

## PERAN DATA SCIENCE DALAM SISTEM REKOMENDASI PADA BISNIS INDUSTRI RETAIL

Data Science merupakan suatu kegiatan mengolah data dengan memanfaatkan ilmu statistik, matematika, algoritma, aljabar, dan pemrograman komputer dimana nantinya dihasilkan suatu informasi. Berbicara mengenai data, dalam dunia industri data adalah komponen yang sangat penting yang dapat digunakan dalam pengambilan sebuah keputusan. Data sangat bermanfaat bagi setiap pembuat keputusan dalam perusahaan. Data science banyak sekali digunakan dalam berbagai industri. Salah satunya adalah bisnis dalam industri retail. Bisnis retail merupakan bisnis yang melibatkan penjualan barang/jasa kepada konsumen dalam jumlah eceran. Pembelian barang/jasa yang dilakukan konsumen ini bertujuan untuk dikonsumsi sendiri, bukan untuk diperjualbelikan kembali. Retail memiliki banyak data transaksi dari konsumennya. Data science berperan penting dalam bisnis retail. Dengan data tersebut, data science dapat digunakan untuk mendapatkan *insight* terkait konsumennya dan tren pasar. *Use Case* data science dalam industri retail yang mampu meningkatkan tingkat kepuasan pelanggan salah satunya adalah *Market Basket Analysis (MBA)*. Dalam MBA bisnis retail akan menganalisa konsumen atau pelanggannya dengan melihat riwayat transaksinya. Dengan mengetahui riwayat transaksi dari konsumen, banyak *insight* yang bisa didapatkan salah satunya adalah rekomendasi produk untuk konsumen. Contohnya jika seseorang baru saja membeli handphone, ada kemungkinan akan membeli casing untuk handphone tersebut. Pada *use case* ini, MBA digunakan untuk **membangun sistem rekomendasi bagi konsumen online shop**. Proses data mining yang akan digunakan nantinya adalah *Cross Industry Standart Process Data Mining (CRISP-DM)*.

Membangun mesin rekomendasi yang baik merupakan bagian tak terpisahkan dari kesuksesan dalam ritel online, karena memberikan rekomendasi yang dipersonalisasi adalah salah satu cara terbaik untuk mendorong penjualan. Data memiliki peran yang penting dalam membangun sistem rekomendasi. Bayangkan ketika menjalankan situs web e-niaga dan ingin membuat sistem yang menawarkan rekomendasi hasil personalisasi. Untuk melakukannya, perlu membuat sistem yang menandai produk berdasarkan karakteristik tertentu, dan juga membuat segmen pengguna berdasarkan preferensi mereka.

Tahapan CRISP-DM dalam membuat sistem rekomendasi pada online shop dijabarkan sebagai berikut. Pada tahap awal perlu adanya **pemahaman terkait bisnis** (*Business Understanding*) online shop.

- a) Tentukan tujuan bisnis terlebih dahulu, yaitu menyediakan dan merekomendasikan suatu produk untuk konsumen.
- b) Memudahkan konsumen online shop untuk mendapatkan produk yang mungkin akan dibeli.
- c) Menentukan strategi untuk menyelesaikan tujuan dengan menggunakan *Market Basket Analysis (MBA)* dengan metode K-Means Clustering untuk mendapatkan kombinasi-kombinasi produk terkait.

Selanjutnya adalah **Data Understanding**, dilakukan pengumpulan data terlebih dahulu untuk mendapatkan data yang valid. Data yang dikumpulkan adalah data hasil transaksi penjualan online shop 1 tahun terakhir, Riwayat transaksi tersebut dipilih agar cakupannya tidak meluas dan tidak terlalu jauh dengan data terbaru. Dalam data transaksi tersebut terdapat atribut no transaksi, tanggal, bulan, customer, umur customer, merek, type, harga jual, harga beli, margin, tanggal iklan, iklan, total iklan, total margin, dan pajak. Ketiga, dilakukan **Preprocessing Data**. Pada tahap ini dilakukan untuk mempersiapkan data serta memproses data mentah sebelum dilakukannya proses data mining. Hal ini dilakukan agar data dapat terintegrasi dengan sistem yang akan digunakan pada saat proses data mining. Pada tahap Preprocessing Data, terdapat 3 tahapan yaitu data cleaning, data selection, dan data transformation. Dalam data selection, dipilih atribut yang tepat agar terintegrasi dengan baik yaitu no transaksi, bulan transaksi, customer, umur customer, merk, tipe, harga jual dan harga beli. Pemilihan atribut tersebut dilakukan karena hanya data tersebut yang diperlukan dalam proses data mining seperti atribut no transaksi digunakan sebagai penomoran setiap transaksi, atribut bulan transaksi diperlukan untuk mengetahui jumlah penjualan setiap bulanya, atribut customer digunakan untuk mengetahui nama customer disetiap transaksinya, atribut umur customer digunakan untuk mengetahui rentang umur berapa saja yang memiliki kecenderungan melakukan transaksi pada produk tertentu, merk dibutuhkan untuk mengetahui nama produk pada setiap transaksinya, atribut type dibutuhkan untuk melengkapi informasi dari atribut merk karena type merupakan type dari merk itu sendiri, atribut harga jual dan harga beli digunakan untuk mengetahui margin pada setiap transaksinya. Dilakukan transformasi data agar data siap disimpan di warehouse.

Selanjutnya adalah **Modelling dengan Machine Learning**. Akan dilakukan pengelompokkan (clustering) dengan metode K-Means Clustering. Hal ini dilakukan untuk memaksimalkan dan meminimalkan kemiripan intra kelas. Metode K-Means Clustering ini termasuk salah satu metode dalam machine learning yang mempunyai sifat tidak terawasi atau tidak memerlukan target output (unsupervised). Ada beberapa tahapan dalam metode ini, tahapannya adalah sebagai berikut.

1. Menentukan sentroid awal secara acak pada data sampel yang digunakan.
2. Mengelompokkan datanya ke dalam suatu cluster dengan cara menghitung jarak terdekat menggunakan rumus *euclidian distance*.

3. Setelah itu menentukan cluster dan centroid baru agar dapat menyesuaikan data dengan centroid sebenarnya.
4. Setelah centroid baru ditemukan, lakukan lagi langkah untuk menghitung *euclidean distance* dengan centroid yang baru secara berulang sampai penentuan cluster menjadi tetap sehingga didapat cluster tetapnya.
5. Kemudian mencari nilai rata-ratanya.

Nilai rata-rata adalah angka representasi atau biasa juga disebut sebagai mean dari suatu kelompok data yang mewakili data secara keseluruhan. Dapat digunakan untuk membandingkan kualitas suatu kelompok dengan kelompok lain, mewakili kondisi suatu kelompok dengan 1 angka dan mendeskripsikan suatu kelompok dengan singkat. Berdasarkan perhitungan yang dihasilkan dari masing-masing variable, didapatkan karakteristik pada masing-masing cluster.

Tahap selanjutnya adalah **pengimplementasian** dari rancangan sistem yaitu menampilkan hasil clustering terhadap produk toko online. Dihasilkan beberapa kluster dengan karakteristik masing-masing. Kemudian dilakukan pengujian terhadap semua data produk dengan memasukkan beberapa titik pusat secara acak kemudian dilakukan perhitungan nilai silhouette coefficient. Dihasilkan perangkat lunak yang dapat mengelompokkan data berdasarkan data transaksi penjualan online shop. Untuk diagram alir tahapan CRISP- DM dalam pembuatan system rekomendasi terlampir.

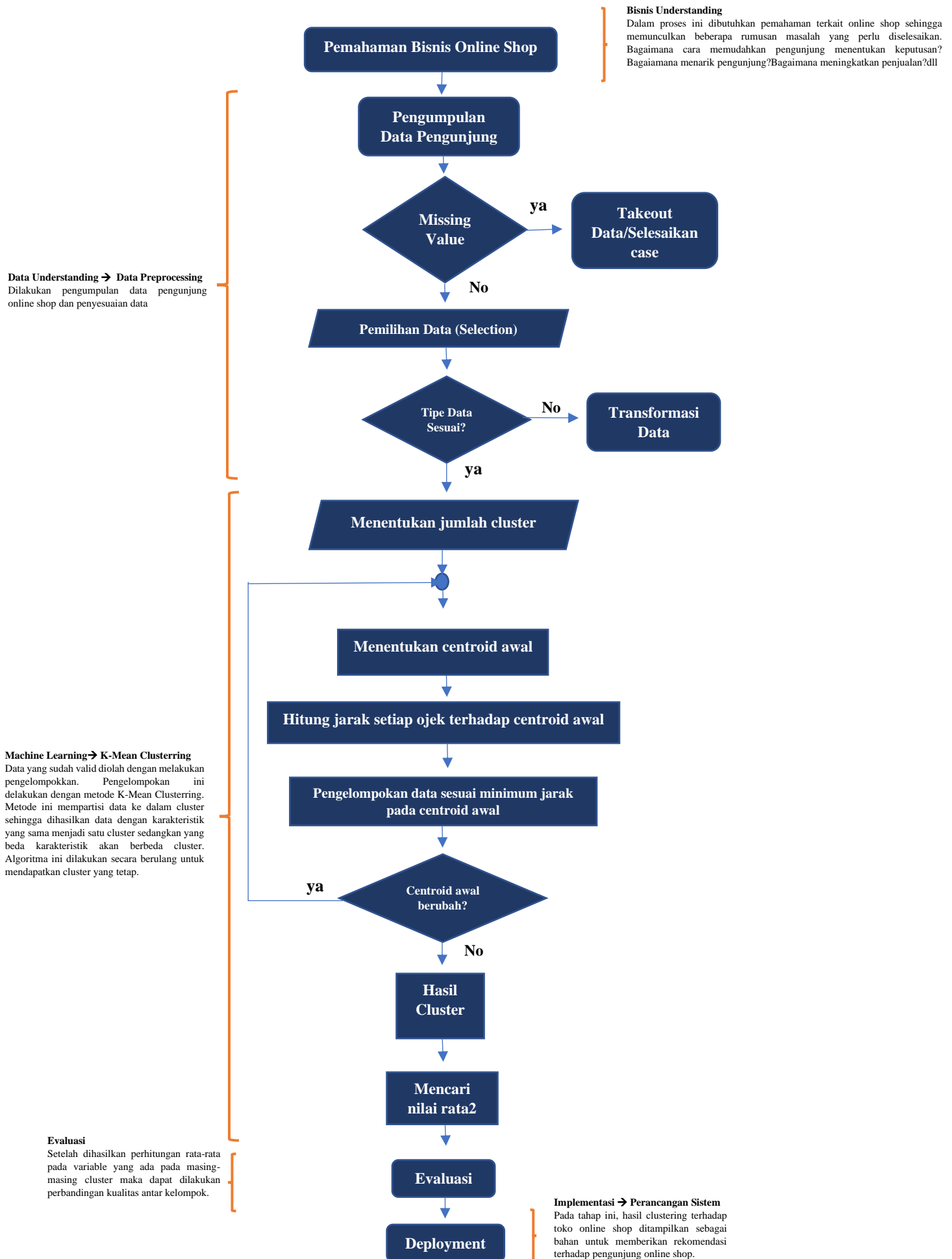
Impact penggunaan data science dalam industri retail yang digunakan dalam membuat sistem rekomendasi adalah dapat mendorong peningkatan penjualan produk pada *online shop*. Selain itu dengan adanya hasil clustering dari tahapan pemodelan pada sistem rekomendasi, dapat diimplementasikan pada toko kelontong, supermarket, minimarket, dan toko-toko kecil lainnya untuk membantu dalam sistem penataan barang di rak toko/ tempat jualan dalam industri retail *offline* dan penataan katalog untuk *online*.

## DAFTAR PUSTAKA

- Bank USA, H. (n.d.). *Memahami Apa Itu Bisnis Retail dan Karakteristiknya*. Retrieved March 18, 2023, from [https://www.hsbc.co.id/1/PA\\_esf-ca-app-content/content/indonesia/personal/offers/news-and-lifestyle/files/articles/html/201906/memahami-apa-itu-bisnis-retail-dan-karakteristiknya.html](https://www.hsbc.co.id/1/PA_esf-ca-app-content/content/indonesia/personal/offers/news-and-lifestyle/files/articles/html/201906/memahami-apa-itu-bisnis-retail-dan-karakteristiknya.html)
- Diana, C. (2020, December 15). KEHADIRAN DATA SCIENCE DI INDUSTRI RITEL. 2020. <https://algorit.ma/blog/data-science/tranformasi-retail-data-science/>
- Harahap, M., Lubis, Y., & Situmorang, Z. (n.d.). *Data Science bidang Pemasaran: Analisis Prilaku Pelanggan*.
- Lubis, F. R., Harahap, M. K., & Husein, A. M. (n.d.). *Analisis Prediktif untuk Keputusan Bisnis: Peramalan Penjualan*.

Naufal, F., Herry Chrisnanto, Y., & Kania Ningsih, A. (2022). Sistem Rekomendasi Penawaran Produk Pada Online Shop Menggunakan K-Means Clustering. *Informatics and Digital Expert (INDEX)*, 4(1), 10–17. <https://doi.org/10.36423/index.v4i1.879>

## Lampiran Tahapan CRISP-DM



**Link Github**

<https://ardiabc.github.io/>

**Akun LinkedIn**

<https://www.linkedin.com/in/ardia-fatma-sari-20483625b/>