

Übungen zu Data Mining 1

Aufgabenblatt 1

Aufgabe 1

Zeigen Sie durch die geeignete Wahl eines Gegenbeispiels, dass die Kullback-Leibler-Divergenz nicht symmetrisch ist.

Aufgabe 2

Pierre-Simon Laplace hat das sogenannte *Indifferenzprinzip* (auch *Prinzip vom unzureichenden Grund* genannt) aufgestellt. Demnach sollte bei unterscheidbaren Ereignissen, die sich gegenseitig ausschließen, die Gleichverteilung angenommen werden, wenn keine weiteren Informationen über die Ereignisse bekannt sind. Argumentieren Sie mit Hilfe der Entropie und der Eigenschaft der Kullback-Leibler Divergenz $D_{KL}(p, q) \geq 0$, warum dies ein sinnvoll gewähltes Prinzip ist.

Aufgabe 3

Beweisen Sie den Zusammenhang

$$D_{KL}(P(X, Y), P(X)P(Y)) = I(X, Y)$$