max planck institut
informatik

Ideen der Informatik

Maschinelles Lernen

Kurt Mehlhorn
Adrian Neumann
Max-Planck-Institut für Informatik

Übersicht

- Lernen: Begriff
- Beispiele für den Stand der Kunst
- Spamerkennung
- Handschriftenerkennung
  - mit und ohne Trainingsdaten
- Neuronale Netzwerke
  - Maschinelles Sehen

Lernen

- Fähigkeit, Verhalten zu verbessern auf Grund von Erfahrungen
- Verallgemeinern von Erfahrungen
- Informatik: Programmieren durch Beispiele anstatt durch Angabe eines Programms
- Ein Lernalgorithmus entwickelt das Programm aus (vielen) Daten

Typische Anwendungen

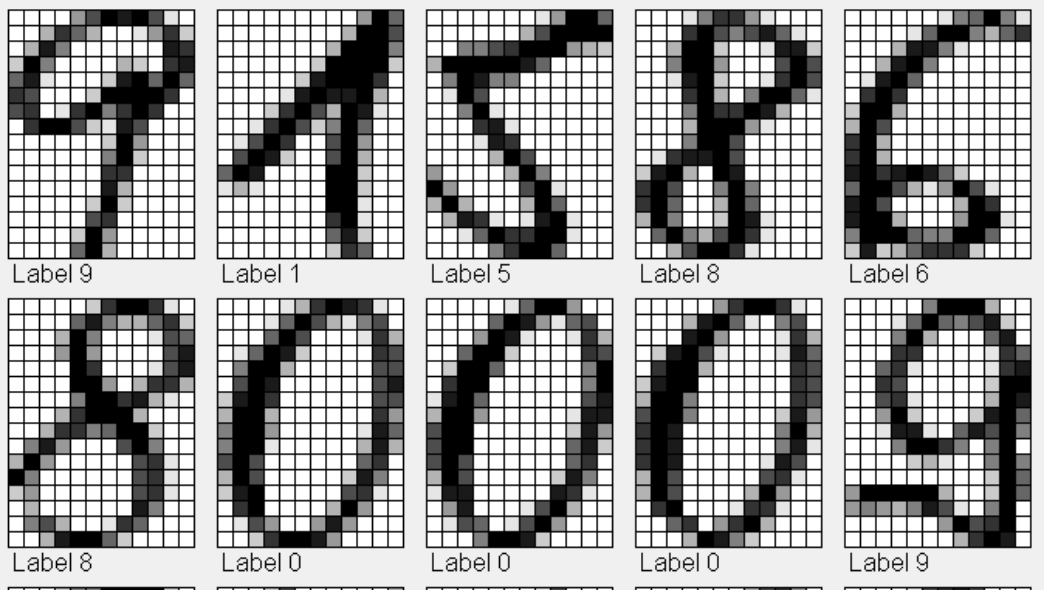
- Klassifikation Spam versus Ham, Ziffernerkennung, Verkehrszeichen-erkennung, Objekte auf Bildern, Identifikation von Personen auf Bildern, Handlungen aus Videosequenzen.
- Robotersteuerung, lerne Autofahren
- Spracherkennung, Sprachsteuerung

Potential

- Suchmaschinen mit Bild/Sprach-Anfragen
- Personenerkennung auf Videos
- XXX mit gesprochener Sprache
- Selbstfahrende Autos
- Bessere Benutzerschnittstellen
- Maschinelle Übersetzung

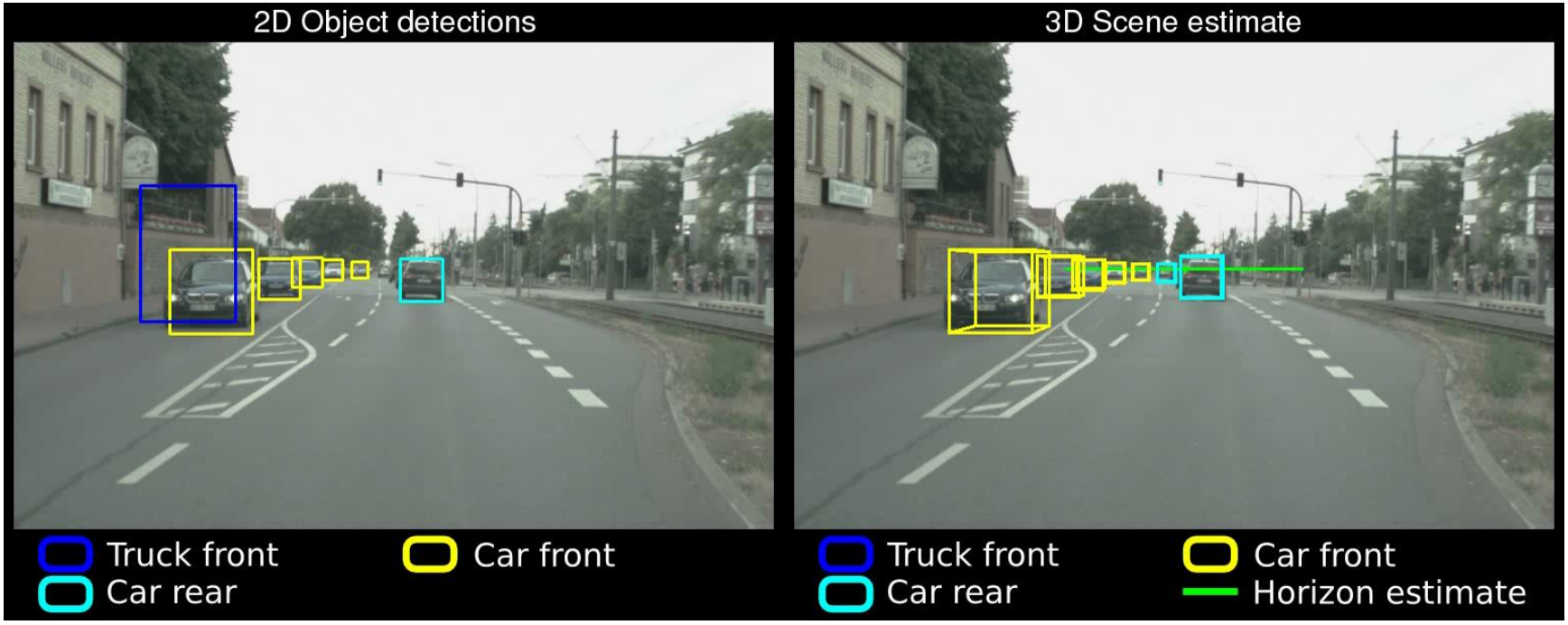
Arten von Lernen

- Supervised: mit Trainingsdaten oder sogar mit Lehrer



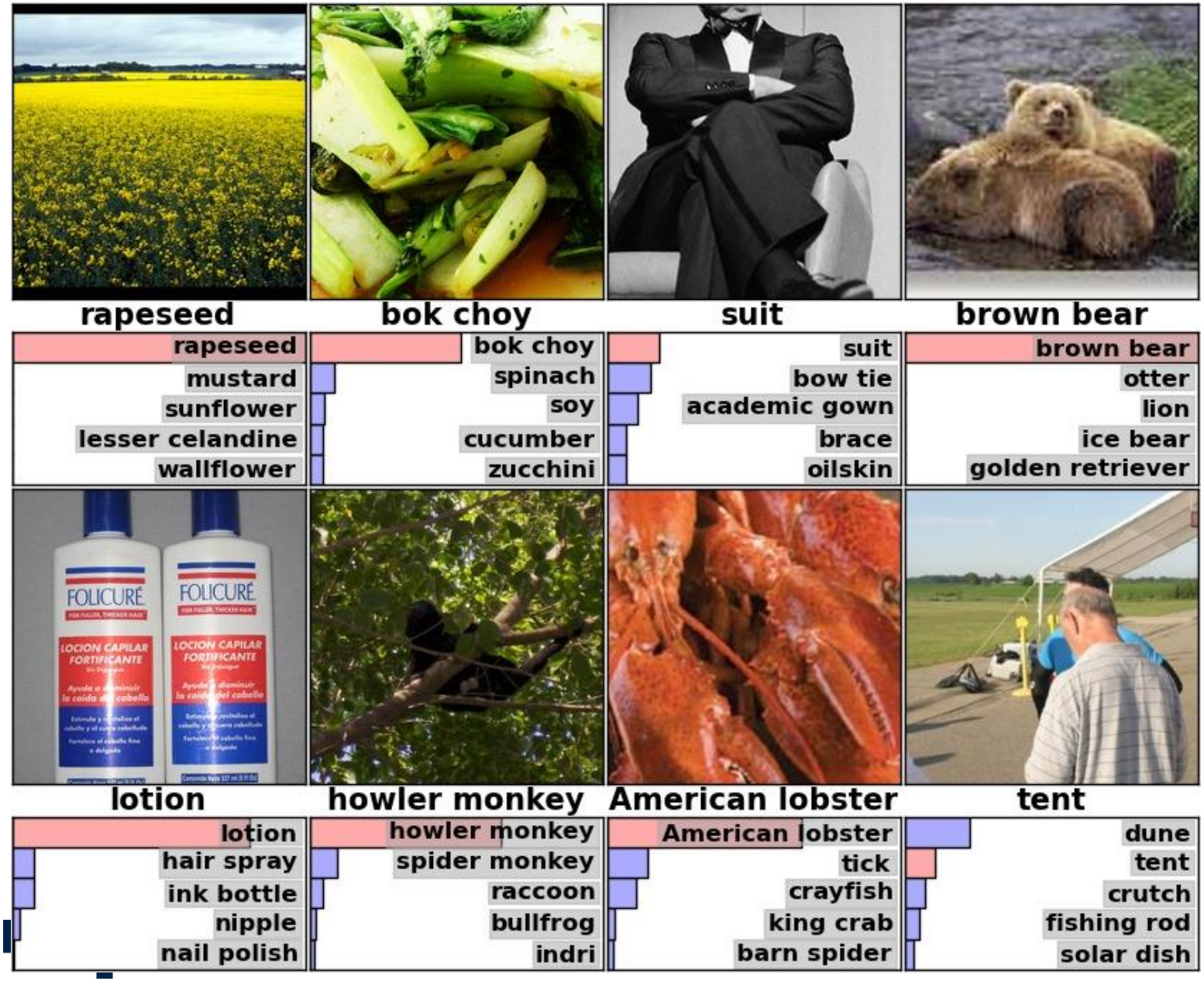
- Unsupervised: ohne Trainingsdaten; dann ist es mehr Entdecken als Lernen

Objekterkennung
Abteilung Schiele: MPI für Informatik



Personenerkennung
Abteilung Schiele: MPI für Informatik

Klassifikation (Krizhevsky et al, 2012)



Suche (Krizhevsky et al, 2012)

Aufgabe

- Was ist der Unterschied zwischen Supervised und Unsupervised Learning?
- Welche der folgenden Aussagen trifft zu?
  - Lernen = Erwerb neuer Fähigkeiten durch Erfahrungen
  - Lernen = Programmierung durch Beispiele
  - Lernen = Programmieren durch Angabe eines Programms

Spamerkennung

Spam = unerwünschte Nachrichten

Ham = erwünschte Nachrichten

Wir lernen einen Bayesschen Filter kennen

Bayessche Regel

(englischer Pfarrer und Mathematiker, 1701 -- 1761)

In einem Sack sind 900 Äpfel und 100 Paprika. Von den Äpfeln sind 10% rot und 90% grün. Bei den Paprika sind es jeweils 50%.

Ich entnehme eine Frucht zufällig. Sie ist rot. Was für eine Frucht ist es?

- Bayes: entscheide dich für den wahrscheinlicheren Fall und den bestimmt man so.

Bayessche Regel

In einem Sack sind 900 Äpfel und 100 Pap-rika. Von den Äpfeln sind 10% rot und 90% grün. Bei den Paprika sind es jeweils 50%.

P(Apfel | rot) = # rote Äpfel / # rote Früchte

(Prozentsatz der Äpfel unter den roten Früchten)

Spam versus Ham (Junk Mail)

- Absenderbasiert
  - email von Bekannten ist kein Spam
  - Schwarze Listen
- Inhaltsbasiert
  - Nutzer klassifiziert emails als gut und schlecht; System lernt daraus; Nutzer muss immer weniger eingreifen

Inhaltsbasierte Filter

- In der Trainingsphase lernen wir
  - Wahrscheinlichkeit von Ham und Spam
  - Jeweils Wahrscheinlichkeiten für Worte
- 70% ist Ham, 30% ist Spam
- Ham

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.1	0.3	0.3	0.1	0.1	0.1

- Spam

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.2	0.1	0.1	0.2	0.3	0.1

Trainingsphase

- Nutzer klassifiziert emails als Spam und Ham (damit beide Wahrscheinlichkeiten)
- Sei n die Gesamtlänge der erwünschten emails (in Worten), sei v die Anzahl der Vorkommen eines bestimmten Wortes

- Wahrscheinlichkeit des Wortes in Ham

$$= v/n$$

Inhaltsbasierte Filter (Bayes Modell)

- Wahrscheinlichkeitsverteilung auf Worten
- Ham

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.1	0.3	0.3	0.1	0.1	0.1

- Spam

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.2	0.1	0.1	0.2	0.3	0.1

- P(Text | Ham) = Produkt der Wahrscheinlichkeiten der Worte im Text

Inhaltsbasierte Filter

- Ham

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.1	0.3	0.3	0.1	0.1	0.1

- Spam

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.2	0.1	0.1	0.2	0.3	0.1

- Viagra Geld Freund
  - P(Text | Ham) = 0.1 x 0.1 x 0.1 =
  - P(Text | Spam) =
- Bei 70% Ham und 30% Spam

Inhaltsbasierte Filter

- Ham

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.1	0.3	0.3	0.1	0.1	0.1

- Spam

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.2	0.1	0.1	0.2	0.3	0.1

- Vorlesung Algorithmus schnell
  - Falls Ham:
  - Falls Spam:
- Bei 70% Ham und 30% Spam

Inhaltsbasierte Filter

- Ham

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.1	0.3	0.3	0.1	0.1	0.1

- Spam

Freund	Vorlesung	Algorithmus	Geld	Viagra	schnell
0.2	0.1	0.1	0.2	0.3	0.1

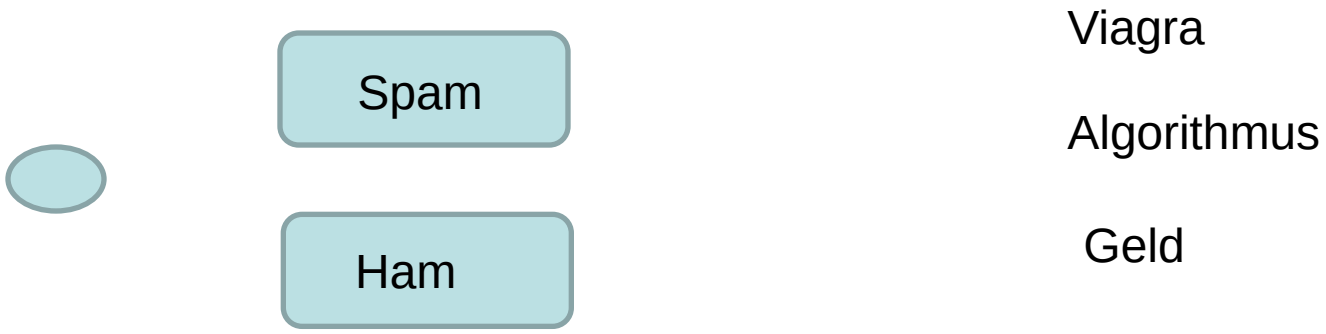
- Viagra Algorithmus schnell
  - Falls Ham: 0.1 x 0.3 x 0.1 = 3/1000
  - Falls Spam: 0.3 x 0.1 x 0.1 = 3/1000
- Bei 10% Ham und 90% Spam

Nutzungsphase

- Nutzungphase: System klassifiziert
- Verteilung wird weiter trainiert (seltene Worte)
- Nutzer kann widersprechen
- Spammer lernen auch dazu: V!agra statt Viagra

Zusammenfassung

- Wir haben Modell, wie Ereignisse (emails) erzeugt werden



- Lernen das Modell in der Trainingsphase
- Geben für jedes Ereignis die wahrscheinlichste Erklärung (Bayes)
- Klassifierung in: Geschäftspost, Privatpost, Spam

Aufgabe

- 5% der Bevölkerung erkranken an der Krankkeit X. Ein Test führt bei gesunden Patienten zu 30% zu einem positiven Ergebnis, bei kranken Patienten zu 100%. Wie hoch ist die Wahrscheinlichkeit bei einem positiven Testergebnis tatsächlich krank zu sein?
- Für das Wort Laufzeit sei die Auftretenshäufigkeit in Ham 5 mal höher als in Spam. Wie klassifizieren sie einen Text, der Algorithmus, Laufzeit und Freund enthält?

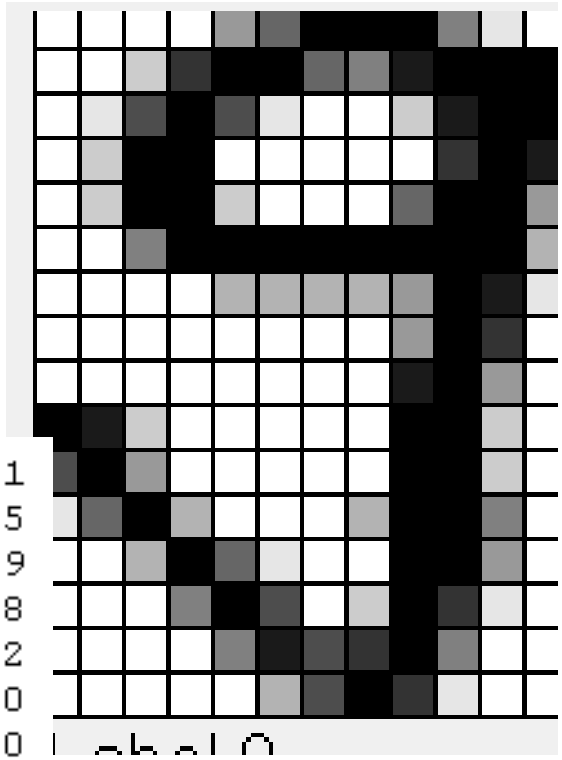
Ziffernerkennung Übersicht

- Darstellung von Bildern in Rechnern
- Trainingsdaten: handgeschriebene Ziffern
- Supervised Learning: mit Label (die Ziffer)
- Unsupervised Learning: ohne Label

Bilder = Matrizen von Zahlen

Ziffer = 12 x 16 Matrix
von Grauwerten in [0,1]

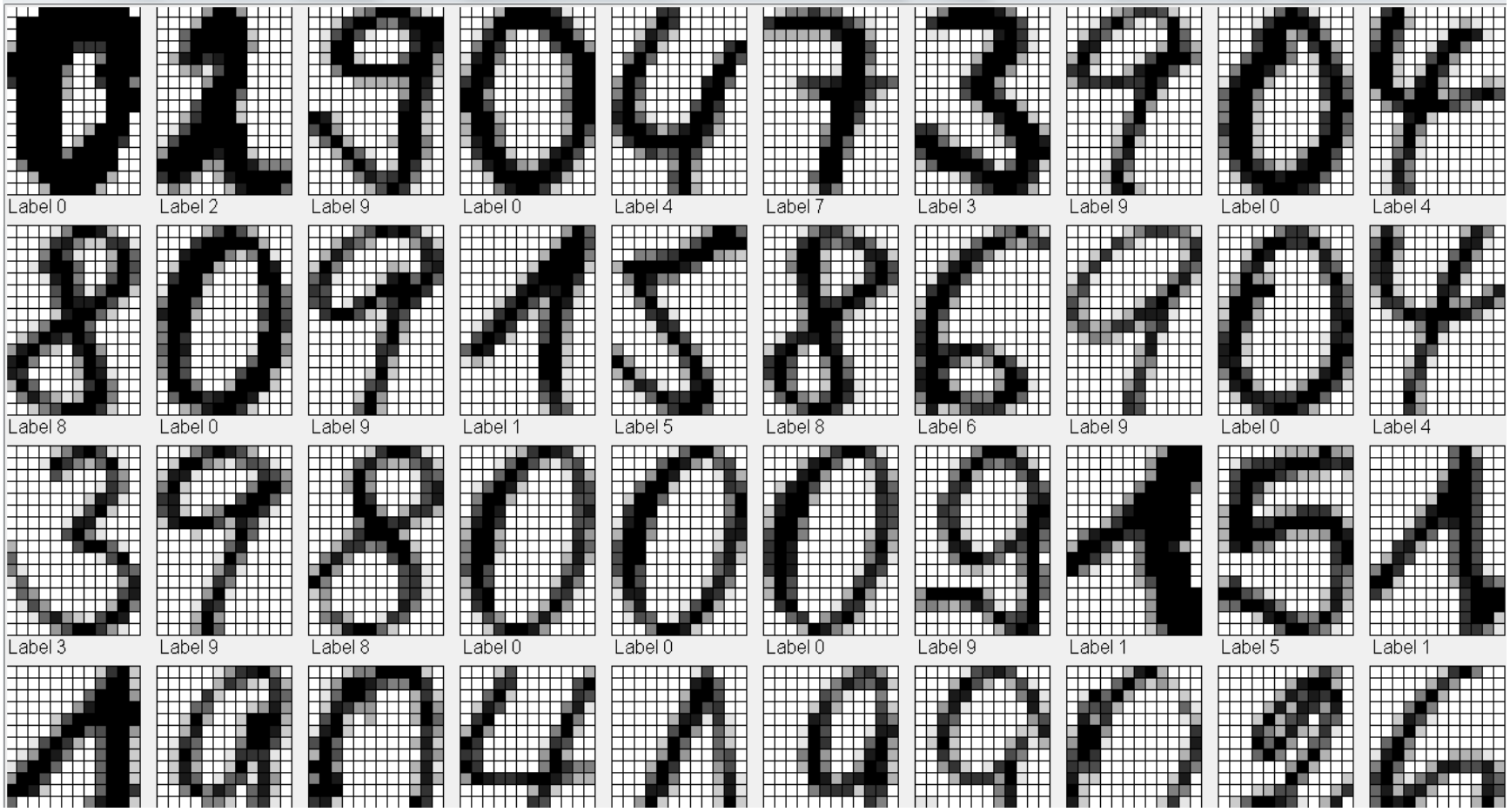
Vektor von Grauwerten
der Länge 192



```
0.0 0.0 0.0 0.2 0.3 0.4 0.8 1.0 1.0 0.7 0.3 0.1
0.0 0.5 0.8 0.9 1.0 1.0 1.0 1.0 1.0 1.0 1.0 0.5
0.3 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 0.9
0.3 1.0 1.0 1.0 1.0 1.0 0.8 1.0 1.0 1.0 1.0 0.8
0.3 1.0 1.0 1.0 1.0 0.8 0.1 0.9 1.0 1.0 0.8 0.2
0.0 0.7 1.0 1.0 1.0 0.8 0.8 1.0 1.0 1.0 0.4 0.0
0.0 0.2 1.0 1.0 1.0 1.0 1.0 1.0 0.9 0.3 0.0 0.0
0.0 0.1 0.7 1.0 1.0 1.0 1.0 0.8 0.1 0.0 0.0 0.0
0.0 0.6 1.0 1.0 1.0 1.0 0.9 0.1 0.0 0.0 0.0 0.0
0.6 1.0 1.0 1.0 1.0 1.0 1.0 0.3 0.0 0.0 0.0 0.0
0.8 1.0 1.0 0.5 0.1 0.7 1.0 0.8 0.2 0.0 0.0 0.0
0.5 1.0 1.0 0.3 0.0 0.0 0.9 1.0 0.9 0.1 0.0 0.0
0.4 1.0 1.0 0.3 0.0 0.0 0.5 1.0 1.0 0.5 0.0 0.0
0.0 0.4 1.0 1.0 0.5 0.3 0.5 1.0 1.0 1.0 0.2 0.0
0.0 0.0 0.5 1.0 1.0 1.0 1.0 1.0 1.0 0.8 0.1 0.0
0.0 0.0 0.0 0.2 0.5 0.7 1.0 1.0 0.9 0.3 0.0 0.0
```

Ihr Gehirn sieht Ziffern, ihr Auge und Computer sehen nur eine Matrix von Grauwerten

Trainingsdaten



Ziemlich gutmütig

Grundidee

- Zwei Bilder repräsentieren die gleiche Ziffer, wenn die Bilder sich ähnlich sind.

- Ähnlich = ähnliche Grauwertverteilung

Ähnlichkeit von Vektoren

- Zwei Vektoren x und y sind ähnlich,
  - wenn d = x – y kurz ist
  - Wenn der aufgespannte Winkel klein ist

Länge eines Vektors x = (x1,…,xn)

$$|| x || = Sqrt(x1^2 + x2^2 + \ldots + xn^2)$$

Winkel zwischen x und y

$$\cos \alpha = x * y / ||x|| * ||y||$$

Verfahren: Nearest Neighbor

Um die Bedeutung des Bildes p zu finden, finde das Trainingsbild x mit dist(p,x) minimal (durch lineare Suche über alle Trainingsdaten)

Gib das Label von x aus

- Erkennungsrate 0.934
- Majority of 3 nearest neighbors 0.945
- Majority of 9 nearest neighbors 0.920

- Mit cos-Distanz 0.940
- Majority of 3 nearest neighbors 0.920

Detaillierte Ergebnisse

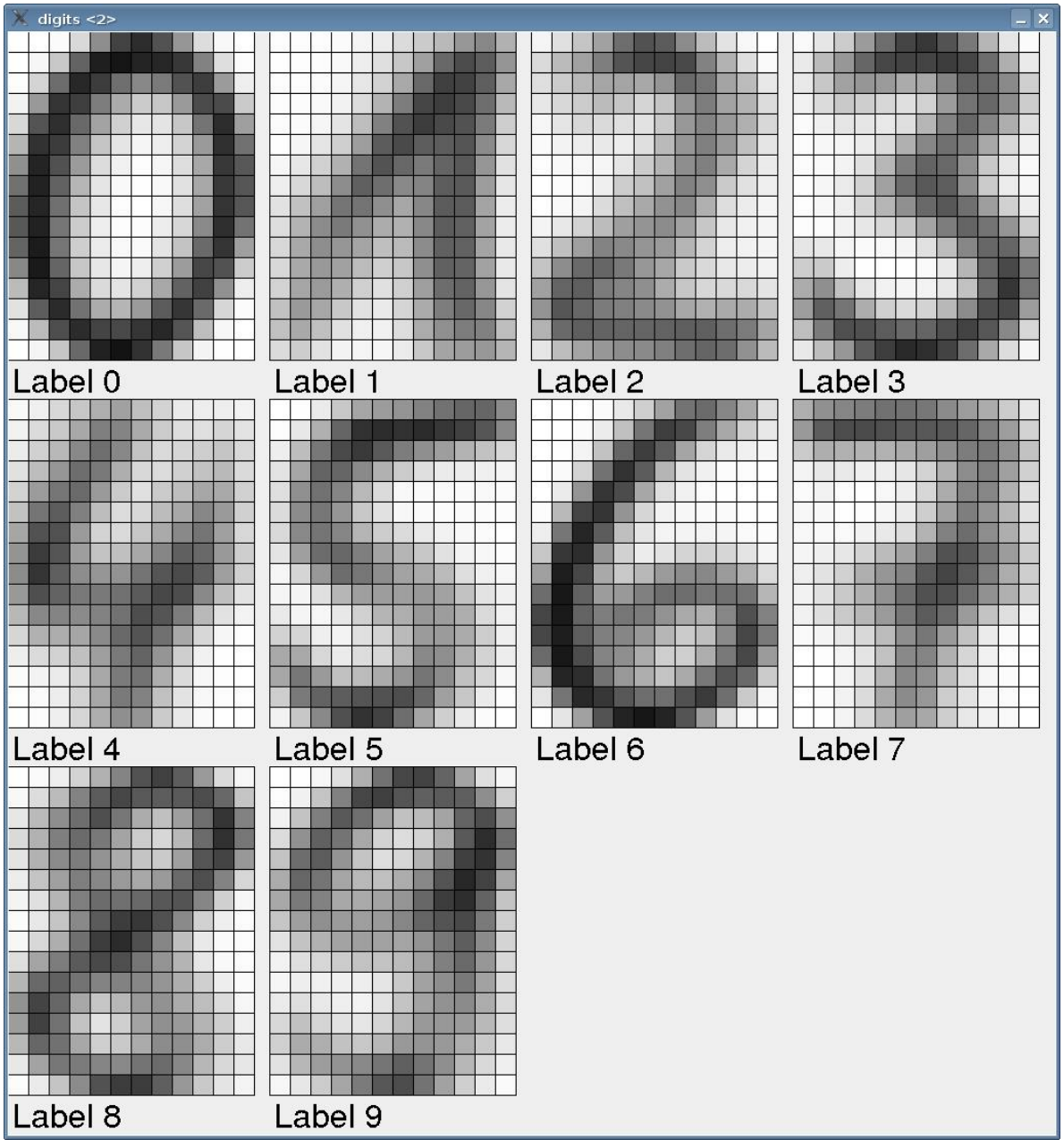
digit = 0 accuracy = 1.0
digit = 1 accuracy = 0.90
digit = 2 accuracy = 0.92
digit = 3 accuracy = 1.0
digit = 4 accuracy = 0.95
digit = 5 accuracy = 0.85
digit = 6 accuracy = 0.84
digit = 7 accuracy = 1.0
digit = 8 accuracy = 0.7
digit = 9 accuracy = 0.94

Klassifzierung ist recht gut, aber sie dauert sehr lang, da jedes Mal ALLE Trainingsdaten angeschaut werden

Klassen → Klassenzentren

- Vorbereitung: Berechne für jede Klasse (Ziffer) das Klassenzentrum durch Durchschnitts. siehe nächste Folie
- Suche: finde das nächstgelegene Zentrum (10 Vergleiche)
- Erkennungsrate: 0.854
- Mit cos-distance 0.894
- Sehr effizient, aber schlechter

Die Klassenzentren

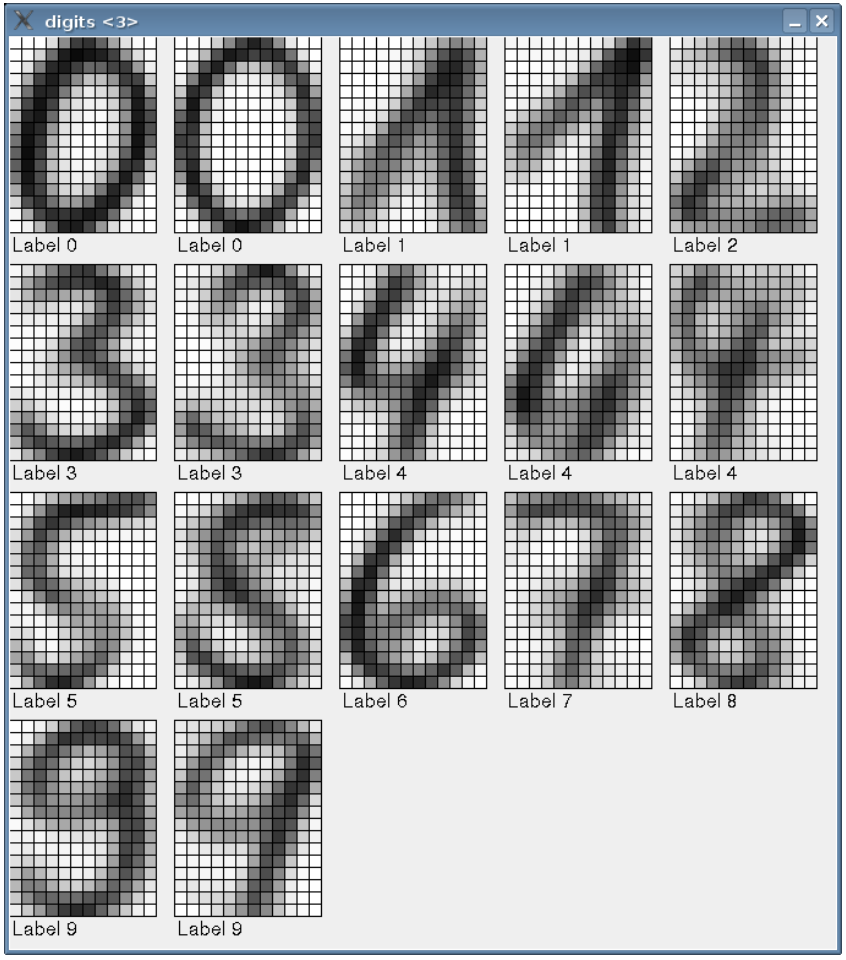


Unsupervised Lernen

- Vorbereiten der Trainingsdaten ist mühsam
- Können wir Klassen entdecken, ohne das uns Klassenlabels gesagt werden?
- Automatische Klassifizierung durch k-means Algorithmus.

- Danach Vergleich mit den Klassenzentren
- k=10, Rate 0.683
- k=17, 0.733
- with cos-distance,
- k=10, 0.728
- k=17, 0.783
- k=30, 0.864

K-means Algorithmus



- Automatische Klassifizierung in 17 Klassen
- Danach (!!!) Zuweisung eines Labels per Hand und Wegwerfen von schlechten Zentren
- Identifiziert die zwei Schreibweisen der Neun und der Eins

k-means Algorithmus

Teilt n Punkte in k Cluster (Haufen) ein.

1. Starte mit k beliebigen (zufälligen) Zentren
2. Weise jeden Punkt dem nächstgelegenen Zentrum zu und bilde so k Cluster
3. Berechne für jeden Cluster seinen Schwerpunkt; das sind die neuen Zentren
4. Gehe nach 2.