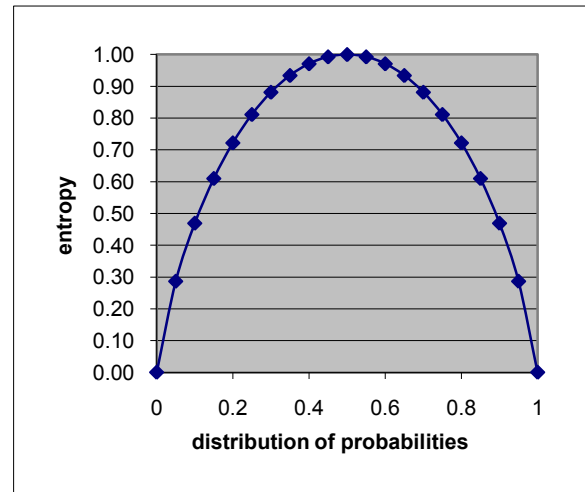


Entropy and Information Gain

probability of class 1	probability of class 2	entropy $E(p_1, p_2) = -p_1 \cdot \log_2(p_1) - p_2 \cdot \log_2(p_2)$
p_1	$p_2 = 1-p_1$	
0	1	0.00
0.05	0.95	0.29
0.10	0.90	0.47
0.15	0.85	0.61
0.20	0.80	0.72
0.25	0.75	0.81
0.30	0.70	0.88
0.35	0.65	0.93
0.40	0.60	0.97
0.45	0.55	0.99
0.50	0.50	1.00
0.55	0.45	0.99
0.60	0.40	0.97
0.65	0.35	0.93
0.70	0.30	0.88
0.75	0.25	0.81
0.80	0.20	0.72
0.85	0.15	0.61
0.90	0.10	0.47
0.95	0.05	0.29
1	0	0.00



$$Gain(S, A) = E(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} E(S_v)$$

attribut A
 set S
 number of examples in the subset
 probability of the "branch"
 number of examples in set S

Training set

Person	Age	Prescription	Astigmatic	Tear_Rate	Lenses
P1	young	myope	no	normal	YES
P2	young	myope	no	reduced	NO
P4	young	hypermetrope	no	reduced	NO
P5	young	myope	yes	normal	YES
P6	young	myope	yes	reduced	NO
P7	young	hypermetrope	yes	normal	YES
P8	young	hypermetrope	yes	reduced	NO
P10	pre-presbyopic	myope	no	reduced	NO
P11	pre-presbyopic	hypermetrope	no	normal	YES
P14	pre-presbyopic	myope	yes	reduced	NO
P17	presbyopic	myope	no	normal	NO
P18	presbyopic	myope	no	reduced	NO
P19	presbyopic	hypermetrope	no	normal	YES
P20	presbyopic	hypermetrope	no	reduced	NO
P21	presbyopic	myope	yes	normal	YES
P22	presbyopic	myope	yes	reduced	NO
P24	presbyopic	hypermetrope	yes	reduced	NO

Testing set

Person	Age	Prescription	Astigmatic	Tear_Rate	Lenses
P3	young	hypermetrope	no	normal	YES
P9	pre-presbyopic	myope	no	normal	YES
P12	pre-presbyopic	hypermetrope	no	reduced	NO
P13	pre-presbyopic	myope	yes	normal	YES
P15	pre-presbyopic	hypermetrope	yes	normal	NO
P16	pre-presbyopic	hypermetrope	yes	reduced	NO
P23	presbyopic	hypermetrope	yes	normal	NO

Petra Kralj Novak

Petra.Kralj.Novak@ijs.si

<http://kt.ijs.si/PetraKralj/IPSKnowledgeDiscovery1011.html>