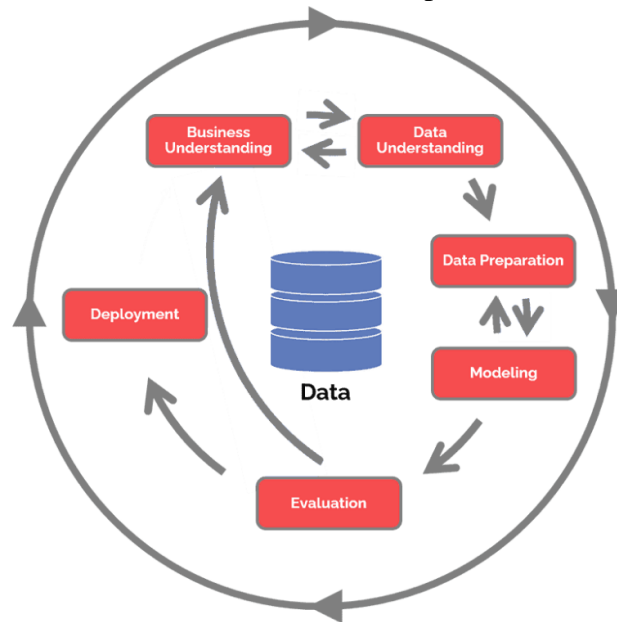


Recommendation engine

Mesin rekomendasi merupakan subkelas dari pembelajaran mesin yang umumnya berurusan dengan peringkat produk atau pengguna. Mesin rekomendasi sering dijumpai di dunia retail. Sistem ini bekerja dengan memprediksi peringkat yang mungkin diberikan pengguna untuk item tertentu. Sistem rekomendasi merupakan hal yang pernah populer untuk diperdebatkan bagaimana sistem ini dapat bekerja. Sistem ini merupakan model yang dibuat oleh perusahaan besar yang tidak mudah ditafsirkan. Item yang direkomendasikan seringkali merupakan item yang pengguna inginkan atau butuhkan tetapi pengguna bahkan belum menyadari bahwa mereka membutuhkannya sampai item tersebut direkomendasikan kepada mereka.



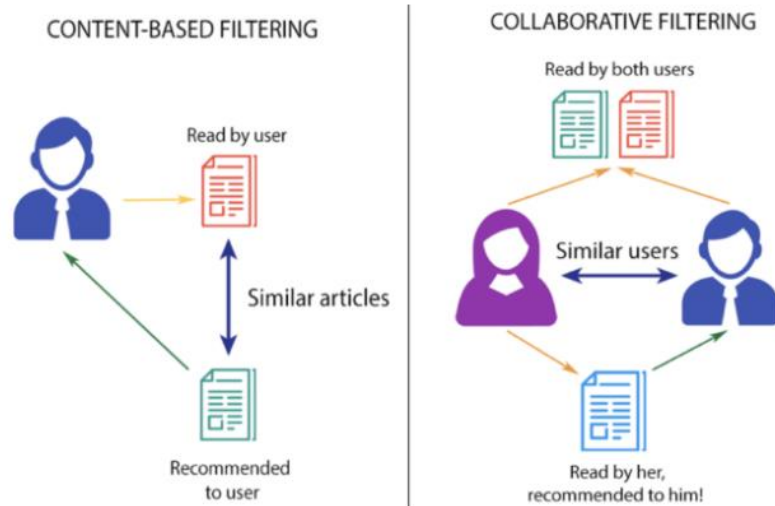
Sumber : wikipedia

Metodologi CRISP-DM merupakan metode yang ditempuh untuk mensukseskan sebuah mesin rekomendasi. Terdapat 6 langkah yang akan ditempuh, diantaranya adalah business understanding, data understanding, data preparation, modeling, evaluation, dan deployment. Business understanding berkaitan dengan pemahaman yang mendalam tentang kebutuhan pelanggan. Tujuan dari dibuatnya mesin rekomendasi adalah untuk menemukan barang yang sesuai dengan preferensi, kebutuhan, dan keinginan pengguna. Dengan adanya mesin rekomendasi ini dapat membantu pengguna serta dapat meningkatkan transaksi jual beli. Selanjutnya, Langkah data understanding, jadi pada langkah ini berfokus untuk mengidentifikasi, mengumpulkan, dan menganalisis kumpulan data untuk mencapai tujuan dari pembuatan suatu proyek. Proses pengumpulan data dilakukan dengan mengambil data yang berupa data implisit maupun eksplisit. Data eksplisit diperoleh dari data yang diinput oleh pengguna seperti ketika memberikan penilaian dan ulasan pada suatu produk. Sedangkan data implisit diperoleh dari aktivitas pengguna. Sebagai contoh aktivitas yang dilakukan pada aplikasi e-commerce, yaitu mencari barang, memilih barang, menampilkan barang, memasukkan barang ke keranjang, sampai membuat pesanan. Data aktivitas

atau perilaku dapat dengan mudah dikumpulkan karena perusahaan dapat menyimpan log aktivitas pengguna di situs perusahaan. Keunggulan dari pendekatan ini adalah aktivitas pengguna sudah pasti ada sebab pengguna memang sudah menggunakan aplikasi, namun kekurangannya adalah lebih sulit menganalisis data. Misalnya, ketika memfilter log yang diperlukan dari log yang kurang diperlukan. Semakin banyak data yang disediakan maka akan semakin baik rekomendasi yang diberikan. Banyaknya data tersebut tentunya membuat proyek rekomendasi menjadi proyek data besar sehingga memerlukan database. Database yang digunakan dapat berupa NoSQL, database SQL standar, atau lainnya.

Proses analisis data dilakukan dengan metode analisis yang berbeda. Beberapa jenis metode, diantaranya adalah real time system, batch analysis, dan near real time analysis. Real time system diperlukan untuk memberikan rekomendasi secara langsung. Batch analysis memungkinkan memproses data secara berkala, sistem ini berfungsi baik untuk mengirim email di kemudian hari. Near real time analysis memungkinkan mengumpulkan data dengan cepat sehingga analisis dapat di update setiap beberapa menit atau detik sehingga sistem ini efektif ketika memberikan rekomendasi selama sesi penelusuran yang sama.

Data preparation sering disebut sebagai “data munging” dimana pada langkah ini menyiapkan kumpulan data akhir untuk pemodelan. Pada langkah ini terdiri atas proses memilih data, membersihkan data, konstruksi data, mengintegrasikan data, dan format data. Selanjutnya, pemodelan untuk membangun sistem rekomendasi dilakukan dengan proses filtering, yaitu metode content based, cluster, collaborative. Metode content based bekerja dengan produk yang direkomendasikan memiliki karakteristik serupa dengan apa yang dilihat ataupun disukai pengguna. Tujuan metode ini adalah untuk membuat profil untuk setiap pengguna dan setiap item. Metode cluster bekerja dengan produk yang direkomendasikan tidak memperdulikan apa yang telah dilakukan pengguna lain sehingga produk rekomendasi berkerja dengan baik. Sedangkan metode kolaboratif bekerja dengan asumsi bahwa ketika dua pengguna menyukai produk yang sama maka dimungkinkan akan menyukai produk yang sama sekarang atau di masa depan. Metode pemfilteran ini memungkinkan membuat prediksi berdasarkan selera pengguna. Selanjutnya integrasi data dilakukan dengan menggunakan algoritma K-Nearest, kesamaan Jaccard, algoritma Dijkstra, kesamaan cosinus. Kesamaan cosinus untuk merekomendasikan berdasarkan item yang paling mirip dengan item yang telah dibeli pengguna. Kesamaan Jaccard dilakukan dengan membandingkan vektor item dan mengembalikan item yang paling mirip. Kesamaan ini hanya berguna ketika vektor mengandung nilai biner.



<http://datameetsmedia.com/an-overview-of-recommendation-systems/>

Rekomendasi dikatakan berhasil jika sistem dapat menampilkan item yang sesuai dengan keinginan atau kebutuhan pengguna. Rekomendasi yang baik adalah masalah tersendiri yang dihadapi banyak perusahaan. Kualitas dari sistem dapat dilihat dari cakupan dan akurasi. Akurasi berkaitan dengan ketepatan sistem dalam menampilkan dari rekomendasi yang benar dari total rekomendasi yang mungkin sementara cakupan mengukur fraksi objek dalam ruang pencarian yang dapat diberikan oleh sistem. Dalam situasi ideal, perusahaan memerlukan reaksi dan respon pengguna untuk dapat meningkatkan tingkat akurasi rekomendasi perusahaan, namun hal ini juga sulit untuk dicapai. Ukuran akurasi statistik umum untuk mengevaluasi akurasi pemberi rekomendasi, diantaranya adalah RMSE (Root Mean Square Error), MAE (Mean Absolute Error), Correlation. MAE adalah yang paling populer dan umum digunakan. Semakin rendah nilai MAE dan RMSE maka semakin akurat mesin rekomendasi memprediksi peringkat pengguna.

Sebuah model tentunya sia-sia jika tidak memiliki daya guna kecuali pengguna dapat merasakan hasilnya. Pada fase ini dilakukan penyebaran, perencanaan pemantauan dan pemeliharaan serta meninjau proyek. Sistem rekomendasi memiliki pengaruh yang cukup besar khususnya di dunia retail. Rekomendasi engine yang akurat tentunya akan memudahkan pengguna dalam menemukan item yang diinginkan. Dengan demikian akan berpengaruh pada transaksi jual beli yang akan meningkat selinear dengan keakuratan sistem. Namun jika sistem kurang akurat, akan memberi pengaruh yang sebaliknya.

Referensi :

Maruti, Techlabs. 2017. How Does a Recommendation Engine Really Work?.
<https://towardsdatascience.com/how-does-a-recommendation-engine-really-work-656bdf12a5fc>

Nick Hotz. 2022. What is CRISP DM?. <https://www.datascience-pm.com/crisp-dm-2/>

Neerja, Doshi. 2018. Recommendation System – Models and Evaluation.
<https://towardsdatascience.com/recommendation-systems-models-and-evaluation-84944a84fb8e>

Roger Chua. 2019. A simple way to explain the recommendation. <https://medium.com/voice-tech-podcast/a-simple-way-to-explain-the-recommendation-engine-in-ai-d1a609f59d97>