



PELATIHAN PENGENALAN DATA SCIENCE UNTUK MENINGKATKAN KEMAMPUAN DALAM PENGOLAHAN DATA

Hairani¹, Ahmad Zuli Amrullah²

Universitas Bumigora^{1,2}

E-mail : hairani@universitasbumigora.ac.id¹, zuli@universitasbumigora.ac.id²

Abstrak

Data science merupakan gabungan ilmu komputer, statistika, dan pengetahuan domain bisnis untuk ekstraksi tumpukan data yang besar menjadi pengetahuan sehingga mendapatkan *pattern* atau pola-pola yang dapat memudahkan pengambil keputusan. Adapun orang menggeluti bidang *data science* disebut *data scientist*. Profesi *data scientist* akhir-akhir ini menjadi profesi yang sangat seksi di abad 21. Sumber daya manusia yang berprofesi sebagai *data scientist* di Indonesia sangat sedikit bila dibandingkan ketersediaan lapangan kerja dibidang *data science*. Dengan kata lain, ketersediaan lapangan kerja *data science* berbanding terbalik dengan ketersediaan SDM yang berprofesi sebagai *data scientist*, dimana jumlah SDM *data scientist* sangat sedikit dibandingkan dengan jumlah lapangan kerja yang berlimpah. Salah satu solusi yang ditawarkan adalah mengadakan pelatihan dan *workshop* untuk pengenalan *data science* untuk meningkatkan SDM bidang *data science* khususnya di Universitas Bumigora. Metode pelaksanaan yang digunakan adalah penyampaian materi tentang *data science* dan simulasi penggunaan metode *data science* dalam kasus *real* menggunakan *Google Colab*. Berdasarkan hasil pelatihan dan *workshop* yang telah dilaksanakan, dapat meningkatkan pemahaman dan kemampuan para peserta untuk menggunakan metode-metode yang ada pada *data science* untuk mengolah data menjadi sebuah pengetahuan.

Kata kunci: *data science*, pelatihan, *workshop*, *data scientist*

Abstract

Data science is a combination of computer science, statistics, and business domain knowledge for extracting large heaps of data into knowledge so as to obtain patterns that can facilitate decision makers. The people in the field of data science are called data scientists. The data scientist profession has recently become a very sexy profession in the 21st century. Human resources who work as data scientists in Indonesia are very few when compared to the availability of employment in the field of data science. In other words, the availability of data science employment is inversely proportional to the availability of human resources who work as data scientists, where the number of HR data scientists is very small compared to the abundant number of jobs. One of the solutions offered is to hold training and workshops for the introduction of data science to improve human resources in the field of data science, especially at Bumigora University. The implementation method used is the delivery of material about data science and simulation of the use of data science methods in real cases using Google Colab. Based on the results of training and workshops that have been carried out, it can increase the understanding and ability of the participants to use existing methods in data science to process data into knowledge.

Keywords: *data science*, training, *workshop*, *data scientist*

Copyright (c) 2020 Hairani, Ahmad Zuli Amrullah

✉Corresponding author:

Address : Perumahan Panorama Alam Orange

Email : Hairani@universitasbumigora.ac.id

Phone : 087839793970

DOI : <https://doi.org/10.31004/abdidas.v1i3.31>

ISSN 2721-9224 (Media Cetak)

ISSN 2721- 9216 (Media Online)

PENDAHULUAN

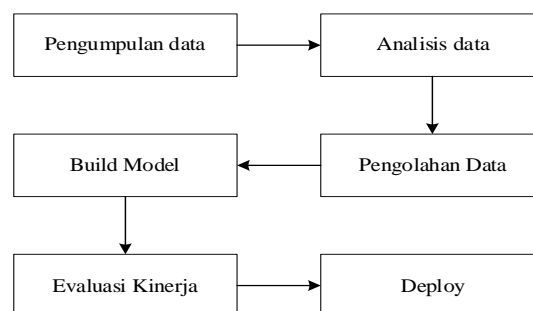
Di era Revolusi Industri 4.0 semua organisasi atau perusahaan menerapkan teknologi informasi dalam menunjang bisnisnya, sehingga lebih efisien dan efektif. Akibat dari penggunaan teknologi informasi tersebut menghasilkan data dalam jumlah besar dan belum dimanfaatkan sepenuhnya. Tumpukan data tersebut dapat diolah menjadi sebuah pengetahuan sehingga dapat memudahkan pengambil keputusan untuk membuat kebijakan kedepannya. Salah satu solusinya adalah menggunakan teknik *data science* atau ilmu data.

Data science merupakan gabungan ilmu komputer, statistika, dan pengetahuan domain bisnis untuk ekstraksi tumpukan data yang besar menjadi pengetahuan sehingga mendapatkan *pattern* atau pola-pola yang dapat memudahkan pengambil keputusan (Igal & Segui, 2017). Adapun orang menggeluti bidang *data science* disebut *data scientist* (Rahman, 2015). Profesi *data scientist* akhir-akhir ini menjadi profesi yang sangat seksi di abad 21. Sumber daya manusia yang berprofesi sebagai *data scientist* di Indonesia sangat sedikit bila dibandingkan ketersediaan lapangan kerja di bidang *data science* (Yusuf, 2018).

Dengan kata lain, ketersediaan lapangan kerja *data science* berbanding terbalik dengan ketersediaan SDM yang berprofesi sebagai *data scientist*, dimana jumlah SDM *data scientist* sangat sedikit dibandingkan dengan jumlah lapangan kerja yang berlimpah. Salah satu kegunaan *data science* adalah dapat melakukan pengolahan data menjadi pengetahuan dengan menggunakan

beberapa teknik seperti mampu memprediksikan data di masa depan menggunakan metode Regresi Linier (Sulistiyono & Sulistiyowati, 2018) dan mampu mengklasifikasikan data menggunakan metode Naïve Bayes (Hairani et al., 2018), C.45 (Gunawan et al., 2020), Support Vector Machine (SVM)(Hairani et al., 2020), k-Nearest Neighbour (k-NN)(Jadhav & Channe, 2016). Tidak hanya itu, kemampuan dalam bidang *data science* dapat mengelompokkan data berdasarkan kemiripannya menggunakan metode K-Means (Yuda Irawan, 2019) (Nugraha & Hairani, 2018).

Untuk menjadi seorang ahli di dalam bidang *data science* dibutuhkan kemampuan seperti statistika, *programming*, dan pengetahuan domain permasalahan. Tidak hanya itu, seorang *data scientist* harus mampu menggunakan metodologi pengolahan data menjadi pengetahuan. Adapun metodologi pengolahan data di bidang *data science* ditunjukkan pada Gambar 1.



Gambar 1: Metodologi Pengolahan *Data Science*

Salah satu solusi yang ditawarkan untuk mengatasi permasalahan kekurangan sumber daya manusia di bidang *data science* adalah mengadakan pelatihan dan *workshop* dalam rangka untuk memperkenalkan bidang *data science*

sehingga dapat meningkatkan kemampuan dalam pengolahan data menjadi sebuah pengetahuan dan juga dapat digunakan sesuai kepentingan. Dengan adanya pelatihan dan *workshop* tersebut diharapkan dapat meningkatkan Sumber Daya Manusia (SDM) untuk menggeluti bidang *data science*. Para peserta yang ditargetkan dalam pelatihan ini adalah mahasiswa Universitas Bumigora dengan jurusan Ilmu Komputer, Statistika, dan Ekonomi.

METODE

Pelaksanaan pelatihan dan *workshop* pengenalan *data science* tersebut dilaksanakan

pada hari Jumat, 20 Desember 2019 bertempat di Aula Universitas Bumigora dengan mengambil tema pengenalan data science untuk kemajuan pariwisata NTB di era Revolusi Industri 4.0. Adapun metode pelaksanaannya dilakukan dengan dua sesi yaitu sesi pertama penyampaian teori *data science* yang disampaikan oleh Ahmad Zuli Amrullah, M.Eng, dan Hairani, M.Eng. Kemudian sesi kedua melakukan simulasi untuk implementasi salah satu teknik *data science* yaitu melakukan prediksi dengan kasus *real* menggunakan metode Regresi Linier berbasis *Google Colab*. Adapun susunan acaranya ditunjukkan pada Tabel 1.

Tabel 1. Susunan Acara

Jam	Kegiatan
08.00 – 08.30	Registrasi peserta seminar
08.30 – 09.00	Sambutan dan pengenalan <i>Data Science Weekend</i>
09.00 - 10.00	Pengenalan <i>Data Science</i>
10.00 - 10.15	Sesi tanya jawab 1
10.15 – 11.15	Potensi penelitian bidang <i>Data Science</i>
11.15 – 11.30	Sesi tanya jawab 2
11.30 – 13.30	Ishoma
13.30 – 14.00	Registrasi peserta <i>workshop</i>
14.00 – 16.00	<i>Workshop</i> dengan <i>Google Colab</i>
16.00 – 16.30	Sesi tanya jawab 3
16.30 – 17.00	Penutupan acara

HASIL DAN PEMBAHASAN

Kegiatan pengabdian dilaksanakan pada tanggal 20 Desember 2019 di Aula Universitas Bumigora. Peserta yang menghadiri kegiatan tersebut sebagian besar mahasiswa Ilmu Komputer, Teknologi Informasi dan umum.

Pelaksanaan kegiatan pengabdian ini diawali dengan penyampaian teori tentang *data science* dan dilanjutkan simulasi menggunakan *Google Colab* untuk implementasi metode *data science* untuk kasus nyata. Adapun dokumentasi acara

pengabdian tersebut, ditunjukkan pada Gambar 2, Gambar 3, Gambar 4, dan Gambar 5.



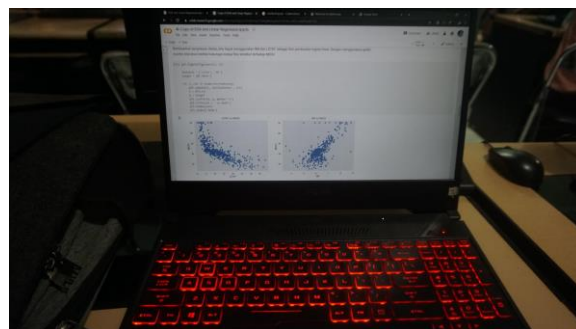
Gambar 2: Pemateri Pengenalan *Data Science*



Gambar 3: Pemateri Tutorial *Data Science*



Gambar 4: Peserta Tutorial *Data Science*



Gambar 5: Hasil Tutorial *Data Science* Oleh Peserta

Berdasarkan hasil pelatihan dan *workshop* yang telah dilakukan dapat memberikan pemahaman dan kemampuan terhadap peserta yang hadir tentang bidang *data science*. Tidak hanya itu, para peserta mampu mengimplementasikan metode-metode yang ada pada *data science* untuk mengolah data menjadi pengetahuan menggunakan *Google Colab* seperti prediksi menggunakan metode Regresi Linier, klasifikasi menggunakan metode Naive Bayes, SVM, k-NN, C4.5, dan lain-lain.

SIMPULAN

Berdasarkan kesimpulan yang dapat ditarik dari hasil pelatihan dan *workshop data science* adalah para peserta yang hadir dapat memberikan pemahaman dan kemampuan terhadap peserta yang hadir tentang bidang *data science*. Tidak hanya itu, para peserta mampu mengimplementasikan metode-metode yang ada pada *data science* untuk mengolah data menjadi pengetahuan menggunakan *Google Colab* seperti prediksi menggunakan metode Regresi Linier, klasifikasi menggunakan metode Naive Bayes, SVM, k-NN, C4.5, dan lain-lain.

UCAPAN TERIMA KASIH

Penulis ucapkan rasa terima kasih kepada Rektor Universitas Bumigora, Kaprodi Jurusan Teknologi Informasi yang telah memberikan izin dalam pelaksanaan pelatihan dan *workshop data science*.

DAFTAR PUSTAKA

- Gunawan, Hanes, & Catherine. (2020). *Information Systems Students' Study Performance Prediction Using Data Mining Approach*. 1–8.
<https://doi.org/10.1109/icic47613.2019.8985718>
- Hairani, H., Nugraha, G., Nurkholis Abdillah, M., & Innuddin, M. (2018). Komparasi Akurasi Metode Correlated Naive Bayes Classifier dan Naive Bayes Classifier untuk Diagnosis Penyakit Diabetes. *InfoTekJar (Jurnal Nasional Informatika Dan Teknologi Jaringan)*, 3(1), 6–11.
<https://doi.org/10.30743/infotekjar.v3i1.558>
- Hairani, H., Saputro, K. E., & Fadli, S. (2020). K-means-SMOTE untuk menangani ketidakseimbangan kelas dalam klasifikasi penyakit diabetes dengan C4.5, SVM, dan naive Bayes. *Jurnal Teknologi Dan Sistem Komputer*, 8(2), 89–93.
<https://doi.org/https://doi.org/10.14710/jtsisko.m.8.2.2020.89-93>
- Igual, L., & Segui, S. (2017). *Introduction to Data Science A Python Approach to Concepts, Techniques and Applications*. Springer.
- Jadhav, S. D., & Channe, H. P. (2016). *Comparative Study of K-NN , Naive Bayes and Decision Tree Classification Techniques*. 5(1), 2014–2017.
- Nugraha, G. S., & Hairani, H. (2018). Aplikasi Pemetaan Kualitas Pendidikan di Indonesia Menggunakan Metode K-Means. *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, 17(2), 13–23.
- Rahman, A. F. (2015). *Profesi Data Scientist, Seperti Apa Sih?*
<https://inet.detik.com/business/d-3079489/profesi-data-scientist-seperti-apa-sih>
- Sulistiyono, S., & Sulistiyowati, W. (2018). Peramalan Produksi dengan Metode Regresi Linier Berganda. *PROZIMA (Productivity, Optimization and Manufacturing System Engineering)*, 1(2), 82.
<https://doi.org/10.21070/prozima.v1i2.1350>
- Yuda Irawan. (2019). Implementation Of Data Mining For Determining Majors Using K-Means Algorithm In Students Of SMA Negeri 1 Pangkalan Kerinci. *Journal of Applied Engineering and Technological Science (JAETS)*, 1(1), 17–29.
<https://doi.org/10.37385/jaets.v1i1.18>
- Yusuf, O. (2018). *Talenta Data Scientist di Indonesia, Permintaan Tinggi Pasokan Kurang*. <https://Tekno.Kompas.Com/>.
<https://tekno.kompas.com/read/2018/03/23/09210087/talenta-data-scientist-di-indonesia-permintaan-tinggi-pasokan-kurang>