

MEMBUAT *RULE BASE* MENGGUNAKAN *DECISION TREE*



www.tyas-tamimy.com



tyas_fa@stmikmpb.ac.id



fitriayuningtyas@stmikmpb.ac.id



081804767700 (Telegram only)



Tyas Tamimy

Knowledge Acquisition dengan Rule Production

- Alternatif untuk melakukan *knowledge acquisition* adalah dengan mengkonversi *database* yang ada ke dalam kumpulan *production rule*.
- *Database* yang ada harus berisikan data-data yang meliputi semua permasalahan yang ada.
- Langkah-langkah membuat *rule production*:
 1. Mengidentifikasi *attribute/ feature, object/ record, label/ target* dari *database/ dataset*
 2. Membuat *Decision Tree*
 3. Membuat *rule* dari *Decision Tree* yang telah dibuat

1. Mengidentifikasi *attribute, object, label* dari *database/ dataset*

| age | spectacle-prescrip | astigmatism | tear-prod-rate | contact-lenses |
|----------------|--------------------|-------------|----------------|----------------|
| young | myope | no | reduced | none |
| young | myope | no | normal | soft |
| young | myope | yes | reduced | none |
| young | myope | yes | normal | hard |
| young | hypermetrope | no | reduced | none |
| young | hypermetrope | no | normal | soft |
| young | hypermetrope | yes | reduced | none |
| young | hypermetrope | yes | normal | hard |
| pre-presbyopic | myope | no | reduced | none |
| pre-presbyopic | myope | no | normal | soft |
| pre-presbyopic | myope | yes | reduced | none |
| pre-presbyopic | myope | yes | normal | hard |
| pre-presbyopic | hypermetrope | no | reduced | none |
| pre-presbyopic | hypermetrope | no | normal | soft |
| pre-presbyopic | hypermetrope | yes | reduced | none |
| pre-presbyopic | hypermetrope | yes | normal | none |
| presbyopic | myope | no | reduced | none |
| presbyopic | myope | no | normal | none |
| presbyopic | myope | yes | reduced | none |
| presbyopic | myope | yes | normal | hard |
| presbyopic | hypermetrope | no | reduced | none |
| presbyopic | hypermetrope | no | normal | soft |
| presbyopic | hypermetrope | yes | reduced | none |
| presbyopic | hypermetrope | yes | normal | none |

- *Dataset contact-lenses*
- *Attribute:*
 1. *age*
 2. *spectacle-prescript*
 3. *astigmatism*
 4. *tear-prod-rate*
- *Label: contact-lense*
- *Object: isi setiap baris/ datanya*

2. Menentukan *Decision Tree*

- Algoritma C4.5 merupakan salah satu algoritma klasifikasi yang digunakan untuk melakukan klasifikasi dan bersifat prediktif, C4.5 merupakan evolusi dari ID3, disajikan oleh penulis yang sama [Quinlan (1993)]
- ID3 menggunakan Information Gain sedangkan C4.5 menggunakan Gain Ratio sebagai kriteria pemisahan (splitting criteria) [C. J. Mantas and J. Abellán, 2014].
- Dasar algoritma C4.5 adalah pembentukan pohon keputusan (decision tree).

2. Menentukan *Decision Tree*

- Layaknya sebuah pohon yang memiliki cabang, akar dan daun (node), dapat disimpulkan bahwa dasar dari algoritma decision tree adalah **pembentukan pohon keputusan hierarkis** dengan **cabang-cabang pohon keputusan sebagai pertanyaan klasifikasi**, **daun-daunnya (*leaf node*) merupakan kelas atau label**, dan **akarnya (*root node*) sebagai node pertama atau node teratas** sehingga pohon keputusan yang terbentuk menyerupai **pohon terbalik**.

Tahapan Algoritma C4.5

1. Siapkan Dataset
2. Pilih atribut sebagai akar
 - 2.1. Menghitung nilai Entropy Total
 - 2.2. Menghitung nilai Information Gain
 - 2.3. Menghitung Split Information
 - 2.4. Menghitung nilai Gain Ratio
3. Buat cabang untuk tiap-tiap nilai
4. Bagi kasus dalam cabang
5. Ulangi proses 2 – 4 untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i$$

2.2. Menghitung nilai Information Gain

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

2.3. Menghitung Split Information

$$SplitInfo(S, A) = - \sum_{i=1}^n \frac{S_i}{S} \log_2 \frac{S_i}{S}$$

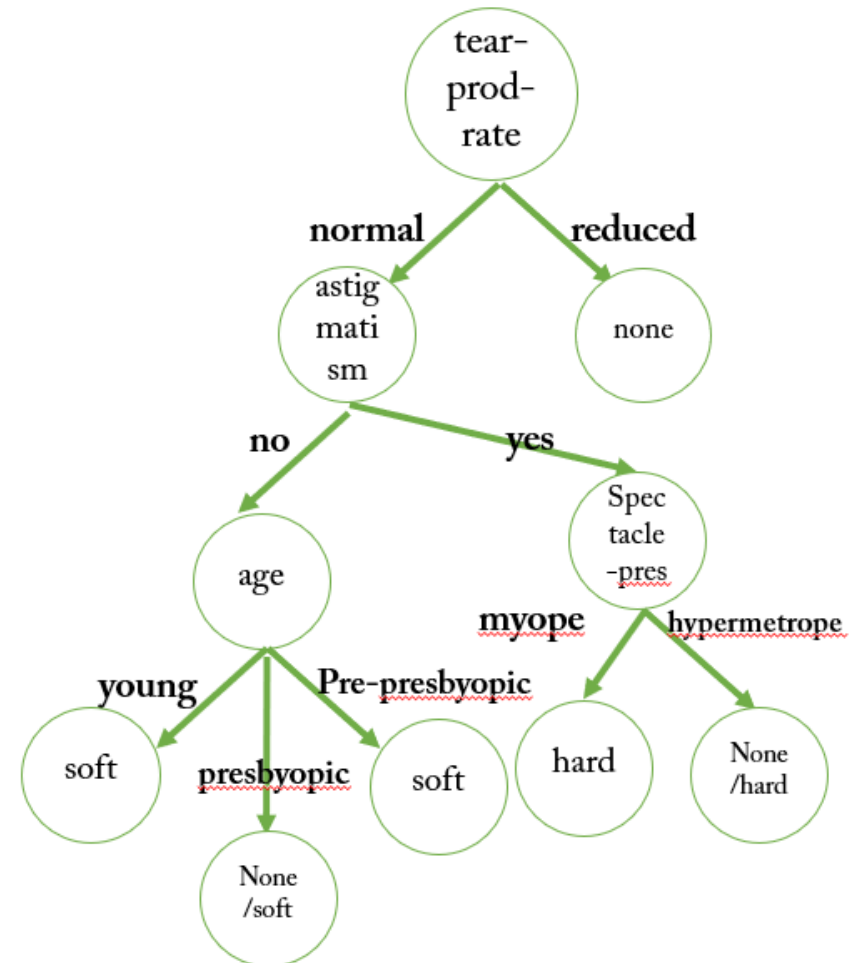
2.4. Menghitung nilai Gain Ratio

$$Gain\ Ratio(S, A) = \frac{Gain(S, A)}{Split\ Information(S, A)}$$

2. Menentukan *Decision Tree*

- Setelah dilakukan perhitungan menggunakan algoritma C4.5 maka didapatkan *decision tree* sebagai berikut:

“Contoh perhitungan algoritma C4.5 akan dibahas pada pertemuan berikutnya”



3. Membuat *Rule* dari *Decision Tree* yang Telah Dibuat

- Rule 1: IF *tear-prod-rate* = *reduce* THEN recommendation = *none*
- Rule 2: IF *tear-prod-rate* = *normal* AND *astigmatism* = *no* AND *age* = *young* THEN recommendation = *soft*
- Rule 3:
- Rule 4:
- Rule 5:
- Rule 6:
- Rule 7:
- Rule 8:
- Rule 9:

TUGAS:

“Silahkan dilanjutkan *rule* berdasarkan *decision tree* yang telah dibuat”