Klassifikation mit Naive Bayes

Naive Bayes Text Classification

- Anwendungen
  - Spam-Nonspam
  - Welcher Autor schrieb welches Dokument?
  - Männlicher/weiblicher AutorIn?
- Textklassifikation:
  - Input: Dokument d;
  - Output: Eine der Klassen $c_1,c_2,\ldots,c_J$
- Bag of Words-Darstellung:
- Naive Bayes:

great	2
love	2
recommend	1
laugh	1
happy	1
...	...

$$P(c \mid d) = \frac{P(d \mid c)P(c)}{P(d)}$$

Bayes-Regel

- Benutze Bayes-Regel: video 1, video 2, video 3
- Maximum a posteriori Klasse

$$c_{MAP} = \underset{c \in C}{\operatorname{argmax}}\, P(c \mid d)$$

$$= \underset{c \in C}{\operatorname{argmax}}\, \frac{P(d \mid c)P(c)}{P(d)}$$

$$= \underset{c \in C}{\operatorname{argmax}}\, P(d \mid c)P(c)$$

Beispiel MLE bei normalverteilten Daten

- Likelihood bei unabhängig Gauss-verteilten Daten

$$L\left(\mu, \sigma^2; x_1, \ldots, x_n\right) = \prod_{j=1}^{n} f_X(x_j, \mu, \sigma^2)$$

$$= \prod_{j=1}^{n} \left(2\pi\sigma^2\right)^{-1/2} \exp\left(-\frac{1}{2}\frac{\left(x_j-\mu\right)^2}{\sigma^2}\right)$$

$$= \left(2\pi\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^{n}\left(x_j-\mu\right)^2\right)$$

- Log davon (Position des Maximum ändert sich nicht):

$$\log L\left(\mu, \sigma^2; x_1, \ldots, x_n\right) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\left(\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{j=1}^{n}\left(x_j-\mu\right)^2$$

- Besitzt ein Maximum bei den Werten

$$\boxed{\begin{aligned}\hat{\mu} &= \frac{1}{n}\sum_{j=1}^{n} x_j\\ \hat{\sigma}^2 &= \frac{1}{n}\sum_{j=1}^{n}\left(x_j-\hat{\mu}\right)^2\end{aligned}}$$

Beispiel MLE bei binomialverteilten Daten

- X=0: Münze ist Kopf
  X=1: Münze ist Zahl
- Wahrscheinlichkeit p für Kopf, 1-p für Zahl
- Wahrscheinlichkeit, bei n Würfen x mal Kopf zu erhalten:

$$f(x|p) = \binom{n}{x} p^x (1-p)^{n-x}$$

- Nehme Log davon:

$$\ln \ell(p|x) = \binom{n}{x} \ln \left[p^x (1-p)^{n-x}\right]$$
$$= x \ln p + (n-x) \ln(1-p)$$

- Wo ist das Maximum?
  → Ableiten nach p

$$0 = \frac{d}{dp} \ln \ell(p|x) = \frac{x}{p} + \frac{n-x}{1-p}$$

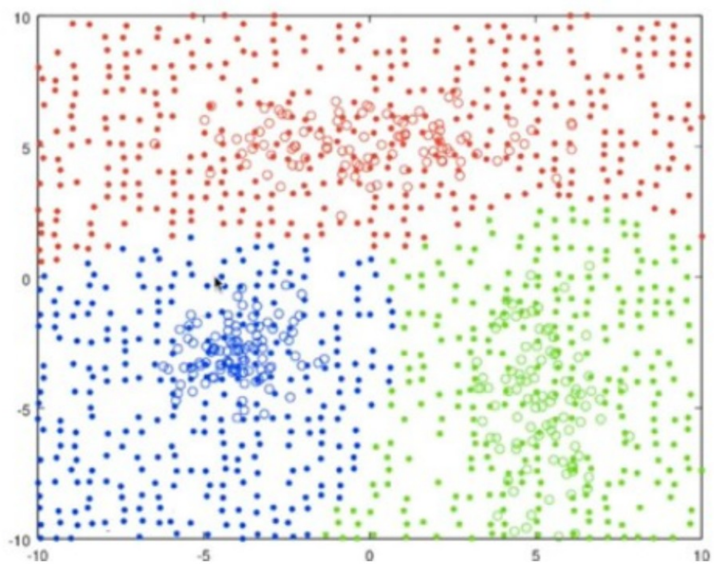
- Dann Auflösen nach p:

$$\frac{p}{x} = -\frac{1-p}{n-x} \Rightarrow (x-n)p = x(p-1)$$
$$\Rightarrow -np = -x$$
$$\boxed{\Rightarrow p = \frac{x}{n}}$$

Likelihood-Grenzen bei Klassifikation

- Likelihood-Grenzen in 2D-Daten:

$$p(c=1|x) \overset{?}{>} p(c=2|x)$$



Klassifizieren nach Bayes

- Bestimme wahrscheinlichste Klasse:

$$c_{MAP} = \underset{c \in C}{\operatorname{argmax}} P(d \mid c) P(c)$$

$$= \underset{c \in C}{\operatorname{argmax}} P(x_1, x_2, \ldots, x_n \mid c) P(c)$$

- P(c): Wahrscheinlichkeit der Klasse (über alle Dokumente)
- Dokument besteht aus Features **unabhängigen** $x_1,\ldots x_n$

$$P(x_1, x_2, \ldots, x_n|c) = P(x_1|c)P(x_2|c)\cdots P(x_n|c)$$

- Für jede Kasse: P(xi|c): Wahrscheinlichkeit der Anwesenheit von Wort i)

$$P(x_1, x_2, \ldots, x_n|c) = P(x_1|c)P(x_2|c)\cdots P(x_n|c)P(c)$$

- Zähler: Wie oft kommt Wort wi in Klasse c vor?

$$P(w_i|c) = \frac{\text{count}(w_i, c)}{\sum_{w \in \text{Vok}} \text{count}(w, c)}$$

- Nenner: Wie viele Wörter enthält Klasse c (Anzahl Wörter in Spam-Dokumenten)
- Insgesamt: Welcher Bruchteil aller Wörter aller Dokumente der Klasse c hat das Wort wi?
  - Mega-Dokument aller Dokumente, welche zur Klasse c gehören
  - Häufigkeit des Wortes wi/ Anzahl Worte im Mega-Dokument

- Problem bei $P(w_i|c) = 0$

$$c_{MAP} = \arg\max_c \hat{P}(c) \prod_i \hat{P}(x_i|c) = 0$$

$$P(w_i|c) = \frac{\text{count}(w_i, c)}{\sum_{w \in \text{Vok}} \text{count}(w, c)}$$

„Add-1-Smoothing!“

$$P(w_i|c) = \frac{\text{count}(w_i, c) + 1}{\sum_{w \in \text{Vok}} (\text{count}(w, c) + 1)}$$

Kochrezept

- 1. Erstelle ein Vokabular aller möglicher Wörter
- 2. Berechne $P(c_j)$ für alle Klassen aus dem Trainingsset

$$P(c) = \frac{\text{Anz. Dok in Klasse c}}{\text{Anz. Dok}}$$

- 3. Für jede Klasse c erstelle Mega-Dokument aller Dokumente der Klasse c: text(c)
- 4. Für das Wort w_i in Klasse c:

$$P(w_i|c) = \frac{\text{count}(w_i, c) + 1}{\sum_{w \in \text{Vok}} (\text{count}(w, c) + 1)}$$

D.h. (Wie oft tritt Wort wi in text(c) auf +1) / (Anzahl Wörter in text(c) + Anzahl Wörter in Vokabular)

Kochrezept 2

- 5. Berechne das Klassenlabel

$$c_{MAP} = \arg\max_{c} \hat{P}(c) \prod_{i} \hat{P}(x_i|c) = 0$$

- arg max ist die Funktion, die jenes c zurück gibt, für welches der Ausdruck rechts davon maximal ist

Beispiel Wetterdaten

- Wird der Ausflug dieses Jahr stattfinden? Bisher galt:

Aussicht	Temp	Feuchtigkeit	Wind	Ausflug findet statt
Sonne	Heiss	Hoch	FALSCH	Nein
Sonne	Heiss	Hoch	WAHR	Nein
Bewölkt	Heiss	Hoch	FALSCH	Ja
Regen	Mild	Hoch	FALSCH	Ja
Bewölkt	Mild	Hoch	FALSCH	Ja
Bewölkt	Kühl	Normal	WAHR	Ja
Sonne	Mild	Normal	WAHR	Nein
Sonne	Mild	Normal	FALSCH	Ja
Sonne	Mild	Normal	FALSCH	Ja
Regen	Kühl	Normal	WAHR	Ja
Regen	Kühl	Normal	WAHR	Ja
Regen	Kühl	Hoch	FALSCH	Nein
Bewölkt	Heiss	Normal	FALSCH	Ja
Regen	Mild	Hoch	WAHR	Nein

Zielvariable

- Dieses Jahr gilt: Es ist sonnig bei kühlenTemperaturen, Es ist feucht und windig

Sonne	Kühl	Hoch	WAHR	?

Beispiel Wetterdaten

- „Vokabular“, Features: x_1=Aussicht, x_2=Temp, x_3=Feuchtigkeit, x_4=Wind
- Satz von Bayes:

$$P(\mathrm{Ja}|x) = \frac{P(x|\mathrm{Ja})P(\mathrm{Ja})}{P(x)}$$

$$P(\mathrm{Nein}|x) = \frac{P(x|\mathrm{Nein})P(\mathrm{Nein})}{P(x)}$$

- **Wähle „Ja“, wenn P(Ja|x) > P(Nein|x), sonst „Nein“**
- P(x) spielt dafür keine Rolle
- Wie gross ist
  - P(Ja)
  - P(Nein)

?

Ausflug findet statt
Nein
Nein
Ja
Ja
Ja
Ja
Nein
Ja
Ja
Ja
Ja
Nein
Ja
Nein

Beispiel Wetterdaten

- „Vokabular", Features: x_1=Aussicht, x_2=Temp, x_3=Feuchtigkeit, x_4=Wind
- Satz von Bayes:

$$P(\text{Ja}|x) = \frac{P(x|\text{Ja})P(\text{Ja})}{P(x)}$$

$$P(\text{Nein}|x) = \frac{P(x|\text{Nein})P(\text{Nein})}{P(x)}$$

- P(Ja) = 9/14
- P(Nein) = 5/14
- **Wie gross ist P(x|Ja)?**

Zielvariable

Ausflug findet statt
Nein
Nein
Ja
Ja
Ja
Ja
Nein
Ja
Ja
Ja
Ja
Nein
Ja
Nein

- Insbesondere für x=

Sonne	Kühl	Hoch	WAHR

Beispiel Wetterdaten

- Featurevektor x: x_1=Aussicht, x_2=Temp, x_3=Feuchtigkeit, x_4=Wind
- Likelihood (="Wahrscheinlichkeit der Daten" x)

$$\begin{aligned} P(x|\text{Ja}) &= P(x_1, x_2, x_3, x_4|\text{Ja}) \\ &= P(x_1|\text{Ja}) \cdot P(x_2|\text{Ja}) \cdot P(x_3|\text{Ja}) \cdot P(x_4|\text{Ja}) \end{aligned}$$

- Welche Bedingung/Annahme erlaubte diesen Schritt?
- Wie berechnen wir nun z.B.

$$P(x_1|\text{Ja})$$

?

- Insbesondere für x_1= $\boxed{\text{Sonne}}$

Beispiel Wetterdaten

- Hilfreich zur Berechnung von $P(x_1|\mathrm{Ja})$

Aussicht (x_1)	Temp (x_2)	Feuchtigkeit (x_3)	Wind (x_4)	Ausflug findet statt
Sonne	Heiss	Hoch	FALSCH	Nein
Sonne	Heiss	Hoch	WAHR	Nein
Bewölkt	Heiss	Hoch	FALSCH	Ja
Regen	Mild	Hoch	FALSCH	Ja
Bewölkt	Mild	Hoch	FALSCH	Ja
Bewölkt	Kühl	Normal	WAHR	Ja
Sonne	Mild	Normal	WAHR	Nein
Sonne	Mild	Normal	FALSCH	Ja
Sonne	Mild	Normal	FALSCH	Ja
Regen	Kühl	Normal	WAHR	Ja
Regen	Kühl	Normal	WAHR	Ja
Regen	Kühl	Hoch	FALSCH	Nein
Bewölkt	Heiss	Normal	FALSCH	Ja
Regen	Mild	Hoch	WAHR	Nein

Aussicht	Ja	Nein
Sonne	2	3
Bewölkt	4	0
Regen	3	2

Temp	Ja	Nein
Heiss	2	2
Mild	4	2
Kühl	3	1

Feucht	Ja	Nein
Hoch	3	4
Normal	6	1

Wind	Ja	Nein
WAHR	3	3
FALSCH	6	2

P(x|c)

- Bedingte Wahrscheinlichkeiten $P(x_1|c_1)$
- Für x=

Sonne	Kühl	Hoch	WAHR

Aussicht	Ja	Nein	
Sonne	2	3	$P(.\|\text{Sonne})$
Bewölkt	4	0	$P(.\|\text{Bewölkt})$
Regen	3	2	$P(.\|\text{Regen})$
	$P(.\|\text{Ja})$	$P(.\|\text{Nein})$	

Aufgabe 1:

- P(Sonne|Ja) = ?
- P(Bewölkt|Nein)= ?
- P(Ja|Regen) = ?

$P(x|c)$

- Bedingte Wahrscheinlichkeiten $P(x_1|c_1)$

Aussicht	Ja	Nein
Sonne	2	3
Bewölkt	4	0
Regen	3	2

- P(Sonne|Ja) = 2/9
- P(Bewölkt|Nein)= 0/5=0
- P(Ja|Regen) = 3/5

Beispiel Wetterdaten

- Zur Erinnerung: x=

Sonne	Kühl	Hoch	WAHR

- $$P(x|\text{Ja}) = P(x_1, x_2, x_3, x_4|\text{Ja})$$
  $$= P(x_1|\text{Ja}) \cdot P(x_2|\text{Ja}) \cdot P(x_3|\text{Ja}) \cdot P(x_4|\text{Ja})$$
- Gesucht ist u.a. $P(x_1|\text{Ja})$
- P(sonnig|Ja) = ?
  P(kühl|Ja) = ?
  P(Hoch|Ja) = ?
  P(WAHR|Ja) = ?

Aussicht	Ja	Nein
Sonne	2	3
Bewölkt	4	0
Regen	3	2

Temp	Ja	Nein
Heiss	2	2
Mild	4	2
Kühl	3	1

Feucht	Ja	Nein
Hoch	3	4
Normal	6	1

Wind	Ja	Nein
WAHR	3	3
FALSCH	6	2

Likelihood P(x|Ja)

- Zur Erinnerung: x=

Sonne	Kühl	Hoch	WAHR

- $$P(x|\text{Ja}) = P(x_1, x_2, x_3, x_4|\text{Ja})$$
$$= P(x_1|\text{Ja}) \cdot P(x_2|\text{Ja}) \cdot P(x_3|\text{Ja}) \cdot P(x_4|\text{Ja}) \cdot P(\text{Ja})$$

- Gesucht ist u.a. $P(x_1|\text{Ja})$

- P(sonnig|Ja) = 2/9
  P(kühl|Ja) = 3/9
  P(Hoch|Ja) = 3/9
  P(WAHR|Ja) = 3/9

- Also insgesamt

$$P(x|\text{Ja}) = \frac{2}{9} \cdot \frac{3}{9} \cdot \frac{3}{9} \cdot \frac{3}{9} = 0.0082$$

Aussicht	Ja	Nein
Sonne	2	3
Bewölkt	4	0
Regen	3	2

Temp	Ja	Nein
Heiss	2	2
Mild	4	2
Kühl	3	1

Feucht	Ja	Nein
Hoch	3	4
Normal	6	1

Wind	Ja	Nein
WAHR	3	3
FALSCH	6	2

Likelihood P(x|Nein)

- Zur Erinnerung: x=

Sonne	Kühl	Hoch	WAHR

- $$P(x|\text{Nein}) = P(x_1|\text{Nein}) \cdot P(x_2|\text{Nein}) \cdot P(x_3|\text{Nein}) \cdot P(x_4|\text{Nein}) \cdot P(\text{Nein})$$

- P(Nein) = 5/14

- P(sonnig|Nein) = ?
  P(kühl|Nein) = ?
  P(Hoch|Nein) = ?
  P(WAHR|Nein) = ?

Aussicht	Ja	Nein
Sonne	2	3
Bewölkt	4	0
Regen	3	2

Temp	Ja	Nein
Heiss	2	2
Mild	4	2
Kühl	3	1

Feucht	Ja	Nein
Hoch	3	4
Normal	6	1

Wind	Ja	Nein
WAHR	3	3
FALSCH	6	2

Beispiel Wetterdaten

- Zur Erinnerung: x=

Sonne	Kühl	Hoch	WAHR

- $$\begin{aligned} P(\text{Nein}|x) =& P(x|\text{Nein})\frac{P(\text{Nein})}{P(x)} \\ =& P(x_1|\text{Nein}) \cdot P(x_2|\text{Nein}) \\ & \cdot P(x_3|\text{Ja}) \cdot P(x_4|\text{Ja}) \cdot \frac{P(\text{Nein})}{P(x)} \end{aligned}$$

- P(Nein) = 5/14
- P(sonnig|Nein) = 3/5
  P(kühl|Nein) = 1/5
  P(Hoch|Nein) = 4/5
  P(WAHR|Nein) = 3/5
- $$P(x|\text{Nein}) = \frac{3}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} \cdot \frac{3}{5} = 0.0576$$
- $$P(\text{Nein}|x) = P(x|\text{Nein})\frac{P(\text{Nein})}{P(x)} = \frac{0.0206}{P(x)}$$

Aussicht	Ja	Nein
Sonne	2	3
Bewölkt	4	0
Regen	3	2

Temp	Ja	Nein
Heiss	2	2
Mild	4	2
Kühl	3	1

Feucht	Ja	Nein
Hoch	3	4
Normal	6	1

Wind	Ja	Nein
WAHR	3	3
FALSCH	6	2

Klassenzuordnung

- P(Ja|x):
  - P(Ja) = 9/14
  - $P(x|\mathrm{Ja}) = \frac{2}{9} \cdot \frac{3}{9} \cdot \frac{3}{9} \cdot \frac{3}{9} = 0.0082$
  - $P(\mathrm{Ja}|x) = P(x|\mathrm{Ja})\frac{P(\mathrm{Ja})}{P(x)} = 0.0053/P(x)$
- P(Nein|x):
  - P(Nein)=5/14
  - $P(x|\mathrm{Nein}) = \frac{3}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} \cdot \frac{3}{5} = 0.0576$
  - $P(\mathrm{Nein}|x) = P(x|\mathrm{Nein})\frac{P(\mathrm{Nein})}{P(x)} = 0.0206/P(x)$
- **P(Nein|x) ist grösser als P(Ja|x), daher klassieren wir**

Sonne	Kühl	Hoch	WAHR

→ **Nein**