Aalto University
School of Electrical
Engineering

CS-C3240 – Machine Learning D

Round 3: From features to classification

Stephan Sigg

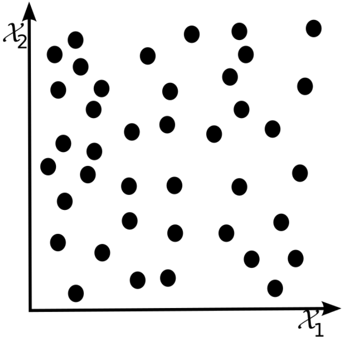
Department of Communications and Networking
Aalto University, School of Electrical Engineering
stephan.sigg@aalto.fi

Version 2.3, January 23, 2022

Semi-supervised learning

Unlabelled training data often easy to obtain

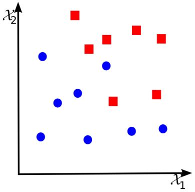
Caveat: labelling the data requires significant manual work



Semi-supervised learning

Increase amount of labelled data:

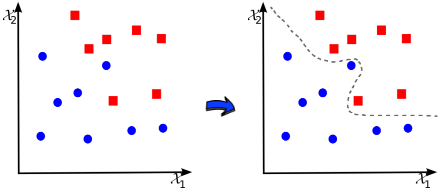
1. Start with labelled data



Semi-supervised learning

Increase amount of labelled data:

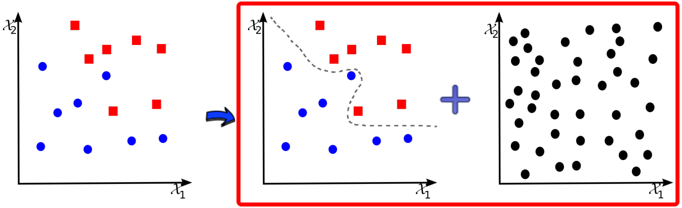
1. Start with labelled data
2. Train the classifier on the labelled data



Semi-supervised learning

Increase amount of labelled data:

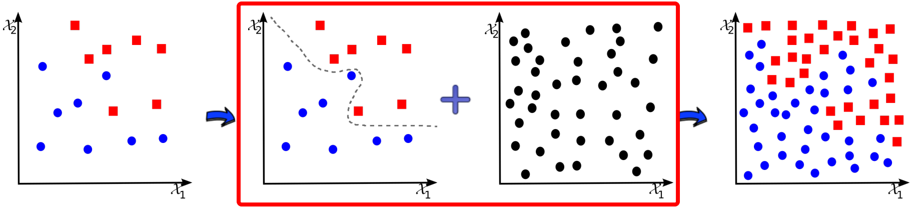
1. Start with labelled data
2. Train the classifier on the labelled data



Semi-supervised learning

Increase amount of labelled data:

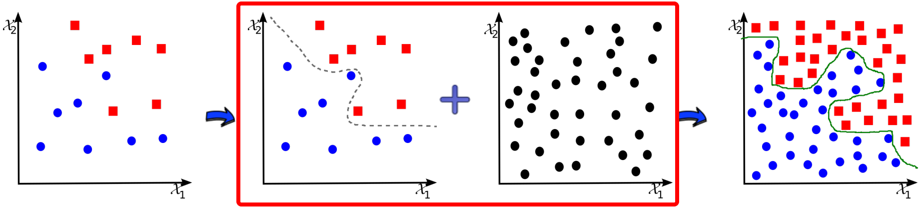
1. Start with labelled data
2. Train the classifier on the labelled data
3. Use the classifier to learn labels for the unlabelled data



Semi-supervised learning

Increase amount of labelled data:

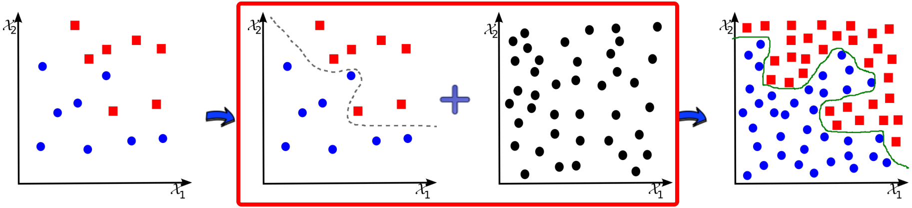
1. Start with labelled data
2. Train the classifier on the labelled data
3. Use the classifier to learn labels for the unlabelled data
4. Train a new classifier on this data



Semi-supervised learning

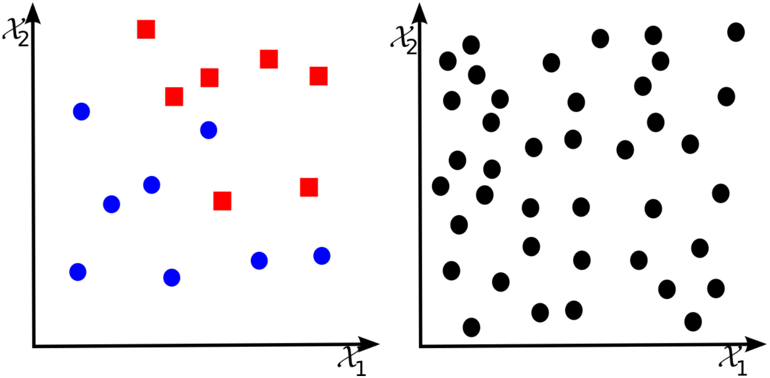
Remarks:

- No guaranteed success → Empirical validation required
- Introducing weights to samples can reduce dependency on learned labels



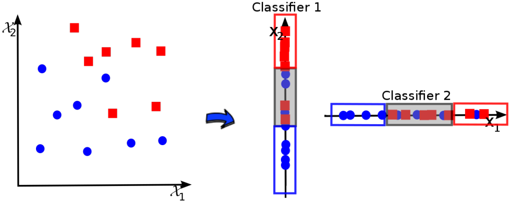
Co-training – Automated labelling

Provided independent feature sub-sets (perspectives), multiple classifiers trained to these sub-sets can iteratively label unlabelled data



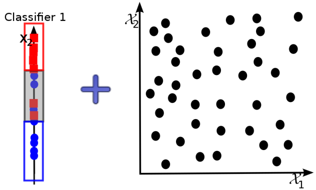
Co-training – Automated labelling

1. Train several classifiers wrt different feature-subsets



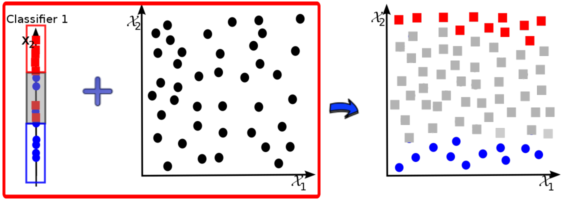
Co-training – Automated labelling

1. Train several classifiers wrt different feature-subsets
2. Apply one of these to the unlabelled data



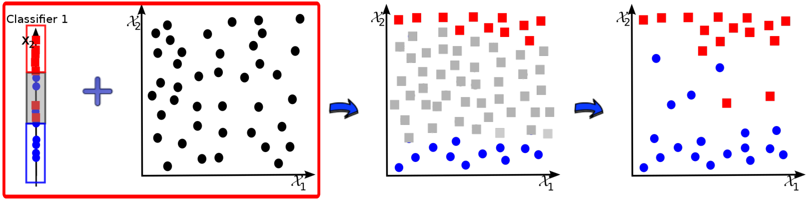
Co-training – Automated labelling

1. Train several classifiers wrt different feature-subsets
2. Apply one of these to the unlabelled data
3. Label those samples with highest probability



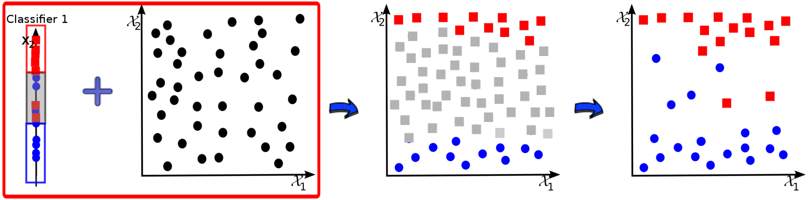
Co-training – Automated labelling

1. Train several classifiers wrt different feature-subsets
2. Apply one of these to the unlabelled data
3. Label those samples with highest probability
4. Combine with the original labelled data



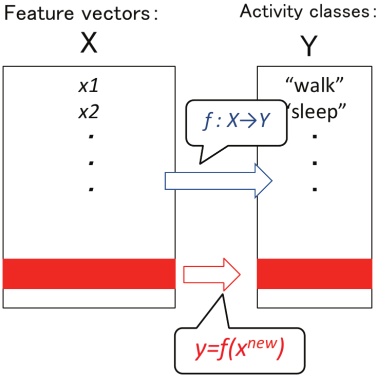
Co-training – Automated labelling

1. Train several classifiers wrt different feature-subsets
2. Apply one of these to the unlabelled data
3. Label those samples with highest probability
4. Combine with the original labelled data
5. Iterate over over all classifiers until convergence reached



Zero-shot learning

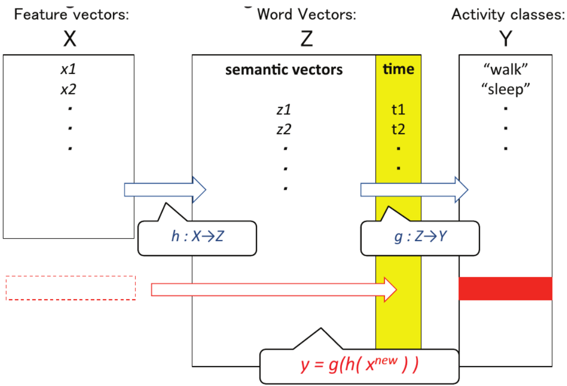
Exploit world-knowledge in order to recognize unknown classes[1]



[1] Picture: Matsuki, Inoue. "Recognizing unknown activities using semantic word vectors and twitter timestamps." 2016 ACM Ubicomp Adjunct.

Zero-shot learning

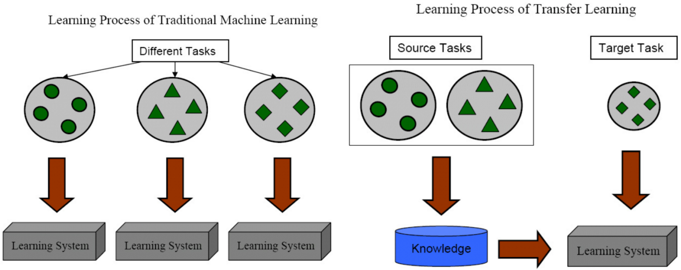
Exploit world-knowledge in order to recognize unknown classes[1]



[1] Picture: Matsuki, Inoue. "Recognizing unknown activities using semantic word vectors and twitter timestamps." 2016 ACM Ubicomp Adjunct.

Transfer learning

Apply a classifier to classes slightly differing from those trained[2]



(a) Traditional Machine Learning (b) Transfer Learning

[2] Picture from: Pan, Yang. "A survey on transfer learning." IEEE Transactions on knowledge and data engineering 22.10 (2010): 1345-1359.

Questions?

Stephan Sigg
stephan.sigg@aalto.fi

Si Zuo
si.zuo@aalto.fi

Literature

- C.M. Bishop: Pattern recognition and machine learning, Springer, 2007.
- R.O. Duda, P.E. Hart, D.G. Stork: Pattern Classification, Wiley, 2001.