

KETENTUAN BPJS KESEHATAN HEALTHKATHON 2022

KATEGORI MACHINE LEARNING

Halo Peserta BPJS Kesehatan Healthkathon 2022,

Kompetisi kategori Machine Learning akan segera dilaksanakan, berikut ini kami sampaikan alur serta syarat dan ketentuan perlombaan yang harus dipatuhi oleh peserta:

- Peserta terbatas untuk mahasiswa aktif maksimal S1
- Peserta yang akan mengikuti kompetisi ini harus melalui tahapan *Registrasi & Basic Test* terlebih dahulu untuk berkompetisi pada tahapan selanjutnya.
- *Tools* pengembangan menggunakan *jupyter notebook* dan menyertakan file *environment setting* yang relevan dengan project yang dikembangkan
- Bahasa pemrograman harus menggunakan bahasa **Python** dan terdapat table of content (daftar isi) pada notebook beserta keterangan/penjelasan yang mudah dipahami untuk setiap tahapan proses sesuai ketentuan/format minimal sebagai berikut:
 - o Data Preparation
 - o Data Cleansing
 - o Exploratory Data Analysis
 - o Feature Engineering
 - o Modeling
 - o Evaluation
- *Library - library* yang digunakan tidak memerlukan lisensi tambahan/berbayar.
- Dalam kategori *Machine Learning* ini terdapat 3 tahapan kompetisi, yaitu:
 - I. Babak Penyisihan:** Submit Hasil Training Fraud Detection Machine Learning Tahap 1(Tentative)
Pada babak penyisihan, semua peserta akan diberikan pembekalan teknis dan akses untuk mendownload data sample BPJS Kesehatan yang dapat di-*download* melalui dashboard peserta. Data yang diberikan oleh BPJS Kesehatan sebanyak +- 11 juta record, data tersebut berupa data mentah yang akan digunakan untuk pengembangan hasil training prediksi *machine learning*,

panitia akan menyiapkan juga data validasi yang digunakan untuk acuan perhitungan *measure*. Peserta mengembangkan sebuah model *machine learning* untuk melakukan prediksi terhadap data yang diberikan selama +/- 3 minggu, submit hasil training prediksi machine learning. Peserta dapat submit 2 kali data hasil training prediksi *machine learning* yang dikembangkan, file yang disubmit berupa csv dengan format kolom sebagai berikut:

- id
- label

Data hasil prediksi machine learning yang di-submit peserta harus minimal mendapatkan hasil metrik dengan minimum hasil yang didapatkan pada *Precision* $\geq 0,55$, *Accuracy* $\geq 0,55$, *Recall* $\geq 0,55$, dan *Specificity* $\geq 0,55$ untuk dapat melanjutkan ke tahap selanjutnya.

II. Babak Kualifikasi Wilayah: Submit Model Machine Learning Tahap 2 (Tentative)

Peserta yang lolos dari tahap 1 diberikan data training dari panitia, peserta dapat men-download data melalui dashboard peserta. Peserta mengembangkan sebuah model *machine learning* dari data training yang telah diberikan selama +/- 5 hari. Model Machine Learning yang diharapkan dapat menghasilkan algoritma yang terbaik untuk melakukan prediktif secara *real-time*. Selanjutnya peserta harus meng-upload hasil *scoring* data validasi berupa file csv dengan format kolom sebagai berikut:

- id
- label

Selain upload csv, peserta juga harus meng-upload file *source code* dalam bentuk zip/rar ke form upload yang disediakan pada dashboard kompetisi. Dalam tahap 2 ini, peserta dapat melakukan submit model *machine learning* sebanyak 3 kali jika belum puas pada nilai yang dihasilkan pada model yang di-submit sebelumnya. Model yang di-upload peserta akan dilakukan verifikasi otomatis *by system* dan poin akan tampil pada *leaderboard*, setelah itu panitia akan melakukan verifikasi source code dan jupyter notebook yang telah disampaikan dengan kesesuaian data yang telah diupload. Peserta yang mengirimkan model lebih dari 1 kali maka nilai tertinggi yang akan tampil di *leaderboard*. Download data validation pada tahap 2 dibuka hanya untuk peserta yang

sudah lolos pada tahap 1. 10 Peserta dengan poin tertinggi akan diundang pada penjurian wilayah. Penjurian Wilayah tentative akan dilaksanakan dibagi sesuai wilayah, seluruh peserta yang masuk 10 besar per wilayah wajib hadir secara daring (melalui video conference). Peserta terbaik per wilayah akan dilombakan ke babak final.

III. Babak Final: Submit Model Machine Learning Tahap Final (Tentative)

Peserta yang lolos ke tahap final diberikan data training kembali dari panitia, peserta dapat *download* data melalui dashboard peserta. Peserta mengembangkan sebuah model *machine learning* dari data training yang telah diberikan selama +- 5 hari, selanjutnya peserta harus *upload* hasil *scoring* data validasi berupa *file* csv dengan format kolom sebagai berikut:

- id
- label

Selain upload csv, peserta juga harus *upload file source code* dalam bentuk zip/rar ke *form upload* yang disediakan pada *dashboard* kompetisi. Dalam tahap final ini, peserta dapat melakukan *submit* model *machine learning* sebanyak 3 kali jika belum puas pada nilai yang dihasilkan pada model yang di-submit sebelumnya. Model yang di-*upload* peserta akan dilakukan verifikasi otomatis *by system* dan poin akan tampil pada *leaderboard*, setelah itu panitia akan melakukan verifikasi source code dan jupyter notebook yang telah disampaikan dengan kesesuaian data yang telah diupload. Peserta yang mengirimkan model lebih dari 1 kali maka nilai tertinggi yang akan tampil di *leaderboard*. *Download data validation* pada tahap final dibuka hanya untuk peserta yang sudah lolos pada tahap 2. Selanjutnya peserta akan mempresentasikan pada babak penjurian final kepada juri untuk hasil model machine learning yang telah dikembangkan. Peserta yang masuk babak final wajib hadir, acara final rencana akan dilaksanakan secara offline di Jabodetabek atau sekitarnya (lokasi dan waktu akan diinfokan lebih lanjut), akomodasi akan ditanggung oleh pihak penyelenggara. Selanjutnya 3 Peserta terbaik akan menjadi pemenang Healthkathon 2020.

Selamat mengikuti kompetisi, ayo tunjukkan skill terbaikmu untuk menjadi yang terbaik pada BPJS Kesehatan Healthkathon 2022 !!!