**Tugas 3**

**CII-2M3 Pengantar Kecerdasan Buatan  
Ganjil 2020/2021**

Diberikan *dataset* (himpunan data) Pima India Diabetes Dataset (PIDD) pada file “Diabetes.csv”. Dataset tersebut berisi 768 objek data (baris). Buatlah lima datasets baru menggunakan skema 5-*fold cross-validation*. Pertama, bagi objek data ke dalam lima *subsets* (sub himpunan) dengan porsi yang sama, masing-masing berisi sat per lima (20%) data. Kemudian, buat lima *dataset* baru dengan komposisi objek-objek data pada *training set* (data latih) dan *testing set* (data uji) sebagai berikut:

1. Baris ke-1 sampai baris ke-614 sebagai *training set* dan sisanya sebagai *testing set*;
2. Baris ke-1 sampai baris ke-461 ditambah baris ke-615 sampai 768 sebagai *training set* dan yang lain sebagai *testing set*;
3. Baris ke-1 sampai baris ke-307 ditambah baris ke-462 sampai 768 sebagai *training set* dan yang lain sebagai *testing set*;
4. Baris ke-1 sampai baris ke-154 ditambah baris ke-308 sampai 768 sebagai *training set* dan yang lain sebagai *testing set*; dan
5. Baris ke-155 sampai sampai 768 sebagai *training set* dan yang lain sebagai *testing set*.

Lakukan analisis, desain, dan implementasi algoritma ***k-nearest neighbor* (kNN)** ke dalam suatu program komputer. Lakukan **seleksi** dan **estimasi** model kNN tersebut menggunakan 5-*fold cross-validation* yang menghasilkan akurasi tertinggi.

**Hal yang dapat diobservasi:**

* Pemilihan ukuran jarak yang digunakan
* Teknik prapemrosesan data
* Teknik rekayasa fitur
* Strategi penggunaan algoritma kNN
* Pemilihan nilai *k* terbaik untuk proses **seleksi** dan **estimasi** model kNN

**Proses yang harus dibangun (bisa berupa fungsi/prosedur):**

* Perhitungan ukuran jarak
* Prapemrosesan data
* Klasifikasi kNN
* Pemilihan nilai *k* terbaik
* Perhitungan rata-rata akurasi kNN menggunakan 5-*fold cross-validation*

**Output** dari sistem adalah:

1. **Nilai *k* terbaik** hasilpembelajaran kNN; dan
2. **Rata-rata akurasi** kNN menggunakan 5-*fold cross-validation*.

**Aturan dan Penilaian**

1. **Pengumpulan**

* Tugas dikumpulkan maksimal hari **Minggu** tanggal **20 Desember 2020** pukul **20.00 WIB** melalui **LMS**.
* Pengumpulan berupa file **.zip** atau **.rar** dan **HANYA** berisi **DUA** **FILE**:

1. *source code* program
2. Laporan Observasi

* Penamaan file zip/rar **HARUS** menggunakan format **KELAS\_NIM**
  + Misal: **IF4107\_1301174068.zip** atau **IFX4301\_1301171111.rar**
  + Misal: **IFIK4102\_1301171234.zip** atau **IF41INT01\_1301174321.rar**
  + Tidak ada tanda hubung (-) atau garis bawah (\_) di penamaan kelas
* Penamaan file dan format source code yang tidak sesuai akan menyebabkan pengumpulan **TIDAK DINILAI**

1. **Source Code Program**

* Source Code harus bisa dijalankan dan terkumpul dalam **SATU FILE** Source Code
* Sistem boleh dibangun menggunakan bahasa pemrograman apapun
* **Tidak diperbolehkan menggunakan library** yang secara langsung melakukan proses kNN atau proses yang harus dibangun/diobservasi
* Penamaan file dan format source code yang tidak sesuai akan menyebabkan pengumpulan **TIDAK DINILAI**
* **JANGAN** mengumpulkan file .exe

1. **Laporan Observasi**

* Laporan maksimal 2 halaman (tidak termasuk sampul)
  + Font: Times New Roman, font size: 12, spacing: 1, dalam format **.pdf**
  + Berisi penjelasan strategi penyelesaian masalah dan analisis hasil ekperimen
  + Tuliskan nilai-nilai **parameter kNN** yang Anda anggap paling optimum
* Berikan **screenshot** hasil running berdasarkan nilai-nilai parameter optimum tersebut.

1. **Detil Penilaian**
   * CLO 2: Source code dan Akurasi kNN.

* **Nilai = E** bagi siapapun yang terbukti melakukan ***KECURANGAN***.
  + Kesamaan kode dan/atau laporan hingga 80%
  + Mengumpulkan kode program dan/atau laporan milik orang lain

Bandung, 16 Desember 2020

(Tim Dosen Kecerdasan Buatan)

**Daftar Asisten Dosen Kecerdasan Buatan Ganjil 2019/2020**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nama Lengkap | NIM | HP | Kelas | Alamat email |
| Muhammad Alfhi Saputra | 1301171205 | 082371806273 | IF-42-01 | alfhisa@student.telkomuniversity.ac.id |
| Angel Metanosa Afinda | 1301174639 | 082129916992 | IF-42-02 | angelmetanosaafinda@student.telkomuniversity.ac.id |
| Anggitha Yohana Clara | 1301174317 | 081267450123 | IF-42-03 | anggithayohana@student.telkomuniversity.ac.id |
| Beladina Elfitri | 1301174046 | 082282794738 | IF-42-04 | beladinaelfitri@student.telkomuniversity.ac.id |
| Rachmansyah Adhi Widhianto | 1301170066 | 089696969147 | IF-42-05 | rachmansyah@student.telkomuniversity.ac.id |
| Anvaqta Tangguh Wisesa | 1301170390 | 082226808866 | IF-42-06 | anvaqta@student.telkomuniversity.ac.id |
| Priscilia Lovita Paelongan | 1301174102 | 082343584884 | IF-42-07 | lovitapaelongan@student.telkomuniversity.ac.id |
| Ahmad Satriamulya | 1301170012 | 085946511309 | IF-42-10 | ahmadrayi@student.telkomuniversity.ac.id |
| Rivaldo Ludovicus Sembiring | 1301174445 | 085361292959 | IF-42-11 | rludovicus@student.telkomuniversity.ac.id |
| Hanif Fadhlurrahman | 1301174609 | 081546132425 | IF-42-12 | haniffadhlurrahman@student.telkomuniversity.ac.id |
| Bimo Satrio Aji | 1301171248 | 085889087773 | IF-42-GAB01 | bimooaji@student.telkomuniversity.ac.id |
| Nadira Putri Arthamevia | 1301170395 | 081212923787 | IF-42-GAB02 | arthamevias@student.telkomuniversity.ac.id |
| Ananda Fitri Karimah | 1301170774 | 082215283845 | IF-42-GAB03 | anandafitri@student.telkomuniversity.ac.id |
| Muhammad Ahsan Athallah | 1301174068 | 082117435016 | IF-42-INT | ahsanathallah@student.telkomuniversity.ac.id |