

Erwartungshorizont Sum of Us 2021

1 Ham of Spam?

Lösungen

(a) 1 : 19.91

Spam

(b) 1 : 30.08

Spam

Kurze Erläuterung

(a) Weil

$$\begin{aligned} &P(\text{Ham} \mid \text{congratulations, click, here, to, see, your, birthday, card}) \\ &\propto \frac{1}{4} \cdot \frac{0+1}{4+2} \cdot \frac{0+1}{4+2} \cdot \frac{1}{4} \cdot \frac{1+1}{4+2} \cdot \frac{1}{4} \cdot \frac{1+1}{4+2} \cdot \frac{0+1}{4+2} \cdot \frac{4}{7} = \frac{1}{2^7 \cdot 3^5 \cdot 7}, \\ &P(\text{Spam} \mid \text{congratulations, click, here, to, see, your, birthday, card}) \\ &\propto \frac{1}{3} \cdot \frac{2+1}{3+2} \cdot \frac{2+1}{3+2} \cdot \frac{2}{3} \cdot \frac{0+1}{3+2} \cdot \frac{1}{3} \cdot \frac{0+1}{3+2} \cdot \frac{0+1}{3+2} \cdot \frac{3}{7} = \frac{2}{5^5 \cdot 7}, \end{aligned}$$

beträgt das zu bestimmende Verhältnis $3.125 : 62.208 \approx 1 : 19.91$.

Dieses ist kleiner als $5 : 95 = 1 : 19$.

(b) Weil

$$\begin{aligned} &P(\text{Ham} \mid \text{congratulations, click, here, to, see, your, birthday, card}) \\ &\propto \frac{1}{2^7 \cdot 3^5 \cdot 7} \cdot \frac{7}{4} \cdot \frac{15}{100} = \frac{1}{5 \cdot 2^{11} \cdot 3^4}, \\ &P(\text{Spam} \mid \text{congratulations, click, here, to, see, your, birthday, card}) \\ &\propto \frac{2}{5^5 \cdot 7} \cdot \frac{7}{3} \cdot \frac{85}{100} = \frac{17}{2 \cdot 3 \cdot 5^6}, \end{aligned}$$

beträgt das zu bestimmende Verhältnis $3125 : 470016 \approx 1 : 150.41$.
Dieses ist kleiner als $5 : 95 = 1 : 19$.

Scores

a)	1 : 19.91	30 Punkte	30 Punkte bei exakt gleichem Verhältnis 20 Punkte für $1 : x$ mit $19 < x < 21$ -10 Punkte für Ham
	Spam	10 Punkte	
b)	1 : 150.41	30 Punkte	30 Punkte bei gleichem Verhältnis 20 Punkte für $1 : x$ mit $149 < x < 152$ -10 Punkte für Ham
	Spam	10 Punkte	

2 Musik

Lösungen

- (a) Chancenverhältnis: $1 : 3.52 : 0.78$
Vorhergesagter Musikgeschmack: Rock
- (b) $k = 4$
Nachbar*innen: 2, 7, 10, 12
Vorhergesagter Musikgeschmack: Rock
- (c) Chancenverhältnis: $1 : 0.21 : 0.57$
Vorhergesagter Musikgeschmack: Pop
- (d) $k = 3$
Nachbar*innen: 1, 5, 13
Vorhergesagter Musikgeschmack: Klassik

Kurze Erläuterung

- (a) Weil

$$P(\text{Pop} \mid M, \text{Rock}, \text{mittel}) \propto \frac{2}{5} \cdot \frac{1}{5} \cdot \frac{2}{5} \cdot \frac{5}{13} = \frac{2^2}{5^2 \cdot 13},$$
$$P(\text{Rock} \mid M, \text{Rock}, \text{mittel}) \propto \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{4}{13} = \frac{3^2}{2^4 \cdot 13},$$
$$P(\text{Klassik} \mid M, \text{Rock}, \text{mittel}) \propto \frac{2}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{4}{13} = \frac{1}{2^3 \cdot 13},$$

beträgt das zu bestimmende Verhältnis $64 : 225 : 50 \approx 1 : 3.52 : 0.78$.

- (b) Die Nutzer*innen 2, 7, 10 und 12 haben alle die höchste gemeinsame Ähnlichkeit mit Josef. Die Ähnlichkeit beträgt $5/6$. Zwei dieser vier Nachbar*innen mögen Rock, eine*r mag Pop und eine*r mag Klassik.
- (c) Weil

$$P(\text{Pop} \mid W, \text{Klassik}, \text{niedrig}) \propto \frac{3}{5} \cdot \frac{1+1}{5+4} \cdot \frac{2+1}{5+3} \cdot \frac{5}{13} = \frac{1}{2^2 \cdot 13},$$
$$P(\text{Rock} \mid W, \text{Klassik}, \text{niedrig}) \propto \frac{1}{4} \cdot \frac{0+1}{4+4} \cdot \frac{2+1}{4+3} \cdot \frac{4}{13} = \frac{3}{2^3 \cdot 7 \cdot 13},$$
$$P(\text{Klassik} \mid W, \text{Klassik}, \text{niedrig}) \propto \frac{2}{4} \cdot \frac{3+1}{4+4} \cdot \frac{0+1}{4+3} \cdot \frac{4}{13} = \frac{1}{7 \cdot 13},$$

beträgt das zu bestimmende Verhältnis $14 : 3 : 8 \approx 1 : 0.21 : 0.57$.

- (d) Die Nutzer*innen 1, 5 und 13 haben die größte gemeinsame Ähnlichkeit mit Lisa. Die Ähnlichkeit beträgt $3/4$. Zwei dieser drei Nachbar*innen mögen Klassik und eine*r mag Pop.

Scores

a)	1 : 3.52 : 0.78 Rock	30 Punkte 10 Punkte	30 Punkte bei gleichem Verhältnis 20 Punkte, wenn die Abrundung auf die erste Dezimalzahl klappt -5 Punkte für Pop oder Klassik
b)	4 2, 7, 10, 12 Rock	10 Punkte 20 Punkte 10 Punkte	5 Punkte abziehen pro Fehler
c)	1 : 0.21 : 0.57 Pop	30 Punkte 10 Punkte	20 Punkte, wenn die Abrundung auf die erste Dezimalzahl klappt
d)	3 1, 5, 13 Klassik	10 Punkte 15 Punkte 10 Punkte	5 Punkte abziehen pro Fehler -5 Punkte für Pop oder Rock

3 No Time to Die

Lösungen

- (a) Umgebung: {Ali, Cas, Eva}
Vorhergesagte Anzahl Sterne: 4.45
- (b) Umgebung: {Dr. No, Goldeneye, Spectre}
Vorhergesagte Anzahl Sterne: 4.65
- (c) Vorhergesagte Anzahl Sterne Cas: 3.27
Vorhergesagte Anzahl Sterne Dim: 3.60
Eher Dim empfehlen

Kurze Erläuterung

- (a) Wegen

$$\begin{aligned}\text{Sim}(\text{Dim}, \text{Ali}) &\approx 0.94, \\ \text{Sim}(\text{Dim}, \text{Bas}) &\approx 0.84, \\ \text{Sim}(\text{Dim}, \text{Cas}) &\approx 0.97, \\ \text{Sim}(\text{Dim}, \text{Eva}) &\approx 0.86,\end{aligned}$$

lautet die Vorhersage:

$$3.67 + \frac{0.94(4 - 3.50) + 0.97(5 - 3.25) + 0.86(3 - 3.00)}{0.94 + 0.97 + 0.86} \approx 4.45.$$

(Das Nicht-Runden von Zwischenergebnissen ändert daran nichts.)

- (b) Wegen

$$\begin{aligned}\text{Sim}(\text{Casino Royale}, \text{Dr. No}) &\approx 0.93, \\ \text{Sim}(\text{Casino Royale}, \text{GoldenEye}) &\approx 0.93, \\ \text{Sim}(\text{Casino Royale}, \text{Spectre}) &\approx 0.93,\end{aligned}$$

lautet die Vorhersage:

$$4.25 + \frac{0.93((4 - 3.00) + (2 - 3.00) + (5 - 3.80))}{3 \cdot 0.93} \approx 4.65.$$

(Das Nicht-Runden von Zwischenergebnissen ändert daran nichts.)

(c) Wegen

$$\text{Sim}(\text{No Time to Die, Dr. No}) \approx 0.87,$$

$$\text{Sim}(\text{No Time to Die, GoldenEye}) \approx 0.96,$$

$$\text{Sim}(\text{No Time to Die, Casino Royale}) \approx 0.97,$$

$$\text{Sim}(\text{No Time to Die, Spectre}) \approx 0.97,$$

lautet die Vorhersage für Cas:

$$3.33 + \frac{0.96(2 - 3.00) + 0.97(5 - 4.40) + 0.97(4 - 3.80)}{0.96 + 0.97 + 0.97} \approx 3.27$$

und die Vorhersage für Dim:

$$3.33 + \frac{0.96(2 - 3.00) + 0.97(5 - 4.40) + 0.97(5 - 3.80)}{0.96 + 0.97 + 0.97} \approx 3.60.$$

(Das Nicht-Runden von Zwischenergebnissen ändert daran nichts.)

Scores

a)	Ali, Cas und Eva 4.45	15 Pkt 30 Pkt	5 Pkt für richtige Namen, -5 Pkt für falsche 20 Pkt wenn im Bereich [4.4, 4.5]
b)	Dr. No, Goldeneye, Spectre 4.65	15 Pkt 30 Pkt	5 Pkt für richtige Namen, -5 Pkt für falsche 20 Pkt, wenn im Bereich [4.6, 4.7]
e)	Vorhersage Cas 3.27 Vorhersage Dim 3.60 Empfehlung an Dim	30 Pkt 30 Pkt 5 Pkt	20 Pkt, wenn im Bereich [3.2, 3.3] 20 Pkt, wenn im Bereich [3.5, 3.7] -5 Pkt für Cas

4 Binge-Watching

Lösungen

- (a) Cluster A: {1, 3, 14}.
Cluster B: {5, 9, 13}.
Cluster C: {2, 6, 12}.
- (b) Cluster B
Science fiction
- (c) Cluster A
Science fiction
- (d) Durchschnittliche Bewertung von Action-Serien: 1.7
Durchschnittliche Bewertung von Drama-Serien: 4.2

Kurze Erläuterung

- (a) Zeichne die Punkte und die entsprechenden Clusterzentren.
- (b) Füge den Punkt (4.0; 4.5) der Zeichnung hinzu. Nutzern in Cluster B gefällt Science fiction besser als Comedy.
- (c) Füge den Punkt (1.3; 4.3) der Zeichnung hinzu. Nutzern in Cluster A gefällt Science fiction besser als Comedy.
- (d) Wir nennen die Durchschnittswerte von Person 17 für Action- und Dramaserien a bzw. d . Dann muss

$$\frac{(1.9; 4.0) + (2.2; 4.8) + (1.9; 4.7) + (1.3; 4.3) + (a, d)}{5} = (1.8; 4.4),$$

gelten, wodurch man a und b berechnen kann.

Scores

a)	Cluster A: 1, 3 und 14 Cluster B: 5, 9 und 13 Cluster C: 2, 6 und 12	10 Punkte 10 Punkte 10 Punkte	1 Fehler = 5 Punkte 1 Fehler = 5 Punkte 1 Fehler = 5 Punkte
b)	Cluster B Science fiction	15 Punkte 10 Punkte	-10 für Comedy
c)	Cluster A Science fiction	15 Punkte 10 Punkte	-10 für Comedy
d)	Action: 1.7 Drama: 4.2	15 Punkte 15 Punkte	keine Spanne keine Spanne