

# Maschinelles Lernen Blatt 5

Nikolas Zeitler, Joshua Hartmann, Alexander Diegel

May 28, 2016

## 1 Fragen zur Vorlesung

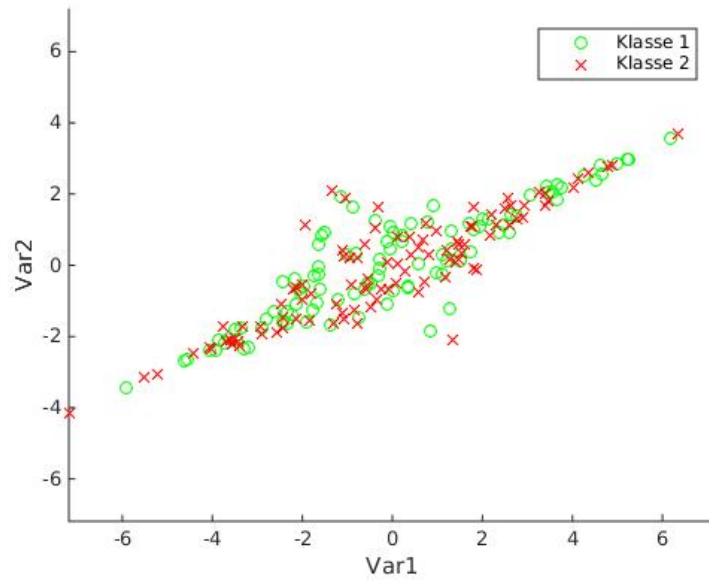
- a) Warum wendet man die Principal Component Analysis (PCA) an? (Zwei Gründe sollten genannt werden.)
- b) Was ist der Unterschied zwischen der PCA und der (Fisher) Linear Discriminant Analysis (LDA)?
- c) Welcher Abstand soll von der LDA maximiert werden?
- d) Warum genügt es nicht, den eben genannten Abstand zu maximieren?
- e) Was beschreibt die within-class scatter matrix (Foliensatz 5, Folie 17)?
- f) Was beschreibt die between-class scatter matrix (Foliensatz 5, Folie 18)?
- g) Man gebe sowohl für PCA als auch LDA an, ob es sich um ein supervised oder unsupervised Lernverfahren handelt.

**PCA:**

**LCA:**

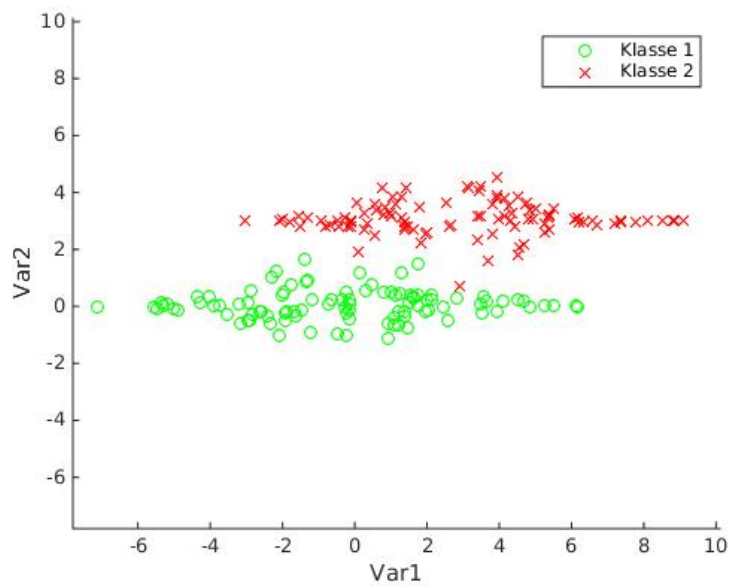
## 2 Principal Component Analysis – 1

(a)



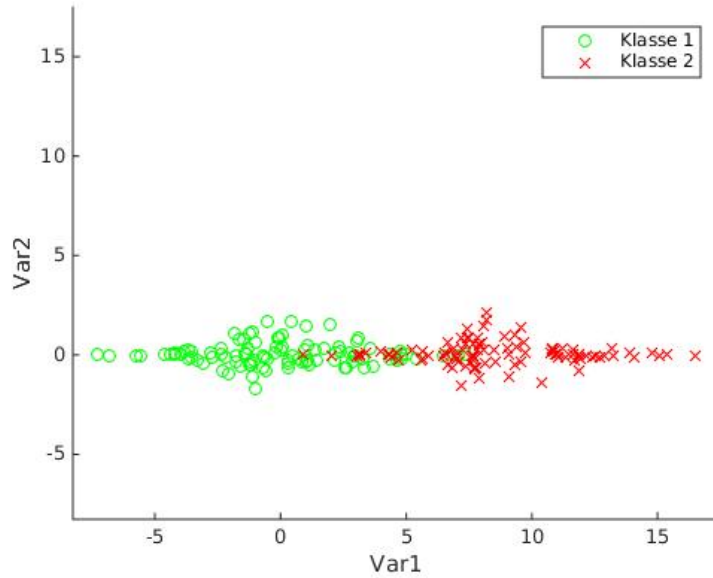
- a) Bietet sich im Allgemeinen eine Dimensionsreduktion auf den PC-Vektor an?
- b) Ist dies für die Klassifikation förderlich?

(b)



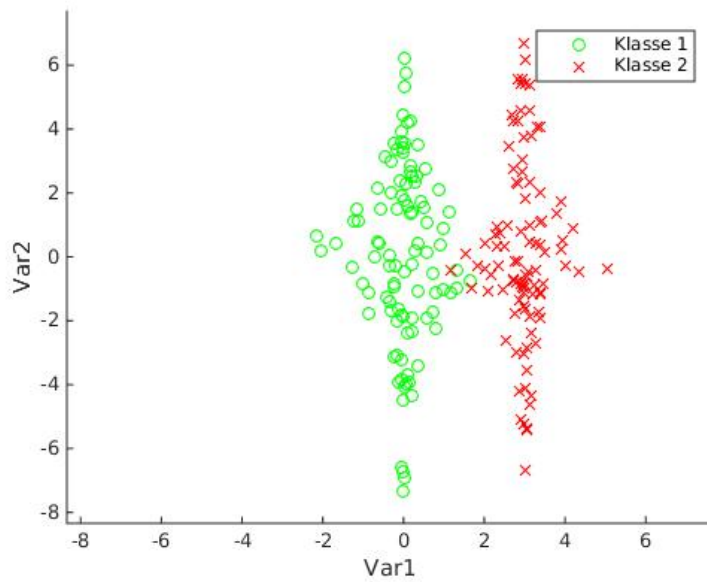
- a) Bietet sich im Allgemeinen eine Dimensionsreduktion auf den PC-Vektor an?
- b) Ist dies für die Klassifikation förderlich?

(c)



- a) Bietet sich im Allgemeinen eine Dimensionsreduktion auf den PC-Vektor an?
- b) Ist dies für die Klassifikation förderlich?

(d)



- a) Bietet sich im Allgemeinen eine Dimensionsreduktion auf den PC-Vektor an?
- b) Ist dies für die Klassifikation förderlich?

### **3 Principal Component Analysis – 2**

Welche PCs enthalten zusammen mindestens 95% der Streuung?

Plotten des Ergebnisses

Ergibt die PCA in diesem Zusammenhang Sinn? Warum oder warum nicht?

### **4 Linear Discriminant Analysis**

Plottet die originalen Datensätze und deren Projektionen und beurteilt bzw. begründet die Resultate.