

PENERAPAN DATA MINING DAN PEMETAAN UNTUK STRATEGI PENCAPAIAN MUTU PENDIDIKAN SMA DI KALIMANTAN SELATAN

(Application of Data Mining and Mapping for the Strategy to Achieving the Quality of High School Education in South Kalimantan)

Desy Ika Puspitasari^{1*}, Al Fath Riza Kholdani²⁾, Tri Wahyu Qur'ana³⁾, Mirza Yogy Kurniawan⁴⁾

^{1,2,3,4)}Universitas Islam Kalimantan MAB Banjarmasin

e-mail: smile4desyka@gmail.com¹, kholdanialfath@gmail.com², twqurana@gmail.com³, mirza@fti.uniska-bjm.ac.id⁴

**) penulis korespondensi*

Abstract. Indonesia is still mapping the quality of its education system to improve the quality of the country's education. The SNP value classification has been carried out by the Education Quality Assurance Institute (LPMP) into two categories: Strengths and Weaknesses. Senior high schools that have not achieved SNP 4 scores are classified as weak, while those that have achieved SNP 4 and SNP are classified as strong. This study uses data mining and mapping approaches to organize the SNP values of high school (SMA) in South Kalimantan district. The study's classification results showed that the Random Forest approach had the highest accuracy, with a precision of 0.969 for neat data and an accuracy of 0.925 for missing value data. Based on the mapping visualization, the differences fluctuate between districts, which occur in the categories of availability of labor, laboratories, and librarians. It is hoped that the Provincial Government, through the Education Office of South Kalimantan Province, will provide competent laboratories, libraries, and librarians to encourage the National Education Standards (SNP). This study provides information about the quality of education through the SNP effectiveness approach.

Keywords: classification, data mining, education, mapping.

1. Pendahuluan

Proses pemetaan mutu Pendidikan (PMP) di Indonesia sampai saat ini masih terus berlangsung. Pemetaan mutu bertujuan untuk melihat bagaimana mutu pendidikan dari sekolah dengan beberapa instrument PMP yang di input dalam aplikasi PMP bersamaan dengan pengisian data sekolah di aplikasi Dapodik [13], [10]. Jenjang SMA menjadi tahapan yang sangat vital bagi seorang siswa dalam menentukan masa depannya, karenanya para siswa harus mendapatkan pelayanan yang baik agar bisa mencapai standar yang diinginkan. Jumlah sekolah jenjang SMA yang terpetakan pada tahun 2018 sebanyak 12.682 sekolah atau sekitar 85,90% dari sebanyak 14.762 sekolah, sedangkan sebanyak 2.082 sekolah atau 14,10% belum terpetakan karena data tidak tersedia. Klasifikasi nilai standar/indikator ini dibagi menjadi 2 kategori yaitu nilai Kekuatan dan Kelemahan. Ruang lingkup penyusunan rekomendasi peningkatan mutu pendidikan difokuskan pada indikator yang lemah. Rekomendasi dirancang meliputi berbagai fokus indikator lemah, kegiatan, sasaran, jumlah, waktu pelaksanaan, anggaran, sumber dana, penanggung jawab, dan jejaring. Standar kompetensi lulusan SMA tahun 2018 sebesar

6,39, yaitu masuk pada tahapan menuju SNP level 4 dan terbilang cukup baik. Capaian SNP yang kurang baik, yaitu pada standar pendidik dan tenaga kependidikan sebesar 3,54 (menuju SNP level 2) serta standar sarana dan prasarana sebesar 4,39 (menuju SNP level 3) [3].

Tabel 1. Kategori Capaian Nilai (Sumber: LPMP Provinsi Kalsel Tahun 2019)

No	Kategori	Rentang Nilai	
		Rendah	Tinggi
1	Menuju SNP 1	0.00	2.07
2	Menuju SNP 2	2.08	3.70
3	Menuju SNP 3	3.71	5.06
4	Menuju SNP 4	5.07	6.66
5	SNP	6.67	7.00

Kategori capaian nilai SNP ditunjukkan oleh tabel 1, analisis kekuatan dan kelemahan berdasarkan capaian pada kategori Menuju SNP 4. Standar/Indikator yang belum mencapai Menuju SNP 4 dikategorikan lemah, sedangkan yang telah mencapai kategori Menuju SNP 4 dan SNP dikategorikan kuat. Penelitian ini bertujuan mengklasifikasi mutu Pendidikan berdasarkan kategori SNP untuk mengetahui pola pemetaannya berdasarkan kategori SNP jenjang SMA sederajat di wilayah Kalsel dengan pengolahan data mining. Ruang lingkup analisis data adalah data yang telah diolah menjadi peta mutu, selanjutnya pemetaan berdasarkan kategori ini diharapkan dapat dimanfaatkan untuk mendorong sekolah guna meningkatkan mutu sebagai persiapan menghadapi akreditasi atau sistem penjaminan mutu eksternal.

Telah banyak penelitian yang menggunakan data mining dalam membantu perusahaan atau Lembaga dengan memanfaatkan keterampilan pelaporan yang ada guna menemukan dan memahami pola yang tersembunyi dalam kumpulan data yang besar [3],[5],[7],[16]. Informasi yang berguna dari data mining dapat digunakan untuk meningkatkan kualitas dan cara penyampaian sistem Pendidikan tinggi [2].

2. Metodologi

2.1 Data Mining untuk Pemetaan dan Prediksi

Data mining untuk pemetaan dan prediksi merupakan data mining (DM) atau secara umum merupakan lingkup *machine learning* (ML) yaitu teknik atau cara-cara guna memprediksi sebuah keadaan berdasarkan pemodelan terhadap suatu data [4] yang selanjutnya dilakukan pemetaan untuk mengetahui potensi suatu wilayah [9].

Salah satu algoritma yang memiliki keuntungan sebagai pengklasifikasi berbasis pohon keputusan yaitu random forest, yang memilih pohon klasifikasi terbaik sebagai pengklasifikasi akhir algoritma, tidak memerlukan validasi silang, dan sepenuhnya dapat diparalelkan [1]. Pengklasifikasi decision tree juga menghasilkan akurasi prediksi yang tinggi [11],[12],[15] jika dibandingkan dengan algoritma lain (seperti KNN, LR, SVM,

dan NB) [6]. SVM juga merupakan algoritma yang berhasil meningkatkan nilai akurasi pada penelitian mengenai klasifikasi [8], [14].

2.2. Proses pengolahan data (Pre-processing data)

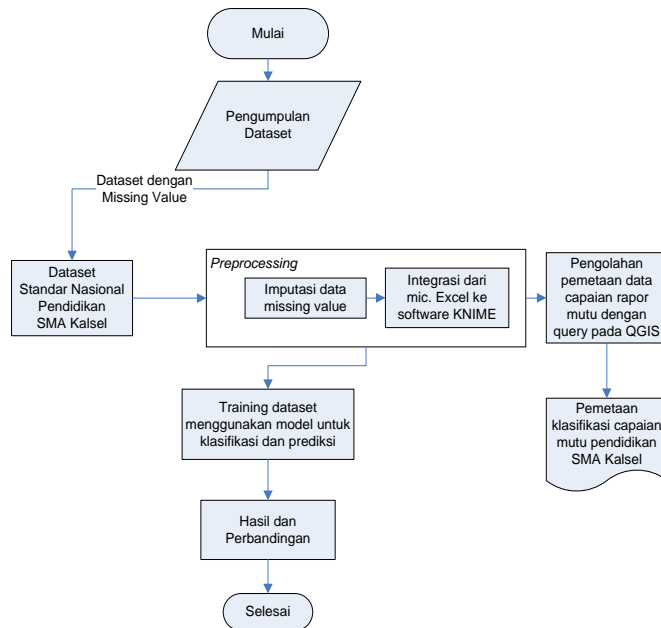
Data yang digunakan dalam penelitian ini ialah data capaian Standar Nasional Pendidikan (SNP/indikator) dengan beberapa kriteria sebagai berikut:

Variabel Indikator Standar Nasional Pendidikan		
Variabel Laten	Indikator	Keterangan
Kompetensi Lulusan	SUB_1_1	Lulusan memiliki kompetensi pada dimensi sikap
	SUB_1_2	Lulusan memiliki kompetensi pada dimensi pengetahuan
	SUB_1_3	Lulusan memiliki kompetensi pada dimensi keterampilan
Isi Pendidikan	SUB_2_1	Perangkat pembelajaran sesuai rumusan kompetensi lulusan
	SUB_2_2	Kurikulum Tingkat Satuan Pendidikan dikembangkan sesuai prosedur
	SUB_2_3	Sekolah melaksanakan kurikulum sesuai ketentuan
Proses Pembelajaran	SUB_3_1	Sekolah merencanakan proses pembelajaran sesuai ketentuan
	SUB_3_2	Proses pembelajaran dilaksanakan dengan tepat
	SUB_3_3	Pengawasan dan penilaian otentik dilakukan dalam proses pembelajaran
Penilaian Pendidikan	SUB_4_1	Aspek penilaian sesuai ranah kompetensi
	SUB_4_2	Teknik penilaian obyektif dan akuntabel
	SUB_4_3	Penilaian pendidikan ditindak lanjuti
	SUB_4_4	Instrumen penilaian menyesuaikan aspek
	SUB_4_5	Penilaian dilakukan mengikuti prosedur
Pendidik dan Tenaga Kependidikan	SUB_5_1	Ketersediaan dan kompetensi guru sesuai ketentuan
	SUB_5_2	Ketersediaan dan kompetensi kepala sekolah sesuai ketentuan
	SUB_5_3	Ketersediaan dan kompetensi tenaga administrasi sesuai ketentuan
	SUB_5_4	Ketersediaan dan kompetensi laboran sesuai ketentuan
	SUB_5_5	Ketersediaan dan kompetensi pustakawan sesuai ketentuan
Sarana dan Prasarana	SUB_6_1	Kapasitas daya tampung sekolah memadai
	SUB_6_2	Sekolah memiliki sarana dan prasarana pembelajaran yang lengkap dan layak
	SUB_6_3	Sekolah memiliki sarana dan prasarana pendukung yang lengkap dan layak
Pembiayaan	SUB_8_1	Sekolah memberikan layanan subsidi silang
	SUB_8_2	Beban operasional sekolah sesuai ketentuan
	SUB_8_3	Sekolah melakukan pengelolaan dana dengan baik
Pengelolaan	SUB_7_1	Sekolah melakukan perencanaan pengelolaan
	SUB_7_2	Program pengelolaan dilaksanakan sesuai ketentuan
	SUB_7_3	Kepala sekolah berkinerja baik dalam melaksanakan tugas kepemimpinan
	SUB_7_4	Sekolah mengelola sistem informasi manajemen

Gambar 1. Variabel Indikator Standar Nasional Pendidikan (Sumber: Anam *et all.*,[3])

Setiap standar nasional Pendidikan memiliki beberapa indikator dan sub indikator. Analisis pola hubungan Standar Nasional Pendidikan SMA di Indonesia dilakukan dengan menggunakan variabel indikator SNP sebagaimana Gambar 1.

Dataset capaian rapor mutu SMA wilayah Kalimantan Selatan pada tahun 2019 mengalami missing value. Missing value merupakan data yang tidak lengkap atau nilai yang hilang dalam kumpulan data. Hal ini dapat menyebabkan keputusan yang salah yang perlu ditangani dengan lebih hati-hati (fokus) untuk memastikan pemilihan fitur dan pengambilan keputusan yang akurat dan tepat [1].



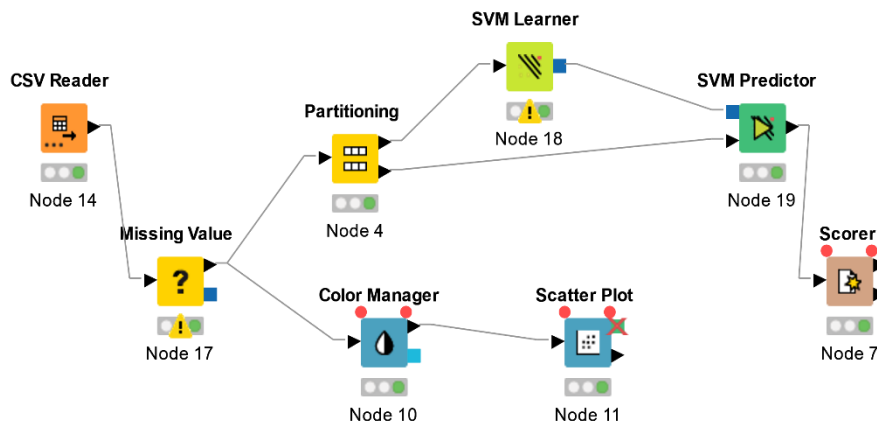
Gambar 2. Alur Metodologi Penelitian

Pengolahan data dilakukan melalui tahap preprocessing terlebih dahulu. Tahapan preprocessing ini yaitu imputasi data missing value menjadi data rapi kemudian diintegrasikan dari Microsoft Excel ke software data mining KNIME.

Data yang telah melalui tahap preprocessing ini dapat diolah dalam pemetaan menggunakan QGIS. Selanjutnya pada kedua dataset dilakukan training menggunakan algoritma klasifikasi, kemudian dilakukan perbandingan hasil akurasi terhadap data missing value dan data rapi.

2.3. Desain Eksperimen

Berikut ini merupakan desain eksperimen penelitian klasifikasi mutu pendidikan berdasarkan nilai SNP menggunakan software Knime:



Gambar 3. Desain Eksperimen Menggunakan Node Algoritma Klasifikasi pada Knime

Pada Gambar 3, merupakan desain eksperimen menggunakan layout node repository pada software Knime yang diimplementasikan dengan algoritma Support Vector Machine (SVM), selanjutnya Random Forest dan Decision Tree untuk mengetahui hasil akurasi klasifikasi. Pada bagian partisi, diuji beberapa parameter data training dan testing. Pengukuran *performance* algoritma dilakukan dengan menggunakan nilai akurasi prediksi, presisi, recall dan f-measure.

3. Hasil dan Pembahasan

Berdasarkan nilai True Negative (TN), False Positive (FP), False Negative (FN), dan True Positive (TP) dapat diperoleh nilai akurasi, presisi recall.

- Akurasi

Merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data (kedua positive dan negative).

$$Accuracy = (TP+TN) / (TP+FP+FN+TN) \tag{1}$$

- Presisi

Menunjukkan jumlah data kategori positif yang diklasifikasikan benar dibagi total data yang diklasifikasikan positif.

$$Precision = (TP) / (TP+FP) \tag{2}$$

- Recall

Menghasilkan berapa rasio data kategori positif yang terklasifikasikan dengan benar oleh sistem.

$$Recall = (TP) / (TP+FN) \tag{3}$$

- F-measure

Sebuah metrik guna mengukur rata-rata presisi and recall untuk menilai kinerja seluruh sistem.

$$F-Measure = 2*(Recall*Precision) / (Recall+Precision) \tag{4}$$

3.1. Perbandingan hasil akurasi

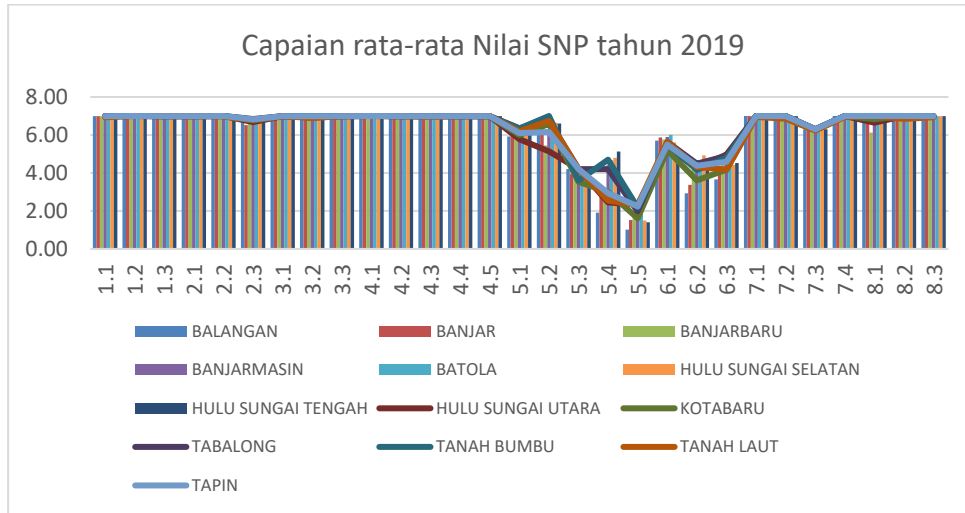
Setelah dilakukan training dan testing pada software Knime, maka dihitung score akurasi, presisi, recall, f-measure, pada masing-masing algoritma, menghasilkan Tabel 2.

Tabel 2. Hasil Perbandingan Akurasi

	SVM		Random Forest		Decision Tree	
	Missing value	Data rapi	Missing value	Data rapi	Missing value	Data rapi
Akurasi	0.903	0.969	0.925	0.969	0.825	0.969
Presisi	0.88	0.952	0.833	0.957	0.75	0.917
Recall	1	1	1	1	1	0.95
F-measure	0.936	0.976	0.909	0.667	0.875	0.957

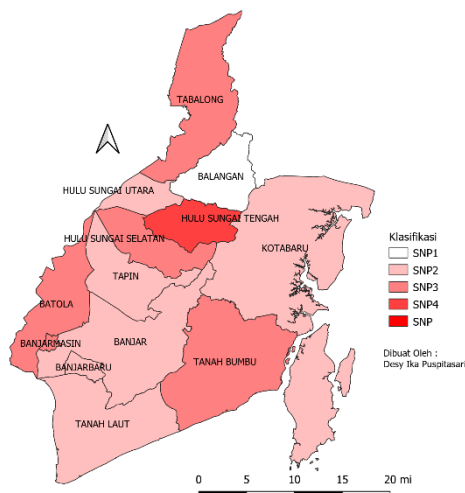
Berdasarkan Tabel 2 hasil perbandingan akurasi tersebut menunjukkan bahwa akurasi dengan nilai tertinggi berada pada data rapi dengan nilai 0.969 menggunakan ketiga algoritma SVM, random forest dan decision tree. Akurasi tertinggi dengan data missing value diperoleh 0.925 menggunakan algoritma random forest.

3.2. Hasil Capaian Rata-rata Nilai SNP pada tahun 2019

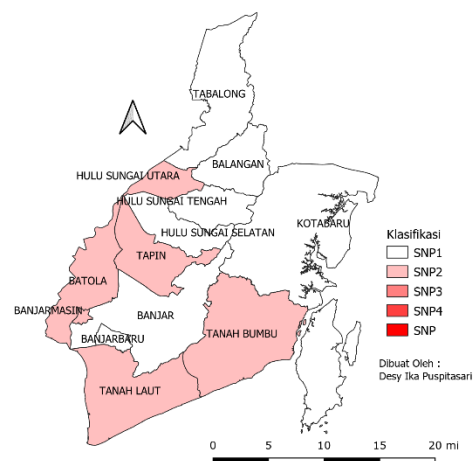


Gambar 4. Visualisasi Rata-rata Nilai Capaian SNP tahun 2019

Gambar 4 merupakan visualisasi rata-rata nilai capaian SNP tahun 2019 jenjang SMA pada provinsi Kalsel, terdapat perbedaan nilai capaian yang fluktuatif antar kabupaten yaitu pada kategori 5.4 (Ketersediaan dan kompetensi laboran sesuai ketentuan) dan kategori 5.5 (Ketersediaan dan kompetensi pustakawan sesuai ketentuan). Berikut ini hasil pemetaan berdasarkan klasifikasi capaian SNP yang merujuk pada kategori 5.4 dan kategori 5.5:



Gambar 5 Pemetaan Klasifikasi Capaian SNP kategori 5.4



Gambar 6 Pemetaan Klasifikasi Capaian SNP kategori 5.5

Berdasarkan pemetaan klasifikasi capaian nilai SNP pada Gambar 5 dan 6, warna merah gelap menunjukkan hasil kabupaten yang masuk dalam kategori menuju SNP 4 dan SNP dengan nilai Kuat. Warna merah terang sampai putih menunjukkan kategori SNP 1 sampai SNP 3 dengan nilai Lemah/rendah. Pada kategori 5.4 (ketersediaan dan kompetensi laboran) di kabupaten Hulu Sungai Tengah paling tinggi dibandingkan SMA di kabupaten lainnya. Hasil lebih lengkap ilustrasi peta diatas dapat dilihat di tabel di bawah ini:

Tabel 3 Hasil klasifikasi Nilai SNP jenjang SMA di Kalsel pada tahun 2019

No	Kabupaten	Capaian Kategori 5.4					Capaian Kategori 5.5				
		SNP 1	SNP 2	SNP 3	SNP 3	SNP 4	SNP 1	SNP 2	SNP 3	SNP 3	SNP 4
1	Hulu Sungai Tengah					✓	✓				
2	Hulu Sungai Utara		✓					✓			
3	Balangan	✓					✓				
4	Tapin		✓					✓			
5	Banjar		✓				✓				
6	Tanah Laut		✓					✓			
7	Banjarmasin			✓				✓			
8	Banjarbaru		✓				✓				
9	Batola (Barito Kuala)			✓				✓			
10	Tanah Bumbu			✓				✓			
11	Hulu Sungai Selatan			✓			✓				
12	Kotabaru		✓				✓				
13	Tabalong			✓			✓				
Total		1	6	5		1	7	6			

Berdasarkan Tabel 3 dapat dilihat bahwa pada capaian kategori 5.4 jumlah kabupaten yang masuk ke dalam kategori SNP 2 dan SNP 3 lebih banyak dibanding kabupaten yang mencapai nilai SNP 4, sedangkan pada capaian nilai kategori 5.5 (ketersediaan dan kompetensi pustakawan) semua kabupaten berada pada nilai rendah, bahkan ada tujuh kabupaten pada level SNP1.

4. Kesimpulan

Hasil analisis menggunakan algoritma SVM, Decision Tree, dan Random Forest menunjukkan bahwa jumlah klasifikasi dengan akurasi tertinggi yaitu algoritma Random Forest 0.969 pada data rapih, dan 0.925 pada data missing value. Pemetaan klasifikasi ini dapat di dimanfaatkan untuk memantau nilai capaian SNP pada tahun berikutnya, apakah dapat naik nilainya ke kategori Kuat atau justru sebaliknya. Melihat capaian yang menurun terjadi pada kategori 5.4 dan kategori 5.5, diharapkan pemerintah provinsi melalui Dinas Pendidikan Provinsi Kalimantan Selatan agar menyediakan tenaga laboran dan kompetensi laboran sesuai ketentuan serta menyediakan pustakawan yang kompeten

agar tercapai Standar Nasional Pendidikan (SNP). Penelitian ini memberikan informasi, dalam pencapaian mutu Pendidikan yang efektif diperlukan kejelasan mengenai sasaran mutu Pendidikan melalui pemetaan capaian nilai SNP berupa klasifikasi capaian nilai SNP.

Ucapan Terima Kasih

Ucapan terima kasih penulis berikan kepada LPPM Universitas Islam Kalimantan Selatan (UNISKA) MAB Banjarmasin atas bantuan dana penelitian tahun anggaran 2021/2022. Tak lupa juga kami ucapkan terima kasih kepada mahasiswa Teknik Informatika FTI UNISKA; Yudha Khalik, Fitriana dan Alya atas bantuannya dalam proses penelitian ini.

Daftar Pustaka

- [1] Abdulkareem, N.M. & Abdulazeez, A.M., (2021), Machine learning classification based on Radom Forest Algorithm: A review, *J. Sci. Bus.*, **5(2)**, 128–142. doi: 10.5281/zenodo.4471118
- [2] Adekitan, A.I. & Noma-Osaghae, E., (2019), Data Mining Approach to Predicting the Performance of First Year Student in a University Using the Admission Requirements, *Educ. Inf. Technol.*, **24(2)**, 1527–1543. doi: 10.1007/s10639-018-9839-7
- [3] Anam, A.K., Tuti, A., Ratnasari, V., (2020), Klasterisasi Mutu Pendidikan SMA di Indonesia, *J. Sains dan Seni ITS*, **9(2)**, 1–4.
- [4] Anwar, M.T. & Permana, D.R.A., (2021), Perbandingan Performa Model Data Mining untuk Prediksi Dropout Mahasiswa, *J. Teknol. dan Manaj.*, **19(2)**, 33–40. doi: 10.52330/jtm.v19i2.34
- [5] Arcinas, M.M., Sajja, G.S., Asif, S., Gour, S., Okoronkwo, E., Naved, M., (2021), Role of Data Mining in Education for Improving Students Performance for Social Change, *Turkish J. Physiother. Rehabil.*, **32(3)**, 6519–6526.
- [6] Charbuty, B. & Abdulazeez, A., (2021), Classification Based on Decision Tree Algorithm for Machine Learning, *J. Appl. Sci. Technol. Trends*, **2(1)**, 20–28. doi: 10.38094/jastt20165
- [7] Fatimah, K., (2018), Pengelompokan Mutu Pendidikan SMP dan Pengaruh Indikator Standar Nasional Pendidikan Terhadap Mutu Kelulusan SMP Negeri di Jawa Timur, *Skripsi*, Institut Teknologi Sepuluh November, Surabaya.
- [8] Irmanda, H.N. & Astriratma, R., (2020), Klasifikasi Jenis Pantun Dengan Metode Support Vector Machines (SVM), *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, **4(5)**, 915–922. doi: 10.29207/resti.v4i5.2313

- [9] Karim, A., *et al.*, (2021), Pemetaan untuk Strategi Dakwah di Kota Semarang Menggunakan Pendekatan Data Mining (Mapping for Da'wah Strategy in Semarang City Using Data Mining Approach), *J. Dakwah Risal.*, **32(1)**, 40-55. doi: 10.24014/jdr.v32i1.12549
- [10] MY, N., (2019), Meraih Mutu Sekolah Melalui Data Mutu: Suatu Tinjauan, *J. Ilm. Univ. Batanghari Jambi*, **19(2)**, 347-352. doi: 10.33087/jiubj.v19i2.677
- [11] Patro, C. & Pan, I., (2021), Decision Tree-Based Classification Model to Predict Student Employability, *Proceedings of Research and Applications in Artificial Intelligence, Advances in Intelligent Systems and Computing*, **1355**. https://doi.org/10.1007/978-981-16-1543-6_32
- [12] Puspitasari, D.I., Riza Kholdani, A.F., Dharmawati, A., Rosadi, M.E., Mega Pradnya Duhita, W., (2021), Stroke Disease Analysis and Classification Using Decision Tree and Random Forest Methods, *Sixth International Conference on Informatics and Computing (ICIC) 2021*, 1-4. doi: 10.1109/ICIC54025.2021.9632906
- [13] Susilo, T.T., (2018), Pemanfaatan Rapor Mutu Untuk Peningkatan Mutu Pendidikan, *Lembaga Penjaminan Mutu Pendidikan Kalimantan Timur*. [Online]. <https://lpmpkaltim.kemdikbud.go.id/2018/08/pemanfaatan-rapor-mutu-untuk-peningkatan-mutu-pendidikan/>
- [14] Utami, D., Kusriani, Sudarmawan, (2021), Analisis Metode Wordtovec Pada Performa Klasifikasi Kategori Berita Menggunakan Metode SVM dan KNN, *J. Explor.*, **11(2)**, 29-34.
- [15] Wang, Y., (2021), Network Education Evaluation System and Its Design Based on Decision Tree Classification Algorithm, *4th International Conference on Information Systems and Computer Aided Education 2021*, 511-515. doi: 10.1145/3482632.3482741
- [16] Zahra, C.A., (2017), Pengaruh Pengelolaan Sekolah Terhadap Kompetensi Lulusan Berdasarkan Standar Nasional Pendidikan Di Sekolah Menengah Pertama Surabaya Menggunakan Metode Structural Equation Modelling-Partial Least Square, *Skripsi*, Institut Teknologi Sepuluh Nopember, Surabaya.