

# STROJNO UČENJE

## 2. Domaća Zadaća

Krešimir Špes

0036419866

ak. god. 2011. / 2012.

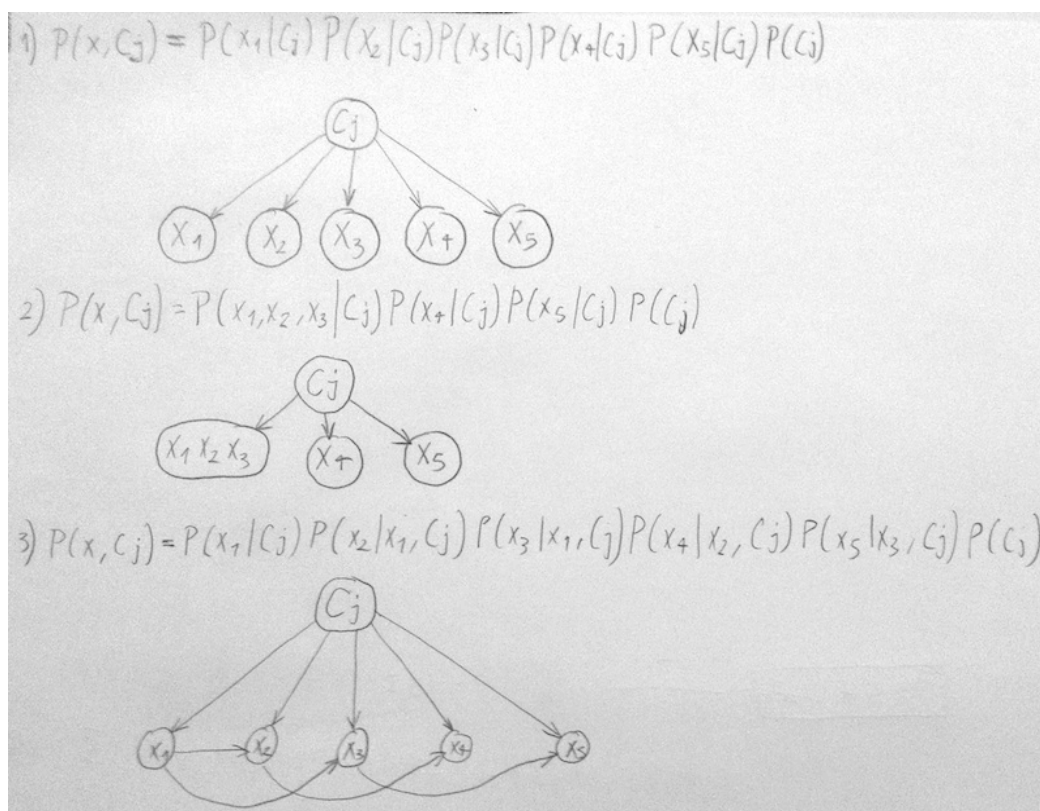
a)

1. **Odabir modela:** Bayesom modeliramo sustav temeljen na vjerojatnosti. Dobar model koji može imati nelinearne granice između klasa ako modeliramo npr polunaivni bayesov klasifikator. Stoga je potencijalno model dosta dobar i prilagodljiv. Što praksa svakako dokazuje.
2. **Definiranje funkcije gubitka:**

$$\mathbb{E}[L] = \sum_{k=1}^K \sum_{j=1}^K \int_{\mathbf{x} \in \mathcal{R}_j} L_{kj} p(\mathbf{x}, \mathcal{C}_k) d\mathbf{x}$$

3. **Optimizacijski postupak:** Iteriranjem kroz primjere za učenje definiramo izglednosti klase koje onda koristimo za klasifikaciju.

b)



c)

formula za broj parametara modela je:

$$K - 1 + \sum_{k=1}^n (K_k - 1)K$$

gdje je K broj klasa a  $K_k$  dimenzionalnost pojedine značajke

1.  $N = (2 - 1) + 2 * [ 2-1 + 2-1 + 3-1 + 3-1 + 3-1 ] = 17$
2.  $N = (2 - 1) + 2*(3*3*3-1) + 2*2*(3-1) = 61$
3.  $N = (2 - 1) + 2*2*3*(3-1) + 2*2*(3-1) = 33$

d) Algoritam nije dobar jer kao heuristiku koristi empirijsku pogrešku. Cilj svakog algoritma za stojno učenje jest što bolja moć generalizacije. Fokusirajući se na empirijsku pogrešku riskiramo prenaučeni model.

Astar je svakako bolji algoritam od uspona na vrh ali i dalje ne daje najoptimalniju odluku. Jedino optimalno dakako je pretražiti sav prostor stanja, ali to je nerijetko vremenski zahtjevno pa čak i nemoguće. No mogli bi reći da je Astar dobra kompromis odluka za ovaj algoritam.

Od poboljšanja bi preporučio kao heuristiku uzeti pogrešku generalizacije ili mix pogreške generalizacije i empirijske pogreške u nekom omjeru. Npr 3:1 u korist pogreške generalizacije.

Možda bi mogli iskoristi i nešto poput genetskog algoritma umjesto Astar-a kako bi dobili bolje rezultate.