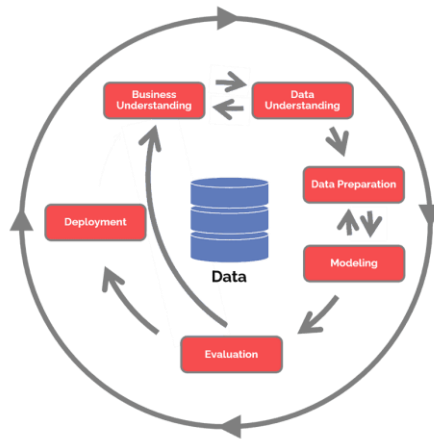


Github link : <https://github.com/kamandanu9/ML-Zenius->

Data Science Use Case in Healthcare

Di zaman yang udah canggih kayak tahun 2022 sekarang, udah banyak banget kemudahan yang sering kita pake. Misalnya pesen makanan dari aplikasi ojek *online*, nonton film favorit tanpa perlu keluar rumah, atau nanya sama komputer virtual yang kerjanya seolah-olah mirip manusia biasa yang banyak dipake di bidang pendidikan dan kesehatan. Nah canggihnya fitur-fitur ini berkat kemajuan teknologi, salah satunya adalah *Artificial Intelligence* (AI) atau kecerdasan buatan. Gampangnya, AI adalah teknologi yang ngelatih komputer atau mesin sehingga punya kemampuan otak mirip sama manusia. AI sendiri masih bisa diklasifikasikan jadi *machine learning* dan *deep learning*. Di artikel ini, gue akan ngebahas bagian *machine learning* yang juga beririsan dengan *data science*, bidang pekerjaan yang lagi naik daun belakangan ini, khususnya gimana penerapan *data science* di industri kesehatan.

Contoh gampang penerapan *data science* di bidang industri kesehatan adalah munculnya fitur *virtual assistance* di aplikasi-aplikasi kesehatan sehingga kita gak perlu nanya langsung ke praktisi kesehatan tapi komputer bakal berusaha ngejawab pertanyaan-pertanyaan kita semaksimal mungkin. Biasanya kita bakal ditanyain terkait keluhan kesehatan yang kita alamin, nah ketika kita selesai ngetik keluhan kita biasanya komputer bakal ngasih rekomendasi artikel yang ditulis dokter atau praktisi kesehatan lainnya terkait keluhan yang kita alamin. Keren, kan? Tapi tentunya sebelum kita bisa nikmatin semua fitur ini pasti ada proses di balik semua ini. *Data scientist* tentu nyiapin semua membantu untuk nyiptain fitur ini. Ada metode untuk nyiapin fitur ini yang bisa dilakukan para *data scientist*. Metode ini berkaitan dengan *data mining* atau dapat diartikan sebagai proses pengumpulan dan pengolahan data yang bertujuan untuk mengekstrak informasi penting pada data. Ada standarisasi pada dunia industri terkait *data mining* yang disebut CRISP-DM. CRISP-DM merupakan akronim dari The Cross Industry Standar Process for Data Mining yang dikembangkan di Eropa^[1]. CRISP-DM terdiri dari 6 tahapan yang bisa lo liat di gambar ini.



Gambar 1. Diagram alir proses *data mining* dengan metode CRISP-DM (Sumber : <https://www.datascience-pm.com/crisp-dm-2/>)

Hal yang pertama dilakukan pada proses *data mining* adalah *business understanding*. Disini *data scientist* berusaha memahami kebutuhan pelanggan. Pada kasus *virtual assistance* di industri kesehatan tentunya fokus bisnis di sektor kesehatan dan kebutuhan pelanggan adalah solusi-solusi kesehatan atas keluhan yang dialaminya. Tahap kedua ialah *data understanding*. Fokus tahap ini adalah untuk mengidentifikasi, mengumpulkan, dan menganalisis *dataset* yang didapatkan. Disini *data scientist* bisa mulai mencocokkan kira-kira kalau ada keluhan atau sesuatu yang mirip dengan A maka bisa diidentifikasi bahwa itu penyakit A yang tentunya datanya dihimpun dari dokter atau praktisi kesehatan terkait. Banyak eksplorasi disini yang dilakukan para *data scientist*. Tahap ketiga ialah *data preparation*. Sesuai namanya, pada tahap ini dilakukan persiapan *dataset* akhir sebelum dimodelkan menjadi suatu algoritma. Tahap keempat yaitu *modelling*. *Modelling* dapat diartikan sebagai suatu bentuk penulisan algoritma berdasarkan masalah yang ada untuk menjawab target bisnis yang telah dijabarkan pada tahap *business understanding*. Bentuk *modelling* dapat berupa klasifikasi, regresi, dll. Para *data scientist* akan mencari model paling efektif yang dapat menjawab kebutuhan pelanggan yang membutuhkan solusi akan keluhan yang mereka alami. Untuk kasus ini, bentuk *modelling* yang cocok adalah klasifikasi yang berbentuk pengelompokan penyakit berdasarkan usia, organisme penyebab, dll. Lalu setelah dimodelkan dalam suatu algoritma dan diujikan pada skala kecil model dievaluasi. Tahap kelima disebut *evaluation*. Model akan terus direview apakah memiliki *error* yang tinggi atau tidak. Bila masih memiliki *error* yang tinggi, maka langkah perlu diulangi dari langkah pertama. Sebaliknya, bila *error* rendah dapat dilanjutkan ke bagian pengimplementasian. Tahap keenam atau terakhir disebut juga *deployment*. Tahap ini berisikan dokumen pengembangan model yang digunakan. Fitur yang udah dirasain pengguna kayak *virtual assistance* yang gue sebut diatas merupakan fitur yang udah melalui berbagai tahap dan udah dideploy^[2].

REFERENSI

- [1]. <https://www.djkn.kemenkeu.go.id/artikel/baca/15134/CRISP-DM-Sebagai-Salah-Satu-Standard-untuk-Menghasilkan-Data-Driven-Decision-Making-yang-Berkualitas.html>
diakses 18 September 2022 11.51 WIB
- [2]. <https://www.datascience-pm.com/crisp-dm-2/> diakses 18 September 2022 15.44 WIB