1. (4 Punkte) In einer Halle kommt es an 1 % der Tage zu einem Brand. Der Rauchmelder schlägt bei einem Brand mit 99 % Wahrscheinlichkeit Alarm, an Tagen ohne Brand gibt es an 2 % der Tage einen Fehlalarm. Heute schlägt der Melder Alarm.
   (a) Benenne Hypothese, Beobachtung, Prior und Likelihood.
   (b) Wie wahrscheinlich brennt es?

2. (5 Punkte) Eine Krankheit tritt bei 2 % der Bevölkerung auf. Ein Test ist bei 95 % der Kranken positiv und bei 3 % der Gesunden (falsch) positiv.
   (a) Wie viele von 10 000 Personen sind krank und positiv getestet, wie viele gesund und positiv getestet?
   (b) Wie wahrscheinlich ist eine positiv getestete Person krank?
   (c) Die Person macht einen zweiten, unabhängigen Test derselben Art, der wieder positiv ist. Wie wahrscheinlich ist sie jetzt krank? (Tipp: Der Posterior aus b) ist der neue Prior.)

3. (6 Punkte) Von 100 E-Mails sind 40 Spam. Gezählt wurde, in wie vielen E-Mails die Wörter „gratis“ und „Gewinn“ vorkommen:

	Spam	kein Spam
Anzahl E-Mails	40	60
enthält gratis	24	6
enthält Gewinn	20	3

(a) Eine neue E-Mail enthält beide Wörter. Berechne mit der naiven Annahme, wie wahrscheinlich sie Spam ist.

(b) Eine E-Mail enthält „gratis“, aber nicht „Gewinn“. Wie wahrscheinlich ist sie Spam?

(c) Was ist an der Annahme naiv?

4. (6 Punkte) Ein naiver Bayes-Klassifikator mit Laplace-Glättung soll Texte den Klassen Sport und Kochen zuordnen. Trainingsdaten:

Text	Klasse
tor tor heute	Sport
sieg heute	Sport
nudeln salz	Kochen
nudeln wasser	Kochen
heute nudeln	Kochen

   (a) Gib die Priors an und die Wortanzahlen je Klasse. Wie groß ist das Vokabular?

   (b) Berechne $P(\text{heute} \mid \text{Sport})$, $P(\text{tor} \mid \text{Sport})$, $P(\text{heute} \mid \text{Kochen})$ und $P(\text{tor} \mid \text{Kochen})$ mit Laplace-Glättung.

   (c) Zu welcher Klasse gehört der Text „heute tor“? Mit welcher Wahrscheinlichkeit?

   (d) Was ändert sich beim Text „heute tor mittag“?

5. (4 Punkte) (a) Warum rechnet man beim naiven Bayes-Klassifikator mit den Logarithmen der Wahrscheinlichkeiten? Warum ändert sich dadurch nicht, welche Klasse gewinnt?

   (b) Wozu dient die Laplace-Glättung? Was würde ohne sie mit einem Text passieren, der ein Wort enthält, das in den Trainingstexten einer Klasse nie vorkommt?

6. (4 Punkte) Gegeben sind die drei Dokumente „der hund bellt“, „der hund schläft“ und „die katze schläft“. Es gilt $\mathrm{IDF} = \ln \frac{1+n}{1+\mathrm{df}} + 1$ (n Anzahl der Dokumente, df Anzahl der Dokumente mit dem Wort).
   (a) Berechne die IDF-Werte für „der“, „hund“ und „bellt“.
   (b) Gib den normierten TF-IDF-Vektor des ersten Dokuments an (Einträge für der, hund, bellt).
   (c) Welches Wort hat im ersten Dokument das größte Gewicht? Warum ist das sinnvoll?

7. (6 Punkte) Schreibe zwei Funktionen für einen naiven Bayes-Klassifikator mit Laplace-Glättung, ohne scikit-learn: `trainiere(texte, klassen)` zählt die Wörter je Klasse (Wörter durch Leerzeichen getrennt), bestimmt das Vokabular und die Priors. `klassifiziere(text, ...)` gibt die wahrscheinlichste Klasse zurück; unbekannte Wörter werden übergangen. Rechne mit Logarithmen. Teste mit den Trainingsdaten „tor tor heute“ (Sport), „sieg heute“ (Sport), „nudeln salz“ (Kochen), „nudeln wasser“ (Kochen), „heute nudeln“ (Kochen) und den Texten „heute tor“, „salz heute“ und „sieg nudeln“.