

Machine Learning exam – summer 2017

First task (Multiple Choice):

- What does w_{ij} in LLE portray? (Answer: Used for reconstruction of X_i by X_j)
- What distribution is used in t-SNE? (Answer: Heavy-tailed distr.)
- Some conditional probabilities to HMM (Answer: State probability only depends on previous state)
- How to find IC? (Answer: Maximize kurtosis of projections)

Second task:

- We have data resembling normal working condition of some factory. This data contains time, much sensor data and annotations by the staff (so a bunch of words for some steps in time)
- How to preprocess the data? (e.g. extract time of day, normalize sensor data, use bag-of-words kernel to create features regarding words)
- What ML algorithm to use? (e.g. SVDD and one-class SVM, you have to write formulas here!)

Third task:

- Given is a CCA problem like $\max v^T Cx$ with some constraints
- First task: Show that this is equivalent to eigenvector problem and display that in block matrices.
- Second task: Show that maximal eigenvalues maximize objective

Fourth task:

- We have a kernel that compares two documents. It does so by comparing for each word in a dictionary if it is present in both documents. We then normalize by the inverse square root of the length of both documents.
- First task: Show that kernel is positive-definite
- Second task: Find feature map
- Third task – Python: We have a weight vector w learned from the data. Write a function that takes this weight vector and a document and outputs the mapping. (Answer: The function has to first create the feature map and then take the scalar product of the weight vector and this feature map)

Fifth task:

- Build an HMM that generates text (Python)
- Build an HMM that takes text as input and generates the model parameters A, B (Answer: Baum-Welch would work, but that's the difficult solution. The easier one is building a model that always moves the same way (e.g. has a shifted identity matrix as A) and outputs only with probability one according to the text).