

# **Tugas Proyek Data Science Kelompok 5 – Classification Kelulusan Mahasiswa**

Anggota :

1. Dian Fetriah
2. Awang Dewan Pratama
3. Akhmad Hasan Arofid
4. Azka Dini Yuntari

# Business Understanding

Klasifikasi terhadap data kelulusan mahasiswa diperlukan dalam hal rujukan dalam menentukan model prediksi kelulusan tepat waktu mahasiswanya berdasarkan faktor seperti status mahasiswa, nilai IPS Semester 1 s.d 8, IPK, dan jenis kelamin.

Dengan mengetahui model prediksi, pihak universitas dapat melakukan Langkah-Langkah untuk meningkatkan tingkat kelulusan tepat waktu dilingkungannya dan menyiapkan strategi penanganan efek keterlambatan mahasiswa

# Data Understanding

## Data Kelulusan Mahasiswa

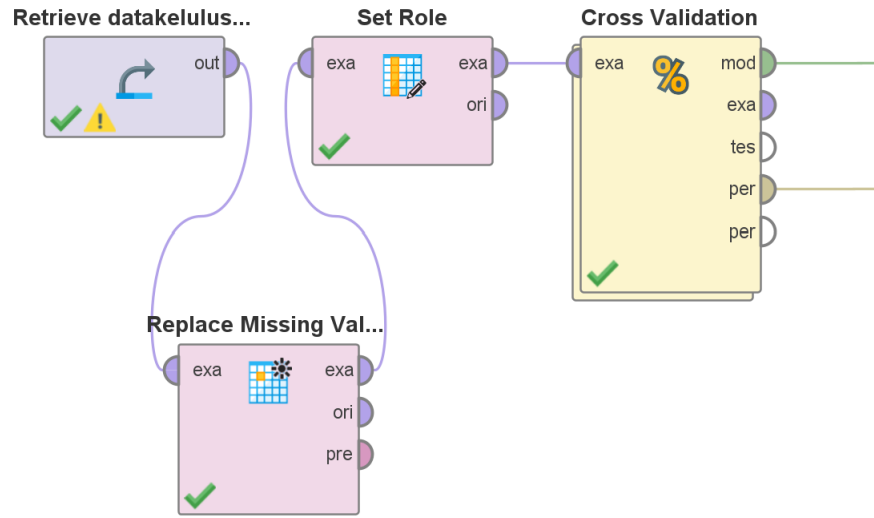
- Nama : Nama Mahasiswa (Polynomial)
- Jenis Kelamin : Keterangan jenis kelamin mahasiswa (Polynomial)
- Status Mahasiswa : status mahasiswa murni atau sambal bekerja (Polynomial)
- Umur : umur mahasiswa (integer)
- Status Nikah : Status pernikahan mahasiswa (Polynomial)
- IPS 1: nilai Indeks Prestasi Semester 1 (real)
- IPS 2 : nilai Indeks Prestasi Semester 2 (real)
- IPS 3 : nilai Indeks Prestasi Semester 3 (real)
- IPS 4 : nilai Indeks Prestasi Semester 4 (real)
- IPS 5 : nilai Indeks Prestasi Semester 5 (real)
- IPS 6 : nilai Indeks Prestasi Semester 6 (real)
- IPS 7 : nilai Indeks Prestasi Semester 7 (real)
- IPS 8 : nilai Indeks Prestasi Semester 8 (real)
- IPK : nilai Indeks Prestasi Kumulatif (real)
- Status Kelulusan : (polynomial)

# Data Preparation

Terdapat Missing Value pada IPS 8 dan IPK

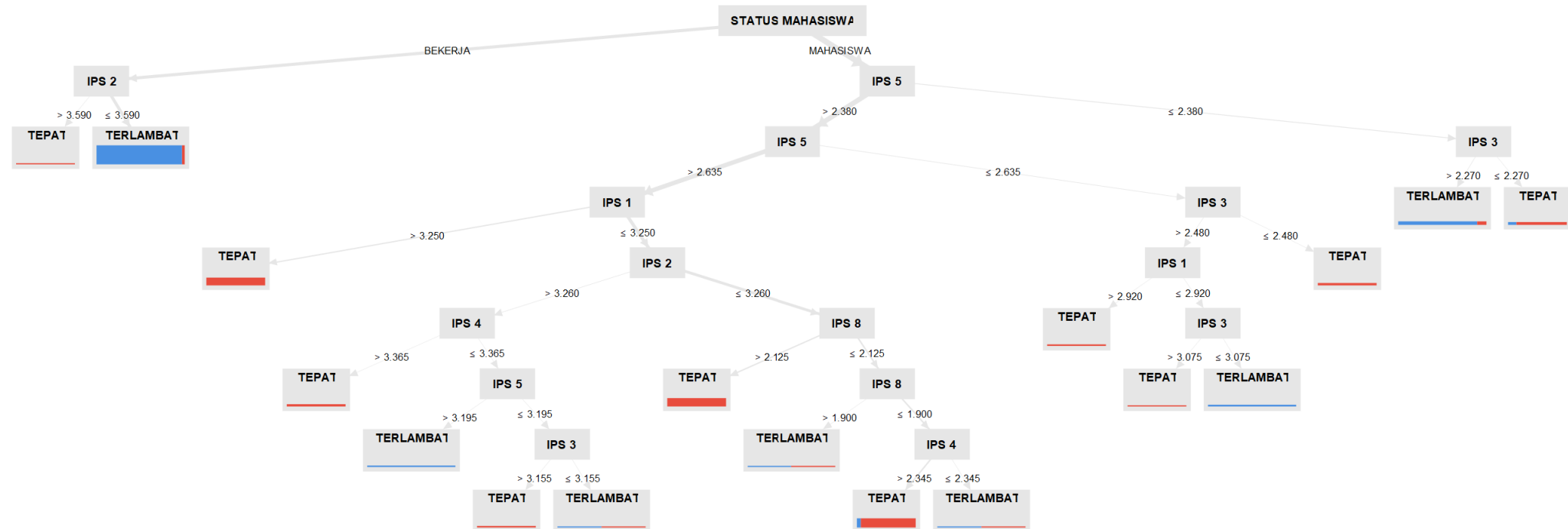
Pengisian Missing Value digunakan menggunakan data average.

# Modeling 1



Algoritma yang digunakan adalah Cross validation

# Evaluasi Model 1



# Evaluasi Model 1

```
STATUS MAHASISWA = BEKERJA
|   IPS 2 > 3.590: TEPAT {TERLAMBAT=0, TEPAT=2}
|   IPS 2 ≤ 3.590: TERLAMBAT {TERLAMBAT=127, TEPAT=4}
STATUS MAHASISWA = MAHASISWA
|   IPS 5 > 2.380
|   |   IPS 5 > 2.635
|   |   |   IPS 1 > 3.250: TEPAT {TERLAMBAT=0, TEPAT=50}
|   |   |   IPS 1 ≤ 3.250
|   |   |   |   IPS 2 > 3.260
|   |   |   |   |   IPS 4 > 3.365: TEPAT {TERLAMBAT=0, TEPAT=10}
|   |   |   |   |   IPS 4 ≤ 3.365
|   |   |   |   |   |   IPS 5 > 3.195: TERLAMBAT {TERLAMBAT=4, TEPAT=0}
|   |   |   |   |   |   IPS 5 ≤ 3.195
|   |   |   |   |   |   |   IPS 3 > 3.155: TEPAT {TERLAMBAT=0, TEPAT=5}
|   |   |   |   |   |   |   IPS 3 ≤ 3.155: TERLAMBAT {TERLAMBAT=1, TEPAT=1}
|   |   |   |   |   |   |   |   IPS 2 ≤ 3.260
|   |   |   |   |   |   |   |   IPS 8 > 2.125: TEPAT {TERLAMBAT=0, TEPAT=54}
|   |   |   |   |   |   |   |   IPS 8 ≤ 2.125
|   |   |   |   |   |   |   |   |   IPS 8 > 1.900: TERLAMBAT {TERLAMBAT=1, TEPAT=1}
|   |   |   |   |   |   |   |   |   IPS 8 ≤ 1.900
|   |   |   |   |   |   |   |   |   |   IPS 4 > 2.345: TEPAT {TERLAMBAT=4, TEPAT=56}
|   |   |   |   |   |   |   |   |   |   IPS 4 ≤ 2.345: TERLAMBAT {TERLAMBAT=1, TEPAT=1}
|   |   |   |   |   |   |   |   |   |   |   IPS 5 ≤ 2.635
|   |   |   |   |   |   |   |   |   |   |   |   IPS 3 > 2.480
|   |   |   |   |   |   |   |   |   |   |   |   |   IPS 1 > 2.920: TEPAT {TERLAMBAT=0, TEPAT=5}
|   |   |   |   |   |   |   |   |   |   |   |   |   IPS 1 ≤ 2.920
|   |   |   |   |   |   |   |   |   |   |   |   |   |   IPS 3 > 3.075: TEPAT {TERLAMBAT=0, TEPAT=2}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   IPS 3 ≤ 3.075: TERLAMBAT {TERLAMBAT=6, TEPAT=0}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   IPS 3 ≤ 2.480: TEPAT {TERLAMBAT=0, TEPAT=11}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   IPS 5 ≤ 2.380
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   IPS 3 > 2.270: TERLAMBAT {TERLAMBAT=17, TEPAT=2}
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   IPS 3 ≤ 2.270: TEPAT {TERLAMBAT=2, TEPAT=12}
```

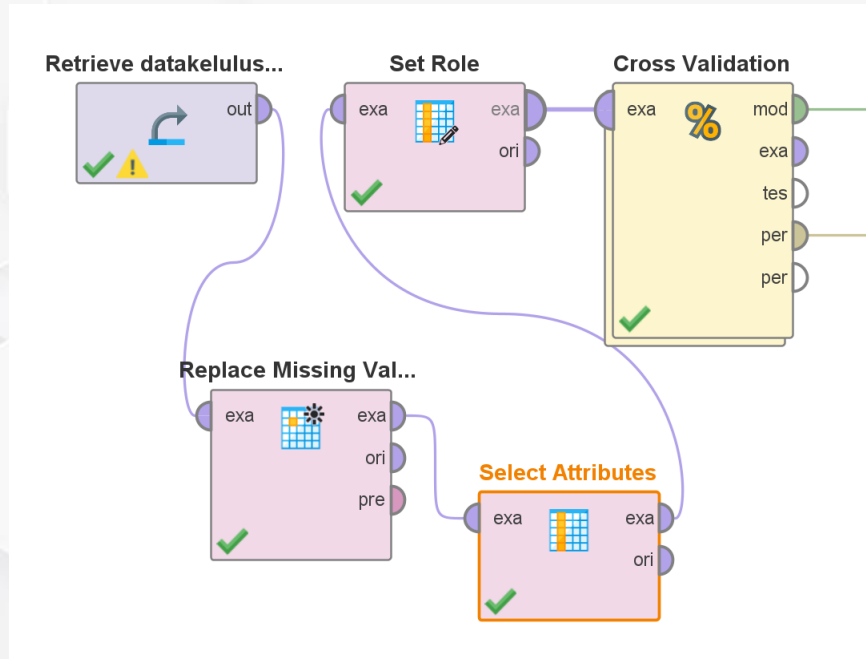
Dari model prediksi ini diketahui:

- Mahasiswa yang bekerja, saat IPS semester 2 nya diatas 3.590 maka diprediksi mahasiswa tersebut akan lulus tepat waktu.

accuracy: 89.70% +/- 5.21% (micro average: 89.71%)

	true TERLAMBAT	true TEPAT	class precision
pred. TERLAMBAT	144	20	87.80%
pred. TEPAT	19	196	91.16%
class recall	88.34%	90.74%	

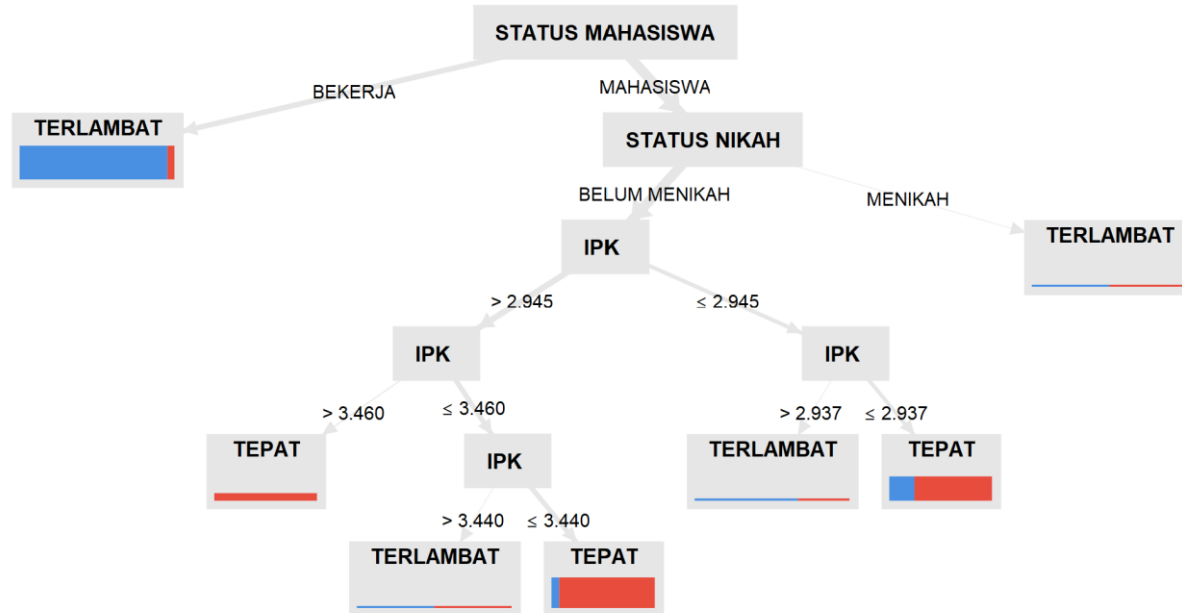
# Modeling 2



Algoritma yang digunakan adalah Cross validation dengan membatasi pada atribut Nama, IPK, Status Kelulusan, Status Mahasiswa, dan Status Nikah



# Evaluasi Model 2



```
STATUS MAHASISWA = BEKERJA: TERLAMBAT {TERLAMBAT=127, TEPAT=6}
STATUS MAHASISWA = MAHASISWA
|   STATUS NIKAH = BELUM MENIKAH
|   |   IPK > 2.945
|   |   |   IPK > 3.460: TEPAT {TERLAMBAT=0, TEPAT=25}
|   |   |   IPK ≤ 3.460
|   |   |   |   IPK > 3.440: TERLAMBAT {TERLAMBAT=1, TEPAT=1}
|   |   |   |   IPK ≤ 3.440: TEPAT {TERLAMBAT=9, TEPAT=111}
|   |   |   |   IPK ≤ 2.945
|   |   |   |   |   IPK > 2.937: TERLAMBAT {TERLAMBAT=2, TEPAT=1}
|   |   |   |   |   IPK ≤ 2.937: TEPAT {TERLAMBAT=23, TEPAT=71}
|   |   |   |   STATUS NIKAH = MENIKAH: TERLAMBAT {TERLAMBAT=1, TEPAT=1}
```

accuracy: 87.86% +/- 4.68% (micro average: 87.86%)

	true TERLAMBAT	true TEPAT	class precision
pred. TERLAMBAT	127	10	92.70%
pred. TEPAT	36	206	85.12%
class recall	77.91%	95.37%	

Berdasarkan Model 2 ini diketahui:

- Mahasiswa yang bekerja sebagian besar terlambat lulus sedangkan
- Mahasiswa yang tidak bekerja, belum menikah dengan IPK 2.945 s.d. 3.44, sebagian besar lulus tepat waktu

# Kesimpulan

- Model 1 : Berdasarkan algoritma Cross Validation diketahui model 1, prediksi kebenaran kelulusan tepat waktu mencapai 90,74%, dan 88,34% untuk kelulusan terlambat, dengan tingkat akurasi 89,70%
- Model 2 : Berdasarkan algoritma Cross Validation diketahui model 2, prediksi kebenaran kelulusan tepat waktu mencapai 95,37%, dan 77,91% untuk kelulusan terlambat, dengan tingkat akurasi 87,86%
- Perlu disimulasi beberapa model berdasarkan variasi atribut untuk dapat diketahui model klasifikasi mana yang memiliki prediksi paling akurat



**GTA**

Government  
Transformation  
Academy



# TERIMA KASIH

#JadiJagoanDigital



Digital Talent Scholarship



digitalent.kominfo



DTS\_kominfo