

Pengembangan Keterampilan Associate Data Scientist melalui Pelatihan dengan RapidMiner

“Skills Development of Associate Data Scientist through Training with RapidMiner”

Egi Safitri^{1*}, Rini Nurlistiani², Hendra Kurniawan³

Institut Informatika dan Bisnis Darmajaya, Lampung, Indonesia^{1,2,3}

egisafitri@darmajaya.ac.id¹, rininurlistiani@darmajaya.ac.id², hendrakurniawan@darmajaya.ac.id³



Riwayat Artikel

Diterima pada 16 Oktober 2024

Revisi 1 pada 29 Oktober 2024

Revisi 2 pada 16 November 2024

Revisi 3 pada 25 November 2024

Disetujui pada 28 November 2024

Abstract

Purpose: This study aims to evaluate the effectiveness of an online Associate Data Scientist training program that utilizes RapidMiner as the primary platform for teaching data science and machine learning. The goal is to assess participants' improvements in data preprocessing, algorithm application, and model evaluation skills.

Methodology/approach: The training program was conducted via Zoom and included interactive lectures, live demonstrations, hands-on exercises, and individual assignments. RapidMiner was used as the main tool throughout the sessions. Participants were evaluated through tasks assigned in each session and a final project that required them to analyze a dataset, apply relevant algorithms, and assess model performance.

Results/findings: The results showed significant improvement in participants' technical understanding and application skills. The average final project score was 87.0, indicating strong competence in data handling, algorithm selection, and model evaluation. Most participants completed the project successfully, demonstrating their readiness to apply data science concepts in real-world scenarios.

Conclusions: The online training effectively bridged the gap between theory and practice, proving that remote learning can deliver quality outcomes in technical education. The combination of RapidMiner and a structured training format enabled participants to gain applicable skills in data science. However, improvements in instructional delivery and interaction are still needed to optimize learning experiences.

Limitations: Challenges included internet connectivity issues and limited real-time interaction, which sometimes hindered learning flow and instructor support.

Contribution: This study provides valuable insights into data science education, proving that online programs with practical tools like RapidMiner can successfully build core competencies in aspiring data professionals.

Keywords: Data Analysis, Data Science Training, Machine Learning, Online Learning, Rapidminer

How to Cite: Safitri, E., Nurlistiani, R, dan Kurniawan, H. (2024). Pengembangan Keterampilan Associate Data Scientist melalui Pelatihan dengan RapidMiner. *Jurnal Pengabdian Kepada Masyarakat*, 5(4), 677-685.

1. Pendahuluan

Dalam beberapa tahun terakhir, perkembangan teknologi digital telah mengubah lanskap industri secara signifikan (Nottbrock, Looy, & Haes, 2023). Salah satu perubahan terbesar adalah munculnya volume data yang sangat besar yang dihasilkan dari berbagai proses bisnis, perangkat teknologi, dan interaksi

digital sehari-hari. Data kini menjadi aset berharga bagi organisasi karena dapat memberikan wawasan untuk pengambilan keputusan strategis dan meningkatkan performa bisnis (Monino, 2021); (Shabbir & Gardezi, 2020). Pengembang sistem/teknologi harus memastikan bahwa teknologi yang dipilih dapat mengakomodasi kebutuhan sistem dan dapat diakses dengan mudah (Taufik, Muhaqiqin, Iلمان, & Sholehurrohman, 2023). Namun, tantangan utama dalam era data ini adalah bagaimana memanfaatkan data secara efektif untuk mendapatkan wawasan yang bermakna. Untuk mengatasi tantangan ini, keahlian dalam *data science* menjadi sangat penting (Grover & Kar, 2017).

Data science adalah disiplin yang menggabungkan statistika, pemrograman, dan pembelajaran mesin untuk mengolah dan menganalisis data secara efisien (Gehlen et al., 2022). Peran *data scientist* semakin dibutuhkan di berbagai sektor, mulai dari keuangan, kesehatan, teknologi informasi, hingga ritel (Mildenberger, Braschler, Ruckstuhl, Vorburger, & Stockinger, 2023); (Tu, Zou, Su, & Zhang, 2023). Seorang *data scientist* memiliki peran krusial dalam menganalisis data dan menerjemahkannya menjadi wawasan yang dapat digunakan untuk pengambilan keputusan yang lebih cerdas dan strategis (Gul & Al-Faryan, 2023); (Rengarajan, Narayanamurthy, Moser, & Pereira, 2022). Beberapa algoritma *data science* yang sering digunakan antara lain *Random Forest* (Suyudi, Sudadio, & Suherman, 2022), algoritma K-Means (Haris, Friadi, Frederick, Huda, & Romdoni, 2024), algoritma decision tree yang merupakan model supervised machine learning dalam data sciences (Setiawan, Nurhatisyah, & Nanra, 2023), serta algoritma Logistic regression yang merupakan algoritma klasifikasi untuk dua kategori kelas (Putra, Suprpto, & Bukhori, 2022).

Meskipun demikian, banyak lulusan universitas belum memiliki keterampilan praktis yang memadai di bidang *data science*. Kurikulum di perguruan tinggi cenderung lebih menekankan aspek teoretis tanpa memberikan cukup ruang bagi mahasiswa untuk berlatih menggunakan alat dan teknik yang relevan dengan kebutuhan industri (Cheng, Adekola, Albia, & Cai, 2022); (Ali, 2022). Kesenjangan antara teori dan praktik ini menjadi masalah signifikan bagi lulusan yang ingin bekerja sebagai *data scientist* atau analis data (Shaharabani & Yarden, 2019). Untuk mengatasi kesenjangan tersebut, pelatihan berbasis alat (*tools*) yang memungkinkan peserta untuk terlibat langsung dalam proses analisis data sangat diperlukan, terutama bagi mahasiswa yang ingin mempersiapkan diri sebagai *associate data scientist* (Lewis & Stoyanovich, 2022). Salah satu alat yang cocok digunakan dalam pelatihan ini adalah *RapidMiner*, sebuah *tools* yang menawarkan antarmuka yang ramah pengguna serta berbagai fitur canggih untuk analisis data, visualisasi, algoritma *machine learning*, dan evaluasi model. Keunggulan utama *RapidMiner* adalah kemampuannya untuk mengotomatisasi banyak proses analisis, sehingga mempermudah pengguna tanpa memerlukan keterampilan pemrograman tingkat lanjut (Arias-Barahona et al., 2023); (Mehmood et al., 2023). *Rapidminer* Pelatihan ini bertujuan untuk membantu mahasiswa dan profesional muda mengembangkan keterampilan praktis yang sesuai dengan kebutuhan industri, serta meningkatkan kepercayaan diri peserta dalam menghadapi tantangan dunia kerja yang semakin berbasis data (Johnson et al., 2021).

Review yang sama pernah dilakukan oleh Putri, Wakhidah, and Utomo (2022) yang melakukan pelatihan media pembelajaran di SMK Hidayah Semarang menggunakan *RapidMiner*. Selain itu Workshop dan Seminar Penerapan *Data Science* di Dunia Pendidikan Menggunakan *RapidMiner* juga dilakukan oleh (A. P. Natasuwarna, 2019); (A. Natasuwarna, Pangihutan, & Ramadani, 2022) dan menghasilkan peningkatan pengetahuan pelajar mengenai *data mining* dan *RapidMiner*. Hal ini menunjukkan pentingnya pelatihan berbasis alat dalam mengembangkan keterampilan *data science*. (Han, Kamber, & Pei, 2012) dalam bukunya *Data Mining: Concepts and Techniques* menjelaskan bahwa meskipun konsep dasar seperti statistika dan *machine learning* diajarkan di universitas, mahasiswa sering kali tidak memiliki keterampilan praktis yang cukup untuk menerapkannya di dunia kerja.

Frank, Hall, and Witten (2016) menyatakan dalam *Practical Machine Learning Tools and Techniques* mengemukakan bahwa mahasiswa kurang percaya diri dalam mengimplementasikan teknik-teknik *machine learning* karena minimnya pengalaman praktis. Studi ini menyoroti pentingnya pelatihan berbasis *tools* untuk meningkatkan keterampilan praktis mahasiswa. Laporan *State of Data Science and Machine Learning*, mengungkapkan bahwa 80% dari *data scientist* di seluruh dunia mempelajari

keterampilan mereka melalui pelatihan praktis yang menggunakan alat dan studi kasus nyata, menunjukkan bahwa pendidikan formal belum sepenuhnya menyediakan pengalaman praktis yang memadai. Selain itu, Vyas and Uma (2018) menemukan bahwa penggunaan platform seperti *RapidMiner* dan *KNIME* secara signifikan meningkatkan keterampilan mahasiswa dalam menerapkan teori analisis data. (Slater, Joksimović, Kovanovic, Baker, & Gasevic, 2017) menggarisbawahi pentingnya *hands-on experience* dalam penggunaan *tools* seperti *RapidMiner*, yang membantu mahasiswa memahami proses analisis data dengan lebih baik. (Rajan, Beymer, Abedin, & Dehghan, 2020) menyoroti bahwa *AI Studio* memudahkan peserta untuk mengaplikasikan algoritma *machine learning* tanpa harus mendalami pemrograman, sehingga memfasilitasi pembelajaran bagi pemula. (Wang et al., 2019) juga menemukan bahwa platform dengan antarmuka intuitif seperti *AI Studio* dapat meningkatkan kemampuan mahasiswa non-teknis dalam memahami konsep *data science* dan mengurangi kecemasan terkait pembelajaran teknis.

Meskipun banyak pelatihan telah dilakukan mengenai penggunaan alat analisis data dalam pelatihan *data science*, masih terdapat kesenjangan yang perlu diperhatikan. Pertama, meskipun *RapidMiner* dan *KNIME* sudah banyak diteliti, penggunaan *RapidMiner* sebagai alat utama dalam pengembangan keterampilan praktis mahasiswa belum banyak diulas. Kedua, pelatihan sebelumnya lebih banyak berfokus pada peningkatan keterampilan teknis tanpa mengeksplorasi secara mendalam dampak penggunaan alat-alat ini terhadap kepercayaan diri dan kesiapan mahasiswa untuk bekerja di industri. Pelatihan ini bertujuan untuk mengatasi kesenjangan tersebut dengan menekankan penggunaan *RapidMiner* sebagai alat utama dalam pelatihan. Dengan pendekatan berbasis aplikasi yang intuitif, pelatihan ini diharapkan dapat memberikan pengalaman praktis yang lebih interaktif dan aplikatif, sehingga dapat membantu peserta mempersiapkan diri secara lebih efektif untuk bekerja di dunia industri berbasis data.

2. Metode Penelitian

Pengabdian ini menggunakan pendekatan kuasi-eksperimental dengan menerapkan pelatihan *Associate Data Scientist* berbasis *RapidMiner*. Pelatihan ini dilaksanakan secara daring melalui platform *Zoom Meeting* dengan total 17 Jam Pelajaran yang terbagi dalam beberapa sesi. Peserta pelatihan adalah mahasiswa Institut Informatika dan Bisnis Darmajaya Bandar Lampung dan Instruktur adalah Dosen yang telah lulus pelatihan menjadi seorang *Trainer* dan memiliki sertifikat BNSP. Pelatihan yang diberikan dilakukan secara intensif kepada peserta dengan kombinasi teori dan praktik menggunakan alat *RapidMiner*, yang dirancang agar peserta dapat mempraktikkan keterampilan analisis data secara langsung selama pelatihan berlangsung.

2.1. Desain Pelatihan

Pelatihan ini dirancang dengan metode pembelajaran campuran yang mencakup penyampaian materi teori dan praktik secara langsung melalui aplikasi *RapidMiner*. Total pelatihan dilakukan selama 17 JP yang terbagi dalam 5 sesi dengan rincian sebagai berikut:

1. Pengantar *Data Science* dan Pengolahan Data (3 JP). Pada sesi ini, peserta diberikan pengenalan konsep dasar *data science*, peran penting data dalam dunia industri, serta langkah-langkah awal pengolahan data menggunakan *RapidMiner*.
2. Eksplorasi Data dan *Preprocessing* (4 JP). Sesi ini berfokus pada eksplorasi data, teknik *preprocessing* data (seperti normalisasi, pengisian nilai yang hilang, dan encoding), serta bagaimana mempersiapkan data untuk proses analisis lebih lanjut.
3. Penerapan Algoritma *Machine Learning* (4 JP). Peserta akan belajar menerapkan berbagai algoritma *machine learning* yang disediakan oleh *RapidMiner*, seperti *Decision Tree*, *Random Forest*, dan *K-Nearest Neighbors (KNN)* untuk tugas klasifikasi dan regresi.
4. Evaluasi Model *Machine Learning* (3 JP). Peserta mempelajari teknik evaluasi model *machine learning* dengan menggunakan berbagai metrik seperti akurasi, *precision*, *recall* dan *F1-score*. Evaluasi model penting untuk menilai performa model dalam memprediksi data baru.
5. Proyek Akhir (3 JP). Pada sesi terakhir, peserta diberikan proyek akhir sebagai tugas kelulusan. Proyek ini mengharuskan peserta untuk menerapkan semua keterampilan yang telah dipelajari selama pelatihan dengan melakukan analisis terhadap dataset nyata dan menyusun laporan hasil analisis.

2.2. Peserta dan Instruktur Pelatihan

Peserta pelatihan berjumlah 30 orang, yang merupakan mahasiswa tingkat akhir Institut Informatika dan Bisnis Darmajaya Bandar Lampung, dari berbagai program studi yang relevan dengan teknologi informasi dan *data science*. Pemilihan peserta didasarkan pada ketertarikan dan motivasi mereka untuk belajar *data science* serta kemampuan dasar dalam memahami analisis data. Setiap peserta diwajibkan memiliki perangkat komputer/laptop dan koneksi internet yang stabil untuk mengikuti pelatihan daring. Kemudian, Instruktur adalah seorang Dosen yang telah memiliki sertifikat *Trainer* dan Kemampuan *Associate Data Scientist* dan telah di akui oleh BNSP.

2.3. Metode Pengajaran

Metode pengajaran yang diterapkan dalam pelatihan ini menggunakan kombinasi beberapa pendekatan yang efektif. Pertama, penjelasan teori diberikan secara interaktif oleh pengajar, yang secara rinci menjelaskan materi serta membuka sesi tanya jawab. Hal ini dilakukan untuk memastikan bahwa peserta memahami konsep yang disampaikan secara mendalam. Selain itu, setiap konsep dilengkapi dengan demonstrasi penerapan menggunakan *RapidMiner*, sehingga peserta dapat melihat secara langsung bagaimana alur kerja analisis data dilakukan dalam praktik. Setelah penjelasan dan demonstrasi, peserta diberikan kesempatan untuk melakukan latihan praktis dengan menggunakan dataset latihan yang disediakan oleh pengajar. Latihan ini memungkinkan peserta untuk langsung mempraktikkan materi yang baru saja dipelajari.

Selain itu, peserta diberikan sedikit game edukasi pembelajaran. Tujuannya untuk menambah minat peserta untuk belajar dan membuat peserta memahami materi dengan baik (Nisya, Wulansari, & Wartariyus, 2023). Selanjutnya, sesi pelatihan diakhiri dengan tugas individu yang harus diselesaikan oleh peserta secara mandiri. Tugas ini dirancang untuk memperkuat pemahaman peserta terhadap materi yang telah diajarkan sebelumnya.

2.4. Evaluasi dan Tugas Akhir

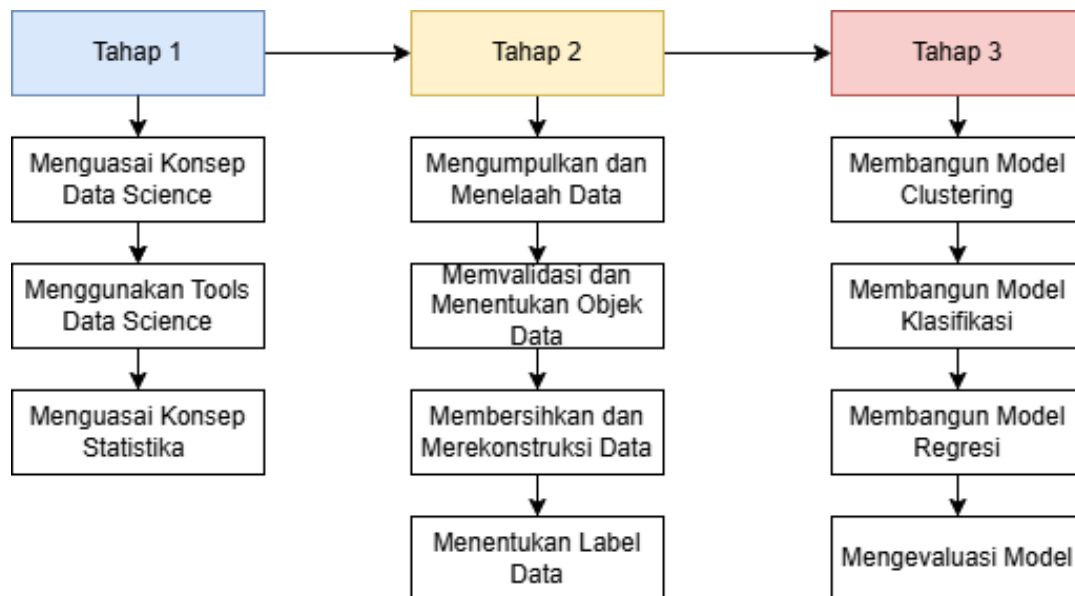
Evaluasi dalam pelatihan ini dilakukan melalui dua pendekatan utama. Pertama, setiap sesi pelatihan diakhiri dengan tugas yang harus diselesaikan oleh peserta, yang berfungsi sebagai evaluasi keterampilan peserta dalam memahami dan menerapkan materi yang diajarkan di setiap sesi. Kedua, evaluasi dilakukan melalui proyek akhir yang diwajibkan sebagai syarat kelulusan. Proyek akhir ini merupakan tugas komprehensif yang mencakup seluruh materi yang telah dipelajari selama pelatihan. Peserta diminta untuk melakukan analisis data secara mandiri menggunakan *RapidMiner*, termasuk memilih dan menerapkan algoritma machine learning yang sesuai serta mengevaluasi hasil analisis tersebut. Proyek akhir ini tidak hanya menguji kemampuan teknis peserta dalam melakukan analisis data, tetapi juga cara mereka menyusun laporan tertulis dan mempresentasikan hasil analisis mereka di sesi akhir pelatihan. Penilaian proyek akhir didasarkan pada beberapa aspek, termasuk akurasi model yang dihasilkan, pemahaman terhadap konsep yang diterapkan, serta kualitas presentasi hasil yang disampaikan oleh peserta.

3. Hasil dan Pembahasan

Pelatihan *Associate Data Scientist* dengan menggunakan *RapidMiner*, yang disampaikan sepenuhnya melalui platform *Zoom Meeting*, bertujuan untuk memberikan pemahaman komprehensif mengenai konsep dasar *data science*, teknik machine learning, serta penerapan *tools* analisis data. Tugas akhir peserta berupa proyek analisis data, di mana mereka diberikan dataset untuk dianalisis dan dipresentasikan sebagai syarat kelulusan.

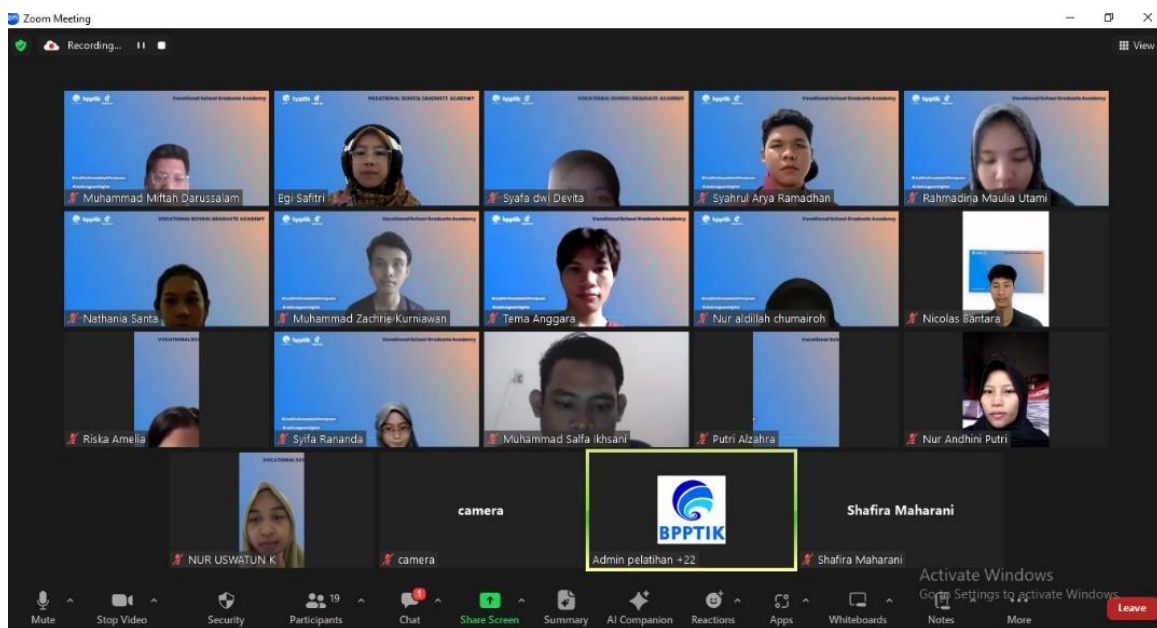
3.1. Pelaksanaan Pelatihan

Pelatihan dilaksanakan selama 17 Jam Pelajaran yang terbagi dalam beberapa tahap dengan pendekatan pengajaran secara daring. Materi disampaikan melalui kombinasi presentasi dan demonstrasi penggunaan *RapidMiner*. Pelatihan *Associate Data Scientist* menggunakan *RapidMiner* dilaksanakan sepenuhnya melalui *Zoom Meeting*. Peserta mengikuti sesi yang mencakup penyampaian materi dan demonstrasi penggunaan untuk analisis data. Sebagai bagian dari evaluasi, proyek akhir berupa analisis dataset yang diberikan menjadi tugas utama peserta. Tahapan pembelajaran dalam pelatihan dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Pembelajaran dalam Pelatihan

Selanjutnya, interaksi aktif antara peserta dan instruktur terlihat dalam setiap sesi, dengan diskusi dan tanya jawab di akhir setiap pertemuan. Gambar 2 berikut menggambarkan suasana pelatihan daring yang menunjukkan partisipasi aktif peserta selama berlangsungnya pelatihan.



Gambar 2. Proses pelatihan dengan *Zoom meeting*

Gambar 2 memperlihatkan interaksi peserta dengan Instruktur pelatihan dengan berdiskusi atau melalui pertanyaan menggunakan fitur *chat* maupun fitur *raise hand* pada aplikasi *Zoom meeting*.

3.2. Proyek Akhir dan Penilaian

Proyek akhir yang diberikan kepada peserta merupakan tugas komprehensif yang mencakup seluruh tahapan analisis data menggunakan *RapidMiner*. Dataset yang digunakan adalah dataset standar yang memerlukan serangkaian langkah preprocessing data, penerapan algoritma machine learning, dan evaluasi performa model. Peserta diharapkan dapat memilih algoritma yang tepat, menginterpretasikan hasil analisis, dan menyusun laporan yang dipresentasikan secara daring pada sesi akhir pelatihan. Penilaian proyek akhir dilakukan berdasarkan empat kriteria utama:

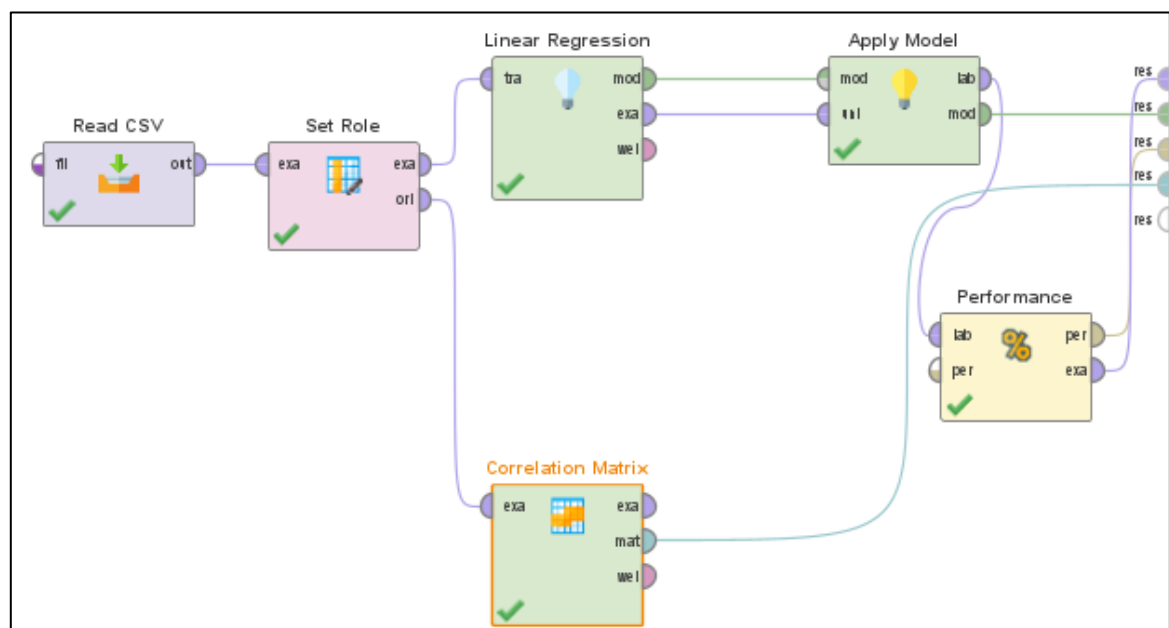
1. Pengolahan Data: Peserta diharapkan mampu melakukan preprocessing data dengan benar, termasuk menangani data yang hilang, melakukan normalisasi, dan encoding.
2. Pemilihan Algoritma: Evaluasi dilakukan berdasarkan ketepatan dalam memilih algoritma yang sesuai dengan dataset yang digunakan. Algoritma yang paling sering digunakan oleh peserta adalah *Linear Regression* dan *Decision Tree*.
3. Evaluasi Model: Kualitas model diukur menggunakan metrik seperti akurasi, precision, dan recall untuk menilai performa model yang dihasilkan.
4. Presentasi Hasil: Peserta mempresentasikan hasil analisis secara daring melalui *Zoom*, dengan fokus pada interpretasi hasil dan rekomendasi perbaikan model.

Tabel berikut menunjukkan rata-rata skor dari proyek akhir yang dinilai berdasarkan kriteria di atas:

Tabel 1. Aspek evaluasi dan rata-rata skor

Aspek Evaluasi	Nilai Rata-rata
Pengolahan data	86.5
Pemilihan Algoritma	88.2
Evaluasi model	84.0
Presensi Hasil	89.3

Sebagian besar peserta berhasil menyelesaikan proyek akhir dengan hasil yang memuaskan, meskipun terdapat beberapa kendala teknis dalam memahami fitur yang lebih kompleks. Rata-rata skor akhir sebesar 87,0 menunjukkan bahwa pelatihan ini secara signifikan meningkatkan keterampilan peserta dalam melakukan analisis data secara efektif. Pada bagian proyek akhir ini, peserta diberikan tugas untuk menjalankan analisis data menggunakan *RapidMiner*, dengan contoh alur kerja yang dapat dilihat pada gambar berikut:



Gambar 3. Proses simulasi dengan *RapidMiner*

Gambar 3 mengilustrasikan beberapa tahapan penting dalam proses analisis data menggunakan *RapidMiner*. Tahapan pertama adalah proses *Read CSV*, di mana dataset diimpor ke dalam platform analisis. Pada langkah ini, peserta diharapkan melakukan preprocessing yang mencakup penanganan data yang hilang dan penyesuaian format data untuk memastikan dataset siap digunakan. Tahap berikutnya, *Set Role*, melibatkan penentuan peran setiap atribut dalam analisis, seperti pemilihan variabel target yang akan diprediksi oleh model.

Selanjutnya, peserta menerapkan algoritma *Linear Regression* untuk membangun model prediktif berdasarkan dataset yang telah disiapkan. Setelah model dikembangkan, langkah *Apply Model* digunakan untuk menerapkan model tersebut pada data uji guna menghasilkan prediksi. Proses ini dilanjutkan dengan *Performance Evaluation*, yang mengevaluasi performa model menggunakan metrik seperti akurasi dan tingkat kesalahan (*error*). Proyek akhir ini dirancang untuk menilai sejauh mana peserta dapat menguasai keseluruhan proses analisis data, mulai dari *preprocessing* hingga evaluasi model. Evaluasi tersebut bertujuan untuk memastikan bahwa peserta tidak hanya memahami konsep machine learning, tetapi juga mampu menerapkannya secara efektif dalam menyelesaikan tugas analisis data yang komprehensif.

3.3. Pembahasan

Hasil pengabdian ini selaras dengan beberapa pengabdian sebelumnya (Aryotejo, Hakim, Firmansah, & Safarizki, 2021); (Kholifah, Gadi, Yuli, Khayati, & Triyanto, 2020); (Sarwosri, Rochimah, Yuhana, Siahaan, & Akbar, 2024) yang menunjukkan bahwa pelatihan yang disampaikan melalui platform *Zoom* tetap dapat efektif dalam meningkatkan keterampilan peserta. Penyampaian materi secara daring memerlukan adaptasi dalam metode pengajaran, di mana instruktur harus memastikan bahwa setiap peserta dapat mengakses materi dan alat yang digunakan dengan baik. Meskipun terdapat beberapa kendala teknis, mayoritas peserta berhasil memahami konsep yang diajarkan dan mampu menerapkannya dalam proyek akhir. Proyek akhir yang diberikan kepada peserta terbukti menjadi metode evaluasi yang efektif dalam mengukur pemahaman dan penerapan keterampilan yang telah dipelajari selama pelatihan. Peserta menunjukkan peningkatan dalam kemampuan mereka untuk melakukan pengolahan data, memilih algoritma yang sesuai, dan mengevaluasi performa model yang mereka buat. Namun, kendala utama yang dihadapi adalah terkait koneksi internet dan interaksi yang terbatas selama pelatihan daring. Hal ini menunjukkan bahwa meskipun metode pengajaran daring efektif, perlu ada penyesuaian lebih lanjut untuk memastikan interaksi yang lebih baik antara instruktur dan peserta serta akses yang lebih lancar terhadap alat yang digunakan. Secara keseluruhan, pelatihan ini memberikan hasil yang positif dalam hal peningkatan keterampilan peserta, meskipun pelaksanaan sepenuhnya daring membutuhkan penyesuaian yang lebih matang di masa depan untuk mengatasi kendala teknis dan memastikan aksesibilitas yang lebih baik.

4. Kesimpulan

Pelatihan *Associate Data Scientist* yang disampaikan sepenuhnya melalui *Zoom Meeting* dengan menggunakan *RapidMiner* berhasil meningkatkan keterampilan peserta dalam *data science* dan *machine learning*. Melalui pengajaran daring dan proyek analisis data sebagai tugas akhir, peserta mendapatkan pengalaman praktis dalam memproses data, menerapkan algoritma, dan mengevaluasi model. Meskipun terdapat kendala teknis seperti koneksi internet dan akses terhadap fitur kompleks, sebagian besar peserta berhasil menyelesaikan proyek dengan baik dan menunjukkan peningkatan pemahaman yang signifikan. Proyek akhir menjadi alat evaluasi efektif untuk mengukur keterampilan peserta secara teknis dan analitis. Hasilnya menunjukkan bahwa pelatihan daring ini mampu menjembatani kesenjangan antara teori dan praktik. Pelatihan daring juga menggarisbawahi pentingnya penyesuaian dalam metode pengajaran dan aksesibilitas alat. Interaksi instruktur dan peserta dapat ditingkatkan, dan solusi teknis harus menjadi prioritas dalam pelatihan serupa di masa depan. Secara keseluruhan, pelatihan ini efektif dalam mempersiapkan peserta menjadi *associate data scientist* yang siap menghadapi tantangan industri berbasis data, meski beberapa aspek masih perlu diperbaiki untuk meningkatkan pengalaman belajar.

Ucapan terima kasih

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada seluruh peserta yang telah berpartisipasi aktif dalam program pelatihan *Associate Data Scientist*. Ucapan terima kasih khusus kami sampaikan kepada tim penyelenggara dan para instruktur atas dedikasi serta keahlian mereka dalam menyelenggarakan sesi pelatihan daring. Kami juga menyampaikan apresiasi kepada tim dukungan teknis yang telah memastikan kelancaran operasional platform *Zoom Meeting* dan penggunaan *RapidMiner* selama pelatihan berlangsung. Akhir kata, kami berterima kasih kepada lembaga dan organisasi yang telah menyediakan sumber daya serta dukungan yang diperlukan untuk mewujudkan pelatihan ini.

Referensi

- Ali, M. G. (2022). A General Perspective about Institutional Rankings, Ranking Framework, Benefits of Rankings and Ranking Methodological Flaws. *International Journal of Educational Research Review*, 7(3), 157-164. doi:<https://doi.org/10.24331/ijere.1067952>
- Arias-Barahona, M. X., Arteaga-Arteaga, H. B., Orozco-Arias, S., Flórez-Ruíz, J. C., Valencia-Díaz, M. A., & Tabares-Soto, R. (2023). Requests Classification in the Customer Service Area for Software Companies Using Machine Learning and Natural Language Processing. *PeerJ Computer Science*, 9. doi:<https://doi.org/10.7717/peerj-cs.1016>
- Aryotejo, G., Hakim, M. M., Firmansah, F., & Safarizki, H. A. (2021). Pelatihan Efisiensi Sumber Daya Sistem Operasi Windows pada Masa Pandemi Covid 19. *Jurnal ABDINUS: Jurnal Pengabdian Nusantara*, 4(2), 238-246. doi:<https://doi.org/10.29407/ja.v4i2.14906>
- Cheng, M., Adekola, O., Albia, J., & Cai, S. (2022). Employability in Higher Education: A Review of Key Stakeholders' Perspectives. *Higher Education Evaluation and Development*, 16(1), 16-31. doi:<https://doi.org/10.1108/HEED-03-2021-0025>
- Frank, E., Hall, M. A., & Witten, I. H. (2016). *The WEKA Workbench "Data Mining: Practical Machine Learning Tools and Techniques" 4th Edition*. Burlington: Morgan Kaufmann Publishers.
- Gehlen, K. P.-v., Höck, H., Fast, A., Heydebreck, D., Lammert, A., & Thiemann, H. (2022). Recommendations for Discipline-Specific FAIRness Evaluation Derived from Applying an Ensemble of Evaluation Tools. *Data Science Journal*, 21(1), 1-21. doi:<https://doi.org/10.5334/dsj-2022-007>
- Grover, P., & Kar, A. K. (2017). Big Data Analytics: A Review on Theoretical Contributions and Tools Used in Literature. *Global Journal of Flexible Systems Management*, 18(3), 203-229. doi:<https://doi.org/10.1007/s40171-017-0159-3>
- Gul, R., & Al-Faryan, M. A. S. (2023). From Insights to Impact: Leveraging Data Analytics for Data-Driven Decision-Making and Productivity in Banking Sector. *Humanities and Social Sciences Communications*, 10(1), 1-8. doi:<https://doi.org/10.1057/s41599-023-02122-x>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques 3rd Edition*. Waltham: Morgan Kaufmann Publishers.
- Haris, Y., Friadi, J., Frederick, A. E. S., Huda, D. N., & Romdoni, M. R. (2024). Clustering Data Stok Penjualan Sparepart Mobil Toyota Bengkel Multi Topindo Menggunakan K-Means. *Jurnal Ilmu Siber dan Teknologi Digital*, 2(2), 109-121. doi:<https://doi.org/10.35912/jisted.v2i2.3308>
- Johnson, M., Jain, R., Brennan-Tonetta, P., Swartz, E., Silver, D., Paolini, J., . . . Hill, C. (2021). Impact of Big Data and Artificial Intelligence on Industry: Developing a Workforce Roadmap for a Data Driven Economy. *Global Journal of Flexible Systems Management*, 22(3), 197-217. doi:<https://doi.org/10.1007/s40171-021-00272-y>
- Kholifah, N., Gadi, A. C. Z., Yuli, S. E., Khayati, E. Z., & Triyanto, T. (2020). Penggunaan Zoom Cloud Meeting sebagai Alternatif Pembelajaran Jarak Jauh. *Prosiding Pendidikan Teknik Boga Busana*, 15(1).
- Lewis, A., & Stoyanovich, J. (2022). Teaching Responsible Data Science: Charting New Pedagogical Territory. *International Journal of Artificial Intelligence in Education*, 32(3), 1-25. doi:<https://doi.org/10.1007/s40593-021-00241-7>
- Mehmood, I., Shahid, S., Hussain, H., Khan, I., Ahmad, S., Rahman, S., . . . Huda, S. (2023). A Novel Approach to Improve Software Defect Prediction Accuracy Using Machine Learning. *IEEE Access*, 11, 63579-63597. doi:<https://doi.org/10.1109/ACCESS.2023.3287326>
- Mildenberger, T., Braschler, M., Ruckstuhl, A., Vorbürger, R., & Stockinger, K. (2023). The Role of Data Scientists in Modern Enterprises-Experience from Data Science Education. *ACM SIGMOD Record*, 52(2), 48-52. doi:<https://doi.org/10.1145/3615952.3615966>
- Monino, J.-L. (2021). Data Value, Big Data Analytics, and Decision-Making. *Journal of the Knowledge Economy*, 12(1), 256-267.
- Natasuwarna, A., Pangihutan, H. B., & Ramadani, S. (2022). Workshop Online Penerapan Data Science pada Dunia Pendidikan. *Seminar Nasional Penelitian dan Pengabdian Kepada Masyarakat CORISINDO*, 186-189.
- Natasuwarna, A. P. (2019). Seminar Pendekatan Data Mining Memprediksi Profil Sosial Masyarakat Menggunakan Aplikasi RapidMiner. *SNPMas: Seminar Nasional Pengabdian pada Masyarakat*, 38-44.

- Nisya, I. S., Wulansari, O. D. E., & Wartariyus, W. (2023). Rancang Bangun Game Edukasi Bencana Alam Menggunakan Metode MDLC. *Jurnal Ilmu Siber dan Teknologi Digital*, 2(1), 23-44. doi:<https://doi.org/10.35912/jisted.v2i1.2374>
- Nottbrock, C., Looy, A. V., & Haes, S. D. (2023). Impact of Digital Industry 4.0 Innovations on Interorganizational Value Chains: A Systematic Literature Review. *Business Process Management Journal*, 29(1), 43-76. doi:<https://doi.org/10.1108/BPMJ-06-2022-0259>
- Putra, T. I. Z. M., Suprpto, S., & Bukhori, A. F. (2022). Model Klasifikasi Berbasis Multiclass Classification dengan Kombinasi Indobert Embedding dan Long Short-Term Memory untuk Tweet Berbahasa Indonesia. *Jurnal Ilmu Siber dan Teknologi Digital*, 1(1), 1-28. doi:<https://doi.org/10.35912/jisted.v1i1.1509>
- Putri, A. N., Wakhidah, N., & Utomo, V. G. (2022). Pemanfaatan Data Mining untuk Media Pembelajaran di SMK Hidayah Semarang. *E-Dimas: Jurnal Pengabdian Kepada Masyarakat*, 13(3), 487-491. doi:<https://doi.org/10.26877/e-dimas.v13i3.5572>
- Rajan, D., Beymer, D., Abedin, S., & Dehghan, E. (2020). Pi-PE: A Pipeline for Pulmonary Embolism Detection using Sparsely Annotated 3D CT Images. *Proceedings of the Machine Learning for Health NeurIPS Workshop*, 116, 220-232.
- Rengarajan, S., Narayanamurthy, G., Moser, R., & Pereira, V. (2022). Data Strategies for Global Value Chains: Hybridization of Small and Big Data in the Aftermath of COVID-19. *Journal of Business Research*, 144, 776-787. doi:<https://doi.org/10.1016/j.jbusres.2022.02.042>
- Sarwosri, Rochimah, S., Yuhana, U. L., Siahaan, D. O., & Akbar, R. J. (2024). Pelatihan Pemrograman Web Dasar untuk Siswa di SMA Negeri 1 Bojonegoro. *Sewagati: Jurnal Pengabdian Kepada Masyarakat*, 8(1), 1053-1060. doi:<https://doi.org/10.12962/j26139960.v8i1.548>
- Setiawan, E., Nurhatisyah, N., & Nanra, S. (2023). Pengontrolan Bahaya Kebakaran Berbasis IOT pada Ruang Server SMFR Balai Monitor Spektrum Frekuensi Radio Kelas II Batam. *Jurnal Ilmu Siber dan Teknologi Digital*, 1(1), 41-51. doi:<https://doi.org/10.35912/jisted.v1i1.1800>
- Shabbir, M. Q., & Gardezi, S. B. W. (2020). Application of Big Data Analytics and Organizational Performance: The Mediating Role of Knowledge Management Practices. *Journal of Big Data*, 7, 1-17. doi:<https://doi.org/10.1186/s40537-020-00317-6>
- Shaharabani, Y. F., & Yarden, A. (2019). Toward Narrowing the Theory–Practice Gap: Characterizing Evidence from in-Service Biology Teachers’ Questions Asked During an Academic Course. *International Journal of STEM Education*, 6, 1-13. doi:<https://doi.org/10.1186/s40594-019-0174-3>
- Slater, S., Joksimović, S., Kovanovic, V., Baker, R. S., & Gasevic, D. (2017). Tools for Educational Data Mining: A Review. *Journal of Educational and Behavioral Statistics*, 42(1), 85-106. doi:<https://doi.org/10.3102/1076998616666808>
- Suyudi, I., Sudadio, & Suherman. (2022). Pengenalan Bahasa Isyarat Indonesia Menggunakan Mediapipe dengan Model Random Forest dan Multinomial Logistic Regression. *Jurnal Ilmu Siber dan Teknologi Digital*, 1(1), 65-80. doi:<https://doi.org/10.35912/jisted.v1i1.1899>
- Taufik, R., Muhaqiqin, Ilman, I. S., & Sholehurrohman, R. (2023). Analisis Informasi Jaringan Homogen dan Heterogen pada Liga Champions UEFA. *Jurnal Ilmu Siber dan Teknologi Digital*, 1(2), 91-110. doi:<https://doi.org/10.35912/jisted.v1i2.1928>
- Tu, X., Zou, J., Su, W. J., & Zhang, L. (2023). What Should Data Science Education Do With Large Language Models?. *Harvard Data Science Review*, 6(1), 1-21. doi:<https://doi.org/10.1162/99608f92.bff007ab>
- Vyas, V., & Uma, V. (2018). An Extensive Study of Sentiment Analysis Tools and Binary Classification of Tweets Using Rapid Miner. *Procedia Computer Science*, 125, 329-335. doi:<https://doi.org/10.1016/j.procs.2017.12.044>
- Wang, D., Weisz, J. D., Muller, M., Ram, P., Geyer, W., Dugan, C., . . . Gray, A. (2019). Human-AI Collaboration in Data Science: Exploring Data Scientists’ Perceptions of Automated AI. *Proceedings of the ACM on Human-Computer Interaction*, 3, 1-24. doi:<https://doi.org/10.1145/3359313>