

Spams filtern mit Bayes

Füllwörter: der, die das, ein, eine, ich, du, Sie, mein, dein, in, für, auf, ist

Wort	Anteil der Ham-Mails, in denen Wort vorkommt	Anteil der Spam-Mails, in denen Wort vorkommt
Geld	2%	4%
Grüße	80%	80%
haben	30%	30%
Konto	1%	5%
Matthias	15%	0,05%
Nigeria	0,1%	5%
überweisen	1%	5%
Universität	2%	0,1%
Urlaub	2%	5%
viele	50%	50%
wie	10%	10%

Aufgabe: Berechne die Spammigkeit folgender Nachrichten:

1. Hast du das Geld für die Universität überwiesen?

2. Wie ist dein Urlaub in Nigeria?

Viele Grüße

Matthias

3. Überweisen Sie das Geld auf das Konto in Nigeria.

Denke dir weitere Sätze aus und berechne deren Spamwahrscheinlichkeit!