



Det Naturvidenskabelige Fakultet



Introduktion til dimensionalitetsreduktion

Nikolaj Overgaard Sørensen
Datalogisk Institut



Dimensionalitätsreduktion

- A'hva for en fisk?
- Dimensioner svarer til målinger af forskellige ting (højde, vægt, temperatur)
- Mange dimensioner gør data uoverskuelig
- Hvordan repræsenteres data



Koordinatsystem

- Data = punkter i koordinatsystem
- Forskellige koordinatsystemer
- Hvordan vælger vi smart
- Find det koordinatsystem hvor data er langs akserne



PCA - Principal Component Analysis

- PCA finder de akser, der beskriver den største spredning i data
- Akserne er ordnet efter deres "vigtighed"
- Vi kan nu vælge de "vigtigste" akser og reducere antal dimensioner

