

Übungsblatt 6 – Musterlösung

6.1 Statistische Modellierung

- (a) Es gibt 2 Klassen: *ADJA* und *NADJA*
- (b) Es gibt 3 binäre Features (jeweils mit den Werten Wahr und Falsch):
- Wort w Artikel
 - Wort $w + 1$ großgeschrieben
 - Wort w endet auf -er, -es, -e, -en, -em
- (c) Es gibt $2^3 = 8$ verschiedene Feature-Vektoren.

Feature-Vektor	Klasse	Beispiel
$\langle 0, 0, 0 \rangle$	ADJA	Das Auto ist <i>rot</i> und schnell
$\langle 0, 0, 0 \rangle$	NADJA	Ich <i>mag</i> das Lied
$\langle 0, 0, 1 \rangle$	ADJA	Das <i>rote</i> schnelle Auto
$\langle 0, 0, 1 \rangle$	NADJA	Ich komme <i>heute</i> nicht
$\langle 0, 1, 0 \rangle$	ADJA	Das <i>rosa</i> Auto
$\langle 0, 1, 0 \rangle$	NADJA	<i>Mein</i> Auto
$\langle 0, 1, 1 \rangle$	ADJA	Das <i>kleine</i> Auto
$\langle 0, 1, 1 \rangle$	NADJA	Artikel können <i>keine</i> Adjektive sein
$\langle 1, 0, 0 \rangle$	ADJA	Artikel können keine Adjektive sein
$\langle 1, 0, 0 \rangle$	NADJA	<i>Ein</i> schöner Tag
$\langle 1, 0, 1 \rangle$	ADJA	Artikel können keine Adjektive sein
$\langle 1, 0, 1 \rangle$	NADJA	<i>Eine</i> kleine Katze
$\langle 1, 1, 0 \rangle$	ADJA	Artikel können keine Adjektive sein
$\langle 1, 1, 0 \rangle$	NADJA	<i>Das</i> Buch
$\langle 1, 1, 1 \rangle$	ADJA	Artikel können keine Adjektive sein
$\langle 1, 1, 1 \rangle$	NADJA	<i>Eine</i> Katze

- (d) Für den Featurevektor $\langle 1, 0, 0 \rangle$ wurden keine Trainingsinstanzen gesehen. Taucht dieser Vektor in neuen Trainingsdaten auf, kann der entsprechenden Instanz keine Klasse zugewiesen werden.
- (e) Klassifikator $K : \Omega \rightarrow \{\text{ADJA}, \text{NADJA}\}$ mit

$$K(f) = \begin{cases} \text{ADJA} & f = \langle 0, 1, 1 \rangle \\ \text{NADJA} & \text{sonst} \end{cases}$$

wobei $f \in \Omega$ ein Feature-Vektor ist.

6.2 Evaluation

(a)

$$Recall = \frac{True\ positives}{True\ positives + False\ negatives}$$

$$Precision = \frac{True\ positives}{True\ positives + False\ positives}$$

$$F = \frac{2PR}{P + R}$$

Für die Klasse „gut gelaunt“ gilt: 990 Patienten sind gut gelaunt. Davon wurden 950 auch als gut gelaunt klassifiziert (also 950 True positives) und entsprechend 40 fälschlicherweise als schlecht gelaunt klassifiziert (also 40 False negatives). Zwei schlecht gelaunte Patienten wurden als gut gelaunt klassifiziert (also 2 False positives).

Für die Klasse „schlecht gelaunt“ gilt: 10 Patienten sind schlecht gelaunt. Davon wurden 8 auch als schlecht gelaunt klassifiziert (also 8 True positives) und entsprechend 2 fälschlicherweise als gut gelaunt klassifiziert (also 2 False negatives). 40 gut gelaunte Patienten wurden als schlecht gelaunt klassifiziert (also 40 False positives).

Konfusionsmatrix:

		Tatsächliche Laune	
		gut	schlecht
Diagnose	gut	950	2
	schlecht	40	8

Werte für „gut gelaunt“:

$$Recall = \frac{950}{950 + 40} = \frac{950}{990} \simeq 0,96$$

$$Precision = \frac{950}{950 + 2} = \frac{950}{952} \simeq 0,998$$

$$F = \frac{2 * 0,998 * 0,96}{0,998 + 0,96} \simeq \frac{1,916}{1,958} \simeq 0,9785$$

Werte für „schlecht gelaunt“:

$$Recall = \frac{8}{8 + 2} = \frac{8}{10} = 0,8$$

$$Precision = \frac{8}{8 + 40} = \frac{8}{48} \simeq 0,17$$

$$F = \frac{2 * 0,17 * 0,8}{0,17 + 0,8} = \frac{0,272}{0,97} \simeq 0,28$$

- (b) Das Modell erzielt schlechtere Werte für die Klasse „schlecht gelaunt“. Zu beachten ist jedoch, dass die geringe Anzahl an tatsächlich schlecht gelaunten Patienten der Grund dafür sein könnte. Da es nur zehn solcher schlecht Gelaunter gibt, reicht eine einzige falsch klassifizierte Person aus, um alle Evaluationsergebnisse deutlich zu verschlechtern. (In unserem Fall senken zwei falsch klassifizierte Personen den Recall-Wert um 0,2.) Gäbe es so viele schlecht gelaunte wie gut gelaunte Patienten, könnten wir die Performanz des Modells noch besser einschätzen.

Wir stellen außerdem fest, dass die Precision für „schlecht gelaunt“ äußerst gering ausfällt. Für „gut gelaunt“ ist die Presicion im Vergleich fast bei einem perfekten Wert. Hier ist aber anzumerken, dass die Precision für „gut gelaunt“ aufgrund der geringen Anzahl schlecht gelaunter Patienten gar nicht schlecht sein kann. Bei der Klasse „schlecht gelaunt“ gibt es umgekehrt mehr Gelegenheiten für einen False positive.

- (c) Ob bei einem Klassifikator Precision ode Recall optimiert werden sollte, ist oft Abwägungssache.

Es ist in unserem Beispiel davon auszugehen, dass die Klinik alle schlecht gelaunten Patienten erkennen will, damit keine einzige Person rückfällig wird. Dementsprechend sollte der Recall des Modells optimiert werden. Eine geringere Präzision würde nur dazu führen, dass die TherapeutInnen ein wenig mehr Zeit mit fälschlicherweise als schlecht gelaunt klassifizierten Patienten verschwenden würden. Für die Brauchbarkeit des Modells ist hier also der Recall (insbesondere der Recall für die Klasse „schlecht gelaunt“) am aussagekräftigsten.