

## Edukasi dan Pelatihan Data Science dan Data Preprocessing

Ibnu Mansyur Hamdani<sup>1</sup>, Nurhidayat<sup>2</sup>, Andi Karman<sup>3</sup>, Nurul Fuady Adhalia H<sup>4</sup>, A. Hermina Julyaningsih<sup>5</sup>

<sup>1</sup>Akademi Komunitas Industri Manufaktur Bantaeng; ibnumansyur@akom-bantaeng.ac.id

<sup>2</sup>Institut Teknologi dan Sains Muhammadiyah Kolaka Utara; nurhidayat@intensmku.ac.id

<sup>3</sup>Universitas Cokroaminoto Palopo; andikarman1@gmail.com

<sup>4</sup>Institut Teknologi Bacharuddin Jusuf Habibie; nurulfuady@ith.ac.id

<sup>5</sup>Universitas Hasanuddin; julyainjuly@gmail.com

Article Info	Abstract
<p><b>Keywords:</b> data science; education; data preprocessing</p> <p><b>Kata Kunci:</b> data science; adukasi; data preprocessing</p> <p><b>Article History</b> Received: 2024-06-05 Reviewed: 2024-06-06 Accepted: 2024-06-18</p>	<p><i>Community service is one form of the tridharma of higher education which aims to make a real contribution to improving the quality of human resources. The Informatics Engineering Study Program, Muhammadiyah Institute of Technology and Science, North Kolaka, carried out an online community service activity via Zoom. This activity, which focused on technology, raised the theme of Data Science and Data Preprocessing, relevant to the Big Data course. The speakers came from the Bantaeng Manufacturing Industry Community Academy, with expertise in Artificial Intelligence and Cloud Computing. The purpose of this activity is to improve students' understanding of the concepts of Data Science and Data Preprocessing, which are important in managing and analyzing data in the era of rapid information technology. The implementation method includes education, training, and active interaction between speakers and participants. The results of the activity showed an increase in students' understanding, with an average pretest score of 67.87 increasing to 81.73 in the posttest. This activity not only enriches students' knowledge but also prepares them to face challenges in the increasingly complex and data-based world of work, as well as supporting the development of science and technology. Thus, this activity provides significant benefits in improving student competence and contributing to the advancement of higher education.</i></p> <hr/> <p><b>Abstrak</b></p> <p>Pengabdian kepada masyarakat merupakan salah satu bentuk tridarma perguruan tinggi yang bertujuan memberikan kontribusi nyata dalam peningkatan kualitas sumber daya manusia. Program Studi Teknik Komputer, Institut Teknologi dan Sains Muhammadiyah Kolaka Utara melaksanakan kegiatan pengabdian masyarakat secara daring melalui Zoom. Kegiatan ini difokuskan pada teknologi, mengangkat tema Data Science dan Data Preprocessing, relevan dengan mata kuliah Big Data. Pemateri berasal dari Akademi Komunitas Industri Manufaktur Bantaeng dengan keahlian di bidang Artificial Intelligence dan Cloud Computing. Tujuan kegiatan ini adalah meningkatkan pemahaman mahasiswa tentang konsep Data Science dan Data Preprocessing dalam pengelolaan dan analisis data di era teknologi informasi yang pesat. Metode pelaksanaan mencakup edukasi, pelatihan, dan interaksi aktif antara pemateri dan peserta. Hasil kegiatan menunjukkan peningkatan pemahaman mahasiswa, dengan nilai rata-rata pretest 67,87 meningkat menjadi 81,73 pada posttest. Kegiatan ini tidak hanya memperkaya pengetahuan mahasiswa tetapi juga mempersiapkan mereka untuk menghadapi tantangan di dunia kerja yang semakin kompleks dan berbasis data, serta mendukung perkembangan ilmu pengetahuan dan teknologi. Dengan demikian, kegiatan ini memberikan manfaat signifikan dalam meningkatkan kompetensi mahasiswa dan berkontribusi pada kemajuan pendidikan tinggi.</p>
 <p>Lisensi: cc-by-sa</p>	
<p>Corresponding Author</p>	<p>Ibnu Mansyur Hamdani Akademi Komunitas Industri Manufaktur Bantaeng; ibnumansyur@akom-bantaeng.ac.id</p>
<p>How to Cite (APA)</p>	<p>Hamdani, I. M., Nurhidayat, N., Karman, A., Adhalia H, N. F., &amp; Julyaningsih, A. H. (2024). Edukasi dan Pelatihan Data Science dan Data Preprocessing. Intisari: Jurnal Inovasi Pengabdian Masyarakat, 2(1), 19–26. <a href="https://doi.org/10.58227/intisari.v2i1.125">https://doi.org/10.58227/intisari.v2i1.125</a></p>

## PENDAHULUAN

Pengabdian kepada masyarakat merupakan salah satu bentuk tridarma perguruan tinggi yang memiliki tujuan untuk memberikan kontribusi nyata dalam peningkatan kualitas sumber daya manusia. Pengabdian kepada masyarakat dapat dilakukan berdasarkan pada keahlian dan dengan tetap memerhatikan kebutuhan masyarakat. Contoh pengabdian yang dapat dilakukan seperti pelatihan pengolahan sumber daya alam (Disman, et al., 2023). Selain itu juga dapat melakukan sosialisasi pencegahan *bullying* pada pelajar (Angreini, et al., 2023). Kegiatan pengabdian kepada masyarakat yang dilaksanakan oleh penulis bertemakan teknologi, di mana telah banyak yang melakukan kegiatan pengabdian dengan tema yang serupa (Wahyudi, et al., 2022; Rachman, et al., 2023; Hamdani, et al., 2023; Putri, et al., 2022; Siregar, et al., 2022). Penulis sekaligus pemateri merupakan instruktur berpengalaman dalam bidang *Artificial Intelligence (AI)* dan *Cloud Computing (AWS)*. Kegiatan ini dilaksanakan tanggal 2 Mei 2024 pada Program Studi Teknik Komputer, Institut Teknologi dan Sains Muhammadiyah Kolaka Utara secara daring melalui platform Zoom.

Kegiatan pengabdian ini bertemakan teknologi. Urgensi dari kegiatan ini sangat relevan dengan mata kuliah Big Data yang telah dipelajari oleh mahasiswa dengan harapan mereka mampu memahami dan menguasai konsep *Data Science* dan *Data Preprocessing*. Dengan perkembangan teknologi informasi yang sangat pesat, kemampuan mengelola dan menganalisis data menjadi semakin penting. Data science telah menjadi komponen kunci dalam pengambilan keputusan berbasis data di berbagai industri (Han, et al., 2023). Oleh karena itu, tambahan ilmu mengenai Data Science tidak hanya akan memperkaya pengetahuan mahasiswa, tetapi juga mempersiapkan mereka untuk menghadapi tantangan di dunia kerja yang semakin kompleks dan data-driven. Selain itu, kegiatan ini juga bertujuan memberikan tambahan materi yang berkaitan dengan mata kuliah Big Data yang telah diprogramkan dalam kurikulum. Mata kuliah ini memerlukan pemahaman yang mendalam tentang bagaimana data yang besar dan kompleks dapat diproses dan dianalisis secara efektif. Pemahaman yang kuat mengenai *Big Data* dan teknik analisisnya sangat penting untuk menghasilkan wawasan yang bermanfaat dari *Big Data*. Dengan adanya pengajaran tambahan dari praktisi yang berpengalaman, diharapkan mahasiswa dapat mengintegrasikan teori yang telah mereka pelajari dengan praktik nyata di lapangan.

Penggunaan platform daring, seperti *Zoom*, dalam kegiatan ini juga memungkinkan akses yang lebih luas dan fleksibel. Pembelajaran daring dapat meningkatkan partisipasi dan keterlibatan mahasiswa secara signifikan (Rahman, et al., 2024). Selain itu, metode ini memungkinkan penyampaian materi yang interaktif dan dinamis, sesuai dengan kebutuhan zaman digital saat ini. Pengajaran Data Science dan Data Preprocessing juga sangat relevan dengan perkembangan terbaru dalam teknologi informasi. Data preprocessing adalah langkah kritis dalam *pipeline data science* yang dapat mempengaruhi hasil akhir analisis secara signifikan. Dengan pemahaman yang mendalam tentang langkah-langkah preprocessing, mahasiswa akan lebih siap menghadapi proyek data yang kompleks. Kegiatan ini juga selaras dengan rekomendasi pendidikan tinggi untuk memperkuat kurikulum dengan keterampilan praktis yang relevan. Keterampilan dalam analisis data dan pemrosesan informasi adalah salah satu keterampilan paling dibutuhkan dalam dunia digital (Yasa, et al., 2022). Selain itu, pendidikan tinggi diharapkan dapat menyediakan pelatihan yang sesuai dengan kebutuhan industri masa kini.

Secara keseluruhan, kegiatan ini diharapkan dapat memberikan manfaat yang signifikan bagi mahasiswa Program Studi Teknik Komputer. Dengan pemahaman yang lebih baik mengenai Data Science dan Data Preprocessing mahasiswa akan lebih siap dalam mengaplikasikan ilmu yang mereka peroleh untuk memecahkan berbagai masalah nyata di industri. Kegiatan ini juga merupakan bentuk komitmen Program Studi Teknik Komputer dalam mendukung perkembangan ilmu pengetahuan dan teknologi, serta berkontribusi pada peningkatan kompetensi mahasiswa.

## **METODE PELAKSANAAN**

Bentuk kegiatan yang dilakukan berupa edukasi dan pelatihan. Pelaksanaan kegiatan pengabdian kepada masyarakat ini dilakukan secara daring melalui platform Zoom dengan target peserta mahasiswa Program Studi Teknik Komputer, Institut Teknologi dan Sains Muhammadiyah Kolaka Utara. Kegiatan pengabdian masyarakat ini dilaksanakan dalam beberapa tahapan, yakni (1) koordinasi dan kerja sama, (2) penentuan tema dan jadwal, (3) *pretest*, (4) pelaksanaan sosialisasi dan pelatihan, (5) serta *posttest* dan evaluasi.

Pada tahap koordinasi dan kerja sama dilakukan kegiatan, antara lain: melakukan rapat koordinasi dengan tim pelaksana dan pemateri untuk menyamakan persepsi dan tujuan kegiatan; menetapkan tim pelaksana yang bertanggung jawab pada setiap aspek kegiatan, seperti teknis pelaksanaan, administrasi, dan dokumentasi; menyusun daftar tugas dan tanggung jawab masing-masing anggota tim pelaksana; dan mendiskusikan terkait penentuan jadwal dan persiapan pelatihan.

Pada tahap penentuan tema dan jadwal dilakukan kegiatan: melakukan *brainstorming* bersama tim pelaksana untuk menentukan tema yang relevan dan bermanfaat bagi peserta; mengkonsultasikan tema yang diusulkan dengan pemateri untuk memastikan kesesuaian dengan keahlian dan pengalaman yang dimiliki; menetapkan tema final dan menginformasikan kepada seluruh tim pelaksana; menyusun jadwal rinci kegiatan, termasuk sesi *pretest*, sesi utama, dan *posttest*; serta menginformasikan jadwal kegiatan kepada semua pihak terkait.

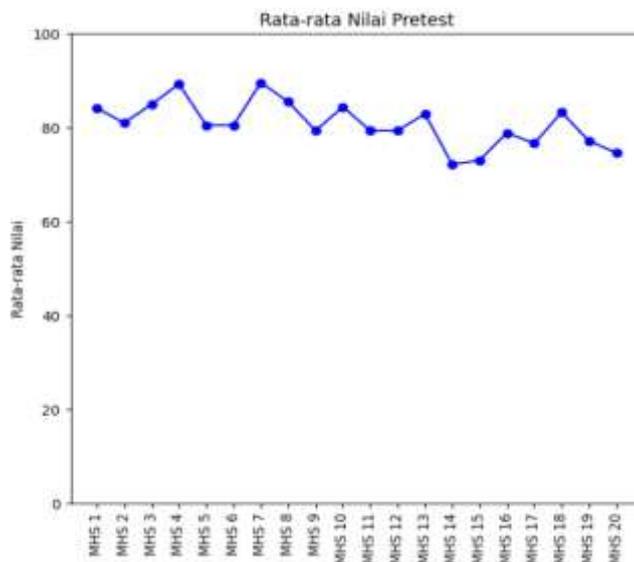
Selanjutnya, pada tahap *pretest* antara lain: menyusun soal *pretest* yang sesuai dengan tema kegiatan; mendistribusikan *pretest* kepada peserta melalui *platform* daring (*google forms*); dan mengumpulkan dan menganalisis hasil *pretest* untuk mengetahui tingkat pemahaman awal peserta.

Pada tahap pelaksanaan sosialisasi dan pelatihan, antara lain: memulai kegiatan dengan sesi pembukaan oleh moderator, yang meliputi pengenalan pemateri dan penjelasan singkat mengenai kegiatan; pemateri menyampaikan materi utama mengenai Data Science dan Data Preprocessing, serta codingan dan contoh aplikasinya; interaksi aktif antara pemateri dan peserta melalui sesi tanya-jawab, diskusi, dan studi kasus; menyediakan materi presentasi dan bahan bacaan tambahan kepada peserta melalui *platform* daring.

Tahap terakhir, *posttest* dan evaluasi, dilakukan kegiatan berikut: menyusun soal *posttest* yang sesuai dengan materi yang telah disampaikan; mendistribusikan *posttest* kepada peserta melalui *google form*; mengumpulkan dan menganalisis hasil *posttest* untuk mengetahui tingkat peningkatan pemahaman peserta; membandingkan hasil *pretest* dan *posttest* untuk mengevaluasi efektivitas kegiatan; mengadakan rapat evaluasi dengan tim pelaksana dan pemateri untuk membahas hasil kegiatan; dan mengumpulkan umpan balik dari peserta mengenai pelaksanaan kegiatan.

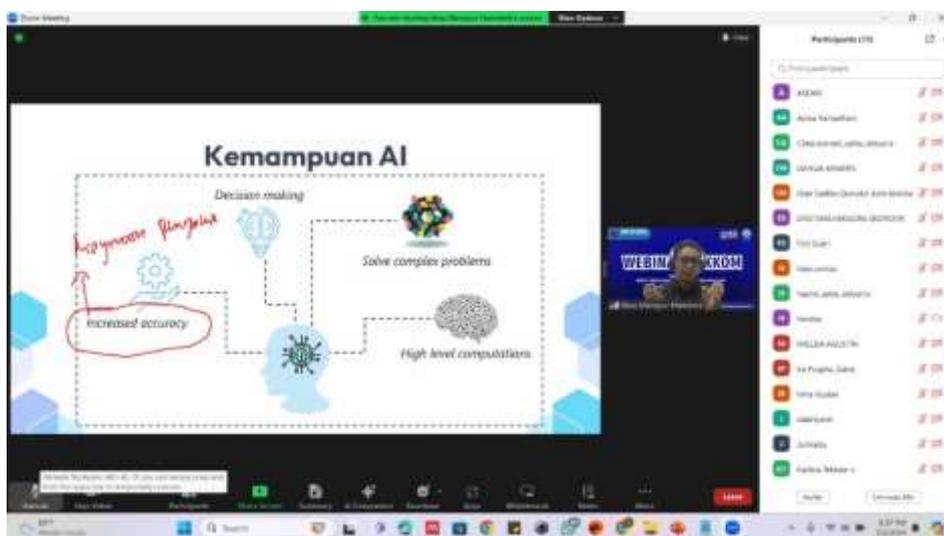
## **HASIL DAN PEMBAHASAN**

Pelaksanaan kegiatan dimulai dengan memberikan *pretest* untuk mengetahui kemampuan awal mahasiswa mengenai Data Science dan Data Preprocessing. Grafik pada Gambar 1 berikut menunjukkan hasil dari *pretest* dengan rata-rata nilai yang diperoleh adalah 67,87. Berdasarkan nilai pada Gambar 1, maka pelatihan yang dilaksanakan berangkat dari asumsi bahwa pemahaman mahasiswa terkait tema yang diambil masih kurang.



Gambar 1. Grafik Hasil Pretest

Pelatihan dimulai dengan penjelasan sejarah *Artificial Intelligence* (AI) dan pemberian contoh-contoh artikel yang menggunakan AI, seperti penelitian perhitungan koloni bakteri menggunakan algoritma *image processing* (Hamdani, et al., 2023), prediksi kanker menggunakan *machine learning* (Islam, et al., 2020), dan sentimen analisis pada twitter (Krishna, 2019). Berdasarkan contoh artikel-artikel tersebut, selanjutnya pelatihan dilanjutkan dengan pemberian pemahaman mengenai kemampuan AI yang menjadi alasan kenapa AI harus digunakan dalam kehidupan sehari-hari.

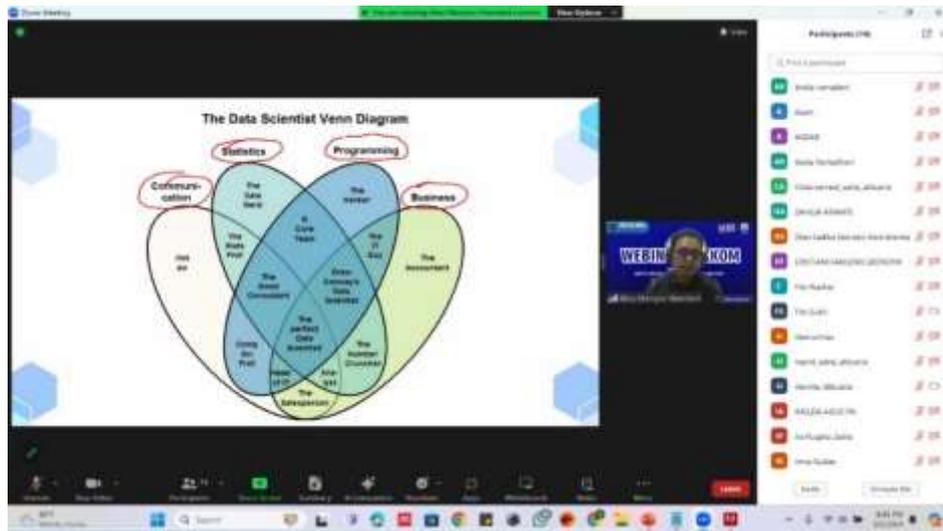


Gambar 2. Penjelasan Materi Kemampuan AI

Pemahaman AI yang telah diperoleh merupakan hal yang paling fundamental sebelum memahami Data Science. Perlu diketahui bahwa Data Science merupakan bidang ilmu yang berfokus untuk mengekstrak informasi berharga dari *big data* dan para data scientist sebaiknya memiliki keterampilan dalam bidang teknologi, matematika, bisnis, dan komunikasi (Krishna, 2019). Mahasiswa diberikan penjelasan mengenai Data Science melalui diagram venn yang akan memudahkan mahasiswa memahami keterampilan yang seharusnya dimiliki untuk menjadi seorang data scientist.

Materi terakhir adalah materi mengenai Data Preprocessing, di mana teknik ini digunakan untuk mengubah data mentah menjadi format yang berguna dan efisien (Varma, et al., 2023). Teknik ini digunakan sebelum pengolahan data. Beberapa proses dalam Data Preprocessing adalah integrasi data, transformasi data, pembersihan data, dan reduksi data (Varma, et al., 2023). Pada materi ini, praktik yang dilakukan lebih berfokus pada proses data cleaning, di mana *data cleaning* meliputi penanganan data redundan, data hilang (*missing value*), dan *outliers*.

Gambaran pelaksanaan pemberian materi terkait penjelasan Diagram Venn Data Science dan penjelasan Data Cleaning dapat dilihat pada Gambar 3 dan 4 berikut.



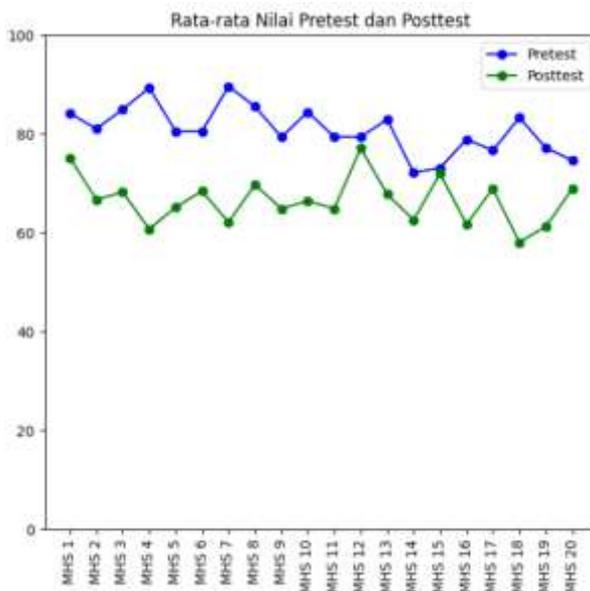
Gambar 3. Penjelasan Diagram Venn Data Science



Gambar 4. Penjelasan Data Cleaning

Berdasarkan materi yang telah dijelaskan, selanjutnya mahasiswa diberikan praktik menggunakan Bahasa Pemrograman *Python* via *Google Collaboratory*. Praktik ini berlangsung secara interaktif, di mana mahasiswa terlihat antusias terhadap contoh-contoh kasus yang diberikan. Pada akhir kegiatan pelatihan, mahasiswa diberikan *posttes* untuk mengetahui pemahaman materi yang diperoleh mahasiswa setelah diberikan penjelasan oleh pemateri. Berdasarkan hasil *posttest* yang telah diberikan kepada mahasiswa, diperoleh gambaran bahwa mahasiswa mengalami peningkatan pemahaman terkait materi yang telah diberikan, di mana

rata-rata nilai yang diperoleh adalah 81,73. Berikut disajikan perbandingan hasil analisis *pretest* dan *posttest* terkait pemahaman mahasiswa, seperti yang tersaji pada Gambar 5 berikut.



Gambar 5. Pretest dan Posttest

## SIMPULAN

Kegiatan pengabdian masyarakat ini dilaksanakan di Program Studi Teknik Komputer, Institut Teknologi dan Sains Muhammadiyah Kolaka Utara, pada tanggal 2 Mei 2024 secara daring melalui Zoom. Fokus utama kegiatan ini adalah pelatihan tentang Data Science dan Data Preprocessing yang relevan dengan mata kuliah Big Data yang dipelajari mahasiswa. Pelatihan ini dipandu oleh pemateri yang berpengalaman dalam bidang *Artificial Intelligence* (AI) dan *Cloud Computing* (AWS). Tujuan dari kegiatan ini adalah untuk meningkatkan pemahaman mahasiswa terhadap Data Science, mempersiapkan mereka menghadapi tantangan dunia kerja yang semakin kompleks dan berbasis data, serta memperkuat keterampilan praktis yang sesuai dengan kebutuhan industri. Kegiatan dilaksanakan melalui beberapa tahapan terstruktur, mulai dari koordinasi dan kerja sama, penentuan tema dan jadwal, *pretest*, pelaksanaan sosialisasi dan pelatihan, hingga *posttest* dan evaluasi. Hasil evaluasi kegiatan menunjukkan peningkatan pemahaman mahasiswa tercermin dari peningkatan nilai rata-rata *pretest* dari 67,87 menjadi 81,73 pada *posttest*. Dengan demikian, secara keseluruhan menunjukkan bahwa kegiatan ini memberikan manfaat signifikan bagi mahasiswa, dan juga meningkatkan kompetensi mereka dalam bidang Data Science dan Data Preprocessing, serta mendukung pengembangan ilmu pengetahuan dan teknologi di lingkungan akademik.

## REFERENSI

- Angreini, D., Tajuddin, A., Purwanto, J., Munaing, & Aswar. (2023). Upaya Mencegah Perilaku Bullying dan Meningkatkan Self Esteem Siswa SMP YP PGRI Disamakan Makassar. *INTISARI Jurnal Inovasi Pengabdian Masyarakat*, 1, 23–32. <https://doi.org/10.58227/intisari.v1i1.31>
- Disman, R., Nadria, A. T., Wahyuni, Nasution, R. A., Rahaditya, D., & Saputra, S. (2023). Optimalisasi Ekonomi Ibu-Ibu PKK Melalui Pemberdayaan dan Pengolahan Coconut Oil di Mandailing Natal. *INTISARI Jurnal Inovasi Pengabdian Masyarakat*, 1, 1–8. <https://doi.org/10.58227/intisari.v1i1.21>
- Hamdani, I. M., Anam, S., Shofianah, N., & Bustamin, S. (2023). Counting Bacterial Colony and

- Reducing noise on Low-Quality Image Using Modified Perona-Malik Diffusion Filter with Sobel Mask Fractional Order. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 12(2), 271–279. <https://doi.org/10.32736/sisfokom.v12i2.1661>
- Hamdani, I. M., Syamsuri, S., Alang, H., & Adhalih, N. F. (2023). Edukasi Mengenai Pentingnya Data Science Untuk Masa Depan. *Jurnal Abdimas Bina Bangsa*, 4(1), 34–40. <https://jabbb.lppmbinabangsa.id/index.php/jabb/article/view/313>
- Han, J., Pei, J., & Tong, H. (2023). *Data Mining: Concepts and Techniques* (S. Merken (ed.); Fourth). Katey Birtcher.
- Islam, M. M., Haque, M. R., Iqbal, H., Hasan, M. M., Hasan, M., & Kabir, M. N. (2020). Breast Cancer Prediction: A Comparative Study Using Machine Learning Techniques. *SN Computer Science*, 1(5), 290. <https://doi.org/10.1007/s42979-020-00305-w>
- Krishna, A. N. P. V. (2019). An efficient approach for sentiment analysis using machine learning algorithm. *Evolutionary Intelligence*, 0123456789. <https://doi.org/10.1007/s12065-020-00429-1>
- Putri, A. N., Wakhidah, N., & Utomo, V. G. (2022). Pemanfaatan Data Mining untuk Media Pembelajaran di SMK Hidayah Semarang. *Educatations - Pengabdian Kepada Masyarakat*, 13(3), 487–491.
- Rachman, F., Alam, I. F., Tay, M. Y., Abidin, M. R., Sefriyandi, R., Nugroho, R. A., Pamulang, U., & Tangerang, K. (2023). Pengenalan Teknologi Software. *Abdi Jurnal Publikasi*, 1(4), 400–403.
- Rahman, E. Y., Payage, N., Pattiruhu, C. M., Shaleh, M., Manado, U. N., Wamena, S. T. T. A., Negara, S. K., & Parepare, I. (2024). Efektivitas Metode Pembelajaran Daring Dalam Mahasiswa Teknologi Informasi Dan Komunikasi. *Jurnal Cahaya Mandalika*, 5, 461–467.
- Siregar, B., Pangruruk, F. A., Siridion, S. T., & Matana, U. (2022). Pengenalan Data Science dan Profesi Data Scientist di SMA Pramita Tangerang. *Jurnal Pengabdian Masyarakat Bestari*, 1(2), 87–96.
- Varma, D., Nehansh, A., & Guide, P. S. (2023). Data Preprocessing Toolkit: An Approach to Automate Data Preprocessing. *International Journal of Scientific Research in Engineering and Management (IJSREM)*, 1–5. <https://doi.org/10.55041/IJSREM18270>
- Wahyudi, J., Qurniati, N., Rohmawan, E. P., & Prawitasari, A. (2022). Pemanfaatan Cloud Computing Untuk Pengamanan Data Bagi Dunia Pendidikan Utilization of Cloud Computing for Data Security for the World of Education. *Jurnal Padamu Negeri*, 3(2), 53–58.
- Yasa, A. D., Kumala, F. N., Andriani, F., & Author, C. (2022). Analysis Of Information Processing Capabilities And Digital Data On Digital Literacy Skills. *Jurnal Pendidikan Dasar Nusantara*, 8(1), 28–37.

