EE488 Special Topics in EE
<Deep Learning and AlphaGo>

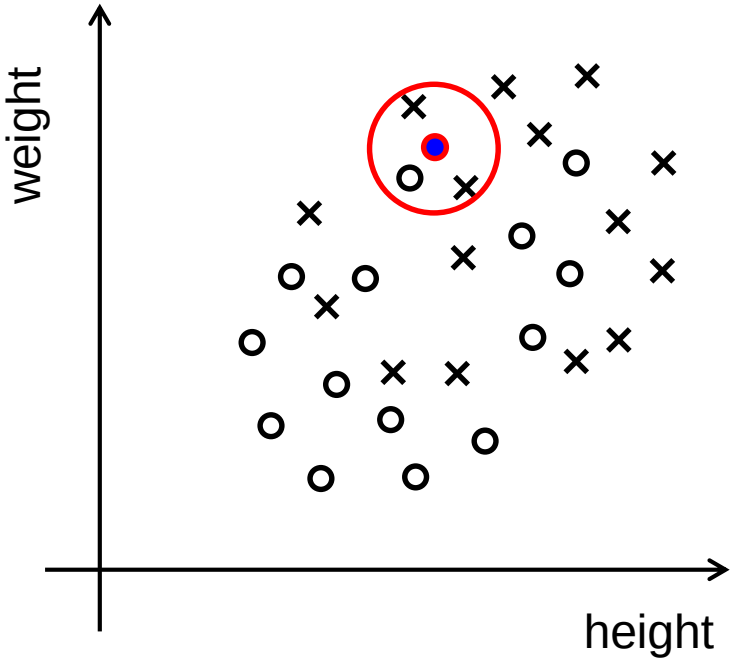
Sae-Young Chung

Lecture 5 Supplementary Material

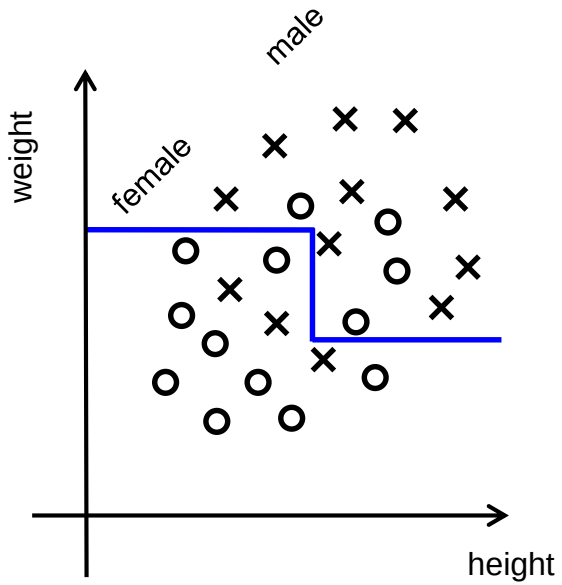
September 25, 2017

k-nearest Neighbors

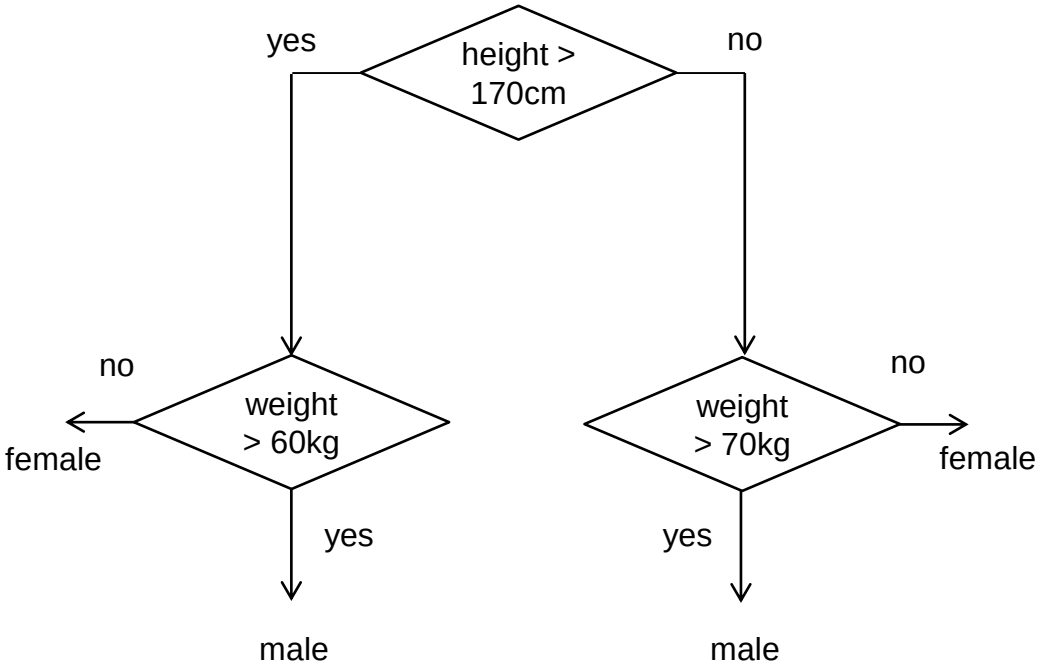
- For each point, look at k nearest neighbors and follow the majority rule



Decision Tree



Decision tree is learned from labeled data



Unsupervised Learning

- Supervised learning examples
  - Regression
  - Classification
  - k-nearest neighbors (non-parametric learning)
  - Decision tree (non-parametric learning)
  - ...
- Unsupervised learning examples
  - Principal component analysis (PCA)
  - k-means clustering
  - ...

Supervised Learning

dog

dog

dog

cat

cat

cat

KAIST

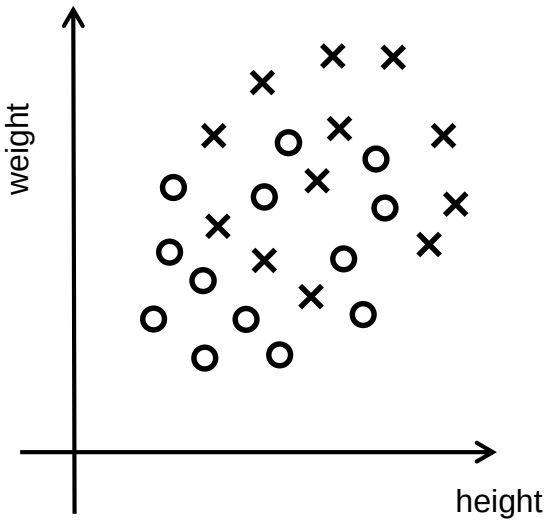
Unsupervised Learning

Semi-supervised Learning

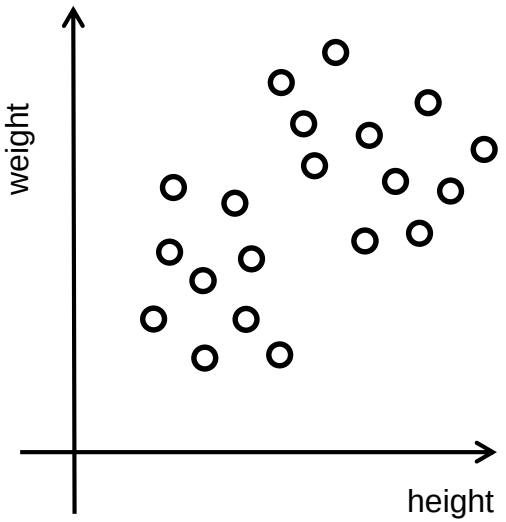
dog

cat

Supervised learning

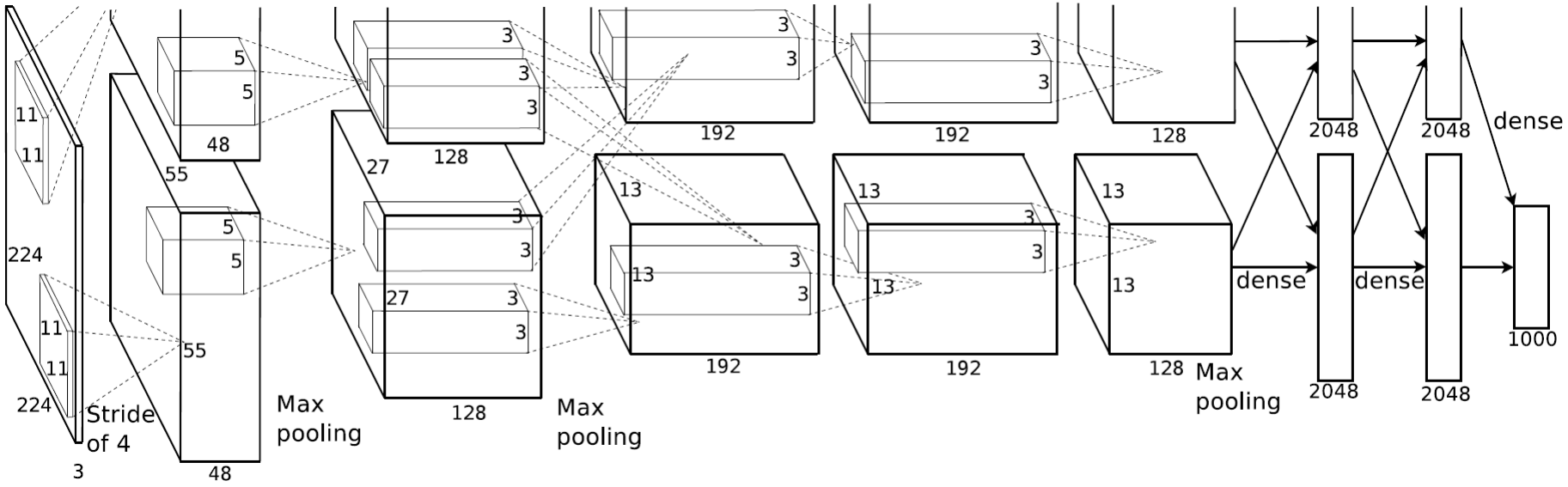


Unsupervised learning

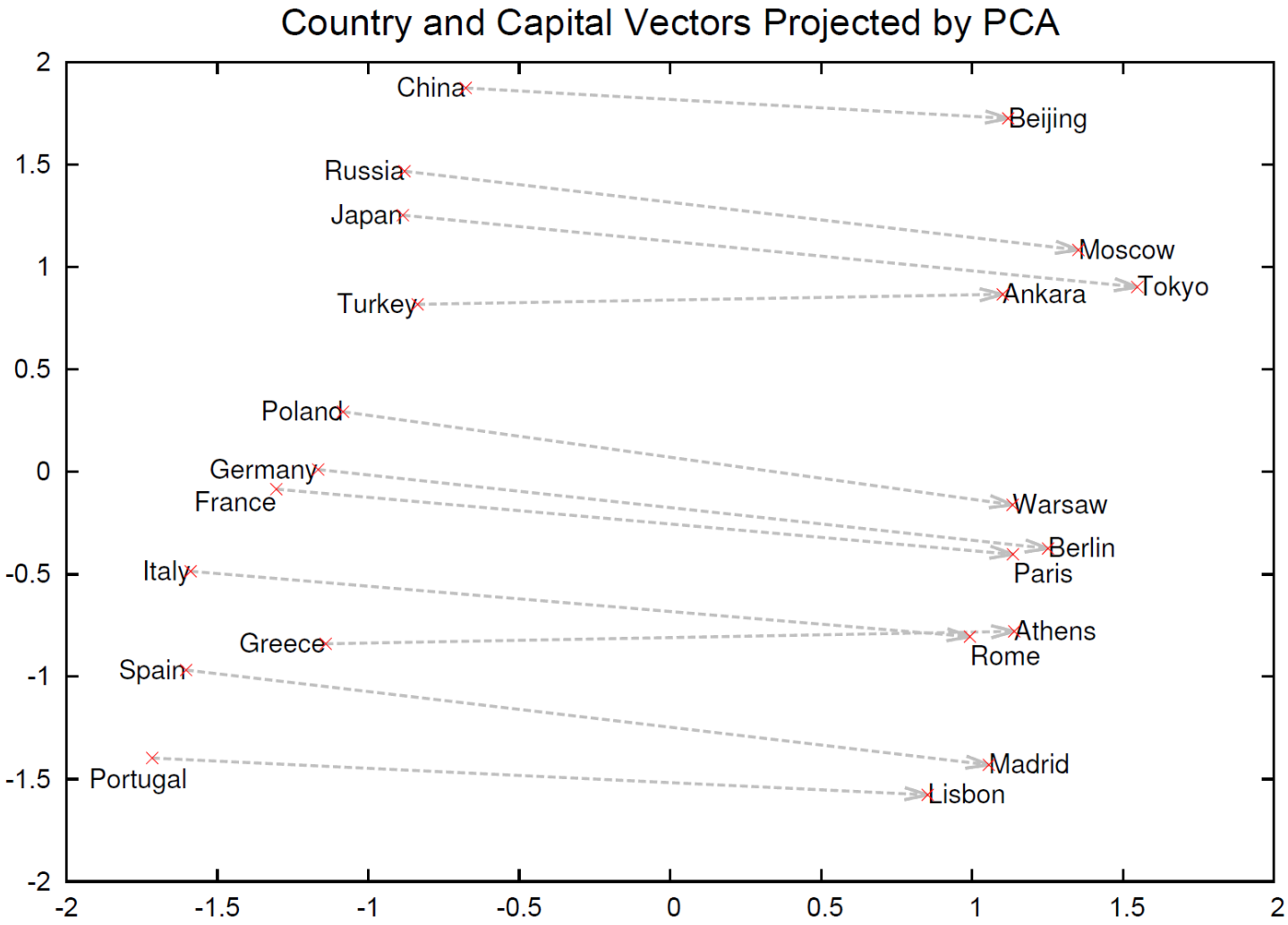


CNN for Image Classification

- Alex Krizhevsky, Ilya Sutskever, Geoffrey Hinton, “ImageNet classification with deep convolutional neural networks”, NIPS 2012



KAIST



Tomas Mikolov, et al., Distributed representations of words and phrases and their compositionality, 2013

Principal Component Analysis

- Data matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$
- $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{W}^T$: SVD
- $\mathbf{X}^T\mathbf{X} = (\mathbf{U}\mathbf{\Sigma}\mathbf{W}^T)^T(\mathbf{U}\mathbf{\Sigma}\mathbf{W}^T) = \mathbf{W}\mathbf{\Sigma}^2\mathbf{W}^T$
- Define $\mathbf{Z} = \mathbf{X}\mathbf{W}$, then

$$\mathbf{Z}^T\mathbf{Z} = \mathbf{W}^T\mathbf{X}^T\mathbf{X}\mathbf{W} = \mathbf{W}^T\mathbf{W}\mathbf{\Sigma}^2\mathbf{W}^T\mathbf{W} = \mathbf{\Sigma}^2$$

- If $\mathbf{X}$ has zero mean, then so does $\mathbf{Z}$. Then, the unbiased estimation of the covariance matrix of $\mathbf{z}$ from the samples $\mathbf{Z}$ is given by

$$\frac{1}{m-1}\mathbf{Z}^T\mathbf{Z} = \frac{1}{m-1}\mathbf{\Sigma}^2$$

Principal Component Analysis

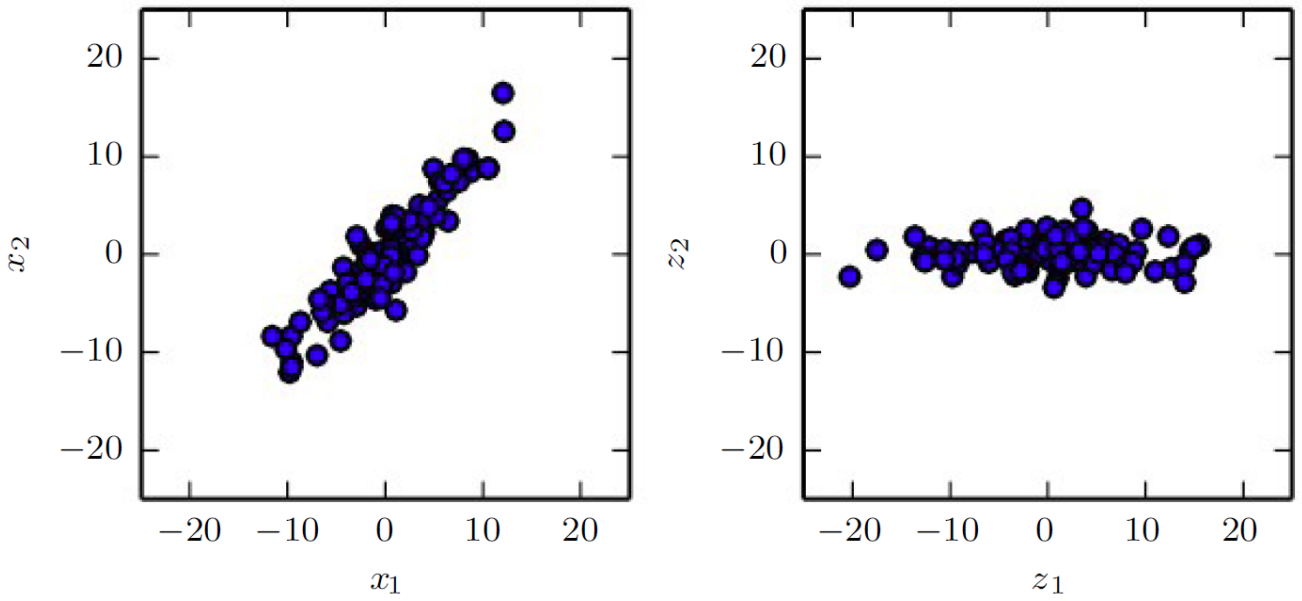


Fig. 5.8