

Metode CRISP-DM terdiri dari beberapa tahap diantaranya yakni:

1. *Business Understanding*

Business Understanding adalah tahap pertama dalam CRISP-DM. Beberapa hal yang dilakukan pada tahap ini seperti memahami kebutuhan serta tujuan dari sudut pandang bisnis selanjutnya mengartikan pengetahuan ke dalam bentuk pendefinisian masalah pada *data science* dan kemudian menentukan rencana serta strategi untuk mencapai tujuan *data science*.

Dalam studi kasus ini ada 2 masalah yang dimiliki oleh pemilik usaha ritel, yakni :

- a. Pemilik usaha ritel ingin menawarkan produk mereka kepada para pelanggan. Agar pelanggan mereka membeli produk yang ditawarkan, maka produk tersebut harus tepat sesuai dengan kebutuhan dan keinginan pelanggan. Oleh karena itu, pemilik usaha perlu memanfaatkan *data science* dalam mengembangkan *product recommendation engine* yang tepat untuk pelanggan mereka.
- b. Pemilik usaha ritel ingin membagi para pelanggan mereka ke dalam beberapa segmen. Pembagian pelanggan menjadi beberapa segmen ini ditujukan agar pemilik usaha dapat menentukan strategi pemasaran yang tepat untuk pelanggannya, seperti pemberian diskon, *cashback* untuk segmen pelanggan baru.

2. *Data Understanding*

Tahapan ini diawali dengan mengumpulkan data, mendeskripsikan data, serta mengevaluasi kualitas data. Tahap ini memberikan fondasi analitik untuk sebuah penelitian dengan membuat ringkasan (*summary*) dan mengidentifikasi potensi masalah dalam data. Jika ada masalah pada tahap ini yang belum terjawab, maka akan mengganggu pada tahap *modeling*.

Untuk menyelesaikan 2 permasalahan di atas, pemilik usaha ritel dapat memanfaatkan data pelanggan mereka, seperti:

- a. Data usia diperlukan untuk menentukan produk yang cocok ditawarkan berdasarkan usia pelanggan dan juga untuk menentukan pelanggan tersebut akan dimasukkan ke segmentasi pelanggan yang mana.
- b. Data jenis kelamin diperlukan untuk menentukan produk yang cocok ditawarkan berdasarkan jenis kelamin pelanggan dan juga untuk menentukan pelanggan tersebut akan dimasukkan ke segmentasi pelanggan yang mana.
- c. Data pekerjaan diperlukan untuk menentukan produk yang cocok ditawarkan berdasarkan pekerjaan, gaji dan kegiatan sehari-hari pelanggan serta untuk

menentukan pelanggan tersebut akan dimasukkan ke segmentasi pelanggan yang mana.

- d. Data alamat tempat tinggal diperlukan untuk menentukan produk yang cocok ditawarkan berdasarkan tempat tinggal pelanggan (daerah dingin atau panas, perkotaan atau perdesaan, dan lain-lain) dan juga untuk menentukan pelanggan tersebut akan dimasukkan ke segmentasi pelanggan yang mana.
- e. Data jenis produk yang sering dibeli diperlukan untuk menentukan produk yang cocok ditawarkan berdasarkan riwayat jenis produk yang pernah dibeli dan juga untuk menentukan pelanggan tersebut akan dimasukkan ke segmentasi pelanggan yang mana.
- f. Data kisaran harga produk yang sering dibeli diperlukan untuk menentukan produk yang cocok ditawarkan berdasarkan kisaran harga produk yang suka dibeli dan juga untuk menentukan pelanggan tersebut akan dimasukkan ke segmentasi pelanggan yang mana.
- g. Data waktu ketika biasanya mereka membeli produk diperlukan untuk menentukan kapan produk perlu ditawarkan dan juga untuk menentukan pelanggan tersebut akan dimasukkan ke segmentasi pelanggan yang mana.

3. *Data Preparation*

Tahap *data preparation* ini terdiri dari beberapa proses diantaranya yakni pembersihan data, penyatuan data, pemilihan atribut, dan transformasi data. Proses ini perlu dilakukan sebelum data mulai digunakan untuk pemodelan. Pada tahap ini apabila terdapat data pelanggan yang tidak lengkap, maka data pelanggan tersebut tidak dapat digunakan dan perlu dibersihkan sehingga tidak mengganggu ketika tahap pemodelan nanti dimulai.

4. *Modeling*

Pada tahap pemodelan data mulai diolah dengan menggunakan algoritma yang telah dipilih. Setelah data dimodelkan dan diolah dengan algoritma maka didapatlah beberapa pola pelanggan sesuai dengan permasalahan yang ingin dipecahkan. Untuk menyelesaikan masalah pertama yakni bagaimana cara menentukan produk yang cocok untuk ditawarkan ke pelanggan, maka perlu melihat data dan pola pembelian produk oleh pelanggan. Misalnya, ketika ada pelanggan dengan umur 18-21 tahun, jenis kelaminnya perempuan, pekerjaannya mahasiswa, alamatnya di jabodetabek, barang yang sering dibelinya adalah peralatan rumah dan ATK, harganya murah dan sering membeli ketika awal bulan, maka pelanggan tersebut cocok untuk

ditawarkan barang seperti mangkok, piring, makanan, minuman, jaket, hoodie untuk usia 18-21, buku dan pena. Barang tersebut diperkirakan cocok karena pelanggan adalah mahasiswa yang sedang berkuliah dan tinggal di kos atau kontrakan di daerah jabodetabek sehingga perlu barang-barang yang berkaitan dengan peralatan rumah dan perkuliahan.

Sedangkan untuk permasalahan kedua, kita juga akan membutuhkan data-data pelanggan sebelumnya. Data tersebut akan diolah dan dimodelkan sehingga nantinya akan menghasilkan beberapa segmen pelanggan. Berikut adalah contoh hasil pembagian pelanggan menjadi beberapa segmen:

- a. Segmen 1 merupakan pelanggan-pelanggan yang bekerja di perkantoran sebagai profesional dan manajer dan memiliki pendidikan yang cukup tinggi, memiliki rumah dan mobil. Berdasarkan informasi yang didapat, pelanggan pada segmen ini belum memiliki bayi, pendapatan tahunan tidak terlalu tinggi tetapi mereka menyukai berbelanja terutama "*fashion*". Fokus konsumsi pada segmen ini adalah kosmetik yang berkualitas tinggi, CD orisinal, majalah dan produk peralatan rumah tangga pada umumnya. Pada umumnya mereka tidak tertarik membeli makanan dan minuman karena penghasilan yang cukup tinggi. Strategi marketing yang cocok untuk segmen ini adalah memberikan kualitas produk dan layanan yang lebih tinggi untuk menarik pelanggan karena pelanggan yang berada disegmen ini tidak terlalu tertarik dengan potongan harga.
- b. Segmen 2 merupakan pelanggan-pelanggan yang bekerja di pabrik sebagai pekerja dan mekanik, dan memiliki penghasilan yang cukup rendah, beberapa memiliki rumah, berbelanja makanan, keperluan rumah tangga, produk balita. sebagian besar merupakan wanita yang suka ke supermarket untuk membeli lebih banyak barang. Pada umumnya mereka tertarik terhadap diskon promosi, sehingga keuntungan tidak terlalu tinggi.
- c. Segmen 3 merupakan pelanggan yang memiliki pendidikan rendah, penghasilan rendah yang hampir sama dengan pekerja pabrik. Pada umumnya mereka membeli komoditas lebih banyak pada saat diskon yang cukup besar. Strategi pemasaran untuk segmen ini adalah memberikan diskon promosi terhadap komoditas tersebut sehingga lebih menarik pelanggan.

5. *Evaluation*

Tahap *evaluation* dilakukan untuk menguji hasil permodelan data yang telah dibuat. Apabila hasil permodelan masih kurang tepat, maka proses perlu diulang dari tahap awal lagi.

Tetapi jika hasil permodelan dirasa sudah tepat maka pengembangan bisa dilanjutkan ke tahap selanjutnya yakni *deployment*.

6. *Deployment*

Deployment merupakan tahap terakhir dimana hasil permodelan yang telah diuji selanjutnya akan mulai disebar dan diterapkan ke seluruh pelanggan.