

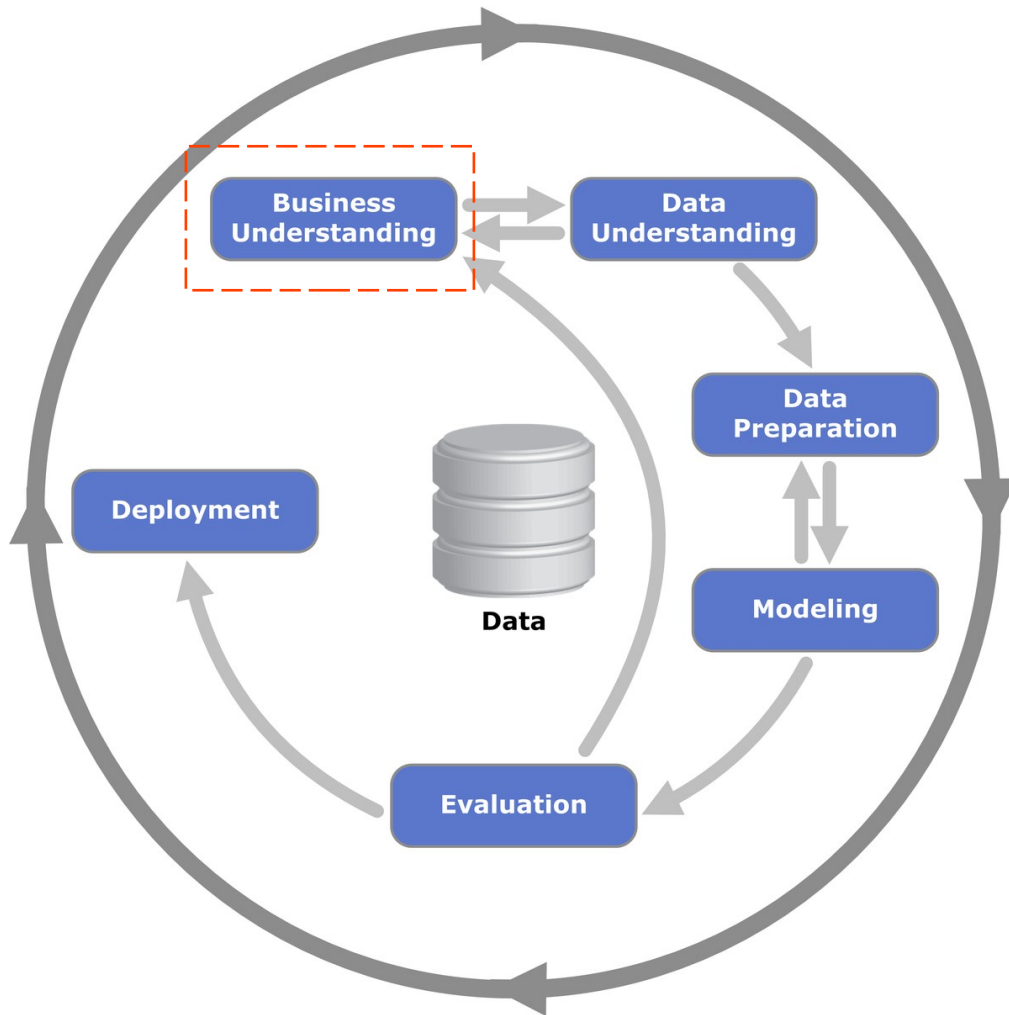
Business Understanding (Pemahaman Bisnis)

Dr. Koko Bachrudin, S.Kom., MMSI.

Outline

- Tahapan Business Understanding
- Proses dan output

Business Understanding



- Pernyataan tujuan bisnis
- Pernyataan tujuan data mining
- Pernyataan kriteria keberhasilan

Berfokus pada pemahaman tujuan dan persyaratan proyek dari perspektif bisnis, kemudian mengubah pengetahuan ini menjadi definisi masalah data mining dan rencana awal yang dirancang untuk mencapai tujuan tersebut.

CRISP-DM: Business Understanding

Business Understanding	Data Understanding	Data Preparation	Modeling	Evaluation	Deployment
Determine Business Objectives Background Business Objectives Business Success Criteria	Collect Initial Data Initial Data Collection Report Describe Data Data Description Report Explore Data Data Exploration Report Verify Data Quality Data Quality Report	Select Data Rationale for Inclusion/Exclusion Clean Data Data Cleaning Report Construct Data Derived Attributes Generated Records Integrate Data Merged Data Format Data Reformatted Data Dataset Dataset Description	Select Modeling Techniques Modeling Technique Modeling Assumptions Generate Test Design Test Design Build Model Parameter Settings Models Model Descriptions Assess Model Model Assessment Revised Parameter Settings	Evaluate Results Assessment of Data Mining Results w.r.t. Business Success Criteria Approved Models Review Process Review of Process Determine Next Steps List of Possible Actions Decision	Plan Deployment Deployment Plan Plan Monitoring and Maintenance Monitoring and Maintenance Plan Produce Final Report Final Report Final Presentation Review Project Experience Documentation
Determine Data Mining Goals Data Mining Goals Data Mining Success Criteria					
Produce Project Plan Project Plan Initial Assessment of Tools and Techniques					

Determine Business Objectives

- Memahami secara detail, dari perspektif bisnis, apa yang ingin dicapai oleh klien
- Temukan faktor-faktor penting, di awal, yang dapat memengaruhi hasil proyek
- Mengabaikan langkah ini akan meningkatkan kegagalan pengembangan proyek data science

What to include in a business objective



A growth-oriented intention



One or more actions



A measurement for success



A timeline to reach success

—zapier

Assess the Situation



- Pencarian fakta secara rinci tentang semua sumber daya, kendala, asumsi, dan faktor lain yang harus dipertimbangkan
- Menguraikan detail-detail spesifik

Inventory of Resource

- **Tugas 1: Riset Sumber Daya Perangkat Keras**
 - Perangkat keras apa yang perlu Anda dukung?
- **Tugas 2: Identifikasi Sumber Data dan Penyimpanan Pengetahuan**
 - Sumber data mana yang tersedia untuk penggalian data? Catat jenis dan format data.
 - Bagaimana data disimpan? Apakah Anda memiliki akses langsung ke gudang data atau basis data operasional?
 - Apakah Anda berencana untuk membeli data eksternal, seperti informasi demografi?
 - Apakah ada masalah keamanan yang mencegah akses ke data yang diperlukan?
- **Tugas 3: Identifikasi Sumber Daya Personel**
 - Apakah Anda memiliki akses ke pakar bisnis dan data?
 - Sudahkah Anda mengidentifikasi administrator basis data dan staf pendukung lainnya yang mungkin diperlukan?

Requirements, Asumptions, dan Constraint

- **Tugas 1: Menentukan Persyaratan:**

- Apakah ada batasan keamanan dan hukum pada data atau hasil proyek?
- Apakah semua orang sepakat dengan persyaratan penjadwalan proyek?
- Apakah ada persyaratan pada penyebaran hasil (misalnya, menerbitkan ke Web atau membaca skor ke dalam basis data)?

- **Tugas 2: Mengklarifikasi Asumsi**

- Apakah ada faktor ekonomi yang dapat memengaruhi proyek (misalnya, biaya konsultasi atau produk kompetitif)?
- Apakah ada asumsi kualitas data?
- Bagaimana sponsor proyek/tim manajemen berharap untuk melihat hasilnya? Dengan kata lain, apakah mereka ingin memahami model itu sendiri atau hanya melihat hasilnya?

- **Tugas 3: Memverifikasi Kendala**

- Apakah Anda memiliki semua kata sandi yang diperlukan untuk akses data?
- Sudahkah Anda memverifikasi semua kendala hukum pada penggunaan data?
- Apakah semua kendala keuangan tercakup dalam anggaran proyek?

Risk dan Contingencies

- Sebaiknya pertimbangkan pula risiko yang mungkin terjadi selama proyek berlangsung. Jenis risiko meliputi:
 - Penjadwalan (Bagaimana jika proyek berlangsung lebih lama dari yang diantisipasi?)
 - Keuangan (Bagaimana jika sponsor proyek menghadapi masalah anggaran?)
 - Data (Bagaimana jika data memiliki kualitas atau cakupan yang buruk?)
 - Hasil (Bagaimana jika hasil awal tidak sedramatis yang diharapkan?)
- Setelah mempertimbangkan berbagai risiko, buatlah rencana darurat untuk membantu mencegah bencana.
- Daftar Tugas
 - Dokumentasikan setiap risiko yang mungkin terjadi.
 - Dokumentasikan rencana darurat untuk setiap risiko.

Cost and Benefit

- Sertakan dalam analisis Anda estimasi biaya untuk:
 - Pengumpulan data dan data eksternal yang digunakan
 - Penerapan hasil
 - Biaya operasional
- Kemudian, pertimbangkan manfaat dari:
 - Sasaran utama yang ingin dicapai
 - Wawasan tambahan yang diperoleh dari eksplorasi data
 - Kemungkinan manfaat dari pemahaman data yang lebih baik

Determining Data Mining Goals



- Mengidentifikasi pelanggan bernilai tinggi berdasarkan data pembelian terkini
- Membangun model menggunakan data pelanggan yang tersedia untuk memprediksi kemungkinan churn untuk setiap pelanggan
- Menetapkan peringkat untuk setiap pelanggan berdasarkan kecenderungan churn dan nilai pelanggan
- Jika tujuan penggalian data ini terpenuhi, maka dapat digunakan oleh bisnis untuk mengurangi churn di antara yang paling

Data Mining Goals

Saat Anda bekerja dengan analis bisnis dan data untuk menentukan solusi teknis bagi masalah bisnis, ingatlah untuk menjaga hal-hal tetap konkret.

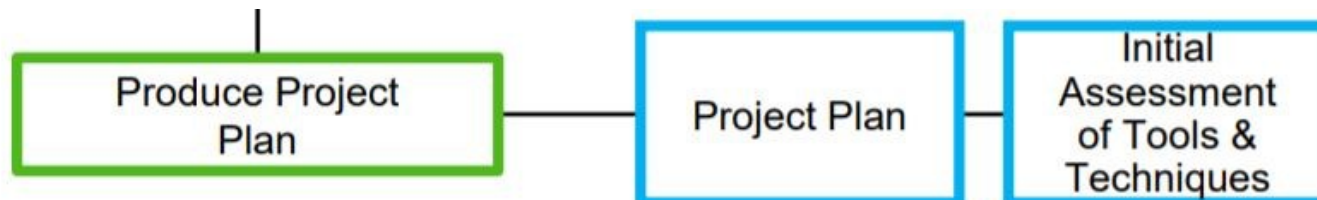
- Jelaskan jenis masalah data mining, seperti clustering, prediksi, atau klasifikasi.
- Dokumentasikan tujuan teknis menggunakan satuan waktu tertentu, seperti prediksi dengan validitas tiga bulan.
- Jika memungkinkan, berikan angka aktual untuk hasil yang diinginkan, seperti menghasilkan skor churn untuk 80% pelanggan yang ada.

Data Mining Success Criteria

Keberhasilan juga harus didefinisikan dalam istilah teknis agar upaya penambangan data Anda tetap pada jalurnya. Gunakan tujuan Data Mining yang ditentukan sebelumnya untuk merumuskan tolok ukur keberhasilan.

- Jelaskan metode untuk penilaian model (misalnya, akurasi, kinerja, dll.).
- Tetapkan tolok ukur untuk mengevaluasi keberhasilan. Berikan angka-angka tertentu.
- Tetapkan pengukuran subjektif sebaik mungkin dan tentukan penentu keberhasilan.
- Pertimbangkan apakah keberhasilan penerapan hasil model merupakan bagian dari keberhasilan penambangan data.

Producing a Project Plan



- Rencana proyek adalah dokumen utama untuk semua pekerjaan penambangan data Anda
- Rencana tersebut dapat memberi tahu semua orang yang terkait dengan proyek tentang tujuan, sumber daya, risiko, dan jadwal untuk semua fase penambangan data

Project Plan

- Saat membuat rencana, pastikan Anda telah menjawab pertanyaan berikut:
 - Sudahkah Anda mendiskusikan tugas proyek dan rencana yang diusulkan dengan semua pihak yang terlibat?
 - Apakah estimasi waktu disertakan untuk semua fase atau tugas?
 - Sudahkah Anda menyertakan upaya dan sumber daya yang diperlukan untuk menerapkan hasil atau solusi bisnis?
 - Apakah titik keputusan dan permintaan tinjauan disorot dalam rencana?
 - Sudahkah Anda menandai fase-fase yang biasanya memerlukan beberapa iterasi, seperti pemodelan?

Assessing Tools and Techniques

- Melakukan pemilihan perangkat dan tools yang sesuai
- Beberapa pertimbangan tools yang digunakan:
 - Ketersediaan support/komunitas
 - Keterbaruan tools dan teknik
 - Biaya yang harus dikeluarkan
 - Proses Development and depolyment

Terima Kasih Atas Perhatiannya

