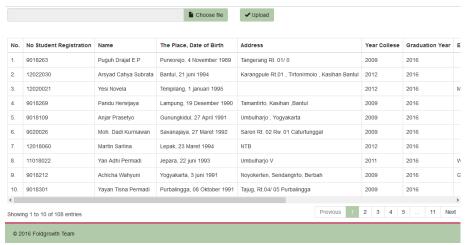
#### **Materi Pelatihan Data Mining**

# (dalam materi ini bisa di linkkan dengan mongodb dan juga interface sederhana)

### 1. Upload Data

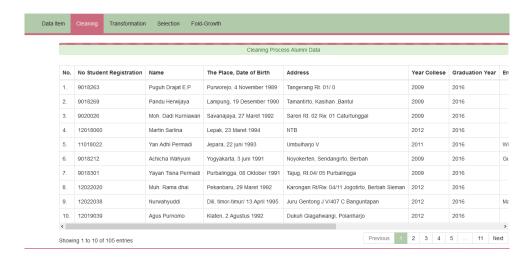
Pada halaman *data awal* admin dapat meng*upload* data alumni dan dapat melihat daftar data alumni yang telah tersedia. Implementasi *user interface* halaman *data awal* dapat dilihat pada Gambar 4.15



Gambar 1. Implementasi hasil load data awal

#### 2. Pembersihan Data

Data yang dipakai dari data alumni tahun 2015-2016 yang berjumlah 102 dilakukan proses yang pertama yaitu pembersihan dan diperoleh 98 data setelah dibersihkan. Pembersihan data ini dilakukan denga menghapus data yang tidak kosong pada atribut Umur, IPK, Toefl, lama studi, gaji pertama bekerja dan lama masa tunggu mencari kerja.



# 3. Integrasi Data

Pada penelitian ini, data tidak memerlukan adanya integrasi data atau penggabungan data dari beberapa database. Karena data alumni yang ada pada universitas ahmad dahlan ini berupa data excel.

### 4. Seleksi Data

Pemilihan data yang sesuai dengan algoritma yang akan digunakan, pemilihan data seleksi merupakan tahap awal dalam pembentukan proses data *mining* pemilihan data seleksi dalam penelitian. Pemilihan atribut dari 13 atribut menjadi 6 atribut. Hanya atribut yang dipakai yang digunakan



#### 5. Transformasi Data

Proses transformasi data yaitu proses pengkategorian atau perubahan data ke format yang sesuai untuk proses dalam data mining. Ada 3 refrensi yang diambil berdasarkan jurnal Borang Sarjana Akreditasi Program Studi Sarjana pada tahun 2012 yaitu IPK, Toefl, dan lama masa tunggu mencari kerja. hasil proses transformasi data atribut adalahsebagai berikut:

a. Nilai IPK dikategorikan menjadi 3 yaituIPK >3.50-4.00= Maksimum, jika IPK >=2.75-3.50= Rata-rata dan jika IPK <2.75= Minimum. Hasil transformasi akan ditunjukkan padaTabel 4.1.

Tabel 4.1. Proses Transformasi IPK

Kategori	Keterangan
IPK Minimum (Min)	Untuk IPK <2.75
IPK Rata-rata (Rat)	Untuk IPK >=2.75 - 3.50
IPK Maksimum (Max)	IPK >3.50 - 4.00

b. Nilai Toefl dikategorikan menjadi 3 yaitu T1= 400-449, T2 = 450-499, dan T3 =
>=500. Hasil transformasi akan ditunjukkan pada Tabel 4.2.

Tabel 4.2 Proses Tranformasi nilai Toefl

Kategori	Keterangan
T1	400 – 449
T2	450 –499
Т3	>=500

c. Waktu mendapat pekerjaandikategorikan menjadi 4 yaitu w1 = 0–2.9bulan, w2 = 3 – 5.9 bulan, w3 = 6 – 8.9 buln, w4 = 9 – 12 bulan. Hasil transformasi akan ditunjukkan pada Tabel 4.3.

Table 4.3 Proses tranformasi waktu mendapat kerja

Kategori	Keterangan
W1	0 -2.9 bulan
W2	3 – 5.9 bulan
W3	6 - 8.9 bulan
W4	9 -12 bulan

d. Data atribut waktu lama studi akan dikatagorikan berdasarkan keputusan Menteri Pendidikan Nasional Nomor 232/U/2000 (2000) tentang Pedoman Penyususnan Kurikulum Pendidikan Tinggi dan Penilaian Hasil Belajar Mahasiswa. Data atribut dikategorikan tepat waktu apabila lama studi<=4 tahun dan tidak tepat waktu apabila lama studi >4 tahun. Data transformasi akan ditunjukkan pada Tabel 4.4.

Table 4.4 Proses tranformasi waktu lama studi

Kategori	Keterangan
Tepat Waktu (TW)	Jika lama studi <=4 tahun
Tidak Tepat Waktu (TTW)	jika lama studi >4 tahun

e. Data transformasi umur akan dikategorikan berdasarkan penelitian yang dilakukan oleh (Purwaningseh, 2010) tentang Analisis pencari kerja menurut kelompok umur. Data atribut umur 15-24 tahun umumnya digolongkanproduktif dalam mencari kerja dan pada umur 25-54 tahun digolongkantidak produktif. Data transformasi umur ditunjukkan pada Tabel 4.5.

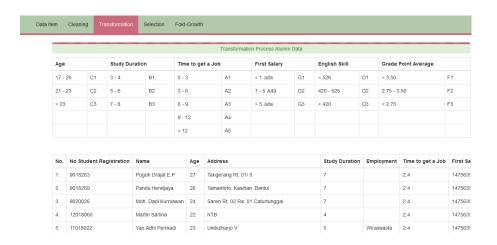
Tabel 4.5 Proses tranformasi umur

Kategori	Keterangan
Usia produktif (UP)	Untuk umur 15-24 tahun
Tidak produktif (TP)	Untuk umur 25-54 tahun

f. Gaji dikatakan diatas UMR apabila nominalnya >=1.572.200 dan gaji dikatakan dibawah UMR apabila nominalnya <1.572.200menurut keputusan Gubernur Daerah Istimewa Yogyakarta no 235/ kep/ 2016 kota yogyakarta tentang penetapan upah minimum kabupaten/kota tahun 2017 ditunjukkan padatabel 4.6.

Table 4.6 Proses tranformasi nominal gaji

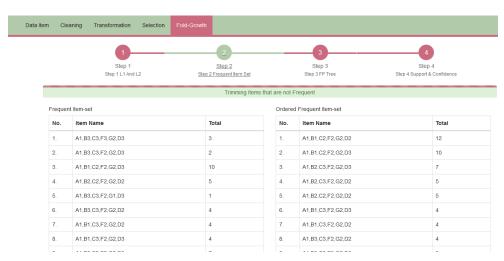
Kategori	Keterangan
Atas UMR (A.UMR)	Gaji >=1.572.200
Bawah UMR (B.UMR)	Gaji <1.572.200



### 6. Proses Asosiasi dengan metodde tertentu dan akurasinya

Metode Asosiasi: FP-Tree, Fold-Growth, Apriori, FOLDRAM, dll

Akurasi : Lift rasio dll



#### 7. Metode Clustering dan akurasi

Metode clustering: K-Means, K-Medoid, AHC, SOM, Fuzzy C-Means dll

Akurasi: Confution matrik, purity tes, dll

Contoh data bisa menggunakan data alumni dengan 6 variabel : umur, IPK, lama masa studi, lama masa tunggu mencari kerja, gaji pertama bekerja dan TOEFL.

#### 8. Metode Klasifikasi dan Akurasi

Metode Klasifikasi: Tree, C4.5, C4.8, SVM, JST, dll

Akurasi: Confution matrix dll

Contoh data bisa menggunakan data alumni dengan 6 variabel : umur, IPK, lama masa studi, lama masa tunggu mencari kerja, gaji pertama bekerja dan TOEFL.

Dengan kelas lama masa tunnggu mencari kerja: cepat, sedang dan lama

## 9. Metode Peramalan dan akurasi

Metode Peramalan : SVR, regresi dll

Akurasi: MAD, MSE dll