

R1

Klasifikátor 1R

Klasifikátor 1R predstavuje veľmi jednoduchú formu klasifikácie inštancií, kde rozhodnutie sa robí na základe jedného vybraného atribútu. Aj keď rozhodovanie na základe jedného atribútu sa zdá byť až priveľmi jednoduché, v mnohých prípadoch takýto prístup dokáže klasifikovať s dostatočne malou chybou. Pre každý atribút sa vytvorí toľko pravidiel, koľko hodnôt tento atribút nadobúda. Pre každú hodnotu sa vytvorí jedno pravidlo. Do dôsledkovej časti pravidla sa vyberie tá trieda, v ktorej sa inštancie s vybranou hodnotou atribútu vyskytujú najčastejšie. V prípade, že dve alebo viac tried má rovnakú početnosť, vyberie sa z nich trieda náhodne. Pre dané pravidlo sa vypočíta chyba, ktorá predstavuje percento inštancií, ktoré majú danú hodnotu atribútu, ale patria do inej triedy. Nakoniec sa vypočíta celková chyba pre sadu pravidiel prislúchajúcich každému atribútu a vyberie sa ten s najmenšou chybou. Každé pravidlo potom zodpovedá jednej hodnote vybraného atribútu.

Trénovacie dáta:

TABUĽKA 2 ZJEDNODUŠENÝ PŘÍKLAD LEKÁRSKEJ DIAGNOSTIKY

#	Bolesť hrdla	Horúčka	Opuchnuté uzliny	Plýnos	Bolesť hlavy	Diagnóza
1	áno	áno	áno	áno	áno	zápal hrdla
2	nie	nie	nie	áno	áno	alergia
3	áno	áno	nie	áno	nie	prechladnutie
4	áno	nie	áno	nie	nie	zápal hrdla
5	nie	áno	nie	áno	nie	prechladnutie
6	nie	nie	nie	áno	nie	alergia
7	nie	nie	áno	nie	nie	zápal hrdla
8	áno	nie	nie	áno	áno	alergia
9	nie	áno	nie	áno	áno	prechladnutie
10	áno	áno	nie	áno	áno	prechladnutie

Po aplikovaní algoritmu na tieto dáta dostaneme nasledovné hodnoty Chyby a Celkovej chyby:

	DIAGNÓZA		ČÍSLA	
	dno	zápal hrdla	3/5	6/10
BOLEST HRDLA	nie	alergia	3/5	6/10
	dno	prechladnutie	1/5	
HORÚČKA	nie	alergia	2/5	3/10
	dno	zápal hrdla	0/3	
OPUCH UŠŤÍN	nie	prechladnutie	3/7	3/10
	dno	prechladnutie	4/8	
PLNY NOS	nie	zápal hrdla	0/2	4/10
	dno	prechladnutie	3/5	
BOLEST HLAVY	nie	zápal hrdla	3/5	6/10
	dno	prechladnutie	3/5	

Výsledkom je klasifikátor – model vo forme nasledovných pravidiel:

- IF HORÚČKA = dno THEN prechladnutie
- IF HORÚČKA = nie THEN alergja
- IF OPUCH UŠŤÍN = dno THEN zápal hrdla
- IF OPUCH UŠŤÍN = nie THEN prechladnutie