

**PENERAPAN *SPLIT FEATURE REDUCTION* PADA *DECISION TREE*
C4.5 UNTUK MENINGKATKAN PREDIKSI PEFORMA BELAJAR
SISWA SMK**

SKRIPSI

diajukan untuk memenuhi syarat dalam memperoleh gelar Sarjana
Pendidikan Program Studi Pendidikan Ilmu Komputer



oleh:

Ali Akbar Jausyani

1807105

**PROGRAM STUDI PENDIDIKAN ILMU KOMPUTER
DEPARTEMEN PENDIDIKAN ILMU KOMPUTER
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN
ALAM
UNIVERSITAS PENDIDIKAN INDONESIA
2022**

Ali Akbar Jausyani, 2022

***PENERAPAN SPLIT FEATURE REDUCTION PADA DECISION TREE C4.5 UNTUK MENINGKATKAN
PREDIKSI PERFORMA BELAJAR SISWA SMK***

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

**PENERAPAN *SPLIT FEATURE REDUCTION* PADA *DECISION TREE C4.5*
UNTUK MENINGKATKAN PREDIKSI PERFORMA BELAJAR SISWA SMK**

Oleh

Ali Akbar Jausyani

1807105

Sebuah skripsi diajukan untuk memenuhi sebagian syarat memperoleh gelar
Sarjana Pendidikan pada Program Studi Pendidikan Ilmu Komputer

© Ali Akbar Jausyani 2022

Universitas Pendidikan Indonesia

Juni 2022

Hak Cipta dilindungi Undang-Undang

Skripsi ini tidak boleh diperbanyak seluruhnya atau Sebagian, dengan dicetak
ulang, difotokopi, atau cara lainnya tanpa izin dari penulis

Ali Akbar Jausyani, 2022

**PENERAPAN *SPLIT FEATURE REDUCTION* PADA *DECISION TREE C4.5* UNTUK MENINGKATKAN
*PREDIKSI PERFORMA BELAJAR SISWA SMK***

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

ALI AKBAR JAUSYANI

PENERAPAN *SPLIT FEATURE REDUCTION* PADA *DECISION TREE C4.5*
UNTUK MENINGKATKAN PREDIKSI PEFORMA BELAJAR SISWA SMK

disetujui dan disahkan oleh:

Pembimbing 1



Dr. Budi Laksono Putro, M.T.

NIP 197607102010121002

Pembimbing 2



Jajang Kusnendar, M.T.

NIP 197506012008121001

Mengetahui

Ketua Program Studi Pendidikan Ilmu Komputer



Dr. Wahyudin, M.T.

NIP 197304242008121001

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penerapan *Split Feature Reduction* pada *Decision Tree C4.5* untuk Meningkatkan Prediksi Performa Belajar Siswa SMK” ini beserta seluruh isinya adalah benar-benar hasil karya saya sendiri. Saya tidak melakukan penjiplakan ataupun pengutipan dengan cara-cara yang tidak sesuai dengan etika ilmu yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung risiko/sanksi apabila di kemudian hari ditemukan adanya pelanggaran etika keilmuan atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Bandung, Juni 2022

Yang membuat pernyataan,



Ali Akbar Jausyani

NIM. 1807105

KATA PENGANTAR

Puji syukur ke hadirat Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Penerapan *Split Feature Reduction* pada *Decision Tree C4.5* untuk Meningkatkan Prediksi Performa Belajar Siswa SMK” dengan baik namun tidak terlepas dari kekurangan.

Skripsi ini ditulis untuk memenuhi sebagian dari syarat memperoleh gelar sarjana Pendidikan Ilmu Komputer Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam Universitas Pendidikan Indonesia.

Dalam proses penyusunan skripsi ini penulis mengalami berbagai kendala, namun atas pertolongan dan Ridha Allah SWT dan bantuan, bimbingan, serta kerja sama dari berbagai pihak kendala tersebut dapat diatasi. Penulis mengucapkan terima kasih kepada pihak yang membantu kelancaran penulisan, termasuk pihak yang telah membagi pengetahuannya sehingga penulis dapat menyelesaikan skripsi ini.

Penulis menyadari skripsi yang ditulis masih jauh dari kata sempurna. Oleh karena itu, kritik dan saran yang membangun penulis nantikan demi perbaikan di masa mendatang. Semoga skripsi ini dapat bermanfaat dan memberikan pembelajaran yang baik untuk penulis dan pembaca khususnya menjadi sumber ilmu pengetahuan.

Bandung, Juni 2022



Ali Akbar Jausyani

NIM. 1807105

UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas rahmat dan karunia-Nya penulis dapat menyelesaikan skripsi ini. Pada penelitian dan penyusunan skripsi ini, penulis banyak mendapat dukungan dan bantuan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih kepada:

1. Kedua orang tua penulis yaitu Bapak Agus Nasihin dan Ibu Yanti Sri Budiarti yang telah mendidik, memberikan doa, dan memberikan dukungan moril dan materiil kepada penulis.
2. Kakak yaitu Iqra Sadra Muhammad, Alia Fatihah Ummulqurani, dan adik yaitu Samudra Aqila Ali Ilmi yang telah memberikan semangat dan dukungan kepada penulis.
3. Bapak Lala Septem Riza, M.T., Ph.D., selaku Kepala Departemen Pendidikan Ilmu Komputer FPMIPA Universitas Pendidikan Indonesia.
4. Bapak Dr. Wahyudin, M.T., selaku Ketua Program Studi Pendidikan Ilmu Komputer FPMIPA Universitas Pendidikan Indonesia.
5. Bapak Erlangga, M.T. selaku pembimbing akademik yang telah membimbing perihal akademis selama perkuliahan.
6. Bapak Dr. Budi Laksono Putro, M.T., selaku dosen pembimbing I yang telah bersedia meluangkan waktu, tenaga, dan pikiran untuk membimbing penulis dalam proses penelitian dan penyusunan skripsi.
7. Bapak Jajang Kusnendar, M.T., selaku dosen pembimbing II yang telah bersedia meluangkan waktu, tenaga, dan pikiran untuk membimbing penulis dalam proses penelitian dan penyusunan skripsi.
8. Bapak dan Ibu Dosen Departemen Pendidikan Ilmu Komputer yang telah memberikan ilmu dan pengalamannya yang bermanfaat selama perkuliahan sebagai bekal pengetahuan.
9. Seluruh guru, staf tata usaha, dan siswa kelas X TMP 2, 3, dan 4 SMK Negeri 2 Bandung yang telah membantu penulis dalam melaksanakan penelitian skripsi.

10. Dimas Anom Priyayi, Sri Lestari, Krisna Milenia, dan Teh Nabila Amanina selaku rekan seperjuangan sidang Juli 2022.
11. Angela Putri Gunawan yang selalu memberikan semangat dan dukungan kepada penulis.
12. Ahmad Afzalulhaq dan Enggar Fahmy Al-Hikam yang telah menjadi sahabat perjuangan dan selalu memberikan semangat dan dukungan selama perkuliahan.
13. Seluruh rekan seperjuangan Pendidikan Ilmu Komputer 2018.
14. Semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini yang tidak dapat penulis sebutkan satu persatu.

Semoga Allah SWT memberikan balasan yang berlipat atas semua kebaikan dari berbagai pihak yang telah membantu penulis dalam menyelesaikan skripsi ini. Amiin.

Bandung, Juni 2022



Ali Akbar Jausyani

NIM. 1807105

PENERAPAN *SPLIT FEATURE REDUCTION* PADA *DECISION TREE C4.5*
UNTUK MENINGKATKAN PREDIKSI PEFORMA BELAJAR SISWA SMK

Oleh

Ali Akbar Jausyani – sansanakbar@upi.edu

1807105

ABSTRAK

Performa belajar siswa yang rendah dapat menyebabkan masalah-masalah besar seperti putus sekolah, maka perlu dilakukan pemantauan terhadap performa belajar siswa agar pendidik dapat mencegah siswa memiliki performa belajar yang rendah. Salah satu cara pemantauan performa belajar siswa adalah dengan cara prediksi. *Educational Data Mining* memiliki banyak teknik yang dapat digunakan dalam memproses data menjadi informasi, salah satunya adalah *Decision Tree C4.5*. *Decision Tree C4.5* merupakan salah satu teknik yang sering digunakan dari berbagai teknik *data mining* dengan salah satu alasannya yaitu mudah dipahami. Namun *Decision Tree C4.5* masih memiliki beberapa kekurangan seperti kemampuan prediksi yang rendah. *Split Feature Reduction* adalah salah satu teknik yang mencoba untuk memperbaiki kekurangan tersebut. Penggunaan teknik ini dapat dilakukan dengan cara mengurangi atribut-atribut siswa yang kurang signifikan dari proses pengolahan data. Maka, penelitian ini mencoba untuk meningkatkan kemampuan prediksi *Decision Tree C4.5* menggunakan teknik *Split Feature Reduction*. Pada penelitian yang dilakukan ini, ditemukan bahwa *Split Feature Reduction* meningkatkan nilai *F1-Score* prediksi *Decision Tree* dalam menentukan performa belajar siswa SMK sebesar 5,32%. Atribut-atribut siswa yang ditemukan berpengaruh terhadap performa belajar siswa SMK adalah kelengkapan tugas, minat terhadap jurusan kelas, pendidikan ayah, akses internet, pendidikan ibu, status ibu bekerja, dan akses komputer. Jadi, penerapan *Split Feature Reduction* benar dapat meningkatkan kemampuan *Decision Tree C4.5* dalam memprediksi performa belajar siswa SMK.

Kata Kunci: *Educational Data Mining*, *Decision Tree C4.5*, *Split Feature Reduction*, Prediksi, Performa Belajar

PENERAPAN *SPLIT FEATURE REDUCTION* PADA *DECISION TREE C4.5*
UNTUK MENINGKATKAN PREDIKSI PEFORMA BELAJAR SISWA SMK

Oleh

Ali Akbar Jausyani – sansanakbar@upi.edu

1807105

ABSTRACT

Low students' performance can lead to a bigger problems like dropout, that is why monitoring students' performance is necessary so educator can prevent students from having low performance. One of many ways to do this is to predict the students' performance. Educational Data Mining has a lot of techniques that can be used, one of the techniques is called Decision Tree. Decision Tree is a popular technique and mostly used because of how easy it is to use. However, Decision Tree still has problems such as low prediction performance. Split Feature Reduction is one of the techniques that can be used to try to fixed those problems. This technique can be used by reducing students' attributes that has low significance from the equation process so that the produced information is clear from insignificant attributes. So, in this research we tried to improve Decision Tree C4.5 prediction capabilities using Split Feature Reduction. In this research, Split Feature Reduction does improve Decision Tree prediction on students' performance with 5,32% difference in F1-Score. Students' attributes that were found to be significant to students' performance are assignments' completion, class major's interest, father's education, internet access, mother's education, working mother status, and computer access. So, implementing Split Feature Reduction into Decision Tree C4.5 does improve the prediction capabilities on students' performance.

Keywords: *Educational Data Mining, Decision Tree C4.5, Split Feature Reduction, Predictions, Students' Performance*

DAFTAR ISI

LEMBAR PERNYATAAN	iv
KATA PENGANTAR	v
UCAPAN TERIMA KASIH.....	vi
ABSTRAK	viii
<i>ABSTRACT</i>	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xv
DAFTAR RUMUS.....	xvi
DAFTAR LAMPIRAN.....	xvii
BAB I	78
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat/Signifikansi Penelitian	4
BAB II.....	6
2.1 Peta Literatur	6
2.2 Performa Belajar.....	6
2.3 <i>Decision Tree</i>	8
2.3.1 <i>Educational Data Mining</i>	8
2.3.2 Karakteristik <i>Decision Tree</i>	9
2.3.3 Algoritma C4.5.....	11
2.3.4 <i>Cross Validation</i>	14

Ali Akbar Jausyani, 2022

PENERAPAN SPLIT FEATURE REDUCTION PADA DECISION TREE C4.5 UNTUK MENINGKATKAN PREDIKSI PERFORMA BELAJAR SISWA SMK

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

2.4	<i>Split Feature Reduction</i>	15
2.5	<i>State of the Art</i>	17
2.5.1	Atribut Siswa yang Berpengaruh pada Performa Belajar	17
2.5.2	Perbandingan Performa Prediksi Teknik <i>Data Mining</i>	19
2.5.3	Pengembangan <i>Decision Tree C4.5</i>	20
BAB III.....		22
3.1	<i>Analyze</i>	23
3.1.1	Instrumen Penelitian.....	23
3.1.1.1	Soal Tes PAS	23
3.1.1.2	Angket.....	26
3.1.1.3	Pedoman Observasi.....	27
3.1.2	Populasi dan Sampel	27
3.2	<i>Design</i>	27
3.3	<i>Develop</i>	28
3.4	<i>Implement</i>	28
3.5	<i>Evaluate</i>	28
BAB IV		32
4.1	<i>Analyze</i>	32
4.1.1	<i>Data Pre-Processing</i>	32
4.2	<i>Design</i>	38
4.2.1	<i>Flowchart</i>	38
4.2.2	<i>Entity Relationship Diagram</i>	40
4.2.3	<i>Storyboard</i>	44
4.3	<i>Develop</i>	52
4.1.2	Pembuatan Model Prediksi <i>Decision Tree C4.5</i> pada Sistem.....	52

4.1.3.1	Model Prediksi <i>Decision Tree C4.5 Unpruned</i>	53
4.1.3.2	Model Prediksi <i>Decision Tree C4.5 Pruned</i>	55
4.4	<i>Implement</i>	56
4.2.1	Penerapan <i>Split Feature Reduction</i> pada <i>Decision Tree C4.5</i>	56
4.2.1.1	Model Prediksi <i>Decision Tree C4.5 Split 1</i>	57
4.2.1.2	Model Prediksi <i>Decision Tree C4.5 Split 2</i>	59
4.2.1.3	Model Prediksi <i>Decision Tree C4.5 Split 3</i>	60
4.2.2	Fungsionalitas Model Prediksi pada Sistem	61
4.5	<i>Evaluate</i>	72
BAB V		75
5.1	Kesimpulan.....	75
5.2	Saran.....	76
DAFTAR PUSTAKA		78
LAMPIRAN		82

DAFTAR GAMBAR

Gambar 2.1 Peta Literatur	6
Gambar 2.2 Tahap Umum <i>Data Mining</i>	9
Gambar 2.3 Tahap <i>Decision Tree C4.5</i>	13
Gambar 2.4 <i>Cross Validation</i>	15
Gambar 2.5 Tahap Penerapan <i>Split Feature Reduction</i> pada <i>Decision Tree C4.5</i>	16
Gambar 3.1 Prosedur Penelitian.....	22
Gambar 4.1 <i>Flowchart</i> Administrator pada Sistem	38
Gambar 4.2 <i>Flowchart</i> Guru pada Sistem.....	39
Gambar 4.3 <i>Flowchart</i> Siswa pada Sistem	40
Gambar 4.4 <i>Entity Relationship Diagram</i> Sistem.....	42
Gambar 4.4 Kode Fungsi <i>Convert Datasets to CSV</i>	52
Gambar 4.5 Kode Fungsi <i>Convert CSV to ARFF</i>	53
Gambar 4.6 Kode Fungsi Pembentukan & Visualisasi Model Prediksi	53
Gambar 4.7 Model Prediksi <i>Decision Tree C4.5 Unpruned</i>	54
Gambar 4.8 Model Prediksi <i>Decision Tree C4.5 Pruned</i>	56
Gambar 4.9 Model Prediksi <i>Decision Tree C4.5 Split 1</i>	58
Gambar 4.10 Model Prediksi <i>Decision Tree C4.5 Split 2</i>	60
Gambar 4.11 Model Prediksi <i>Decision Tree C4.5 Split 3</i>	61
Gambar 4.12 <i>Chart</i> Perbandingan Data Aktual dan Prediksi	65
Gambar 4.13 Pengujian <i>Decision Tree C4.5 Unpruned</i> pada WEKA.....	66
Gambar 4.14 Pengujian <i>Decision Tree C4.5 Unpruned</i> pada LMS.....	66
Gambar 4.15 Pengujian <i>Decision Tree C4.5 Pruned</i> pada WEKA	67
Gambar 4.16 Pengujian <i>Decision Tree C4.5 Pruned</i> pada LMS	67
Gambar 4.17 Pengujian <i>Decision Tree C4.5 Split 1</i> pada WEKA.....	68

Ali Akbar Jausyani, 2022

PENERAPAN SPLIT FEATURE REDUCTION PADA DECISION TREE C4.5 UNTUK MENINGKATKAN PREDIKSI PERFORMA BELAJAR SISWA SMK

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Gambar 4.18 Pengujian <i>Decision Tree C4.5 Split 1</i> pada LMS.....	68
Gambar 4.19 Pengujian <i>Decision Tree C4.5 Split 2</i> pada WEKA.....	69
Gambar 4.20 Pengujian <i>Decision Tree C4.5 Split 2</i> pada LMS.....	69
Gambar 4.21 Pengujian <i>Decision Tree C4.5 Split 3</i> pada WEKA.....	70
Gambar 4.22 Pengujian <i>Decision Tree C4.5 Split 3</i> pada LMS.....	70
Gambar 4.23 Notifikasi <i>Early Warning</i> pada LMS	71
Gambar 4.24 Alasan Prediksi Performa Belajar Siswa pada LMS	72

DAFTAR TABEL

Tabel 3.1 Kategori Validitas Instrumen Soal Tes PAS.....	24
Tabel 3.2 Kategori Reliabilitas Instrumen Soal Tes PAS	25
Tabel 3.3 Kategori Indeks Kesukaran Instrumen Soal Tes PAS	26
Tabel 3.4 Kategori Daya Pembeda Instrumen Soal Tes PAS	26
Tabel 3.5 <i>Confusion Matrix</i>	29
Tabel 4.1 Klasifikasi Atribut Data	34
Tabel 4.2 Atribut dan Nilai Data.....	34
Tabel 4.3 <i>Dataset</i> Siswa.....	36
Tabel 4.4 <i>Storyboard</i> Administrator pada Sistem.....	44
Tabel 4.5 <i>Storyboard</i> Guru pada Sistem.....	47
Tabel 4.6 <i>Storyboard</i> Siswa pada Sistem.....	49
Tabel 4.6 <i>Confusion Matrix Decision Tree C4.5 Unpruned Cross Validation</i>	55
Tabel 4.7 <i>Confusion Matrix Decision Tree C4.5 Pruned</i>	56
Tabel 4.8 Himpunan Bagian Atribut.....	57
Tabel 4.9 <i>Confusion Matrix Decision Tree C4.5 Split 1 Cross Validation</i>	59
Tabel 4.10 <i>Confusion Matrix Decision Tree C4.5 Split 2 Cross Validation</i>	60
Tabel 4.11 <i>Confusion Matrix Decision Tree C4.5 Split 3 Cross Validation</i>	61
Tabel 4.12 Aktual dan Hasil Prediksi	62
Tabel 4.13 <i>Confusion Matrix</i> Pengujian <i>Decision Tree C4.5 Unpruned</i>	65
Tabel 4.14 <i>Confusion Matrix</i> Pengujian <i>Decision Tree C4.5 Pruned</i>	66
Tabel 4.15 <i>Confusion Matrix</i> Pengujian <i>Decision Tree C4.5 Split 1</i>	67
Tabel 4.16 <i>Confusion Matrix</i> Pengujian <i>Decision Tree C4.5 Split 2</i>	69
Tabel 4.17 <i>Confusion Matrix</i> Pengujian <i>Decision Tree C4.5 Split 3</i>	70
Tabel 4.18 Evaluasi Model Prediksi	72

DAFTAR RUMUS

Rumus 2.1 Nilai <i>Entropy</i>	13
Rumus 2.2 Nilai <i>Gain</i>	13
Rumus 2.3 Nilai <i>Split Info</i>	13
Rumus 2.4 Nilai <i>Gain Ratio</i>	14
Rumus 3.1 Korelasi <i>Product Moment</i>	24
Rumus 3.2 <i>Spearman-Brown</i>	25
Rumus 3.3 Indeks Kesukaran.....	25
Rumus 3.4 Daya Pembeda	26
Rumus 3.5 Nilai <i>Accuracy</i> Model Prediksi.....	29
Rumus 3.6 Nilai <i>Precision (Positive)</i> Model Prediksi.....	30
Rumus 3.7 Nilai <i>Precision (Negative)</i> Model Prediksi.....	30
Rumus 3.8 Nilai <i>Recall (Positive)</i> Model Prediksi	30
Rumus 3.9 Nilai <i>Recall (Negative)</i> Model Prediksi.....	30
Rumus 3.10 Nilai <i>F1-Score (Positive)</i> Model Prediksi	30
Rumus 3.11 Nilai <i>F1-Score (Negative)</i> Model Prediksi	30
Rumus 3.12 Nilai <i>Error-Rate</i> Model Prediksi	31

DAFTAR LAMPIRAN

Lampiran 1	Soal Tes PAS	82
Lampiran 2	Hasil PAS Siswa	92
Lampiran 3	Validasi Guru Soal PAS	95
Lampiran 4	Uji Validitas Butir Soal PAS	109
Lampiran 5	Uji Validitas Nilai PAS terhadap Nilai PTS.....	111
Lampiran 6	Uji Reliabilitas.....	113
Lampiran 7	Uji Tingkat Kesukaran	115
Lampiran 8	Uji Daya Pembeda.....	117
Lampiran 9	Instrumen Angket Data Diri Siswa.....	119
Lampiran 10	Pedoman Observasi	122

DAFTAR PUSTAKA

- Abdullah, P. M. (2015). Metodologi Penelitian Kuantitatif. In *Aswaja Pressindo*.
- Abu Zohair, L. M. (2019). Prediction of Student's performance by modelling small dataset size. *International Journal of Educational Technology in Higher Education*, 16(1).
<https://doi.org/10.1186/s41239-019-0160-3>
- Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37(January), 13–49.
<https://doi.org/10.1016/j.tele.2019.01.007>
- Arifin, Z. (2017). Kriteria Instrumen dalam Suatu Penelitian. *Jurnal THEOREMS (The Original Research of Mathematics)*, 2(1), 28–36.
- Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers and Education*, 113, 177–194. <https://doi.org/10.1016/j.compedu.2017.05.007>
- Berrar, D. (2018). Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology*, 1, 542–545.
- Charbuty, B., & Abdulazeez, A. (2021). Classification Based on Decision Tree Algorithm for Machine Learning. *Journal of Applied Science and Technology Trends*, 2(01), 20–28. <https://doi.org/10.38094/jastt20165>
- Chen, F., Li, X., & Liu, L. (2013). *Improved C4.5 Decision Tree Algorithm Based on Sample Selection*. 779–782.
- Cherfi, A., Nouria, K., & Ferchichi, A. (2018). Very Fast C4.5 Decision Tree Algorithm. *Applied Artificial Intelligence*, 32(2), 119–137.
<https://doi.org/10.1080/08839514.2018.1447479>
- Czibula, G., Mihai, A., & Crivei, L. M. (2019). S PRAR: A novel relational association rule mining classification model applied for academic performance prediction. *Procedia Computer Science*, 159, 20–29. <https://doi.org/10.1016/j.procs.2019.09.156>
- Dabhade, P., Agarwal, R., Alameen, K. P., Fathima, A. T., Sridharan, R., & Gopakumar, G. (2021). Educational data mining for predicting students' academic performance using machine learning algorithms. *Materials Today: Proceedings*, 47, 5260–5267.
<https://doi.org/10.1016/j.matpr.2021.05.646>
- Fernandes, E., Holanda, M., Victorino, M., Borges, V., Carvalho, R., & Erven, G. Van. (2019). Educational data mining: Predictive analysis of academic performance of

- public school students in the capital of Brazil. *Journal of Business Research*, 94, 335–343. <https://doi.org/10.1016/j.jbusres.2018.02.012>
- Hamsa, H., Indiradevi, S., & Kizhakkethottam, J. J. (2016). Student Academic Performance Prediction Model Using Decision Tree and Fuzzy Genetic Algorithm. *Procedia Technology*, 25, 326–332. <https://doi.org/10.1016/j.protcy.2016.08.114>
- Kaur, P., Singh, M., & Josan, G. S. (2015). Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector. *Procedia Computer Science*, 57, 500–508. <https://doi.org/10.1016/j.procs.2015.07.372>
- Khan, A., & Ghosh, S. K. (2021). Student performance analysis and prediction in classroom learning: A review of educational data mining studies. In *Education and Information Technologies* (Vol. 26, Issue 1). Education and Information Technologies. <https://doi.org/10.1007/s10639-020-10230-3>
- Križanić, S. (2020). Educational data mining using cluster analysis and decision tree technique: A case study. *International Journal of Engineering Business Management*, 12, 1–9. <https://doi.org/10.1177/1847979020908675>
- Kumar, V., & Minz, S. (2014). Feature Selection: A literature Review. *The Smart Computing Review*, 4(3). <https://doi.org/10.6029/smarter.2014.03.007>
- Li, J., Cheng, K., Wang, S., Morstatter, F., Tang, J., & Liu, H. (2017). Feature selection: A Data Perspective. *Intelligent Systems Reference Library*, 50(6), 1–45. https://doi.org/10.1007/978-3-319-10247-4_7
- Malini, J., & Kalpana, Y. (2021). Investigation of factors affecting student performance evaluation using education materials data mining technique. *Materials Today: Proceedings*. <https://doi.org/10.1016/j.matpr.2021.05.026>
- Matzavela, V., & Alepis, E. (2021a). Decision tree learning through a Predictive Model for Student Academic Performance in Intelligent M-Learning environments. *Computers and Education: Artificial Intelligence*, 2, 100035. <https://doi.org/10.1016/j.caeai.2021.100035>
- Matzavela, V., & Alepis, E. (2021b). Decision tree learning through a Predictive Model for Student Academic Performance in Intelligent M-Learning environments. *Computers and Education: Artificial Intelligence*, 2, 100035. <https://doi.org/10.1016/j.caeai.2021.100035>
- Maulana, M. F., & Defriani, M. (2020). Logistic Model Tree and Decision Tree J48 Algorithms for Predicting the Length of Study Period. *PIKSEL : Penelitian Ilmu*

- Komputer Sistem Embedded and Logic*, 8(1), 39–48.
<https://doi.org/10.33558/piksel.v8i1.2018>
- Mengash, H. A. (2020). Using data mining techniques to predict student performance to support decision making in university admission systems. *IEEE Access*, 8, 55462–55470. <https://doi.org/10.1109/ACCESS.2020.2981905>
- Mienye, I. D., Sun, Y., & Wang, Z. (2019). Prediction performance of improved decision tree-based algorithms: A review. *Procedia Manufacturing*, 35, 698–703.
<https://doi.org/10.1016/j.promfg.2019.06.011>
- Muslim, M. A., Rukmana, S. H., Sugiharti, E., Prasetyo, B., & Alimah, S. (2017). Optimization of C4.5 algorithm-based particle swarm optimization for breast cancer diagnosis. *Journal of Physics: Conference Series*, 983(1).
<https://doi.org/10.1088/1742-6596/983/1/012063>
- Muslim, Much Aziz, Nurzahputra, A., & Prasetyo, B. (2018). Improving accuracy of C4.5 algorithm using split feature reduction model and bagging ensemble for credit card risk prediction. *2018 International Conference on Information and Communications Technology, ICOIACT 2018, 2018-Janua(1996)*, 141–145.
<https://doi.org/10.1109/ICOIACT.2018.8350753>
- Namoun, A., & Alshanqiti, A. (2021). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences (Switzerland)*, 11(1), 1–28. <https://doi.org/10.3390/app11010237>
- Okfalisa, Budianita, E., Yuliani, R., Suanmali, L., Megawati, Rusnedy, H., & Saktioto. (2022). Forecasting company financial distress: C4.5 and adaboost adoption. *Engineering and Applied Science Research*, 49(3), 300–307.
<https://doi.org/10.14456/easr.2022.31>
- Patil, V., Suryawanshi, S., Saner, M., & Patil, V. (2017). Student Performance Prediction Using Classification Data Mining Techniques. *International Journal for Research in Emerging Science and Technology*, 4(4), 15–18.
- Priyam, A., Gupta, R., Rathee, A., & Srivastava, S. (2013). *Comparative Analysis of Decision Tree Classification Algorithms*. 334–337.
- Rosmansyah, Y., Putro, B. L., Putri, A., Utomo, N. B., & Suhardi. (2022). A simple model of smart learning environment. *Interactive Learning Environments*.
<https://doi.org/10.1080/10494820.2021.2020295>
- Sari, B. N. (2016). Implementasi Teknik Seleksi Fitur Information Gain Pada Algoritma

- Klasifikasi Machine Learning Untuk Prediksi Performa Akademik Siswa. *Seminar Nasional Teknologi Informasi Dan Multimedia 2016, March*, 55–60.
<http://semnas.amikom.ac.id/document/pdf/1482.pdf>
- Spencer, D., & Temple, T. (2021). Examining students' online course perceptions and comparing student performance outcomes in online and face-to-face classrooms. *Online Learning Journal*, 25(2), 233–261. <https://doi.org/10.24059/olj.v25i2.2227>
- Sukmadinata, P. D. N. S. (2012). *Metode Penelitian Pendidikan* (8th ed.). PT Remaja Rosdakarya.
- Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers and Education*, 143. <https://doi.org/10.1016/j.compedu.2019.103676>
- Yuan, Z., & Wang, C. (2016). *An Improved Network Traffic Classification Algorithm Based on Hadoop Decision Tree*. 53–56.
- Yuliansyah, H., Imaniati, R. A. P., Wirasto, A., & Wibowo, M. (2021). Predicting Students Graduate on Time Using C4.5 Algorithm. *Journal of Information Systems Engineering and Business Intelligence*, 7(1), 67.
<https://doi.org/10.20473/jisebi.7.1.67-73>
- Yulianto, L. D., Triayudi, A., & Sholihati, I. D. (2020). Implementation Educational Data Mining For Analysis of Student Performance Prediction with Comparison of K-Nearest Neighbor Data Mining Method and Decision Tree C4.5. *Jurnal Mantik*, 4(1), 441–451.
- Yusup, F. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Tarbiyah : Jurnal Ilmiah Kependidikan*, 7(1), 17–23.
<https://doi.org/10.18592/tarbiyah.v7i1.2100>
- Zebari, R., Abdulazeez, A., Zeebaree, D., Zebari, D., & Saeed, J. (2020). A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction. *Journal of Applied Science and Technology Trends*, 1(2), 56–70. <https://doi.org/10.38094/jastt1224>
- Zurada, J. (2010). Does Feature Reduction Help Improve the Classification Accuracy Rates? A Credit Scoring Case Using a German Data Set. *Review of Business Information Systems (RBIS)*, 14(2), 35–40. <https://doi.org/10.19030/rbis.v14i2.496>