

# Multimodale Analyse

## Signale

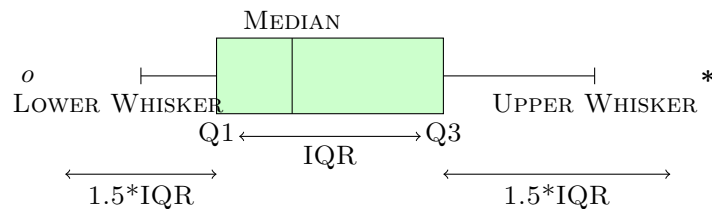
**Abtastsignal** Werte eines abgetasteten analogen Signals zu bestimmten Zeitpunkten

**Quantisierung** Binäre Repräsentation aller Abtastwerte

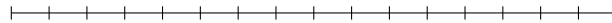
**Digitales Signal** Quantisiertes Abtastsignal

**Nyquist-Frequenz**  $f_a = 2 \cdot f_{max}$

## Klassifikation



**Boxplot**



**Konfusionsmatrix**  $x$  : Klassenzugehörigkeit ,  $y$  : Klassifikationsergebnis

**Overall accuracy**  $\frac{\text{\#korrekt klassifizierte Instanzen}}{\text{\#alle Instanzen}}$

**Weighted accuracy**  $\sum_{i=1}^N \frac{1}{N} \frac{\text{\#korrekt als Klasse } i \text{ klassifizierte Instanzen}}{\text{\#alle Instanzen in Klasse } i}$

**tp/fn/fp/fn** True/False Negatives/Positives

**Precision**  $\frac{tp}{tp + fp}$

**Recall**  $\frac{tp}{tp + fn}$

**M-Estimate**  $\frac{n_{i,j} + m \cdot p}{n_j + m}$

## Naive Bayes - Wahrscheinlichkeitstheorie

$D$  Daten

$h$  Hypothese

$P(h)$  Warsch. dass  $h$  erfüllt ist

$P(D)$  Warsch. dass  $D$  beobachtet werden

$P(D|h)$  Warsch. dass man  $D$  beobachtet wenn  $h$  erfüllt ist

$P(h|D)$  Warsch. dass  $h$  bei  $D$  gilt

$h_{MAP}$   $\operatorname{argmax} P(h|D), h \in H$

$h_{LM}$   $\operatorname{argmax} P(D|h), h \in H$

# Hidden Markov Modelle

Modell  $\lambda$

Zeitpunkt  $t$

Zustand  $S_i, i = |S|$

Beobachtung  $o_i, i = |O|$

Zustandssequenz  $Q = q_1, q_2, \dots, q_t$

Beobachtungssequenz  $O = o_1, o_2, \dots, o_t$

Forward Algorithmus  $t = 1 : \alpha_1(i) = P(o_1, o_2, \dots, o_t, q_t = S_i | \lambda)$

$$= \pi_{S_i} * b_{S_i}(o_1), i = 1, \dots, |S|$$

$t + 1 : \alpha_{t+1}(j) = P(o_1, o_2, \dots, o_{t+1}, q_{t+1} = S_j | \lambda)$

$$= \left[ \sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(o_{t+1})$$

$$1 \leq t \leq T - 1, 1 \leq j \leq N$$

Wahrscheinlichkeit daß

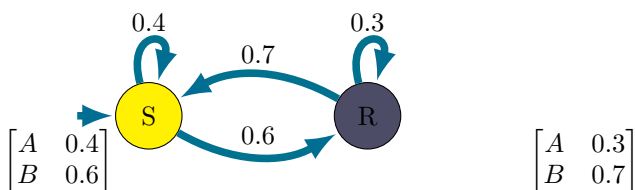
O beobachtet wird

Backward Algorithmus  $t = 1 : \beta_1(i) = \sum_{j=1}^N a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)$

$$1 \leq t \leq T - 1, 1 \leq i \leq N$$

Wahrscheinlichkeit daß

man sich in Zustand  $S_i$   
befindet und beobachtet wird



## Neuronale Netze

### Spracherkennung

**Word Error Rate**  $\frac{\text{Anzahl der fehlerhafte Worte}}{\text{Anzahl der Worte im korrekten Transkript}}$

**Ersetzungsfehler (S)**

**Auslassungsfehler (D)**

**Einfügefehler (I)**

**Sprecherunabhängigkeit** Dialekte/Nicht-Muttersprachler/etc..

**Vokabular** Sorgfältige Kontrolle der Vokab.größe/Behandlung von Worten, die sich nicht im Vokab. befinden

**Sprechweise** Isolierte Einzelwörter/mehrere Einzelwörter/fließend gesprochene "normale" Sprache

**Umgebung** Hintergrundgeräusche/Gespräche (Crosstalk)

$V$  Lexikongröße

$N$  Anzahl der Worte im Text

**N-Gram** Unigram  $P(w_i) = \frac{\text{count}(w_i)}{N} \xrightarrow{\text{adding-one}} \frac{\text{count}(w_i) + 1}{N + V}$

Bigram  $P(w_i | w_{i-1}) = \frac{\text{count}(w_{n-1} w_n)}{\text{count}(w_{n-1})} \xrightarrow{\text{adding-one}} \frac{\text{count}(w_{n-1} w_n) + 1}{\text{count}(w_{n-1}) + V}$

Trigram  $P(w_i | w_{i-1}, w_{i-2})$

$P(w_i, w_{i+1}, w_{i+2}, w_{i+3}, w_{i+4}) = P(w_i | < start >) P(w_{i+1} | w_i) P(w_{i+2} | w_{i+1}) P(w_{i+3} | w_{i+2}) P(w_{i+4} | w_{i+3}) P(w_{i+5} | w_{i+4})$

**Perplexität**  $PP(w_1, \dots, w_T) = \frac{1}{T \sqrt{P(w_1, \dots, w_T)}}$

**Absolute-discounting**