**Soal Kuis Data mining 2020**

**Paralel 1 Lab 1**

Kode soal: Ganjil

1. Bukalah data pokemon1.csv! Gabungkan data pokemon1.csv dengan pokemon2.csv. Lampirkan screenshot code R
2. Lakukan eksplorasi data tersebut menggunakan kode R. Lampirkan screenshot code R serta hasilnya. Jelaskan hasil yang didapat.
3. Lakukan tahapan praproses terhadap data pokemon. Lampirkan screenshot code R serta hasilnya. Praproses berupa:
   1. Pengisian missing value
   2. Reduksi data yang diambil sebanyak 200 data secara acak
4. Lakukan tahapan clustering terhadap data hasil praproses! Tahapan yang dilakukan pada clustering yaitu clustering menggunakan K-Means dengan k=4. Visualisasikan *cluster* dengan menggunakan *scatterplot* pada atribut **Attack** dan **Defense**.Tampilkan screenshot code R serta hasilnya. Jelaskan hasil yang didapat.

Ketentuan Pengumpulan:

1. Jawaban dituliskan dalam file .doc atau .docx dilengkapi dengan **nama, nim, tipe soal, dan nomor meja**
2. Laporan dan kode R diunggah dalam LMS sesuai jadwal yang ditentukan.
3. Ketentuan nama fail laporan: Kuis1\_Paralel<kelas paralel>\_NIM\_<Tipe Soal>.docx, contoh: Kuis1\_Paralel1\_G64160039\_Ganjil.docx
4. Ketentuan nama Zip: Kuis1\_Paralel<kelas paralel>\_NIM\_<Tipe Soal>.zip, contoh: Kuis1\_Paralel1\_G64160039\_Ganjil.zip