

# ecision\_Tree\_Feature\_Engineering\_Kemampuan\_Beradaptasi\_Siswa.pdf

*by Julian Espinoza*

---

**Submission date:** 28-Apr-2024 08:58PM (UTC+0700)

**Submission ID:** 2225858097

**File name:** ecision\_Tree\_Feature\_Engineering\_Kemampuan\_Beradaptasi\_Siswa.pdf (351.06K)

**Word count:** 3710

**Character count:** 24154

# Klasifikasi Tingkat Kemampuan Adaptasi Siswa dalam Pembelajaran Online Menggunakan Decision Tree

*Classification of Students' Adaptability Level in Online Learning Using Decision Tree*

Asmaul Lailiyah<sup>1</sup>, Vinna Rahmayanti Setyaning Nastiti<sup>2</sup>, Evi Dwi Wahyuni<sup>3</sup>, Christian Sri Kusuma Aditya<sup>4</sup>

<sup>1</sup>Informatika, Universitas Muhammadiyah Malang

E-mail: <sup>1</sup>asmaullailiyah@webmail.umm.ac.id, <sup>2</sup>vinastiti@umm.ac.id, <sup>3</sup>evidwi@umm.ac.id,

<sup>4</sup>christianskaditya@umm.ac.id

## Abstrak

*Kemajuan dalam ilmu pengetahuan dan teknologi mendorong adaptasi terhadap pemanfaatan teknologi di berbagai sektor, seperti komunikasi, pendidikan, dan informasi. Terutama dalam konteks teknologi pendidikan, dapat diamati bahwa pembelajaran online sedang mendapatkan popularitas yang signifikan di berbagai lembaga pendidikan. Oleh karena itu, penting untuk mengeksplorasi seberapa baik peserta didik dapat beradaptasi dengan lingkungan pembelajaran online. Memprediksi tingkat adaptasi peserta didik memiliki signifikansi yang besar bagi pendidik dan pengembang platform pembelajaran online, dengan tujuan meningkatkan efisiensi dan kualitas pengalaman belajar. Penelitian ini menggunakan dataset "Students Adaptability Level in Online Education" dengan menerapkan pendekatan Algoritma Decision Tree. Hasil penelitian memperoleh akurasi sebesar 95%, meningkat 7,44% dari penelitian sebelumnya yang hanya memperoleh akurasi sebesar 87,56% dengan menggunakan algoritma yang sama tanpa Feature Engineering. Hal ini menunjukkan bahwa Feature Engineering memegang peranan penting dalam klasifikasi tingkat kemampuan adaptasi siswa untuk mendapatkan hasil yang baik dengan akurasi yang tinggi.*

**Kata kunci:** Pembelajaran Mesin, *Decision Tree*, *Feature Engineering*, Kemampuan Beradaptasi Siswa

## Abstract

*Advances in science and technology encourage adaptation to the utilization of technology in various sectors, such as communication, education, and information. Especially in the context of educational technology, it can be observed that online learning is gaining significant popularity in various educational institutions. Therefore, it is important to explore how well learners can adapt to the online learning environment. Predicting learners' adaptation level has great significance for educators and developers of online learning platforms, with the aim of improving the efficiency and quality of the learning experience. This research uses a dataset from Kaggle by applying the Decision Tree Algorithm approach. The research results obtained an accuracy of 95%, an increase of 7.44% from previous research which only obtained an accuracy of 87.56% using the same algorithm without Feature engineering. This shows that Feature Engineer play an important role in classifying students' adaptation levels to get good results with high accuracy.*

**Keywords:** *Machine Learning, Decision Tree, Feature Engineering, Student Adaptability*

## 1. PENDAHULUAN

Perkembangan ilmu pengetahuan dan teknologi telah mendorong kita untuk menyesuaikan diri dengan penggunaan teknologi dalam berbagai bidang, termasuk komunikasi, pendidikan, dan informasi. Salah satu hal yang menonjol adalah pendidikan, di mana

pembelajaran online telah mendapatkan popularitas yang signifikan di banyak lembaga pendidikan [1], [2]. Pembelajaran online adalah inovasi pendidikan yang melibatkan penggunaan teknologi informasi dalam proses belajar-mengajar. Pada masa sekarang, kebermaknaan pembelajaran online meningkat karena perkembangan teknologi telah mendorong sejumlah besar sekolah untuk mengadopsi metode pembelajaran online [3].

Tantangan dalam berbagai aspek infrastruktur teknologi yang tidak merata, kesenjangan akses internet, serta kemampuan pendidik dan peserta didik untuk menyesuaikan diri dengan metode pembelajaran online menjadi hambatan dalam proses adaptasi pembelajaran online. Peserta didik yang mampu beradaptasi dengan baik terhadap lingkungan pembelajaran online dapat mendapatkan pendidikan yang berkualitas, menguasai materi secara mandiri, dan mengembangkan keterampilan teknologi dan digital. Karena itu, penting untuk memahami sejauh mana peserta didik mampu beradaptasi dengan lingkungan pembelajaran online. Kemampuan untuk menilai kemampuan beradaptasi peserta didik menjadi kunci bagi pendidik dan pengembang platform pembelajaran online guna meningkatkan kualitas dan efektivitas pengalaman belajar [4], [5].

Klasifikasi Decision Tree adalah teknik pembelajaran mesin yang dapat digunakan untuk memprediksi tingkat kemampuan beradaptasi siswa dalam pembelajaran online. Kemampuan beradaptasi adalah kemampuan untuk menyesuaikan diri dengan situasi baru dan mengatasi tantangan dalam lingkungan belajar [6]. Pembelajaran online adalah bentuk pendidikan yang mengandalkan internet dan teknologi digital untuk menyampaikan konten dan interaksi [7]. Pembelajaran online dapat menimbulkan berbagai kesulitan bagi siswa, seperti masalah teknis, kurangnya motivasi, isolasi, dan gangguan. Oleh karena itu, penting untuk mengidentifikasi faktor-faktor yang mempengaruhi kemampuan beradaptasi siswa dalam pembelajaran online dan memberikan dukungan dan bimbingan yang tepat [5]. Dalam konteks pendidikan, *Machine Learning* menjadi pilihan alternatif dalam mengolah data mentah [8]. Melibatkan penggunaan beragam data terkait dengan bidang pendidikan, yang kemudian diolah dan diinterpretasikan menjadi berbagai atribut yang kemudian digunakan sebagai dasar untuk algoritma model [9].

Salah satu cara untuk mendekati masalah ini adalah dengan menggunakan klasifikasi *Decision Tree* untuk menganalisis data yang dikumpulkan dari siswa yang berpartisipasi dalam pembelajaran online. Klasifikasi *Decision Tree* adalah metode yang menggunakan struktur pohon di mana setiap *node* mencakup tes pada suatu fitur, setiap cabang merepresentasikan hasil tes, dan setiap node daun mencerminkan label kelas. [10]. Pohon ini dibangun dengan membagi data secara rekursif berdasarkan metrik yang mengukur perolehan informasi atau pengurangan ketidakmurnian setelah pemisahan. Tujuannya adalah untuk mempermudah pengambilan keputusan dalam merancang strategi pembelajaran yang lebih efektif [11].

Algoritma *Decision Tree* [12] menjadi pilihan populer dalam klasifikasi karena kemampuannya menangani data kategorikal dan numerik, menangani nilai yang hilang, dan menangkap hubungan non-linear. Klasifikasi ini juga mudah diinterpretasikan dan divisualisasikan, karena memberikan penjelasan yang jelas tentang logika di balik klasifikasi [13].

*Feature Engineering* adalah proses di mana data mentah diubah atau diproses untuk membuat fitur-fitur baru yang lebih informatif dan relevan untuk tujuan analisis atau model prediktif [14]. Tujuannya adalah untuk meningkatkan kualitas data, meningkatkan kinerja model, dan membantu model dalam menangkap pola atau informasi yang mungkin tersembunyi dalam data [15], [16]. *Feature Engineering* berkontribusi secara signifikan terhadap kesuksesan model *machine learning* dan dapat membuat perbedaan besar dalam kualitas prediksi dan analisis data [17]. *Feature Engineering* dengan *Decision Tree* melibatkan penggunaan prediksi model *Decision Tree* sebagai fitur baru untuk model lain. Hal ini dapat membantu menangkap interaksi dan pola yang kompleks yang mungkin tidak dapat ditangkap oleh fitur asli.

Tujuan utama dalam penelitian ini adalah untuk mengkalsifikasikan kemampuan adaptasi siswa dan mencapai tingkat akurasi evaluasi tertinggi dengan menggunakan pendekatan *Decision Tree*. Suzan M, dkk [4] menerapkan beberapa algoritma *machine learning*, seperti *decision tree*, *random forest*, *naive Bayes*, *support vector machine*, *K-nearest neighbour*, dan jaringan syaraf tiruan, untuk memprediksi tingkat kemampuan beradaptasi siswa dalam pendidikan daring

berdasarkan faktor sosio-demografi mereka. Penelitian ini menggunakan dataset siswa dari berbagai tingkatan (sekolah, perguruan tinggi, dan universitas) di Bangladesh dan menemukan bahwa random forest mencapai akurasi terbaik sebesar 89,63% dan mengungguli algoritme lainnya. Pada penelitian ini algoritma *Decision Tree* menunjukkan hasil akurasi sebesar 87,56%.

Hark Söylemez N, dkk [2] mengevaluasi kinerja kursus pembelajaran online menggunakan algoritma *Decision Tree* berdasarkan umpan balik dan faktor kepuasan siswa, seperti konten kursus, instruktur, interaksi, penilaian, dan dukungan. Penelitian ini menggunakan dataset siswa dari sebuah universitas di Cina dan menerapkan algoritma C4.5 untuk membangun *Decision Tree*. Penelitian ini juga menganalisis pentingnya setiap faktor dan memberikan beberapa saran untuk meningkatkan kualitas pembelajaran online.

Alkayyali, dkk [18] membandingkan kinerja teknik *machine learning* dan *deep learning*, seperti *decision tree*, *random forest*, *naive bayes*, *support vector machine*, *K-nearest neighbour*, dan jaringan syaraf tiruan, untuk memprediksi tingkat kemampuan beradaptasi siswa dalam pembelajaran elektronik berdasarkan faktor sosio-demografi mereka. Penelitian ini menggunakan dataset siswa dari repositori Kaggle dan menemukan bahwa algoritma deep learning yang diusulkan mencapai akurasi terbaik sebesar 94,67% dan mengungguli algoritma lainnya.

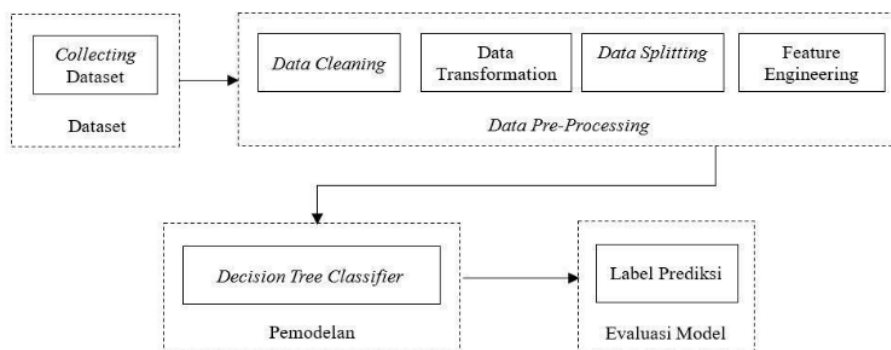
Sri, dkk [19] membandingkan hasil akurasi tingkat adaptasi siswa pada pembelajaran online menggunakan metode *Support Vector Machine* tanpa optimasi dan metode *Support Vector Machine* menggunakan optimasi *GridSearchCV*. Didapatkan hasil metode *Support Vector Machine* menggunakan optimasi *GridSearchCV* jauh lebih unggul dengan akurasi sebesar 95%, sedangkan yang tidak dilakukan optimasi hanya mendapat akurasi sebesar 80%.

Kumar, D [13] membahas keunggulan *Decision Tree Classification* dibandingkan metode klasifikasi lainnya. Makalah ini menyimpulkan bahwa *Decision Tree Classification* adalah salah satu metode yang terkenal untuk klasifikasi data, dan fitur yang paling signifikan adalah kemampuannya untuk mengubah masalah pengambilan keputusan yang rumit menjadi proses yang sederhana, sehingga menemukan solusi yang dapat dimengerti dan lebih mudah diinterpretasikan.

Pada penelitian ini menggunakan algoritma *Decision Tree* dengan karakteristik yang terletak pada kemampuannya untuk menyederhanakan masalah pengambilan keputusan yang kompleks menjadi suatu proses yang lebih mudah dipahami, untuk mengklasifikasikan tingkat kemampuan beradaptasi siswa dalam pembelajaran online, didukung oleh *Feature Engineering* untuk meningkatkan akurasi pada kasus adaptabilitas siswa.

## 2. METODE PENELITIAN

Penelitian ini menggunakan algoritma *Decision Tree* untuk mengklasifikasikan tingkat kemampuan beradaptasi siswa dalam pembelajaran online. Metode penelitian ini didasarkan pada penelitian [19] dan meliputi langkah-langkah sebagai berikut:



Gambar 1. Alur Penelitian

### 2.1 Dataset

Dataset yang digunakan dalam penelitian ini berasal dari Kaggle dengan judul “*Students Adaptability Level in Online Education*”. Dataset diperoleh dari hasil gabungan survei yang dilakukan di lembaga pendidikan di Bangladesh, India baik secara konvensional maupun daring [4]. Terdiri dari 1.205 data yang diambil dan diproses untuk merekayasa atribut – atribut yang memiliki korelasi dengan kemampuan adaptasi Siswa. Dataset terdiri dari 14 atribut, antara lain *Gender, Age, Education Level, Institution Type, IT Student, Location, Load-shedding, Financial Condition, Internet Type, Network Type, Class Duration, Self Lms, Device, dan Adaptivity Level*. Dengan target fitur untuk menemukan tingkat kemampuan beradaptasi Siswa dalam pembelajaran online.

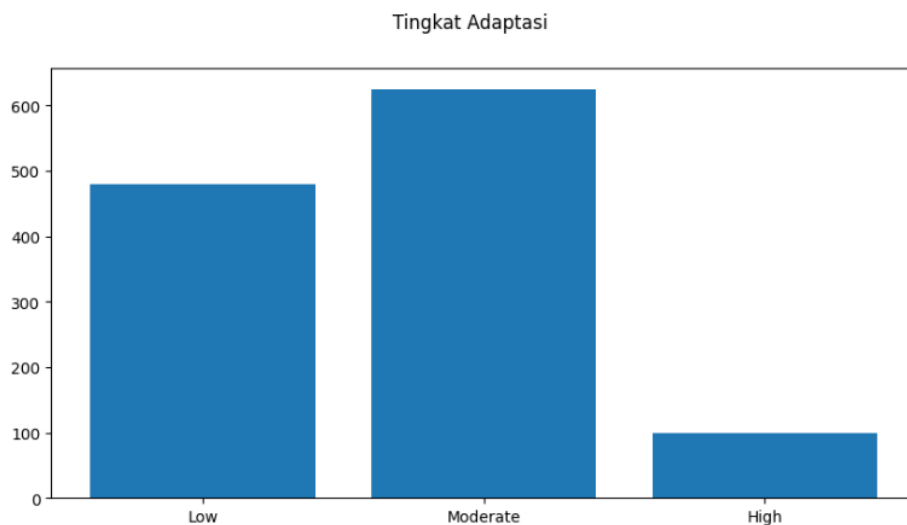
Tabel 1. Atribut dan Nilai Dataset

| N  | Atribut                    | Deskripsi                              | Nilai                                 |
|----|----------------------------|--|---------------------------------------|
| 0  |                            |  |                                       |
| 1  | <i>Gender</i>              | Jenis kelamin                          | <i>Boy dan Girl</i>                   |
| 2  | <i>Age</i>                 | Rentang umur                           | 21-25, 11 -15, 16-20, 1-5 dan 26-30   |
| 3  | <i>Education Level</i>     | Tingkat pendidikan                     | <i>School, University dan College</i> |
| 4  | <i>Institution Type</i>    | Tipe institusi                         | <i>Non Government dan Government</i>  |
| 5  | <i>IT Student</i>          | Status IT                              | <i>True dan False</i>                 |
| 6  | <i>Location</i>            | Status tinggal di Kota                 | <i>True dan False</i>                 |
| 7  | <i>Load-shedding</i>       | Tingkat beban                          | <i>Low dan High</i>                   |
| 8  | <i>Financial Condition</i> | Kondisi finansial keluarga             | <i>Poor, Mid dan Rich</i>             |
| 9  | <i>Internet Type</i>       | Tipe jaringan internet                 | <i>Mobile Data dan Wifi</i>           |
| 10 | <i>Network Type</i>        | Tipe konektivitas jaringan internet    | <i>4G, 3G dan 2G</i>                  |
| 11 | <i>Class Duration</i>      | Durasi kelas sehari-hari               | 1-3, 3-6 dan 0                        |
| 12 | <i>Self Lms</i>            | Status institusi memiliki lms          | <i>True dan False</i>                 |
| 13 | <i>Device</i>              | Perangkat yang digunakan saat di kelas | <i>Mobile, Computer dan Tab</i>       |
| 14 | <i>Adaptivity Level</i>    | Tingkat adaptasi                       | <i>Moderate, Low dan High</i>         |

Tabel diatas mendeskripsikan atribut dan nilai dataset yang digunakan dalam penelitian ini. Dapat dilihat dataset terdiri dari 14 atribut, yaitu *Gender, Age, Education Level, Institution Type, IT Student, Location, Load-shedding, Financial Condition, Internet Type, Network Type,*

*Class Duration*, *Self Lms*, *Device*, dan *Adaptivity Level* dengan deskripsi dan masing- masing nilai seperti pada tabel diatas.

EDA (*Exploratory Data Analysis*) merujuk pada proses analisis data yang dilakukan secara eksploratif untuk memahami karakteristik, pola, dan informasi penting dalam dataset sebelum melakukan analisis yang lebih mendalam atau membangun model. Tujuan utama EDA adalah untuk memperoleh wawasan awal tentang struktur data, menemukan anomali, dan merumuskan hipotesis yang dapat membimbing analisis lebih lanjut. Berikut adalah grafik distribusi tingkat adaptabilitas siswa yang beragam dan merupakan data target yang digunakan dalam penelitian ini



Gambar 2. *Exploratory Data Analysis* Tingkat Adaptasi

Grafik yang ditampilkan adalah grafik distribusi tingkat adaptabilitas siswa yang beragam dan merupakan data target yang digunakan dalam penelitian ini. Mayoritas siswa memiliki adaptabilitas pada tingkat sedang dengan jumlah data lebih dari 600, yang merupakan kategori terbesar, menunjukkan bahwa sebagian besar siswa mampu beradaptasi secara wajar terhadap tantangan. Sejumlah signifikan siswa berada pada tingkat adaptabilitas rendah dengan jumlah data dibawah 500, mengindikasikan bahwa siswa mengalami tingkat adaptasi yang rendah pada adaptasi pembelajaran online. Sejumlah 100 siswa menunjukkan adaptabilitas tinggi, menandakan mereka memiliki keterampilan adaptasi yang lebih maju. Kesimpulannya, data ini memperlihatkan keragaman dalam kemampuan adaptasi siswa yang harus diperhatikan oleh pendidik untuk membantu setiap siswa berkembang sesuai dengan kemampuan adaptasi mereka.

## 2.2 Data Preprocessing

*Data Preprocessing* dalam penelitian ini meliputi *data cleaning*, *data transformation*, *splitting data*, dan *Feature Engineering*. Langkah pertama adalah melakukan pengecekan terhadap setiap atribut untuk menentukan apakah terdapat nilai yang masuk dalam kategori missing value. Penelitian ini melakukan pengecekan missing value menggunakan teknik *programmatic*, hasilnya menunjukkan bahwa tidak ada nilai yang hilang, yang mengindikasikan bahwa dataset memiliki kualitas yang baik .

Langkah selanjutnya adalah *Data Transformation* yang bertujuan untuk mengubah skala pengukuran dari nilai asli ke bentuk nilai yang berbeda. Pada dataset yang digunakan, seluruh atribut memiliki data kategorikal, sehingga dilakukan transformasi data dengan menerapkan

teknik transformasi menggunakan Standard Scaler untuk menormalisasikan data menjadi data numerik.

Dataset yang telah didapatkan setelah adanya proses *Data Transformation* akan dibagi menjadi *Data Training* dan juga *Data Testing* dengan rasio 8:2. *Data Training* digunakan untuk melatih data pada model dan kemudian *Data Testing* digunakan dalam proses evaluasi guna menentukan kinerja model. Selanjutnya adalah menentukan objek data, dimana pembagian terdiri dari data fitur (x) yang berisi semua kolom kecuali Adaptasi Siswa, dan data target (y) yang berisi hanya kolom Adaptasi Siswa. Tabel dibawah ini merupakan hasil dari pembagian *Data Training* dan *Data Testing*.

Tabel 2. Pembagian *Data Training* dan *Data Testing*

| Keterangan | <i>Data Training</i> | <i>Data Testing</i> | Jumlah |
|------------|----------------------|---------------------|--------|
| Presentase | 80%                  | 20%                 | 100%   |
| Jumlah     | 964                  | 241                 | 1205   |

Tabel diatas menjelaskan bahwa presentase *Data Training* sejumlah 80% dan *Data Testing* sejumlah 20%, sesuai dengan pembagian dataset dengan rasio 8 : 2. Dengan rincian data keseluruhan berjumlah 1.205 data dibagi menjadi *Data Training* dengan jumlah 964 data dan *Data Testing* berjumlah 241 data.

*Feature Engineering* adalah langkah penting dalam pemrosesan data yang melibatkan pembuatan fitur – fitur baru yang lebih informatif dan relevan untuk tujuan analisis atau model prediktif. Tujuannya adalah untuk meningkatkan kualitas data, meningkatkan kinerja model dan membantu model dalam menangkap informasi yang mungkin tersembunyi dalam data. Hasil implementasi *Feature Engineering* menggunakan teknik *One Hot Encoding* adalah penambahan jumlah atribut yang semula berjumlah 14 atribut menjadi 36 atribut.

### 2.3 Decision Tree Classifier

Karakteristik utama dari *Decision Tree Classifier* terletak pada kemampuannya untuk menyederhanakan masalah pengambilan keputusan yang kompleks menjadi suatu proses yang lebih mudah dipahami. Hal tersebut memungkinkan untuk menemukan solusi yang lebih mudah diinterpretasikan. *Decision Tree* adalah model prediktif dalam *Machine Learning* yang digunakan untuk membuat keputusan atau prediksi berdasarkan serangkaian aturan atau pertanyaan yang hierarkis. Model ini memvisualisasikan struktur yang mirip dengan pohon, memiliki titik-titik cabang (*node*) yang menunjukkan keputusan atau pertanyaan, dan ujung-ujung daun (*leaf*) yang mencerminkan hasil atau kelas.

Penelitian ini mengadopsi algoritma *Decision Tree* sebagai metode utama untuk menganalisis tingkat adaptasi Siswa dalam konteks pembelajaran online. Pertama – tama dataset diperoleh dari Kaggle dengan judul “*Students Adaptability Level in Online Education*”, yang terdiri dari 1.205 data. Selanjutnya dilakukan konstruksi model *Decision Tree*, dimana algoritma membagi data berdasarkan atribut yang paling informatif untuk memaksimalkan perbedaan antara kelas atau hasil objek yang berbeda. Model ini menghasilkan *Decision Tree* yang memberikan pandangan terstruktur mengenai faktor – faktor yang mempengaruhi tingkat adaptasi Siswa. Proses pelatihan melibatkan penggunaan *Data Training* untuk mengajarkan model pola – pola yang terdapat dalam data, sehingga model dapat membuat prediksi yang akurat.

Selanjutnya, evaluasi model dilakukan dengan menggunakan *Data Testing* yang tidak pernah dilihat sebelumnya. Metrik evaluasi, seperti akurasi, presisi, recall, dan F1-Score, digunakan untuk mengevaluasi kinerja model *Decision Tree*. Hasil dari model tersebut dianalisis untuk mendapatkan wawasan tentang faktor – faktor kunci yang mempengaruhi tingkat adaptasi Siswa dalam pembelajaran online. Dengan demikian, hasil analisis *Decision Tree* didukung dengan *Feature Engineering* dapat memberikan kontribusi yang berarti dalam mendukung pengambilan keputusan dan pengembangan strategi pembelajaran online yang lebih efektif.

### 3. HASIL DAN PEMBAHASAN

Tahap ini membahas hasil yang berkaitan dengan metode yang digunakan, termasuk pengolahan data dan evaluasi akhir yang meliputi akurasi, recall, precision, dan F1 score. Perbandingan hasil evaluasi model dapat dilihat pada tabel dibawah ini.

Tabel 3. Performa model *Decision Tree* tanpa *Feature Engineering*

| Report    | Adaptivity Level |          |       | Accuracy |
|-----------|------------------|----------|-------|----------|
|           | Low              | Moderate | High  |          |
| Precision | 0.87             | 0.89     | 0.83  | 0.87     |
| Recall    | 0.90             | 0.90     | 0.65  |          |
| F1-Score  | 0.89             | 0.89     | 0.73  |          |
| Support   | 52.00            | 56.00    | 13.00 |          |

Tabel 4. Performa model *Decision Tree* dengan *Feature Engineering*

| Report    | Adaptivity Level |          |       | Accuracy |
|-----------|------------------|----------|-------|----------|
|           | Low              | Moderate | High  |          |
| Precision | 0.98             | 0.92     | 1.00  | 0.95     |
| Recall    | 0.98             | 0.98     | 0.69  |          |
| F1-Score  | 0.98             | 0.95     | 0.82  |          |
| Support   | 52.00            | 56.00    | 13.00 |          |

Berdasarkan tabel 3 dan tabel 4 diatas didapatkan perbandingan performa yang signifikan antara performa model *Decision Tree* tanpa *Feature Engineering* dengan performa model *Decision Tree* menggunakan *Feature Engineering*. Dengan akurasi yang mencakup nilai precision, recall, f1-score, dan support pada model *Decision Tree* tanpa *Feature Engineering* pada jurnal [20] mendapatkan hasil akurasi sebesar 87,56%. Peningkatan akurasi sebesar 7,44% didapatkan dari model *Decision Tree* menggunakan *Feature Engineering* pada penelitian ini yang mendapatkan hasil akurasi sebesar 95%. Penerapan metode, *splitting data*, dan penggunaan dataset yang sama, namun dengan perbedaan dalam penambahan *Feature Engineering* pada tahap *preprocessing*, menghasilkan perbedaan hasil yang signifikan. Dengan adanya peningkatan akurasi yang signifikan, berarti penggunaan *Feature Engineering* lebih baik dibandingkan dengan algoritma *Decision Tree* tanpa *Feature Engineering*.

Penelitian ini menunjukkan bahwa penerapan *Feature Engineering* pada algoritma *Decision Tree* memiliki dampak pada kinerja model *Machine Learning*. Temuan ini dapat berfungsi sebagai parameter evaluasi dalam membuat keputusan terkait langkah-langkah yang harus diambil dalam konteks pembelajaran online, sekaligus sebagai langkah awal bagi institusi pendidikan dalam menentukan langkah yang diperlukan ke depannya.

Dilihat dari "*recall* di *Adaptivity Level high* bernilai rendah" mengindikasikan bahwa pada tingkat adaptivitas yang tinggi, nilai *recall*-nya rendah. *Recall* (sensitivitas) adalah metrik evaluasi kinerja model yang mengukur kemampuan model untuk mengidentifikasi semua contoh positif sebenarnya dalam kelas positif. Oleh karena itu, rendahnya nilai *recall* pada tingkat



adaptivitas yang tinggi dapat menunjukkan adanya tantangan atau kelemahan yang perlu diatasi dalam mengoptimalkan kinerja model pada kasus-kasus yang lebih kompleks atau bervariasi.

#### 4. KESIMPULAN DAN SARAN

Berdasarkan hasil dan analisis pada penelitian ini, metode *Decision Tree* telah sukses diimplementasikan dengan baik. Hasil penelitian menunjukkan bahwa implementasi metode *Decision Tree* untuk memprediksi tingkat adaptasi siswa pada pembelajaran online mencapai akurasi sebesar 87,56%. Sementara itu, penggunaan metode *Decision Tree* dengan menggunakan *Feature Engineering* untuk memprediksi tingkat adaptasi siswa pada pembelajaran online menghasilkan akurasi sebesar 95%. Hal ini menandakan bahwa penggunaan *Feature Engineering* memberikan hasil yang lebih baik pada algoritma *Decision Tree*, terbukti dari peningkatan akurasi sebesar 7,44%. Penelitian ini mengindikasikan bahwa penggunaan *Feature Engineering* pada algoritma *Decision Tree* memberikan peningkatan terhadap kinerja model *Machine Learning*.

Rekomendasi dari hasil penelitian ini adalah untuk mendalami pemahaman terhadap faktor-faktor yang memengaruhi tingkat adaptasi siswa dalam pembelajaran daring, mengeksplorasi algoritma *Decision Tree* dan juga teknik *preprocessing* data untuk mengoptimalkan *Feature Engineering* dan mencapai model yang lebih baik. Penelitian ini juga memberikan peluang untuk dikembangkan lebih lanjut guna mencapai performa yang lebih unggul.

#### DAFTAR PUSTAKA

- [1] C. HARK, H. OKUMUŞ, and T. UÇKAN, "Adaptation to Online Education: An Educational Data Mining Application," *Computer Science*, Nov. 2022, doi: 10.53070/bbd.1199055.
- [2] N. Hark Söylemez, "An Investigation on Students' Level of Adjustment to Online Education with C5.0 Decision Tree Algorithm," *Shanlax International Journal of Education*, vol. 11, no. 2, pp. 64–73, Mar. 2023, doi: 10.34293/education.v11i2.5783.
- [3] B. B. Locke, "Online education in the post-COVID era," *Nature Electronics*, vol. 4, no. 1. Nature Research, pp. 5–6, Jan. 01, 2021. doi: 10.1038/s41928-020-00534-0.
- [4] M. M. H. Suzan, N. A. Samrin, A. A. Biswas, and M. A. Pramanik, "Students' Adaptability Level Prediction in Online Education using Machine Learning Approaches," in *2021 12th International Conference on Computing Communication and Networking Technologies, ICCCNT 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ICCCNT51525.2021.9579741.
- [5] B. Sekeroglu, K. Dimililer, and K. Tuncal, "Student performance prediction and classification using machine learning algorithms," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, 2019, pp. 7–11. doi: 10.1145/3318396.3318419.
- [6] G. H. Wang, J. Zhang, and G. S. Fu, "Predicting student behaviors and performance in online learning using decision tree," in *Proceedings - 2018 7th International Conference of Educational Innovation through Technology, EITT 2018*, Institute of Electrical and Electronics Engineers Inc., Jul. 2018, pp. 214–219. doi: 10.1109/EITT.2018.00050.
- [7] A. Malik *et al.*, "Forecasting students' adaptability in online entrepreneurship education using modified ensemble machine learning model," *Array*, vol. 19, Sep. 2023, doi: 10.1016/j.array.2023.100303.
- [8] I. Pan, A. Mukherjee, and V. Piuri, "Advances in Intelligent Systems and Computing 1355 Proceedings of Research and Applications in Artificial Intelligence." [Online]. Available: <http://www.springer.com/series/11156>

- [9] Ş. Aydoğdu, "Predicting student final performance using artificial neural networks in online learning environments," *Educ Inf Technol (Dordr)*, vol. 25, no. 3, pp. 1913–1927, May 2020, doi: 10.1007/s10639-019-10053-x.
- [10] V. Matzavela and E. Alepis, "Decision tree learning through a Predictive Model for Student Academic Performance in Intelligent M-Learning environments," *Computers and Education: Artificial Intelligence*, vol. 2, Jan. 2021, doi: 10.1016/j.caeai.2021.100035.
- [11] A. B. Raut and A. A. Nichat, "Students Performance Prediction Using Decision Tree Technique," 2017. [Online]. Available: <http://www.ripublication.com>
- [12] G. Koulinas, P. Paraschos, and D. Koulouriotis, "A machine learning-based framework for data mining and optimization of a production system," in *Procedia Manufacturing*, Elsevier B.V., 2021, pp. 431–438. doi: 10.1016/j.promfg.2021.10.059.
- [13] D. Kumar, "Biographical notes: Priyanka received her Bachelor of Technology in Computer Science and Engineering (CSE) and Master of Technology in CSE from GJUS&T," 2020.
- [14] B. Rajoub, "Characterization of biomedical signals: Feature engineering and extraction," in *Biomedical Signal Processing and Artificial Intelligence in Healthcare*, Elsevier, 2020, pp. 29–50. doi: 10.1016/B978-0-12-818946-7.00002-0.
- [15] W. Long, Z. Lu, and L. Cui, "Deep learning-based feature engineering for stock price movement prediction," *Knowl Based Syst*, vol. 164, pp. 163–173, Jan. 2019, doi: 10.1016/j.knosys.2018.10.034.
- [16] P. W. Khan and Y. C. Byun, "Genetic algorithm based optimized feature engineering and hybrid machine learning for effective energy consumption prediction," *IEEE Access*, vol. 8, pp. 196274–196286, 2020, doi: 10.1109/ACCESS.2020.3034101.
- [17] C. Fan, Y. Sun, Y. Zhao, M. Song, and J. Wang, "Deep learning-based feature engineering methods for improved building energy prediction," *Appl Energy*, vol. 240, pp. 35–45, Apr. 2019, doi: 10.1016/j.apenergy.2019.02.052.
- [18] Z. K. D. Alkayyali, B. S. Abu-Nasser, A. M. Taha, Q. M. M. Zarandah, and S. S. Abu-Naser, "Prediction of Student Adaptability Level in e-Learning using Machine and Deep Learning Techniques," 2022. [Online]. Available: [www.ijeais.org/ijaar](http://www.ijeais.org/ijaar)
- [19] C. Sri, K. Aditya, V. Rahmayanti, S. Nastiti, Y. Aminuddin, and M. A. Rofiq, "Prediction of Student Adaptability Level in Online Education Using Support Vector Machine with GridSearchCV Optimization."
- [20] M. M. H. Suzan, N. A. Samrin, A. A. Biswas, and M. A. Pramanik, "Students' Adaptability Level Prediction in Online Education using Machine Learning Approaches," in *2021 12th International Conference on Computing Communication and Networking Technologies, ICCCNT 2021*, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ICCCNT51525.2021.9579741.

# ecision\_Tree\_Feature\_Engineering\_Kemampuan\_Beradapt...

---

## ORIGINALITY REPORT

---

**14%**

SIMILARITY INDEX

**11%**

INTERNET SOURCES

**4%**

PUBLICATIONS

**5%**

STUDENT PAPERS

---

## MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

---

2%

★ Submitted to University of California, Los Angeles

Student Paper

---

Exclude quotes  On

Exclude matches  Off

Exclude bibliography  On