi*

1. PENDAHULUAN

Pendidikan merupakan suatu sistem yang didalamnya mengandung unsur penting yang saling berkaitan dengan erat yang mempunyai tujuan untuk menghasilkan taraf SDM yang kompeten demi kemajuan bangsa [1]. Pendidikan identik dengan proses belajar mengajar sehingga kualitas dari suatu lembaga pendidikan harus sangat diperhatikan. Ada beberapa indikator kualitas lembaga pendidikan, yaitu prestasi akademik siswa, kurikulum, ketersediaan fasilitas dan sumber daya serta kualitas tenaga pengajar. Kualitas dan pengelolaan lembaga pendidikan dapat dipengaruhi oleh prestasi siswa karena siswa yang berhasil cenderung memiliki lingkungan belajar yang lebih baik yang disediakan oleh sekolah mereka [2].

Menurut Novauli, prestasi belajar ialah hasil belajar siswa yang merupakan salah satu gambaran atau cerminan keberhasilan kemampuan siswa [3]. Setiap siswa mempunyai tingkat kemampuan dan keterampilan yang berbeda-beda sehingga mereka membutuhkan pengajaran yang berbeda. Oleh karena itu, tingkat kemampuan siswa harus diketahui terlebih dahulu agar lembaga pendidikan dapat memberikan pengajaran yang sesuai dengan keterampilan dan

kemampuan siswa tersebut [4]. Namun, dari hasil wawancara dengan salah satu guru yang ada di SMAS Katolik Santo Yoseph Denpasar, terdapat permasalahan dimana pihak sekolah kesulitan dalam mengidentifikasi kemampuan dan kebutuhan siswa sehingga siswa tidak mendapatkan perhatian dan bantuan yang tepat. Maka dari itu, perlu dilakukan prediksi prestasi siswa agar lembaga pendidikan dapat memberikan bimbingan dan pengajaran yang tepat dengan kualitas yang baik terhadap siswa sesuai dengan kemampuannya.

Keberhasilan akademik seorang siswa merupakan hasil interaksi berbagai faktor yang mempengaruhi. Menurut Slameto, (2003:5-7) yang dikutip dari Jurnal Fakultas Ekonomi UNNES, prestasi belajar dipengaruhi oleh dua faktor, yaitu faktor internal yang berasal dari mahasiswa itu sendiri dan juga faktor eksternal yang berasal dari luar kepribadian mahasiswa. Faktor internal terdiri dari faktor jasmani, psikologis, dan kelelahan. Sedangkan faktor eksternal terdiri dari lingkungan keluarga, lingkungan sekolah, dan lingkungan masyarakat [5].

Kedisiplinan merupakan salah satu faktor yang masuk kedalam faktor psikologis. Oemar Hamalik (2013:115) berpendapat bahwa "Salah satu faktor

psikologis adalah masalah disiplin." [6]. Disiplin sangat penting untuk pengajaran dan pembelajaran. Proses pendidikan tidak dapat mewujudkan potensinya secara penuh tanpa kesadaran akan kepatuhan terhadap aturan yang telah ditetapkan. Siswa membutuhkan sikap disiplin untuk memperkuat diri melalui latihan agar terbiasa mematuhi dan mengembangkan pengendalian diri. Disiplin melalui pengenalan diri lebih tahan lama dan terasa senang daripada disiplin melalui kontrol orang lain [2]. Slameto (2010:67) mengemukakan bahwa "Siswa harus menerapkan sikap disiplin dalam belajar agar lebih maju baik di sekolah, perpustakaan dan di rumah." Dari sudut pandang ini dapat dijelaskan bahwa kedisiplinan dapat memberikan kesempatan kepada siswa untuk belajar dengan cara yang lebih tinggi dan kemajuan yang dicapai akan meningkatkan prestasi siswa [7].

Selain faktor internal yaitu kedisiplinan, ada juga faktor eksternal yaitu lingkungan keluarga yang dapat mempengaruhi prestasi siswa. Menurut Slameto (2003:61) dalam Jurnal *Economica Education* oleh Husnan Jamil mengatakan bahwa, keluarga merupakan institusi pertama dan terpenting. Keluarga besar dan sehat berarti keluarga merupakan lembaga pendidikan yang kecil, tetapi ditentukan untuk pendidikan yang besar, yaitu pendidikan bagi bangsa, negara dan dunia [8]. Tingkat ekonomi termasuk kedalam faktor lingkungan keluarga. Kemampuan ekonomi keluarga yang berbeda akan mempengaruhi kegiatan belajar siswa. Fasilitas belajar yang disediakan orang tua akan sangat bergantung pada kemampuan keuangan mereka, untuk memastikan bahwa siswa memiliki akses ke infrastruktur dan fasilitas yang mereka butuhkan untuk meningkatkan prestasi akademik mereka.

Menurut Gerungan (2004:196), kondisi sosial ekonomi keluarga tidak diragukan lagi dapat berdampak pada perkembangan anak. Jika diperhatikan, dalam situasi kecukupan ekonomi, anak-anak dihadapkan pada lingkungan fisik yang lebih luas di rumah, dan mereka memiliki lebih banyak kesempatan untuk mengembangkan keterampilan yang tidak dapat mereka kembangkan tanpa infrastruktur yang memadai [9]. Situasi keuangan keluarga berdampak langsung pada kegiatan pendidikan yang dilakukan anak. Keadaan keuangan orang tua yang tidak memadai dapat menghambat kemampuan anak-anak mereka untuk belajar karena banyaknya kebutuhan pendidikan seperti makanan, kesehatan, pakaian dan fasilitas pendidikan yaitu ruang kelas, meja, lampu, buku serta kursi dan masih banyak lagi yang tidak terpenuhi. Keadaan ini akan selalu menurunkan semangat anak untuk belajar, yang akan mempengaruhi prestasi belajarnya [9].

Tingkat ekonomi pada SMAS Katolik Santo Yoseph Denpasar terbelang variatif. Dari data yang didapat, rata-rata rentang pendapatan orang tua berkisar 5 sampai 20 juta. Namun beberapa orang tua memiliki pendapatan dibawah 5 juta dan lebih dari 20 juta. Terjadi jarak yang cukup jauh dalam hal ini. Atas

hal tersebut, penelitian lebih lanjut diperlukan untuk mengetahui prediksi prestasi siswa berdasarkan kedisiplinan dan tingkat ekonomi orang tua. Salah satu teknik yang digunakan untuk mengetahui prediksi prestasi siswa adalah data mining.

Data mining adalah proses penambahan nilai tambah dari sekumpulan data berupa knowledge atau pengetahuan yang sebelumnya tidak diketahui secara manual [10]. *Data mining* adalah suatu proses penggalian informasi yang terpendam atau tersembunyi untuk memperoleh pengetahuan dan menemukan pola pada tumpukan data dalam database yang biasanya berskala besar [4]. *Data mining* sendiri memiliki beberapa metode salah satunya menggunakan metode *forecasting* atau peramalan. *Forecasting* adalah peramalan atau prediksi dari satu atau lebih peristiwa yang terjadi di masa depan untuk mencari hubungan dan pola di dalam system [11]. *Forecasting* yang akurat adalah informasi yang dibutuhkan untuk keputusan manajemen. Adanya *forecasting* ini bertujuan untuk menentukan keputusan untuk masa yang akan datang. Salah satu algoritma dalam *forecasting* adalah regresi linier berganda.

Algoritma Regresi Linier Berganda adalah metode prediksi yang menggunakan lebih dari dua faktor yang dapat mempengaruhi hasil yang ingin ditemukan [12]. Regresi linier berganda menggunakan kumpulan data interval serta terdapat lebih dari satu prediktor. Kumpulan data yang disebut di atas berlaku untuk semua jenis variabel, terutama variabel dependen [13]. Singkatnya, analisis regresi linier adalah teknik data mining yang digunakan untuk mengetahui apakah ada hubungan antara variabel yang ingin diprediksi (variabel dependen) dengan variabel lain (variabel independen) [14].

2. TINJAUAN PUSTAKA

Prestasi merupakan hasil dari usaha yang telah dilakukan pada suatu kegiatan yang telah diselesaikan atau dijalankan. Sementara itu, belajar adalah proses perubahan perilaku seseorang yang terjadi melalui pengalaman dan latihan akibat interaksi dengan lingkungan sekitarnya [15]. Prestasi belajar adalah hasil belajar yang diperoleh atau dicapai oleh siswa melalui pengalaman dan latihan [16].

Slameto (2003:5-7) menyatakan bahwa ada dua unsur yang mempengaruhi prestasi belajar yaitu faktor internal seperti faktor jasmani, psikologis dan kelelahan serta faktor eksternal yang terdiri dari lingkungan keluarga, lingkungan sekolah dan lingkungan masyarakat [5].

Faktor internal adalah kondisi dan kemampuan dari dalam diri siswa dalam memahami pembelajaran. Faktor internal tersebut terdiri sebagai berikut: Faktor Jasmani, Faktor Psikologis dan Faktor Kelelahan. Kedisiplinan masuk ke dalam faktor psikologis karena merupakan suatu perilaku pengendalian diri yang dimana melibatkan kemampuan seseorang untuk mengendalikan pikiran, emosi, dan perilaku mereka sendiri. Hal ini sejalan dengan Oemar Hamalik

(2013:115) yang mengatakan bahwa "Faktor psikologis dalam belajar salah satu diantaranya adalah masalah disiplin" [6].

Disiplin sangat penting untuk mencapai tujuan pendidikan. Selain pengaruh lingkungan keluarga, sekolah, dan masyarakat, faktor penting yang mempengaruhi seberapa baik prestasi belajar siswa adalah disiplin. Kondisi disiplin dipengaruhi dan dibentuk oleh berbagai tindakan yang mewakili nilai tertentu antara lain kesetiaan, kepatuhan, keteraturan dan ketaatan. Ketika seseorang telah mengembangkan rasa disiplin, sikap atau tindakannya tidak lagi dianggap sebagai beban. Bahkan jika seseorang tidak mengikuti aturan yang ada, dia menganggap aktivitas itu sebagai beban [17].

Faktor eksternal merupakan aspek lingkungan luar siswa yang mempengaruhi hasil belajar siswa. Faktor eksternal tersebut terdiri dari: Faktor Keluarga, Faktor Sekolah dan Faktor Masyarakat. Anak berkembang di lingkungan pertamanya, yaitu keluarga mereka. Perkembangan dan pertumbuhan anak sangat dipengaruhi oleh keluarga. Faktor keluarga merupakan faktor yang lebih berpengaruh terhadap tumbuh kembang anak. Perhatian orang tua terhadap anak dan kondisi ekonomi keluarga semuanya berdampak pada prestasi akademik siswa. Tetapi, kondisi ekonomi orang tua yang lebih berpengaruh terhadap prestasi siswa dimana hal ini didapat dari penelitian Sefti Wiri Febriana [18].

Data mining adalah proses penambahan nilai tambah dari sekumpulan data berupa pengetahuan yang sebelumnya tidak diketahui secara manual [10]. *Data mining* adalah suatu proses penggalian informasi yang terpendam atau tersembunyi untuk memperoleh pengetahuan dan menemukan pola pada tumpukan data dalam database yang biasanya berskala besar [4]. Terdapat berbagai macam metode, teknik atau algoritma dalam data mining yaitu *forecasting*, *description*, *classification*, *association*, *estimation* dan *clustering* [19].

Metode forecasting pada *data mining* mengacu pada teknik atau proses yang digunakan untuk memprediksi nilai atau perilaku masa depan dari suatu data atau variabel. Tujuan utama dari metode *forecasting* adalah untuk memprediksi apa yang akan terjadi di masa depan berdasarkan pola atau kecenderungan yang terdapat pada data historis dimana hasilnya berupa persamaan.

Forecasting adalah peramalan atau prediksi satu atau lebih kejadian di masa depan, atau meneliti data historis untuk menemukan hubungan, kecenderungan, dan pola dalam suatu sistem. Definisi lain dari peramalan adalah proses memperkirakan permintaan di masa mendatang, termasuk memperkirakan jumlah, kualitas, waktu, dan lokasi barang atau jasa yang akan dibutuhkan untuk memenuhi permintaan [11].

Ada banyak algoritma yang biasa dipakai dalam metode *forecasting*, diantaranya analisis regresi, *neural network*, model *time series* dan lain-lain. Regresi Linier Berganda adalah metode peramalan

yang menggunakan lebih dari dua faktor yang dapat mempengaruhi hasil sehingga dapat dicari hasil yang maksimal [12]. Regresi linier berganda menggunakan kumpulan data interval serta terdapat lebih dari satu prediktor. Kumpulan data yang disebut di atas berlaku untuk semua jenis variabel, terutama variabel dependen [13]. Singkatnya, analisis regresi linier adalah teknik data mining yang digunakan untuk mengetahui apakah ada hubungan antara variabel yang ingin diprediksi (variabel dependen) dengan variabel lain (variabel independen) [14].

Persamaan regresi mengungkapkan hubungan linier antara variabel dependen dan variabel independen. Variabel terikat diberi lambang Y dan variabel bebas diberi lambang X . Jika terdapat lebih dari satu variabel maka penulisannya adalah X_1, X_2 sampai dengan X_n . Secara umum, rumus persamaan regresi linier berganda yaitu $Y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n$.

Orange data mining adalah paket perangkat lunak pemrograman visual berbasis komponen yang membantu pemrosesan data visual pada *machine learning*, *data mining* dan analisis data. Pemrograman visual diimplementasikan menggunakan antarmuka yang menghubungkan widget yang telah ditentukan sebelumnya atau yang dirancang pengguna untuk membentuk alur kerja. *Orange Data Mining* relatif mudah digunakan karena didasarkan pada GUI (*Graphical user interface*) yang membuatnya dapat diakses bahkan oleh pengguna awam sekalipun. Adanya *Python*, bahasa skrip dan lingkungan pemrograman generasi baru, memungkinkan skrip *data mining* yang sederhana namun berkualitas tinggi [20].

3. METODE PENELITIAN

Metode yang digunakan di dalam penelitian ini adalah KDD (*Knowledge Discovery in Database*). KDD adalah metode yang digunakan dalam penelitian ini. KDD adalah teknik untuk mendapatkan informasi berupa data dari basis data yang sudah ada. Ada tabel yang terhubung satu sama lain dalam basis data. Informasi dari data yang dikumpulkan melalui proses KDD dapat berfungsi sebagai database untuk pengambilan keputusan [19].

Knowledge Discovery in Database (KDD) dan *data mining* adalah kata-kata yang sering digunakan secara bergantian untuk menggambarkan proses pengambilan informasi tersembunyi dari basis data yang besar. Istilah KDD dan *data mining* memiliki keterkaitan satu sama lain, tetapi memiliki konsep yang berbeda. *Data mining* adalah salah satu fase atau tahapan dalam proses KDD [21].

Secara umum, proses KDD dapat digambarkan sebagai berikut:

3.1. Data Selection

Langkah pertama sebelum penggalian informasi di *Knowledge Discovery in Database* (KDD) yaitu, data harus diseleksi atau dipilih dari data operasional.

Data yang dipilih untuk proses data mining disimpan terpisah dari basis database operasional dalam sebuah file [19].

3.2. Pre-processing/Cleaning

Sebelum memulai proses data mining, penting untuk membersihkan data, yang merupakan inti dari *Knowledge Discovery in Database* (KDD). Pembersihan termasuk menghapus duplikat data, memeriksa data yang tidak konsisten, dan memperbaiki masalah data seperti salah cetak. Dilakukan juga proses “memperkaya” data terkini dengan tambahan data atau informasi yang berguna dan diperlukan untuk *Knowledge Discovery in Database* (KDD), seperti data atau informasi eksternal lainnya [21].

3.3. Transformation

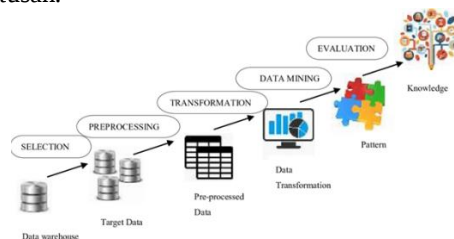
Transformation merupakan proses mengubah data yang dipilih untuk membuat data tersebut sesuai dalam data mining. Prosesnya sangat bergantung pada database, model informasi atau jenis yang akan dicari. Definisi lain dari transformasi adalah pencarian fitur yang berguna untuk merepresentasikan data tergantung pada tujuan yang ingin dicapai [19].

3.4. Data Mining

Proses menemukan pola atau informasi yang menarik pada data terpilih dengan menggunakan alat atau metode tertentu disebut data mining. Teknik, metode, dan algoritma dalam data mining sangat berbeda. Pendekatan atau algoritma yang tepat ditentukan oleh tujuan dan keseluruhan proses KDD [21].

3.5. Interpretation/Evaluation

Tahapan ini dikenal sebagai interpretasi dalam proses *Knowledge Discovery in Database* (KDD). Pola informasi yang dibuat sebagai bagian dari proses penambangan data harus disajikan dengan cara yang dapat dipahami oleh pihak yang berkepentingan. Langkah ini menentukan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau teori sebelumnya [19]. Pada tahapan ini juga akan menghasilkan knowledge atau pengetahuan yang nantinya akan berguna pada saat pengambilan keputusan.



Gambar 1. Tahapan KDD

4. HASIL DAN PEMBAHASAN

Proses atau tahapan KDD terdapat pada penjelasan berikut:

4.1. Data Selection

Data yang dipilih untuk proses *data mining* di dalam penelitian ini adalah data siswa SMAS Katolik Santo Yoseph Denpasar yang diseleksi dari Tahun Angkatan 2019/2020, 2020/2021, 2021/2022 kelas X, XI dan XII sebanyak total 3262 data, dimana jumlah data dari tahun ajaran 2019/2020 sebanyak 1096, jumlah data dari tahun 2020/2021 sebanyak 1080 dan jumlah data dari tahun ajaran 2021/2022 sebanyak 1086 data. Data yang diambil berupa data penghasilan orang tua per bulan, data ketidakhadiran siswa pada semester 1 dan semester 2 serta nilai rapor siswa.

4.2. Pre-processing/Cleaning

Setelah melakukan *data selection*, yang dilakukan selanjutnya adalah melakukan proses peninjauan pada data inkonsisten, dan memperbaiki masalah data termasuk kesalahan pencetakan serta menghapus duplikasi data. Jumlah data siswa yang di seleksi berdasarkan data excel yang diperoleh dari SMAS Katolik Santo Yoseph Denpasar pada Tahun Angkatan 2019/2020, 2020/2021, 2021/2022 adalah sebanyak 3262 data terpilih dan data tersebut akan di proses pada tahap *preprocessing/cleaning*.

Proses cleaning terhadap 3262 data ini dilakukan dengan menghapus beberapa atribut dan data yang tidak dapat digunakan seperti *data null*. *Data null* merupakan data *error* atau tidak sah dikarenakan data tersebut bernilai nol atau kosong. *Data null* yang dihapus berupa data penghasilan orang tua per bulan, data rapor siswa dan data ketidakhadiran siswa. Selain itu, proses pembersihan ini juga menghilangkan variabel yang tidak diperlukan pada data excel nilai rapor siswa. Variabel yang dihapus yaitu mata pelajaran peminatan pada setiap jurusan. Total data setelah dilakukan proses *cleaning* adalah sebanyak 2672 data siswa.

Tabel 1. Pre-processing/cleaning

Tahun	Sebelum Cleaning	Data Null			Sesudah Cleaning
		Penghasilan Orang Tua	Ketidakhadiran Siswa	Nilai Rapor	
2019/2020	1096	158	2	28	908
2020/2021	1080	189	1	10	880
2021/2022	1086	191	0	11	884
TOTAL	3262	538	3	49	2672
		590			

4.3. Transformation

Tahap selanjutnya yaitu *transformation*. Tahap ini dilakukan untuk mengubah data yang dipilih agar data tersebut bisa diolah secara lebih mudah. Tahap ini juga disebut normalisasi. Normalisasi data dilakukan dengan mengubah nilai X1 menjadi bilangan pecahan, membagi nilai X2 dengan 1000000 dan membagi nilai Y dengan 1000.

Tabel 2. Transformation

No	Sebelum Cleaning			Setelah Cleaning		
	Kehadiran Siswa	Penghasilan Orang Tua	Total Nilai Siswa	Kehadiran Siswa	Penghasilan Orang Tua	Total Nilai Siswa
1	95,4%	10.500.000	796	0,954	10,5	0,796
2	98,7%	8.000.000	806	0,987	8,0	0,806
3	99,3%	3.500.000	799	0,993	3,5	0,799
4	98,7%	5.000.000	779	0,987	5,0	0,779
5	100,0%	15.000.000	841	1,000	15,0	0,841
...						
2672	99,5%	4.000.000	851	0,995	4,0	0,851

4.4. Data Mining

Tahap berikutnya adalah *data mining*. Pada langkah ini, peneliti mengidentifikasi pola atau informasi yang menarik dari data yang dipilih dan kemudian menentukan algoritma mana yang akan digunakan untuk menguji data tersebut. Penelitian ini menggunakan *tools Orange Data Mining* untuk melakukan penelitian terhadap siswa SMAS Katolik Santo Yoseph Denpasar dalam memprediksi prestasi siswa berdasarkan kedisiplinan dan tingkat ekonomi orang tua. Penelitian ini menggunakan dataset sebanyak 2672 yang dibagi menjadi dua data, yaitu *data training* dan *data testing* dengan perbandingan 90% *data training* dan 10% *data testing*, 80% *data training* dan 20% *data testing*, serta 70% *data training* dan 30% *data testing*.

4.5. Interpretation/Evaluation

Tahap terakhir adalah *interpretation/evaluation*. Pada fase ini, model data yang dihasilkan oleh proses data mining harus disajikan dengan cara yang dapat dipahami oleh pihak yang berkepentingan. *Tools* yang digunakan untuk mengimplementasikan data mining adalah *Orange Data Mining*. Tahapan ini bertujuan untuk mengetahui apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau teori sebelumnya.

Sebagai evaluasi, peneliti menggunakan tingkat akurasi untuk mengukur seberapa baik model regresi linier berganda menggunakan MAE, MSE dan RMSE. Perbandingan data dalam penelitian ini dilakukan dengan menganalisis 90% *data training* dan 10% *data testing*, 80% *data training* dan 20% *data testing*, serta 70% *data training* dan 30% *data testing*.

Tabel 3. Hasil Tingkat Akurasi

Tingkat Akurasi	Data Training	Data Testing	Hasil Akurasi
MAE	90%	10%	0,027
	80%	20%	0,026
	70%	30%	0,026
MSE	90%	10%	0,001
	80%	20%	0,001
	70%	30%	0,001
RMSE	90%	10%	0,033
	80%	20%	0,033
	70%	30%	0,032

Kesimpulan dari tingkat akurasi yang paling baik antara MAE, MSE dan juga RMSE adalah menggunakan perbandingan data 70% dan 30% karena hasil dari tingkat akurasi tersebut yang paling kecil dimana tingkat akurasi MAE sebesar 0,026; tingkat akurasi MSE sebesar 0,001; dan tingkat akurasi RMSE sebesar 0,032. Karena perbandingan data training 70% dan 30% data testing adalah yang paling akurat, maka persamaan umum yang didapat dalam penelitian ini adalah sebagai berikut:

$$1. Y = 0,668603 + 0,150065x_1 + 8,54719e-05x_2$$

Keterangan:

Y = Prestasi siswa siswi (variabel terikat)

x_1 = Kedisiplinan (variabel bebas)

x_2 = Tingkat ekonomi orang tua (variabel bebas)

a = Konstanta (Intercept)

b = Koefisien variabel bebas

5. KESIMPULAN DAN SARAN

Prediksi prestasi siswa SMAS Katolik Santo Yoseph Denpasar berdasarkan kedisiplinan dan tingkat ekonomi orang tua menghasilkan persamaan Regresi Linier Berganda sebagai berikut:

$$2. Y = 0,668603 + 0,150065x_1 + 8,54719e-05x_2$$

Persamaan tersebut menunjukkan adanya hubungan yang positif antara kedisiplinan siswa, tingkat ekonomi orang tua terhadap prestasi siswa. Kedisiplinan mempunyai pengaruh yang lebih signifikan terhadap prestasi siswa dibandingkan dengan tingkat ekonomi orang tua, sebagaimana tercermin dalam koefisien regresi. 2. Tingkat akurasi prediksi terhadap prestasi siswa SMAS Katolik Santo Yoseph Denpasar berdasarkan kedisiplinan dan tingkat ekonomi orang tua menggunakan metode regresi linier berganda yang paling tinggi yaitu menggunakan perbandingan data 70% *data training* dan 30% *data testing* dengan nilai MAE sebesar 0,026, nilai MSE sebesar 0,001 dan nilai RMSE sebesar 0,032.

Berdasarkan hasil temuan penelitian yang telah dijelaskan diatas, terdapat beberapa saran yang dapat diusulkan dari peneliti untuk meningkatkan prediksi prestasi siswa, yaitu: pertama, memberi perhatian khusus kepada siswa yang memiliki persentase

kehadiran yang kecil terutama siswa yang sering alpa dengan cara mengetahui akar permasalahan, memberi pembinaan atau konseling sesuai kebutuhan, memantau siswa secara terus-menerus serta konsultasi dan bekerjasama dengan orang tua. Kedua, saran bagi peneliti selanjutnya adalah menambah variabel lainnya yang berhubungan dengan prestasi siswa seperti motivasi, dukungan keluarga atau kualitas pengajar. Ketiga, melakukan perbandingan dengan algoritma forecasting lainnya seperti *Random Forest* dan *Neural Network*.

DAFTAR PUSTAKA

- [1] K. Sinaga and K. Handoko, "IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI KELULUSAN SISWA DENGAN METODE NAÏVE BAYES," *Jurnal Comasie*, vol. 04, p. 97, 2021.
- [2] H. Susanto and S. , "DATA MINING UNTUK MEMPREDIKSI PRESTASI SISWA BERDASARKAN SOSIAL EKONOMI, MOTIVASI, KEDISIPLINAN DAN PRESTASI MASA LALU," *Jurnal Pendidikan Vokasi*, vol. 4, p. 223, 2014.
- [3] L. T. Yulianingsih and A. Sobandi, "Kinerja mengajar guru sebagai faktor determinan prestasi belajar siswa," *JURNAL PENDIDIKAN MANAJEMEN PERKANTORAN*, vol. 2, no. 2, p. 160, 2017.
- [4] N. Pradita, "Penerapan Data Mining sebagai Cara untuk Memprediksi Prestasi Siswa Berdasarkan Status Ekonomi dan Kedisiplinan Menggunakan Metode Regresi Linier Berganda," *Jurnal Ilmiah Mahasiswa Manajemen, Bisnis dan Akuntansi*, vol. 4, no. 5, p. 685, 2022.
- [5] M. and C. A. Vemilia, "PENGARUH FAKTOR INTERN DAN FAKTOR EKSTERN TERHADAP PRESTASI BELAJAR EKONOMI," *Jurnal Pendidikan Ekonomi*, vol. 4, p. 273, 2009.
- [6] P. Herlambang, Z. Rohmad and S. Subagya, "PENGARUH KEAKTIFAN DALAM KEGIATAN EKSTRAKURIKULER, MINAT BACA, DAN KEDISIPLINAN BELAJAR TERHADAP PRESTASI BELAJAR SISWA SMA NEGERI 1 KARTASURA," *Universitas Sebelas Maret Surakarta*, p. 6.
- [7] L. A. Matoka, W. Hulukati and M. B. Smith, "ANALISI TENTANG DISIPLIN BELAJAR SISWA KELAS VIII DI SMP NEGERI 2 KOTA GORONTALO," p. 5.
- [8] H. Jamil and F. I. Azra, "PENGARUH LINGKUNGAN KELUARGA DAN MOTIVASI BELAJAR SISWA TERHADAP HASIL BELAJAR AKUNTANSI SISWA KELAS X SMK NEGERI 1 SOLOK SELATAN," *Journal of Economic and Economic Education*, vol. 2, no. 2, p. 88, 2014.
- [9] L. N. Chotimah, H. M. Ani and J. Widodo, "PENGARUH STATUS SOSIAL EKONOMI ORANG TUA TERHADAP PRESTASI BELAJAR SISWA," *Jurnal Pendidikan Ekonomi*, vol. 11, no. 1, p. 76, 2017.
- [10] M. Prasajo, S. and J. Triwidiati, "Prediksi Prestasi Siswa SMK Masuk Pasar Kerja Menggunakan Teknik Data Mining (Studi Kasus SMKN 1 Kota Agung Timur Tanggamus, Lampung)," *Seminar Nasional Hasil Penelitian dan Pengabdian Masyarakat*, p. 135, 2021.
- [11] R. B. Manalu, "SISTEM APLIKASI FORECASTING TINGKAT KELULUSAN MAHASISWA DENGAN ALGORITMA APRIORI (STUDI KASUS: STM IK BUDIDARMA MEDAN)," *Majalah Ilmiah INTI*, vol. 12, p. 224, 2017.
- [12] E. Triyanto, H. Sismoro and A. D. Laksito, "IMPLEMENTASI ALGORITMA REGRESI LINEAR BERGANDA UNTUK MEMPREDIKSI PRODUKSI PADI DI KABUPATEN BANTUL," *RABIT (Jurnal Teknologi dan Sistem Informasi Univrab)*, vol. 4, p. 66, 2019.
- [13] I. G. D. P. Wiwaha and N. W. Suryathi, "PENGARUH PARTISIPASI ANGGARAN DAN GAYA KEPEMIMPINAN TERHADAP KINERJA MANAJERIAL: STUDI KASUS PADA SELURUH LPD DI DESA DINAS DALUNG," *Jurnal Ilmiah Akuntansi & Bisnis*, vol. 3, p. 176, 2018.
- [14] P. P. S. Ramadhan and N. Safitri, "Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada BPS Deli Serdang," *Sains dan Komputer (SAINTIKOM)*, vol. 18, p. 56, 2019.
- [15] I. K. Indrawan, N. N. Parwati and I. P. Suryawan, "PENINGKATAN PRESTASI BELAJAR MATEMATIKA SISWA MELALUI PEMBELAJARAN ARIAS BERBANTUAN GEOGEBRA," *Jurnal Pendidikan dan Pembelajaran Matematika Indonesia*, vol. 7, no. 1, p. 80, 2018.
- [16] F. Mawarni and Y. Fitriani, "Peningkatan Prestasi Belajar Bahasa Indonesia Materi Pokok Teks Eksposisi di Kelas X IPA 2 SMA Negeri 1 Sembawa Kabupaten Banyuasin," *Jurnal Pembelajaran Bahasa Dan Sastra Indonesia*, vol. 9, no. 2, p. 3, 2019.
- [17] I. Ernawati, "PENGARUH LAYANAN INFORMASI DAN BIMBINGAN PRIBADI TERHADAP KEDISIPLINAN SISWA KELAS XII MA COKROAMINOTO WANADADI BANJARNEGARA TAHUN AJARAN 2014/2015," *Jurnal Bimbingan dan Konseling*, vol. 1, no. 1, p. 5, 2016.
- [18] S. W. Febriana and W. Rohmah, "1PENGARUH KONDISI SOSIAL EKONOMI ORANG TUA DAN PERHATIAN ORANG TUA TERHADAP PRESTASI BELAJAR," *Jurnal Pendidikan Ilmu Sosial*, vol. 24, no. 1, p. 6, 2014.

- [19] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *Jurnal Edik Informatika*, vol. 2, no. 2, p. 215.
- [20] M. S. Nawawi, F. Sembiring and A. Erfina, "Implementasi Algoritma K-Means Clustering Menggunakan Orange Untuk Penentuan Produk Busana Muslim Terlaris," *Seminar Nasional Teknologi Informasi dan Komunikasi*, p. 792, 2021.
- [21] M. Alfariqi and A. Finandhita, "IMPLEMENTASI DATA MINING METODE CLUSTERING PADA TOKO BUKU BI-OBSES BANDUNG DENGAN PENERAPAN ALGORITMA K-MEANS," *Jurnal Ilmiah Komputer dan Informatika*, p. 2.
- [22] [17] N. W. Utami dan I. G. J. E. Putra, "Text Minig Clustering Untuk Pengelompokan Topik Dokumen Penelitian Menggunakan Algoritma K-Means Dengan Cosine Similarity," *Jurnal Informatika Teknologi dan Sains*, 2022.
- [23] [18] P. A. C. Dewi, "Pengembangan Perangkat Model Pembelajaran Tutor Sebaya Berbantuan Komik Matematika untuk Meningkatkan Kemampuan Pemecahan Masalah Matematika Siswa," *Wahana Matematika Dan Sains: Jurnal Matematika, Sains, Dan Pembelajarannya*, vol. 14, no. 1, 2020.