

SILABUS TRAINING DATA SCIENCE

Data Visualization using Python

Data yang sudah dibersihkan, divalidasi, dan dianalisa tentunya tidak akan dapat memberikan pandangan atau informasi yang memadai untuk dipresentasikan, karena format datanya masih dalam bentuk tabular. Disinilah pentingnya peranan visualisasi data untuk dapat memberikan informasi, dan pandangan yang memadai untuk dianalisa dan dijadikan sebagai dasar pengambilan keputusan. Dengan beragamnya format visual seperti grafik, chart, dan diagram mampu memberikan sebuah "storytelling" yang mendeskripsikan hasil analisa data.

Materi training ini membahas tentang implementasi bahasa Python untuk penggunaan package-package yang berhubungan dengan visualisasi data. Beberapa library visual yang dibahas yaitu Pandas, Matplotlib, Seaborn, Folium, dan Plotly Dash. Kemampuan visualisasi data ini merupakan skill yang wajib dimiliki untuk memvisualisasikan hasil analisa data dan hasil model yang dibuat di Machine Learning.

Setelah mengikuti training ini peserta diharapkan dapat memahami konsep dan mengimplementasikan bahasa Python untuk kebutuhan visualisasi data yang dapat membantu dalam pengambilan keputusan bisnis organisasi atau perusahaan.

Materi training disusun dari berbagai sumber dan media pembelajaran. Training disampaikan dalam bentuk pengenalan konsep, teori, dan praktek dalam porsi yang seimbang, serta dipandu oleh trainer atau instruktur yang telah berpengalaman di bidangnya.

Durasi dan Waktu

- Durasi Pelaksanaan: 2 (Dua) Hari
- Waktu Pelaksanaan: 09.00 - 15.00 WIB

Prasyarat

- Peserta harus sudah memiliki pengetahuan dasar pemrograman Python seperti: penggunaan library, variabel, tipe data, struktur data (list, tuple, dictionary, set, array), struktur keputusan, struktur perulangan, pembuatan fungsi, dan yang lainnya.

Outline Materi

■ Instalasi dan Konfigurasi Tools

Awal pertemuan training ini akan diawali dengan pembahasan tentang instalasi dan konfigurasi tools yang akan digunakan selama training. Tools yang akan digunakan yaitu JupyterLab yang dapat diinstal melalui [Pip](#) (packager installer for Python) atau distribusi [Anaconda](#). Ekstensi JupyterLab dan Google Colab juga dapat diinstal di [Visual Studio Code](#) sehingga dapat dijadikan sebagai pilihan utama tools lainnya yang dapat digunakan.

■ Pengenalan Library Visualisasi Data

Pada sesi ini dibahas tentang beragam library python yang dapat digunakan untuk visualisasi data seperti Pandas, Matplotlib, Seaborn, Folium, dan Plotly Dash beserta kegunaan dan perbedaannya.

■ Penggunaan Library Pandas

Pandas memiliki fitur untuk plotting data secara builtin dan terintegrasi langsung dengan dataframe dan bermanfaat untuk eksplorasi visual di tahap awal. Beragam plot yang dapat digunakan yaitu berupa: line chart, scatter plot, bar plot, histogram, dan box plot.

■ Matplotlib

Pandas hanya memberikan fitur visualisasi yang sederhana, untuk keperluan yang lebih kompleks dan penggunaan jenis plot yang tidak tersedia di pandas maka kita harus menggunakan library lain yang menyediakan kebutuhan tersebut yaitu Matplotlib. Beberapa teknik dan jenis visualisasi yang dibahas di sesi ini yaitu area plot, bubble plot, sub plot, dan pengaturan atribut plot seperti title, label, warna, pengaturan interval label, dan yang lainnya. Sub plot digunakan untuk menampilkan beberapa plot sekaligus dalam satu objek grafik.

■ Seaborn

Seaborn memberikan visualisasi yang lebih estetik atau menarik dibandingkan dengan matplotlib. Selain itu juga seaborn memberikan kemudahan untuk membuat plot multivariate dengan pair plot yang dapat memvisualisasikan korelasi antar beberapa variabel dalam satu plot. Regression plot disediakan di seaborn untuk dapat menampilkan line chart sekaligus dengan scatter plot dalam satu grafik. Seaborn dibangun di atas library matplotlib sehingga dapat disebut sebagai perluasan dari matplotlib.

■ Folium

Untuk keperluan pembuatan map yang dapat digunakan untuk memetakan data terhadap lokasi latitude dan longitude atau geospatial diperlukan library seperti Folium. Folium menyediakan beberapa fitur interaktif seperti popup label untuk

memberikan teks pada suatu titik, marker untuk memberikan penanda lokasi tertentu, label posisi mouse untuk mendapatkan koordinat, serta pembuatan garis yang dapat digunakan untuk menghubungkan antar titik dengan garis.

■ Plotly Dash

Pada sesi terakhir ini dibahas tentang pembuatan dashboard interaktif yang berbasis web. Dashboard interaktif memberikan visualisasi yang dinamis yang dapat difilter outputnya berdasarkan nilai dari variabel-variabel yang digunakan. Semua jenis plot yang sudah dibahas pada sesi sebelumnya dapat digunakan dalam dashboard sesuai dengan kebutuhan.

Note

- Regular Training dan Online Training tidak ada minimum kuota peserta.
- In-House Training dan Off-Site Training minimum kuota peserta 8 orang.
- Online Training menggunakan software Zoom Meeting.
- Hubungi kami untuk biaya training serta pembuatan Surat Penawaran.