



max planck institut
informatik

Ideen und Konzepte der Informatik

Künstliche Intelligenz

Maschinelles Lernen

Kurt Mehlhorn

Übersicht

- Künstliche Intelligenz und Maschinelles Lernen: Begriff
- Beispiele für den Stand der Kunst
- Spamerkennung, Bayessche Inferenz
- Handschriftenerkennung
 - mit und ohne Trainingsdaten
- Gesichts- und Personenerkennung
- Neuronale Netzwerke, Teil 2 der Vorlesung
 - Maschinelles Sehen

Künstliche Intelligenz und M-Lernen

- Künstliche Intelligenz = intelligentes Verhalten von Maschinen.
 - Schachspielen, Sprache verstehen, Autofahren, Rätsel lösen, Gefühle verstehen, ..., alles Verhalten, das Menschen ausmacht.
- Lernen = Fähigkeit, Verhalten zu verbessern aufgrund von Erfahrungen, Verallgemeinern von Erfahrungen
aus Daten wird Verhalten.
- Zieht in den Alltag ein: Verkaufsempfehlungen, Sprachschnittstelle, personalisierte Nachrichten,

Riesiges Potential

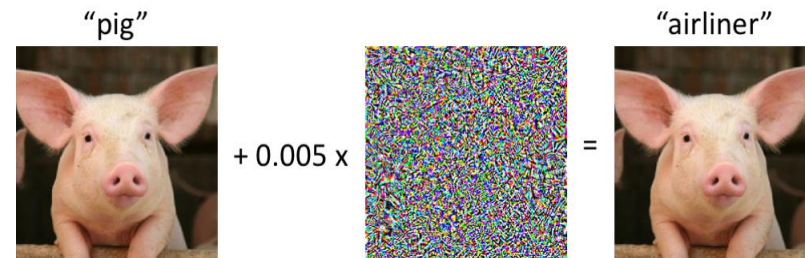
- Suchmaschinen mit Bild / Sprach-Anfragen
- Personenerkennung, Überwachung
- Selbstfahrende Autos, Roboter
- Bessere Benutzerschnittstellen, Sprache, Gesten
- Maschinelle Übersetzung
- Sachbearbeiter
- Persönliche Assistenten

“If a typical person can do a mental task with less than one second of thought, we can probably automate it using AI either now or in the near future.”

— **Andrew Ng**

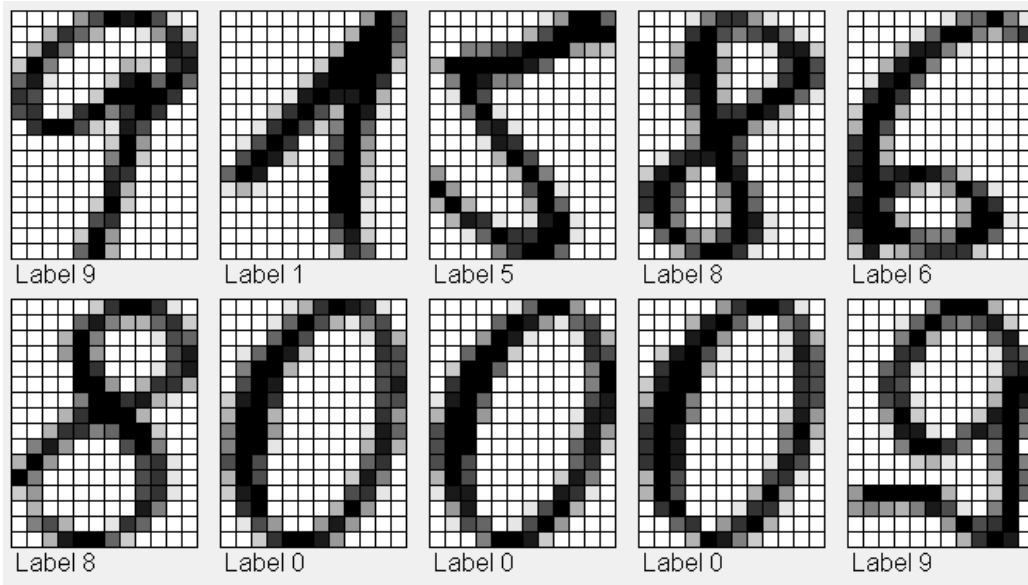
Aber auch große Gefahren

- Missbrauch (Überwachungsstaat à la China, Verlust der Privatsphäre, Google, Facebook,)
- Lernende Verfahren können in der Regel ihre Entscheidungen nicht erklären.
- Gewaltige Änderungen in der Arbeitswelt, z.B., Sachbearbeiter, Lastwagenfahrer, ...
- Verfahren werden überschätzt.
- „Die KI hat ...“



Arten von Lernen

- **Supervised:** mit Trainingsdaten oder sogar mit Lehrer



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

Der Stand der Kunst

- Personenerkennung
- Sprachübersetzung
- Klassifikation von Bildern
- Suchen in Bildersammlungen
- Spiele, Planen und strategisches Verhalten

Personenerkennung

Abteilung Schiele: MPI für Informatik



Übersetzung durch DeepL

Noch immer bedrohen die Buschbrände in Australien zahlreiche Küstenorte. Die Behörden haben nun eine Sperrzone eingerichtet, die Urlauber unverzüglich räumen müssen. Es sei die "größte Massenverlegung" der Region. Die Bilder aus dem Militärhubschrauber zeigen brennende Wälder soweit das Auge reicht. Hochhaushohe Flammenwände, dichter Qualm verdunkelt den Himmel. Die Buschbrände in Australien haben bis zu 1300 Häuser zerstört, aber so genau weiß man es nicht. In viele Regionen kann nicht einmal die Feuerwehr

The bush fires in Australia are still threatening numerous coastal towns. The authorities have now set up a restricted zone, which holidaymakers must evacuate immediately. It is said to be the "largest mass relocation" in the region. The pictures from the military helicopter show burning forests as far as the eye can see. Skyscraper-high walls of flames, dense smoke darkens the sky. The bush fires in Australia have destroyed up to 1300 houses, but it is not known for sure. In many regions not even the fire brigade can reach.

Klassifikation (Krizhevsky et al., 2012)



rapeseed

rapeseed
mustard
sunflower
lesser celandine
wallflower



bok choy

bok choy
spinach
soy
cucumber
zucchini



suit

suit
bow tie
academic gown
brace
oilskin



brown bear

brown bear
otter
lion
ice bear
golden retriever



lotion

lotion
hair spray
ink bottle
nipple
nail polish



howler monkey

howler monkey
spider monkey
raccoon
bullfrog
indri



American lobster

American lobster
tick
crayfish
king crab
barn spider



tent

dune
tent
crutch
fishing rod
solar dish



Suche (Krizhevsky et al., 2012)

Linke Spalte: Suchanfrage, andere Spalten: ähnliche Bilder im Fundus



Spiele

- AlphaGo schlägt besten menschlichen Spieler 3:0.
- Hat sich das Spielen selbst beigebracht.
 - Anfangswissen = Regeln von Go + gelernt werden soll eine Funktion von Stellungen nach Zügen.
 - Danach Spiel gegen sich selbst für 40 Tage.
- Hat Muster und Spielzüge gefunden, die Menschen bisher verborgen blieben.
- Schachmeister nutzen Computer für das Training.
- Schach hat nichts von seiner Faszination verloren.

Bayesscher Filter + Spamerkennung

Spam, Junk = unerwünschte Nachrichten

Ham = erwünschte Nachrichten

Bayesscher Filter zur Unterscheidung von Ham und Spam.
Auch im Alltagsleben sehr nützlich, z. B., zum Bewerten
von Testergebnissen.

Spam (ursprünglich): Spiced Ham, Markenname für sehr
verbreitetes Dosenfleisch, Monty Python Video

Bayes'sche 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.

Bayessche Regel im Alltag

- 5% der Bevölkerung erkranken an der Krankheit X. Es gibt einen Test:
 - Gesunde: 90% negativ, 10% positiv.
 - Kranke: 0% negativ, 100% positiv.

Sie bekommen einen positiven Befund. Was bedeutet das?

Bayes'sche Regel

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%.

$$\text{Bruchteil der Äpfel unter den roten Früchten} = \frac{\# \text{ rote Äpfel}}{\# \text{ rote Früchte}}$$

$$P(\text{Apfel} \mid \text{rot}) = \text{Bruchteil der Äpfel unter den roten Früchten}$$

Bayes'sche Regel, math. Formel

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%.

$$P(\text{Apfel} \mid \text{rot}) = \frac{P(\text{Apfel und rot})}{P(\text{rot})} =$$

$$\frac{P(\text{rot} \mid \text{Apfel}) P(\text{Apfel})}{P(\text{rot} \mid \text{Apfel}) P(\text{Apfel}) + P(\text{rot} \mid \text{Paprika}) P(\text{Paprika})}$$

Spam versus Ham (Junk-Mail)

- Absenderbasiert
 - E-Mail von Bekannten ist kein Spam
 - Schwarze Listen
- Inhaltsbasiert
 - Nutzer klassifiziert E-Mails 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	Vorle- sung	Algorith- mus	Geld	Viagra	schnell
	0.1	0.3	0.3	0.1	0.1	0.1

- Spam

	Freund	Vorle- sung	Algorith- mus	Geld	Viagra	schnell
	0.2	0.1	0.1	0.2	0.3	0.1

Trainingsphase

- Nutzer klassifiziert E-Mails als Spam und Ham (damit beide Wahrscheinlichkeiten)
- Sei n die Gesamtlänge meiner guten E-Mails (in Worten), sei v die Anzahl der Vorkommen eines bestimmten Wortes
- Wahrscheinlichkeit des Wortes in Ham $= \frac{v}{n}$

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

- $P(\text{Text} | \text{Ham}) = \text{Produkt der Ham-Wahrscheinlichkeiten der Worte im Text}$

Inhaltsbasierte Filter

- Ham

	Freund	Vorle- sung	Algorith- mus	Geld	Viagra	schnell
	0.1	0.3	0.3	0.1	0.1	0.1

- Spam

	Freund	Vorle- sung	Algorith- mus	Geld	Viagra	schnell
	0.2	0.1	0.1	0.2	0.3	0.1

- Text = Viagra Geld Freund
 - $P(\text{Text} \mid \text{Ham}) = 0.1 \cdot 0.1 \cdot 0.1 =$
 - $P(\text{Text} \mid \text{Spam}) =$
- 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

- Text = Viagra Geld Freund

- $P(\text{Text} \mid \text{Ham}) = 0.1 \cdot 0.1 \cdot 0.1 =$

- $P(\text{Text} \mid \text{Spam}) =$

- Bei 70% Ham und 30% Spam ist $P(\text{Spam} \mid \text{Text}) = \dots$

Inhaltsbasierte Filter

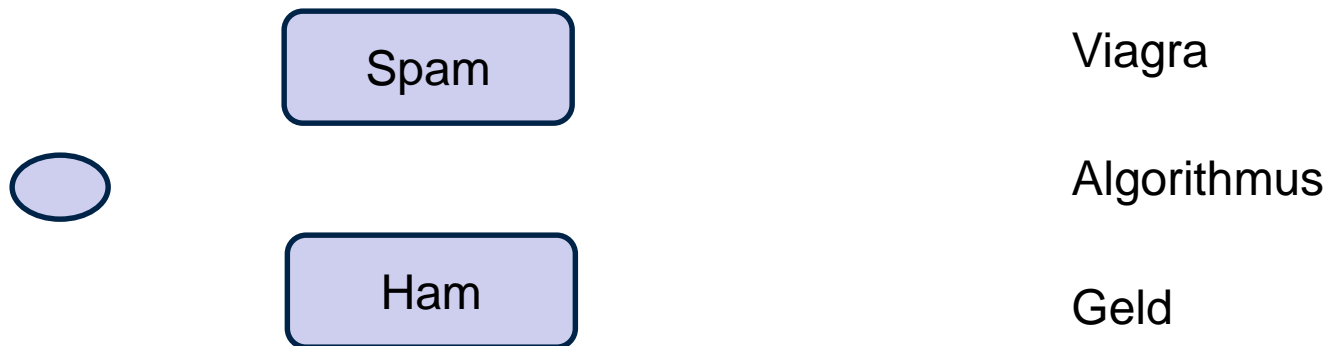
- Viagra Geld Freund
 - $P(\text{Text} | \text{Ham}) = 0.1 \cdot 0.1 \cdot 0.1 = 1/1000$
 - $P(\text{Text} | \text{Spam}) = 12/1000$
- Bei 70% Ham und 30% Spam erhalten wir
- $$P(\text{Spam} | \text{Text}) = \frac{P(\text{Text} | \text{Spam})P(\text{Spam})}{(P(\text{Text} | \text{Ham})P(\text{Ham}) + P(\text{Text} | \text{Spam})P(\text{Spam}))} =$$
$$12 \cdot 3 / (1 \cdot 7 + 12 \cdot 3) = 36/43 \approx 0.84.$$

Nutzungsphase

- Nutzungsphase: 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 (E-Mails) erzeugt werden



- Lernen Modell in der Trainingsphase und geben dann für jedes Ereignis die wahrscheinlichste Erklärung (Bayes).
- Klassifizierung in: Geschäftspost, Privatpost, Spam.

Judea Pearl: Turing Award 2012, Causal Inference



Ziffernerkennung Übersicht

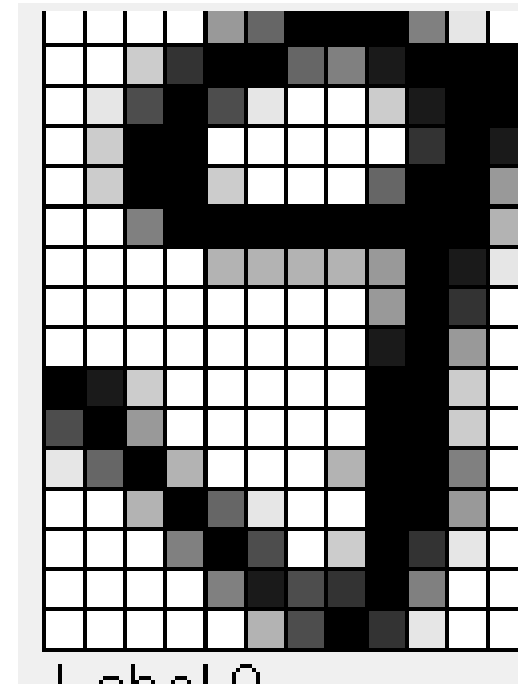
- Darstellung von Bildern in Rechnern
- Trainingsdaten: handgeschriebene Ziffern
- Supervised Learning: mit Label (die Ziffer)
- Unsupervised Learning: ohne Label

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, 10000 Beispiele

Grundidee

- Zwei Bilder repräsentieren die gleiche Ziffer, wenn die Bilder sich ähnlich sind.
- Ähnlich = ähnliche Grauwertverteilung.
- Ähnlichkeit von Vektoren der Grauwerten:
 - Euklidische Distanz zweier Vektoren
 - Cos-Distanz, Winkel zwischen Vektoren

Ähnlichkeit von Vektoren

- Zwei Vektoren x und y sind ähnlich,
 - wenn $x - y$ kurz ist
 - wenn der aufgespannte Winkel klein ist

- Länge eines Vektors $x = (x_1, \dots, x_n)$

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

- Winkel zwischen x und y

$$\cos \alpha = \frac{x \cdot y}{\|x\| \cdot \|y\|}$$

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) und gib das Label von x aus.

- Erkennungsrate bei Euklidischer Distanz 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

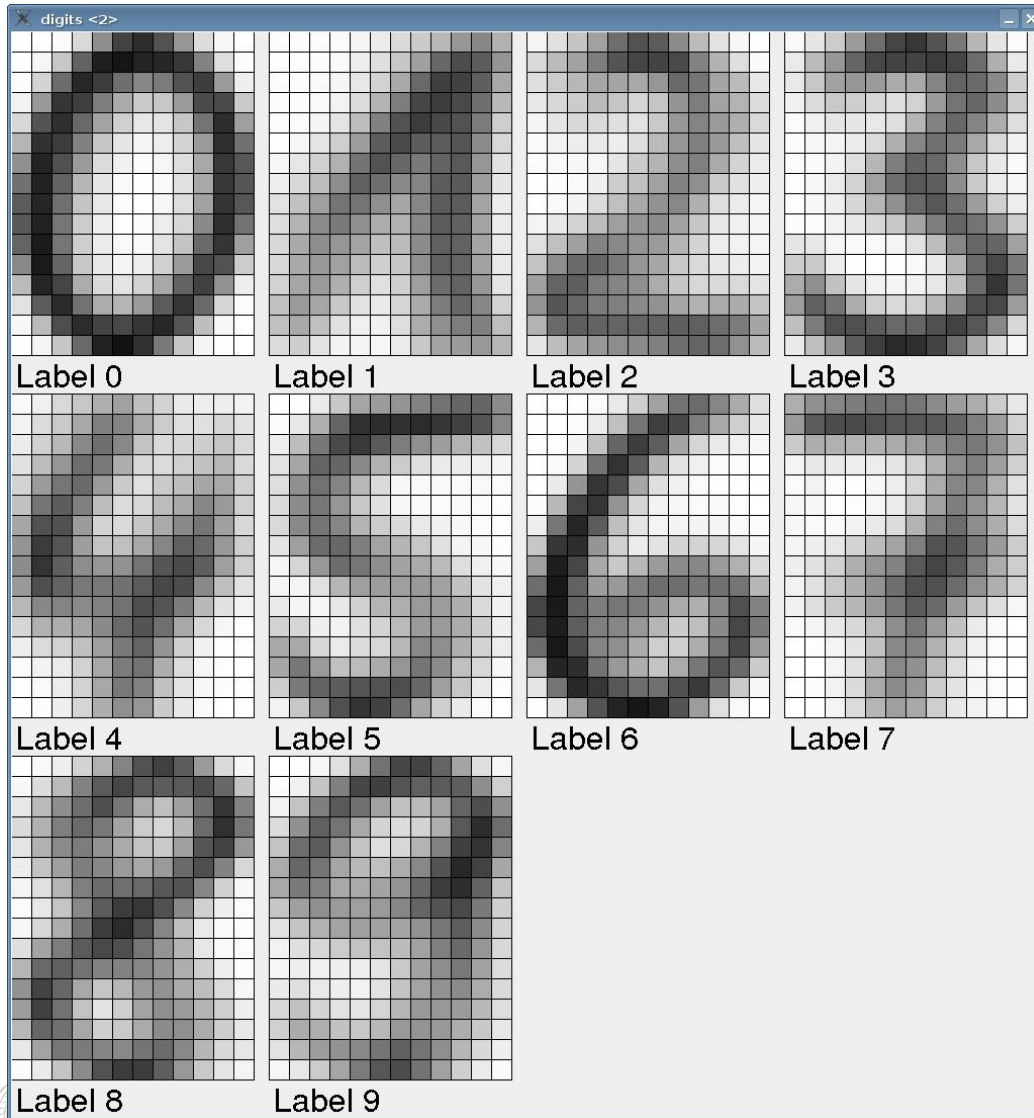
digit = 0 accuracy = 1.0
digit = 1 accuracy = 0.90
digit = 2 accuracy = 0.92
digit = 3 accuracy = 1.0
digit = 4 accuracy = 0.95
digit = 5 accuracy = 0.85
digit = 6 accuracy = 0.84
digit = 7 accuracy = 1.0
digit = 8 accuracy = 0.7
digit = 9 accuracy = 0.94

Klassifizierung 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 Durchschnittsbildung (*siehe nächste Folie*)
- Suche: Finde das nächstgelegene Zentrum (10 Vergleiche)
- Erkennungsrate bei Euklidischer Distanz: 0.854
- Mit cos-distance 0.894
- Sehr effizient, aber schlechter

Die Klassenzentren

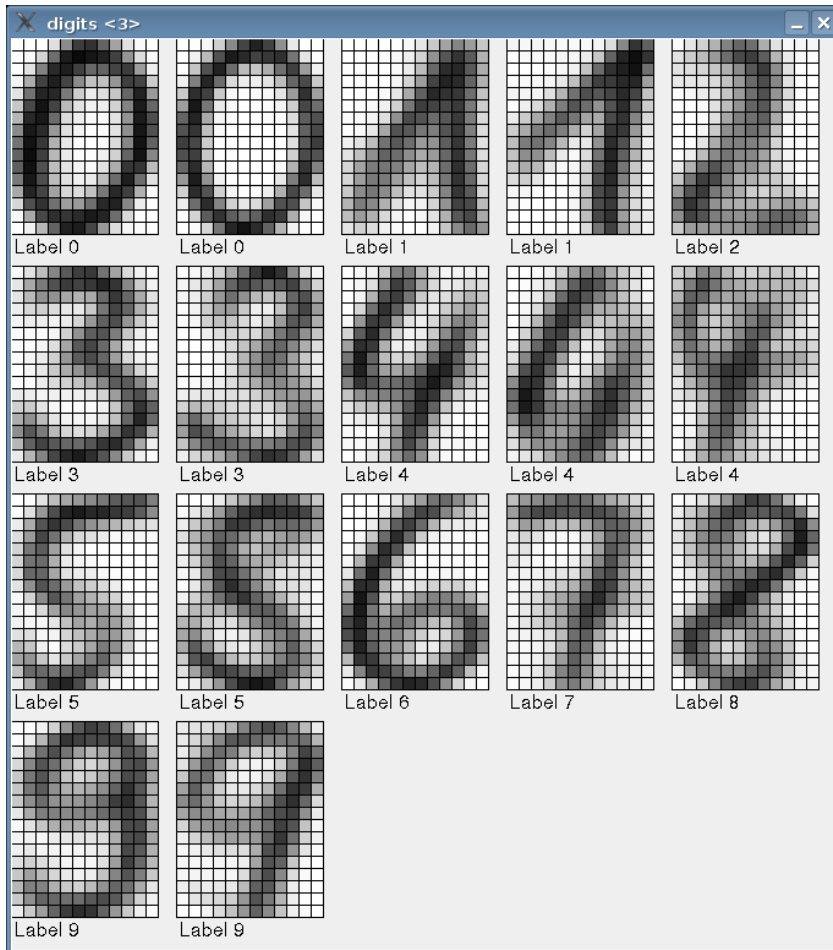


Durchschnitt
(Schwerpunkt einer
Klasse) =
Summe der Elemente
der Klasse / Anzahl der
Elemente

Unsupervised Lernen

- Vorbereiten der Trainingsdaten ist mühsam
- Können wir Klassen entdecken, ohne dass 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.

Gesichtserkennung

- Gesicht → Vektor
 - Augenfarbe, Abstand der Augen, Abstand Nase – Augen, Hautfarbe, Abstand Ohren, und so weiter, dann wie bei Ziffernerkennung.
- Aufgaben
 - Zwei Bilder: gleiche Person?
 - Klassifiziere Bilder nach abgebildeter Person
 - Finde Gesicht in einer Datenbank.



Personen



Zugehörig



Fotos von diesem Tag anzeigen