



Data Science

Pokok Bahasan : *Description of Each Phase Analytics Lifecycle*

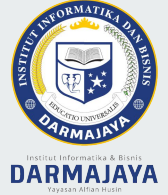
Dosen Pengampu : Hary Sabita, ST., MTI

18 April 2022

Outline :

1. *Phase Discovery Detail*
2. *Phase Data Preparation Detail - next*
3. *Phase Model Planning Detail - next*
4. *Phase Model Execution Detail - next*
5. *Phase Communicate Results Detail - next*
6. *Phase Operationalize Detail - next*

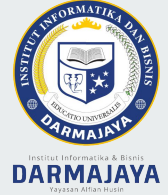
Next - Apa tujuan mempelajari ini ??



Tujuan yang akan dicapai :

1. Mahasiswa akan memahami secara detail deskripsi dari fase discovery
2. Mahasiswa mampu memahami setiap permasalahan pada fase ini

Next - Fase Discovery Overview



Phase Discovery - Overview

1. Data Scientist harus belajar dan menyelidiki masalah
2. Belajar mengembangkan konteks dan pemahaman
3. Belajar tentang sumber data yang dibutuhkan dan tersedia untuk proyek
4. Belajar merumuskan hipotesis awal bahwa nantinya dapat diuji dengan data.

Next - Apa saja yang harus dipahami??

Memahami Domain Bisnis

1. Data Scientist dapat menerapkan keilmuan komputasi dan komputatif yang mendalam serta dapat diterapkan di berbagai disiplin ilmu.
2. Data Scientist memiliki pengetahuan mendalam tentang metode, teknik, dan cara untuk menerapkan heuristik ke berbagai masalah bisnis dan konseptual.
3. Dapat menentukan seberapa banyak pengetahuan bisnis atau domain yang dibutuhkan.
4. Memastikan tim memiliki keseimbangan yang tepat antara pengetahuan domain dan keahlian teknis.

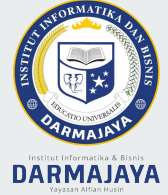
Memahami Sumber Daya

1. **Teknologi;** pertimbangkan alat dan teknologi yang tersedia yang akan digunakan tim dan jenis sistem yang diperlukan untuk fase selanjutnya untuk mengoperasikan model.
2. **Alat dan Sistem;** Apa saja yang dapat mendukung proyek analisis ini
3. **Data;** Jenis data yang tersedia bagi tim untuk proyek tersebut. Pertimbangkan apakah data yang tersedia cukup untuk mendukung tujuan proyek.
4. **Orang;** Agar proyek berhasil dalam jangka panjang, jenis keterampilan dan peran apa yang dibutuhkan untuk penerima model yang sedang dikembangkan?

Next - Membingkai Masalah

Memmingkai Masalah

1. Perlu dengan jelas mengartikulasikan situasi saat ini dan tantangan utamanya.
2. Mengidentifikasi tujuan utama proyek, mengidentifikasi kebutuhan apa dicapai dalam istilah bisnis, dan mengidentifikasi apa yang perlu dilakukan untuk memenuhi kebutuhan.
3. Mempertimbangkan tujuan dan kriteria keberhasilan proyek.



Identifikasi Stakeholder

1. Siapa saja yang terlibat pada bisnis tersebut
2. Apa peran mereka
3. Apakah mereka memiliki tujuan dalam bisnis tersebut

Next - Mengembangkan Hipotesis Awal

Mengembangkan Hipotesis Awal

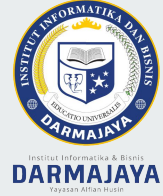
1. Membentuk ide-ide yang didapat dan mengujinya dengan data.
2. Membandingkan jawabannya dengan hasil percobaan atau tes untuk menghasilkan kemungkinan solusi tambahan untuk masalah.
3. Mengeksplorasi beberapa data awal untuk menginformasikan diskusi dengan stakeholder selama tahap pembentukan hipotesis.

Next - Mengidentifikasi Sumber Data Potensial

Mengidentifikasi Sumber Data Potensial

1. Identifikasi sumber data: cari sumber data untuk diuji.
2. Capture sumber data agregat: Ini memungkinkan tim untuk mendapatkan gambaran singkat tentang data dan melakukan eksplorasi lebih lanjut tentang daerah tertentu.
3. Tinjau data mentah: Dapatkan data awal.
4. Mengevaluasi struktur data dan alat yang diperlukan: Tipe dan struktur data menentukan alat mana yang tim dapat digunakan untuk menganalisis data.
5. Cakupan jenis infrastruktur data yang diperlukan untuk jenis masalah ini: Selain alat diperlukan, data mempengaruhi jenis infrastruktur yang diperlukan, seperti penyimpanan disk dan jaringan kapasitas kerja.

Next - Mengidentifikasi Sumber Data Potensial



Next..

Pertemuan berikutnya akan membahas secara rinci phase data preparation...



TERIMA KASIH



Tugas Kelompok (Tugas 4)

1. Mempresentasikan data yang sudah didapatkan..