

DATA MINING



GUIDEBOOK

REVISI 14.01.2020

PENJELASAN UMUM

Data mining merupakan salah satu lomba dalam JOINTS 2020, yang pada tahun ini mengusung tema “Technology Driven: Make Society Better for Live”. Tema ini memiliki arti yaitu berkembangnya teknologi membuat society kita semakin berkembang dan semakin baik. Selain itu, juga memudahkan kita untuk saling terhubung dan terintegrasi satu sama lain.

Data mining sendiri berfokus pada pengolahan data mentah (raw data) dengan cara menemukan pola menarik atau informasi dengan menggunakan metode tertentu yang kemudian divisualisasikan sehingga dapat digunakan sebagai solusi atau pengetahuan lebih lanjut dari sebuah permasalahan data. Lomba ini mengharapkan peserta dapat melakukan proses olah data (Data Mining) terhadap dataset yang diberikan dengan hasil yang memiliki tingkat akurasi setinggi-tingginya.

KETENTUAN PESERTA

1. Setiap peserta merupakan mahasiswa aktif dari universitas yang terdaftar di Indonesia.
2. Setiap tim terdiri atas 2-3 orang yang berasal dari universitas yang sama.
3. Setiap peserta hanya boleh terdaftar ke dalam satu tim.
4. Setiap peserta wajib menyelesaikan administrasi pendaftaran terakhir pada tanggal 20 Maret 2020.
5. Hanya terdapat 1 (satu) kategori dalam Data Mining JOINTS 2020.
6. Pendaftaran yang sah hanya melalui official website joints.id. Waspada! penipuan.
7. Panitia berhak mendiskualifikasi peserta yang melakukan pelanggaran tanpa pemberitahuan sebelumnya.
8. Terus pantau website JOINTS dan OA untuk berita/perubahan setelah guidebook ini dipublikasikan.

BIAYA PENDAFTARAN

Adapun biaya pendaftaran untuk mengikuti Data Mining JOINTS 2020 adalah sebagai berikut beserta ketentuan waktu melunasi pembayaran:

- a. Early bird I:
(30 Januari – 14 Februari)
Mahasiswa : Rp 60.000,-
- b. Early bird II:
(15 Februari – 28 Februari)
Mahasiswa : Rp 75.000,-
- c. Reguler:
(29 Februari – 20 Maret)
Mahasiswa : Rp 100.000,-

HADIAH PEMENANG

Juara I : Rp 3.500.000,00

Juara II : Rp 2.500.000,00

Juara III : Rp 1.500.000,00

Pemenang juga akan mendapatkan sertifikat, plakat, dan hadiah menarik lainnya

TIMELINE LOMBA

Keterangan	Tanggal	Waktu	Tempat
Registrasi Data Mining JOINTS dibuka	30 Januari 2020		Online
Registrasi Data Mining JOINTS ditutup	20 Maret 2020		Online
Pembagian Dataset dan Soal (<u>Babak Penyisihan</u>)	23 Maret 2020		Online
Submisi dan Pengumpulan Makalah (<u>Babak Penyisihan</u>)	23-29 Maret 2020		Online
Pengumuman Finalis	4 April 2020		Online
Pengerjaan Dataset dan Soal (<u>Babak Final</u>)	18 April 2020		Lokasi Babak Final
Presentasi Hasil Pengerjaan (<u>Babak Final</u>)	19 April 2020		Stage

MATERI UMUM PERLOMBAAN

1. Prediksi Dataset

Dalam setiap tahap lomba, peserta akan diberikan sebuah dataset. Setiap peserta diminta untuk membuat sebuah prediksi berdasarkan dataset yang diberikan.

2. Makalah

Pembuatan makalah bertujuan untuk menguji pemahaman peserta mengenai metode-metode *Data Mining* yang digunakan oleh peserta dalam melakukan prediksi terhadap dataset yang diberikan. Pembuatan makalah juga bertujuan sebagai bukti apakah peserta mengerjakan prediksi tersebut sendiri.

3. Presentasi

Presentasi akan dilakukan oleh peserta di babak Final. Presentasi bertujuan untuk menguji kemampuan peserta dalam menjelaskan prediksi dataset, serta metode dan teknik yang digunakan oleh peserta. Presentasi juga akan menguji pemahaman peserta mengenai *library* dan *tools* yang digunakan serta membuktikan apakah peserta mengerjakan prediksi tersebut sendiri.

MEKANISME PERLOMBAAN

1. Babak Penyisihan

- a. Peserta akan diberikan sebuah dataset yang telah ditentukan oleh panitia melalui Kaggle. Platform Kaggle akan dibuka pada 23 Maret 2020 dan ditutup pada 29 Maret 2020. Pengumpulan proposal juga dilaksanakan pada 29 maret 2020 melalui website JOINTS.
- b. Peserta diharapkan dapat mengolah data yang ada serta membuat prediksi seakurat mungkin berdasarkan data yang sudah diolah tersebut.
- c. Peserta diharapkan untuk membuat sebuah makalah yang berisi:
 - i. Judul makalah
 - ii. Latar belakang
 - iii. Tujuan dan manfaat yang diperoleh dari proses Data Mining
 - iv. Metode Data Mining (Software, Dataset, Algoritma)
 - v. Analisis
 - vi. Kesimpulan
 - vii. Dokumentasi
- d. Peserta boleh menggunakan tools atau library apa saja.
- e. 10 tim terbaik akan dinyatakan lolos babak penyisihan berdasarkan sistem penilaian yang ditetapkan, dan akan bertanding di babak final.

2. Babak Final

- a. 10 tim terbaik akan dinyatakan lolos babak penyisihan berdasarkan sistem penilaian yang ditetapkan, dan akan bertanding di babak final
- b. Pada final hari pertama peserta akan diberikan waktu lima jam untuk mengerjakan sebuah dataset.
- c. Peserta akan diberikan sebuah dataset yang berbeda dengan dataset penyisihan.
- d. Peserta diharapkan dapat mengolah data yang ada serta membuat prediksi seakurat mungkin berdasarkan data yang sudah diolah tersebut.
- e. Peserta diharapkan untuk membuat sebuah makalah (**dalam bentuk powerpoint**) untuk dipresentasikan pada hari selanjutnya, yang berisi:
 - i. Judul makalah
 - ii. Latar belakang
 - iii. Tujuan dan manfaat yang diperoleh dari proses *Data Mining*
 - iv. Metode *Data Mining* (Software, Dataset, Algoritma)
 - v. Analisis
 - vi. Kesimpulan
 - vii. Dokumentasi
- f. Peserta boleh menggunakan tools atau library apa saja.
- g. Peserta diharapkan membawa laptop masing-masing untuk keperluan mengerjakan dataset maupun presentasi.

- h. Pada hari kedua peserta diharapkan untuk mempresentasikan semua hal yang dilakukan dalam menentukan strategi *Data Mining* pada hari berikutnya.
- i. Presentasi akan dilakukan di depan panel juri yang akan mengkaji dan menilai.
- j. Juri mendapatkan hak untuk mengajukan beberapa pertanyaan untuk menguji peserta.
- k. Peserta akan diberikan waktu sepuluh menit untuk melakukan presentasi dan diharapkan untuk memaksimalkan waktu yang diberikan.
- l. Dewan juri akan diberikan waktu sepuluh menit untuk mengajukan beberapa pertanyaan, peserta diharapkan dapat menjawab pertanyaan tersebut sebaik mungkin.
- m. Keputusan dewan juri tidak dapat diganggu gugat.

SISTEM PENILAIAN

1. Penilaian utama untuk lomba Data Mining adalah seberapa akuratnya prediksi (hasil olah data) yang telah dibuat oleh peserta.
2. Makalah yang dibuat oleh peserta (untuk babak Penyisihan), diharapkan dapat membuktikan pemahaman peserta mengenai metode pengolahan data yang digunakan, serta untuk membuktikan bahwa peserta benar-benar mengerjakan prediksi tersebut sendiri.
3. Dalam sesi presentasi, peserta diharapkan dapat menjelaskan metode/strategi data mining yang digunakan dengan sebaik mungkin. Peserta juga diharapkan mampu menjawab pertanyaan yang diberikan oleh juri sebaik mungkin.

KETENTUAN LOLOS

1. Babak Penyisihan

No	Kriteria	Bobot (%)
1	Akurasi Prediksi	60
2	Makalah	40

Peserta yang lolos babak penyisihan adalah peserta dengan skor tertinggi berdasarkan akurasi prediksi peserta dan makalah, dengan bobot yang dituliskan di atas.

2. Babak Final

No	Kriteria	Bobot (%)
1	Hasil Prediksi	40

2	Presentasi	60
---	------------	----

Pada babak Final, peserta akan dinilai berdasarkan hasil prediksi peserta dan nilai yang diberikan oleh juri pada saat sesi presentasi, dengan bobot yang dituliskan di atas.

KETENTUAN DISKUALIFIKASI

1. Peserta yang terdaftar dalam lebih dari satu tim.
2. Tim yang tidak melengkapi syarat administrasi.
3. Tim yang melakukan tindak kecurangan dan pelanggaran serius terhadap plagiarisme.
4. Tim yang terlambat melakukan konfirmasi kehadiran saat tahap selanjutnya.
5. Tim yang hadir dengan anggota tidak lengkap tanpa konfirmasi alasan yang jelas.

CONTACT PERSON

Alfi

Line : faqihafiframdani

WA : 089633994684