

Notes Eivind.

What is classification?

- Labels on data entries, documents. Examples
 - Binary
 - Multiclass
- Feature representation of documents
- Train the model by supervised learning

Typical Methods:

SVM:

- Find a line/ high dimensional plane to separate the data.
- If The data is not linearly separable, use kernel with higher dimensions

Random Forrest / Decision Tree:

- Build a tree structure on the attributes/features, and have labels on the leaf nodes.
- Performance is dependent on topology(how you split)
- Solution: several trees with randomness injected

ANN:

- Hot topic due to amazing results on image recognition, and fast GPU
- Inspired by the brain
- Network by summation nodes and weights
- Propagate the input vector thru the network, and tune the weights so you get the correct output classification.