

# Demo Day Proposal

**Machine Learning with TensorFlow Training**

**Professional Academy Digital Talent Scholarship 2022**

|                      |  |
|----------------------|--|
| <b>Group Number</b>  | HC_1   |
| <b>Dataset</b>       | Healthcare   |
| <b>Name - DTS ID</b> | <ol style="list-style-type: none"><li>1. Anthonius Adi Nugroho (152236035100-32)</li><li>2. Alfonsus Haryo Sangaji (152236035101-664)</li><li>3. Alma Vita Sophia (152236035101-224)</li><li>4. Alda Raharja (152236035101-478)</li><li>5. Zia Firnanda (152236035100-942)</li></ol> |

## **Selected Theme: Health Care**

### **Title of the Project: Disease Classification Using Neural Network**

#### **Latar belakang masalah**

Data merupakan kumpulan catatan atas fakta yang terjadi. Data merupakan bentuk jamak dari bahasa latin yang berarti “sesuatu yang diberikan”. Sumber dari data sendiri bisa dari internal maupun dari eksternal. Dalam penggunaan sehari-hari data berarti suatu pernyataan yang diterima apa adanya oleh sistem pencatatan ataupun oleh database jika berbicara mengenai suatu sistem. Catatan sehari-hari mengenai data ini disebut juga dataset dan kumpulan dataset bisa disebut juga *data science*.

Data merupakan bahan mentah dari sebuah solusi dari teknologi yang bernama AI (*Artificial Intelligence*). Perkembangan teknologi AI sangat pesat khususnya dalam bidang data. Data dipandang mulai berguna dalam membantu menentukan pilihan. Beberapa tahun terakhir perkembangan teknologi mempercepat proses terjadinya inovasi di berbagai lini, termasuk inovasi teknologi di bidang kesehatan. Salah satu platform memiliki banyak peminat perkembangan teknologi kesehatan masa depan adalah kecerdasan buatan atau *artificial intelligence* (AI). Perkembangan dalam dunia medis ini dapat membantu manusia untuk mendapatkan insight baru yang yang membantu pekerja medis.

Dalam bidang kedokteran, seorang dokter dalam upaya melakukan diagnosa pada pasien berdasarkan hasil anamnesis dan pemeriksaan fisik. Diagnosis yang dilakukan oleh seorang dokter memungkinkan dokter tersebut untuk memetakan daftar gejala medis dan kemudian membandingkannya dengan data lain. Hasilnya adalah penentuan penyakit atau kondisi apa yang diderita oleh pasien. Dari gambaran tersebut, pada dasarnya seorang dokter sudah mempunyai data gejala medis beserta jenis penyakit serta saran pengobatan yang dapat diberikan berdasarkan hasil diagnosis.

Cara dokter mendigagnosis penyakit dengan membandingkan antara gejala dan jenis penyakit dapat digunakan untuk mengembangkan sebuah aplikasi *machine learning*. Di bidang kedokteran, kecerdasan buatan dapat diaplikasikan dalam hal diagnosis, tatalaksana, dan prediksi risiko. Kecerdasan buatan diharapkan dapat melakukan hal-hal berikut:

- Membantu klinisi mendiagnosis suatu penyakit dan mengoptimalkan proses tata laksana
- Mengurangi angka misdiagnosis dan meningkatkan efisiensi diagnostik
- Mengenali hasil pencitraan dan memberi informasi diagnostik pencitraan yang lebih akurat

- Menyediakan hasil analisis prediksi pasien yang lebih akurat menggunakan analisis big data
- Mendukung penelitian obat-obatan dan meningkatkan efisiensi pengembangan obat baru

Pada penelitian ini mencoba melakukan klasifikasi 41 penyakit dari 132 gejala yang dialami oleh pasien. Data gejala penyakit memiliki bobot sudah disediakan yang akan menjadi acuan dalam menentukan penyakit yang dialami oleh pasien. Dalam menentukan penyakit yang dialami oleh pasien, dilakukan pelatihan menggunakan data yang sudah ada dengan neural network. Hasil dari pelatihan akan menghasilkan model yang dapat mengklasifikasikan 41 penyakit dengan gejala yang di inputkan.

### **Pernyataan Masalah**

1. Bagaimana menentukan penyakit berdasarkan gejala yang dialami untuk selanjutnya memberikan saran pengobatan.

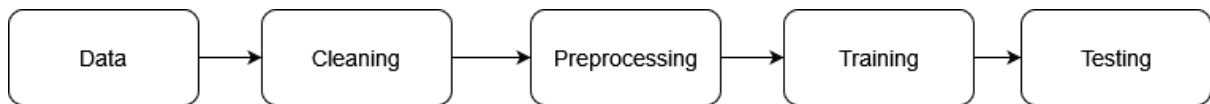
### **Pertanyaan Penelitian**

1. Bagaimana Menentukan 41 penyakit berdasarkan gejala yang yang dialami dan memberikan saran pengobatan yang diberikan
2. Apakah algoritma yang terbaik untuk membangun sebuah *machine learning* yang dapat menentukan jenis penyakit berdasarkan gejala yang dialami serta memberikan saran pengobatan yang diberikan

### **Project Scope & Deliverables**

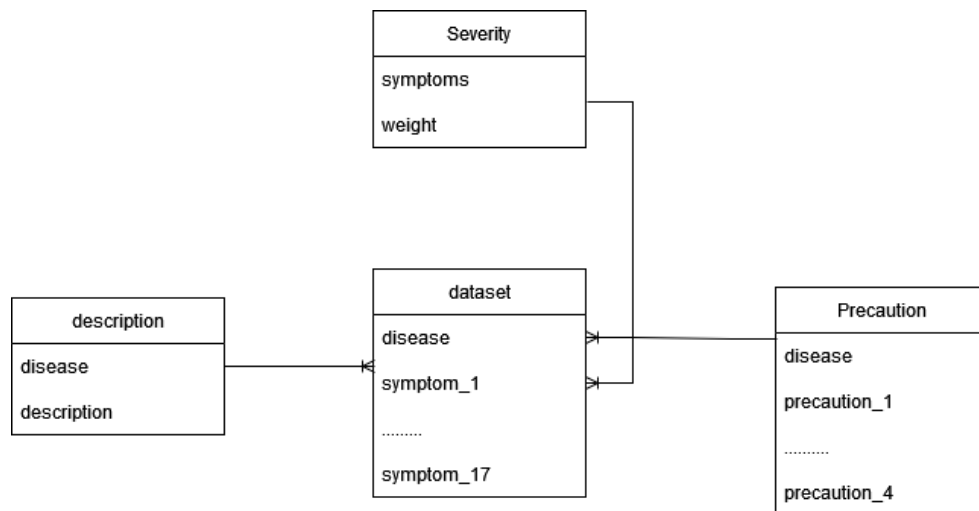
Ruang lingkup proyek ini juga terbatas pada data penyakit yang ada di dalam dataset yang diperoleh yaitu sebanyak 41 jenis penyakit dan juga berdasarkan data gejala yang diperoleh. Proyek ini juga hanya sebatas untuk pembelajaran dalam membangun sebuah *machine learning* di bidang kesehatan dan tidak ditujukan untuk menggantikan saran dari tenaga kesehatan. Adapun rangkaian metode dalam membangun penelitian ini.

- a) Masing masing anggota membuat data cleaning dan preprocessing, hasil data yang sudah clean dari masing-masing anggota akan didiskusikan dan akan disepakati yang akan digunakan untuk tahap selanjutnya.
- b) Menentukan model yang disepakati untuk dilakukan *testing* dan *training*.
- c) Tahap selanjutnya akan dilakukan pelatihan neural network menggunakan tensorflow dengan arsitektur yang akan disepakati pada zoom meeting grup.
- d) Hasil akhir dari penelitian ini dapat melakukan klasifikasi penyakit dari gejala yang di inputkan oleh user. Performa model akan diukur saat melakukan testing. Penjelasan diatas dapat digambarkan dalam flowchart seperti dibawah ini.



Gambar 1 Tahapan Pengolahan Dataset

Dataset yang dimiliki memiliki 4 tabel, yaitu tabel dataset, precaution, description dan severity. Tabel dataset berisi record dari pasien, tabel precaution berisi pencegahan dari disease yang dialami pasien, tabel description berisi deskripsi dari penyakit/disease dan tabel severity merupakan tabel yang berisi bobot dari symptom/gejala yang dialami. Relasi antar tabel dapat dibuat kardinalitasnya seperti gambar dibawah ini.



Gambar 2 Relasi Tabel Dataset

### Project Schedule:

Kegiatan proyek dilakukan mulai tanggal 10 Juli 2022. Berikut ini adalah *Time Schedule* dalam pengerjaan Proyek *Machine Learning Healthcare* ini.

| No | Nama Kegiatan   | JULI 2022 |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |       |
|----|---|-----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------|
|    |   | 10        | 11     | 12     | 13     | 14     | 15     | 16     | 17     | 18     | 19     | 20     | 21     | 22     | 23     | 24     | 25     | 26     | 27    |
| 1  | Zoom Perkenalan Tim & Diskusi Dataset                 | Kuning    |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |       |
| 2  | Tim Melakukan Data Cleaning & Eksplorasi Data         |           | Coklat | Coklat | Coklat |        |        |        |        |        |        |        |        |        |        |        |        |        |       |
| 3  | Zoom Diskusi Hasil Data Cleaning & Rancangan Proposal |           |        |        |        | Kuning |        |        |        |        |        |        |        |        |        |        |        |        |       |
| 4  | Pengerjaan Proposal                                   |           |        |        |        |        | Coklat | Coklat | Coklat | Coklat | Coklat |        |        |        |        |        |        |        |       |
| 5  | Finalisasi Proposal                                   |           |        |        |        |        |        |        |        |        |        | Coklat | Biru   |        |        |        |        |        |       |
| 6  | Pembuatan Model                                       |           |        |        |        |        |        |        |        |        |        |        | Coklat | Coklat | Coklat |        |        |        |       |
| 7  | Zoom Evaluasi Model                                   |           |        |        |        |        |        |        |        |        |        |        |        |        | Coklat |        |        |        |       |
| 8  | Pembaharuan Model & Deploy Model                      |           |        |        |        |        |        |        |        |        |        |        |        |        |        | Coklat |        |        |       |
| 9  | Evaluasi dan Diskusi Model                            |           |        |        |        |        |        |        |        |        |        |        |        |        |        | Kuning |        |        |       |
| 10 | Rancangan Laporan Akhir                               |           |        |        |        |        |        |        |        |        |        |        |        |        |        | Kuning |        |        |       |
| 11 | Penyempurnaan Model & Laporan Akhir                   |           |        |        |        |        |        |        |        |        |        |        |        |        |        |        | Coklat | Coklat |       |
| 12 | Pembuatan Video                                       |           |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        | Hijau  |       |
| 13 | Evaluasi dan Pertemuan Akhir                          |           |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        |        | Hijau |

**Keterangan**

**Kuning** : Zoom Meeting

**Biru** : Final Proposal

**Coklat** : Pengerjaan

**Hijau** : Finalisasi Proyek

**Risk and Issue Management Plan:**

- Fitur yang tidak ada bobotnya
- Fitur yang memiliki salah penamaan
- Fitur yang memiliki bobot ganda atau redundant
- Banyaknya data duplikasi pada dataset (variasi kasus setiap penyakit dalam dataset sedikit). Data pada dataset.csv terdapat **4920** dan variasi data hanya **304**.

**Project Reference:**

1. <https://www.yarsi.ac.id/2022/04/24/melihat-peran-artificial-intelligence-dalam-dunia-kesehatan/>
2. <https://www.kaggle.com/datasets/itachi9604/disease-symptom-description-dataset>
3. <https://www.alomedika.com/peran-artificial-intelligence-dalam-kedokteran-kardiovaskular>
4. <https://www.harianhaluan.com/teknologi/pr-103357406/9-manfaat-machine-learning-di-bidang-kesehatan>
5. [https://skillslab.fk.uns.ac.id/wp-content/uploads/2018/08/MANUAL-IPPA\\_2018-smt-1.pdf](https://skillslab.fk.uns.ac.id/wp-content/uploads/2018/08/MANUAL-IPPA_2018-smt-1.pdf)