

## STATISTIKA DAN TEORI PROBABILITAS UNTUK PENGUATAN LITERASI SAINS DAN NUMERASI MELALUI SOFTWARE RAPIDMINER PADA MATA KULIAH STATISTIKA PEMBELAJARAN MESIN BAGI SISWA SEKOLAH MENENGAH DAN MAHASISWA MATEMATIKA

Yulia Resti<sup>1\*</sup>, Des Alwine Zayanti<sup>1</sup>, Endang Sri Kresnawati<sup>1</sup>, Ning Eliyati<sup>1</sup>, Irsyadi Yani<sup>2</sup>

<sup>1</sup>Jurusan Matematika FMIPA Universitas Sriwijaya

<sup>2</sup>Jurusan Teknik Mesin Fakultas Teknik Universitas Sriwijaya

Corresponding author: yulia\_resti@mipa.unsri.ac.id

**ABSTRAK:** Mata kuliah Statistika Pembelajaran Mesin merupakan salah satu mata kuliah pada bidang minat Statistika dan Komputasi di Program Studi S1 Matematika FMIPA Universitas Sriwijaya. Mata kuliah ini juga merupakan mata kuliah pilihan pada kelompok bidang keahlian Konstruksi di Program Studi S1 Teknik Mesin Fakultas Teknik Universitas Sriwijaya, namun dengan nama mata kuliah Pembelajaran Mesin. Mata kuliah ini menggunakan RapidMiner sebagai salah satu software dalam proses pengayaan topik pembelajarannya. Software ini bersifat open source, yang dapat diartikan bahwa kode sumber atau kode dasar program komputer yang mendukungnya dapat digunakan oleh banyak orang. Di era teknologi informasi sekarang ini, mengenal software open source merupakan salah satu cara untuk meningkatkan kemampuan literasi sains dan numerasi. Indonesia merupakan salah satu negara berkembang yang memiliki kemampuan literasi sains dan numerasi rendah. Tujuan dari kegiatan Pengabdian pada Masyarakat (PPM) dengan skema Perkuliahan Desa ini adalah untuk menguatkan kemampuan literasi sains dan numerasi peserta didik melalui software RapidMiner. Penguatan literasi sains dan numerasi diukur berdasarkan pra-test dan post-test terkait pengetahuan dasar tentang statistika dan probabilitas, serta tugas kelompok yang menggunakan RapidMiner. Hasil kegiatan menunjukkan bahwa literasi sains dan numerasi khalayak sasaran meningkat signifikan.

**Kata Kunci:** literasi sains dan numerasi, open source, RapidMiner, statistika dan probabilitas

**ABSTRACT:** *The Machine Learning Statistics course is one of the courses in the field of interest in Statistics and Computing in the Bachelor of Mathematics Study Program FMIPA Sriwijaya University. This course is also an elective course in the Construction expertise group in the Mechanical Engineering Undergraduate Study Program, Faculty of Engineering, Sriwijaya University, but with the name of the Machine Learning course. This course uses RapidMiner as one of the software in the process of enriching its learning topics. This software is open source, which means that the source code or basic code of the computer program that supports it can be used by many people. In the current era of information technology, getting to know opensource software is one way to improve scientific literacy and numeracy skills. Indonesia is a developing country that has low scientific literacy and numeracy capabilities. The aim of Community Service (PPM) activities with the Village Lecture scheme is to strengthen students' scientific literacy and numeracy skills through RapidMiner software. Strengthening scientific literacy and numeracy is measured based on pre-test and post-test related to basic knowledge of statistics and probability, as well as group assignments using RapidMiner. The results of the activity showed that the scientific literacy and numeracy of the target audience increased significantly.*

**Keywords:** *scientific literacy and numeracy, open source, RapidMiner, statistics and probability*

### PENDAHULUAN

Di Program Studi S1 Matematika, FMIPA, Universitas Sriwijaya, Statistika Pembelajaran Mesin (Statistical Machine Learning) merupakan mata kuliah pilihan di semester ganjil untuk bidang peminatan Statistika dan Komputasi. Mata kuliah ini juga merupakan mata kuliah pilihan pada kelompok bidang keahlian Konstruksi di Program Studi S1 Teknik Mesin Fakultas

Teknik Universitas Sriwijaya, namun dengan nama mata kuliah Pembelajaran Mesin (Machine Learning). Kedua mata kuliah tersebut memiliki silabus yang hampir sama. Pada kedua mata kuliah pilihan tersebut, terdapat topik yang membahas tentang metode-metode analisis data yang menggunakan software open source dalam proses pengayaannya. Salah satu software opensource-nya adalah software RapidMiner. Kelebihan software open source di antaranya adalah memerlukan biaya yang lebih

sedikit (bahkan umumnya gratis), dapat dimodifikasi dan dikembangkan sendiri oleh pengguna, dan dapat membuat software versi sendiri. Selain itu, jika terjadi masalah pada software open source, ada kemungkinan dapat diselesaikan sendiri tanpa bantuan pengembang. Namun begitu, jika dilihat dari kegunaannya, software open source sedikit lebih sulit jika dibandingkan dengan software closed source. Hal tersebut karena panduan penggunaan untuk software open source lebih ditujukan untuk pengembang dan kurang ramah untuk pengguna awam. Di era teknologi informasi sekarang ini, mengenal software open source merupakan salah satu cara untuk meningkatkan kemampuan literasi sains dan numerasi. Indonesia merupakan salah satu negara berkembang yang memiliki kemampuan literasi sains dan numerasi rendah. Kedua literasi tersebut menjadi unsur penting untuk mencapai tujuan pendidikan yang dicanangkan pemerintah.

Sekolah Menengah Atas (SMA) dan Sekolah Menengah Kejuruan (SMK) Kartini, merupakan sekolah swasta yang berada sekitar 1 km di belakang kampus Universitas Sriwijaya. Sekolah ini memberikan ekstrakurikuler berupa pengetahuan internet dan dasar-dasar komputer. Pengetahuan ini merupakan modal dasar untuk memahami software open source dan materi dasar pada mata kuliah Statistika Pembelajaran Mesin. Untuk dapat memahami metode-metode analisis data yang digunakan pada software Rapidminer yang diberikan pada mata kuliah tersebut, pemahaman terhadap konsep-konsep statistika dan teori probabilitas merupakan hal yang sangat penting. Konsep tersebut dimulai dari himpunan dan operasi himpunan kejadian dan ruang sampel, kejadian tunggal, kejadian majemuk, probabilitas kejadian tunggal dan kejadian majemuk, kejadian dan probabilitas bersyarat, pemahaman tentang data & strukturnya proses pengumpulan, penyajian, pengolahan, dan analisis data, pengambilan keputusan, jenis-jenis teknik sampling dan resampling, hingga validasi model.

Tujuan dari kegiatan Pengabdian pada Masyarakat (PPM) dengan skema Perkuliahan Desa ini adalah untuk menguatkan kemampuan literasi sains dan numerasi peserta didik melalui software RapidMiner. Penguatan literasi sains dan numerasi diukur berdasarkan pra-test dan post-test terkait pengetahuan dasar tentang statistika dan teori probabilitas, serta tugas kelompok yang menggunakan RapidMiner.

## MATERI DAN METODE PELAKSANAAN

RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan pembelajaran mesin yang dapat diintegrasikan pada produknya sendiri. RapidMiner sebelumnya bernama YALE (Yet Another Learning Environment), dimana versi awalnya mulai dikembangkan pada tahun 2001 oleh Ralf Klinkenberg,

Ingo Mierswa, dan Simon Fischer di Artificial Intelligence Unit dari University of Dortmund. RapidMiner didistribusikan di bawah lisensi AGPL (GNU Affero General Public License) versi 3. Hingga saat ini telah ribuan aplikasi yang dikembangkan menggunakan RapidMiner di lebih dari 40 negara. RapidMiner ditulis dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi. RapidMiner menyediakan GUI (Graphic User Interface) untuk merancang sebuah pipeline analitis. GUI ini akan menghasilkan file XML (Extensible Markup Language) yang mendefinisikan proses analitis keinginan pengguna untuk diterapkan ke data. File ini kemudian dibaca oleh RapidMiner untuk menjalankan analisis secara otomatis. Terkait penggunaan Rapidminer untuk mata kuliah Statistika Pembelajaran Mesin, konsep dasar statistika dan probabilitas yang diberikan terdiri dari himpunan dan operasi himpunan, kejadian dan ruang sampel, kejadian tunggal, kejadian majemuk, probabilitas kejadian tunggal dan kejadian majemuk, kejadian dan probabilitas bersyarat, pemahaman tentang data & strukturnya proses pengumpulan, penyajian, pengolahan, dan analisis data, pengambilan keputusan, jenis-jenis teknik sampling dan resampling, hingga validasi model.

Sebelum pemberian materi baik yang menyajikan konsep-konsep statistika dan teori probabilitas maupun pengenalan tentang RapidMiner, khalayak sasaran terlebih dulu diberikan pra-test untuk mengetahui kemampuan peserta terhadap materi yang akan disampaikan. Untuk menilai hasil kegiatan ini, dilakukan tes tertulis dan tugas kelompok. Tes tertulis diberikan dalam bentuk pra-test dan post-test tentang konsep dasar statistik, statistika, probabilitas, dan perhitungan probabilitas kelas. Masing-masing pra-test dan post-test terdiri dari 20 soal dan setiap soal yang dijawab dengan benar memperoleh skor 5 dan yang dijawab salah tidak memperoleh skor sehingga skor tertinggi yang dapat diperoleh oleh peserta adalah 100. Kriteria dan indikator pencapaian tujuan serta keberhasilan kegiatan PPM ini adalah ketika hasil post-test minimal 70 % khalayak sasaran berhasil memperoleh nilai dengan rata-rata minimal 70 dan ada peningkatan yang signifikan dari pra-test maka dapat disimpulkan bahwa tujuan kegiatan telah tercapai. Secara spesifik, tahapan pelaksanaan kegiatan pengabdian pada masyarakat skema Perkuliahan Desa terdiri dari:

1. Pra-test tentang konsep-konsep dasar statistika dan probabilitas seperti populasi dan sampel, jenis-jenis data, statistika deskriptif dan statistika inferensia, peluang bersyarat, teorema Bayes, dan metode naïve Bayes.
2. Penyampaian materi terkait konsep-konsep dasar statistika dan probabilitas seperti populasi dan sampel, jenis-jenis data, statistika deskriptif dan statistika inferensia, peluang bersyarat, teorema Bayes, dan metode naïve Bayes.

3. Post-test tentang konsep-konsep dasar statistika dan probabilitas seperti populasi dan sampel, jenis-jenis data, statistika deskriptif dan statistika inferensia, peluang bersyarat, teorema Bayes, dan metode naïve Bayes.
4. Penyampaian materi terkait pengenalan dan eksplorasi software Rapidminer dan contoh penerapannya untuk literasi sains dan numerasi.
5. Tugas kelompok menggunakan software Rapidminer Selanjutnya dilakukan uji beda rata-rata antara kedua kelompok dengan hipotesis nol bahwa rata-rata nilai tes peserta pelatihan pada saat pra-test dan post-test tidak berbeda. Perbedaan rata-rata keduanya untuk menunjukkan bahwa kegiatan yang dilakukan memberikan manfaat bagi peserta kegiatan, dan apabila yang terjadi adalah sebaliknya, diperlukan upaya untuk memberikan pengayaan yang lebih banyak sehingga peserta memperoleh manfaat yang signifikan dimana kesignifikan tersebut dapat diukur dengan jelas.

**HASIL KEGIATAN**

Literasi sains dan numerasi sangat diperlukan masyarakat modern untuk menghadapi berbagai permasalahan ilmu pengetahuan dan teknologi. Selain itu, literasi sains dan numerasi juga berguna dalam menunjang pembangunan berkelanjutan (Udompong & Wongmanich, 2014). Menurut Kementerian Pendidikan dan Kebudayaan (2017), kemampuan literasi sains dan numerasi merupakan salah satu faktor penting bagi kemajuan sebuah negara dalam menjalani kehidupan di era globalisasi. RapidMiner merupakan software yang berdiri sendiri untuk analisis data dan pembelajaran mesin yang dapat diintegrasikan pada produknya sendiri.

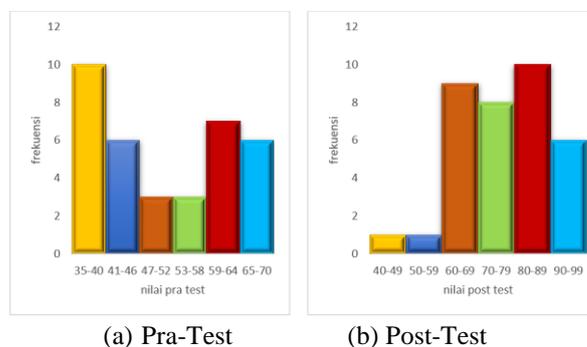
Terkait tujuan kegiatan yaitu untuk meningkatkan kemampuan literasi sains dan numerasi bagi peserta didik melalui software open source RapidMiner dimana peserta didik terdiri dari siswa SMA dan SMK Terpadu Kartini dan juga mahasiswa yang mengambil mata kuliah pilihan Statistika Pembelajaran Mesin secara umum berjalan lancar dan sesuai dengan rencana. Selain khalayak sasaran, turut hadir pada kegiatan ini adalah unsur pimpinan sekolah dan guru-guru SMA dan SMK Terpadu Kartini, serta tim pengabdian pada masyarakat yang terdiri dari 3 orang dosen dan 8 orang mahasiswa.

Hasil pra-test dan post-test terhadap 35 orang peserta kegiatan diringkaskan dalam bentuk statistik dan histogram yang masing-masing disajikan pada Tabel 1 dan Gambar 1. Tabel 1 menunjukkan bahwa hasil pra-test memiliki rentang nilai 35 – 70 dengan rata-rata 50,86 dan standar deviasi 11,54. Hasil post-test memiliki nilai minimum dan nilai maksimum yang lebih baik daripada nilai pra-test, begitu juga dengan nilai rata-ratanya, namun nilai post-test memiliki standar deviasi yang lebih lebar. Fakta ini menginformasikan bahwa sudah cukup

banyak mahasiswa yang mengalami peningkatan nilai dari pra-test ke post-test yang cukup banyak, namun masih ada peserta kegiatan yang mengalami peningkatan nilai yang sedikit dari pra-test ke post-test.

Tabel 1. Resume Penilaian Peserta Kegiatan PPM Perkuliahan Desa

	Nilai Pra-test	Nilai Post-test
Minimum	35,00	40,00
Quartil ke-1	40,00	65,00
Median	50,00	75,00
Rata-rata	50,86	74,43
Modus	60,00	85,00
Quartil ke-3	60,00	85,00
Maksimum	70,00	95,00
Standar deviasi	11,54	12,93



Gambar 1. Histogram Hasil Penilaian Peserta Pelatihan: (a) Pra-Test, (b) Post-Test

Gambar 1 (a) menunjukkan bahwa distribusi nilai pra-test untuk kelompok nilai 35-40 adalah yang paling banyak, namun modus nilai tes tidak berada pada kelompok nilai ini melainkan pada kelompok nilai 59-64. Dua kelompok nilai dengan distribusi paling sedikit adalah kelompok nilai 47-52 dan 53-58. Kelompok yang mengandung nilai rata-rata dan median berada dalam salah satu kelompok dengan distribusi nilai paling sedikit yaitu 47-52. Artinya cukup banyak peserta yang mendapatkan nilai pratest dalam kelompok nilai kedua terbesar, namun lebih banyak lagi peserta yang mendapatkan nilai yang berada dalam kelompok nilai terkecil. Ini berakibat nilai rata-rata berada pada kelompok nilai dengan distribusi paling sedikit dan berada pada kelompok nilai menengah ke bawah.

Gambar 1 (b) menunjukkan bahwa walaupun nilai pra-test dan post-test memiliki jumlah kelas interval sama, namun panjang kelas interval nilai post-test lebih lebar daripada pra-test. Deskripsi ini juga diindikasikan oleh standar deviasinya. Berbeda dengan hasil pra-test, pada post-test modus berada pada kelas interval dengan distribusi nilai yang paling banyak yaitu kelompok nilai 80-89. Sementara itu, rata-rata dan median nilai post-test berada pada kelompok nilai menengah ke atas.

Kedua Gambar 1(a) dan 1(b) mengindikasikan bahwa terdapat peningkatan nilai yang signifikan antara

dari pra-tes ke post-test. Indikasi ini selaras dengan hasil uji beda rata-rata kedua kelompok nilai pra-test dan post-test seperti yang ditampilkan pada Tabel 2. Dengan menggunakan uji Z, tingkat signifikansi 5 % dan hipotesis nol bahwa rata-rata nilai kelompok pra-test dan post-test adalah sama, Tabel 2 menginformasikan bahwa terdapat perbedaan signifikan antara rata-rata kedua nilai tes tersebut dan nilai rata-rata pra-tes ke post-test dengan masing-masing variansi sebesar 133,07 dan 167,31 tersebut meningkat secara signifikan.

Tabel 2. Hasil Pengujian hipotesis rata-rata dua kelompok

Statistik	Parameter
Variansi pra-test	133,07
Variansi post-test	167,31
Z-hitung	-7,99
Z-tabel	1.64
p-value	0,00

## KESIMPULAN

Kegiatan yang bertujuan untuk meningkatkan kemampuan literasi sains dan numerasi bagi peserta didik melalui software open source RapidMiner ini telah dilaksanakan dengan hasil yang cukup memuaskan. Berdasarkan hasil kegiatan yang dilakukan diperoleh kesimpulan bahwa peserta didik yang merupakan khalayak sasaran kegiatan dapat memahami materi yang diberikan dan memiliki kemampuan yang merata terkait penggunaan Rapidminer untuk masing-masing kelompok siswa dan mahasiswa. Selain itu terdapat peningkatan kemampuan literasi sains dan numerasi yang signifikan. Hal ini ditunjukkan oleh hasil pengujian uji beda rata-rata baik antara pra-test dan post-test, antar kelompok siswa, dan antar kelompok mahasiswa. Dari hasil pengujian uji beda rata-rata ini dapat disimpulkan bahwa tujuan kegiatan pengabdian pada Masyarakat skema Perkuliahan Desa telah tercapai.

## UCAPAN TERIMA KASIH

Penelitian/publikasi artikel ini dibiayai oleh Anggaran DIPA Badan Layanan Umum Universitas Sriwijaya Tahun Anggaran 2023 Nomor SP DIPA-023.17.2.67751512023, tanggal 30 November 2022. Sesuai dengan SK Rektor Nomor 0004/UN9/SK.LP2M/2023 Tanggal 20 Juni 2023.

## DAFTAR PUSTAKA

- Andriani, N., Saparini, Ikhsan, H. 2018. Kemampuan Literasi Sains Fisika Siswa SMP Kelas VII di Sumatera Selatan Menggunakan Kerangka PISA (Program for International Student Assessment). Berkala Ilmiah Pendidikan Fisika. Vol 6 No 3, pp. 278 – 291.
- Aprillia, D., Baskoro, D.A., Ambarwati, L., Wicaksana, I.W. S. 2013. Belajar Data Mining dengan RapidMiner.
- Badan Pusat Statistik Kota Palembang, 2021. Kota Palembang dalam Angka.
- Detik News, 2019. <https://news.detik.com/berita/d-4818251/gantikan-un-apa-itu-asesmen-kompetensi-minimum-dan-survei-karakter>
- Direktorat Guru dan Tenaga Kependidikan, 2020. <https://http://pgdikmen.kemdikbud.go.id/read-news/bersiap-menuju-asesmen-kompetensi-minimum-/#>.
- Tribunnews.com, 2021. <https://www.tribunnews.com/nasional/2021/03/22/kemendikbud-tingkat-literasi-siswa-indonesia-di-peringkat-pisa-masih-rendah>.
- Kementerian Pendidikan dan Kebudayaan, 2017. Panduan Gerakan Literasi Nasional.
- Kompas, 2020. <https://www.kompas.com/edu/read/2020/03/12/150556671/tahun-depan-un-diganti-asesmen-kompetensi-minimum-ini-contoh-soalnya>.
- Kompas, 2020, 10 September. Jenis-jenis Penyakit Degeneratif yang Perlu Diwaspadai. <https://health.kompas.com/read/2020/09/10/193400068/jenis-jenis-penyakit-degeneratif-yang-perlu-diwaspadai?page=all>
- Pedoman dan Standar Mutu Kegiatan Pengabdian kepada Masyarakat Sumber Dana DIPA LPPM Unsri, LPPM Unsri, 2017.
- Pedoman Kegiatan Pengabdian LPPM Unsri, LPPM Unsri, 2021.
- Teja, M. 2014. Pelindungan Terhadap Anak Telantar di Panti Asuhan. Info Singkat Kesejahteraan Sosial, Vol. VI, No. 05/I/P3DI/Maret/2014
- Udompong, L. & Wongmanich, S. 2014. Diagnosis of the scientific literacy characteristics of primary students. *Procedia-Social and Behavioral Sciences*, Vol. 116, 5091 – 5096.