Abstract. This question list is a collection of topics that enumerate different targets for the exam questions. By no way they correspond to actual questions for any test, but they are representative candidates.

1 General Questions

Question 1. Introduce the notion of Machine Learning and discuss the following related topics:

- Training process
- Performance evaluation
- Error rate, precision and recall

Try to explain how and why the latter are connected with the first topic.

Question 2. Discuss the notion of learning algorithm and discuss at least two examples.

Question 3. Discuss the notion of feature in example-based learning. Why is it important? Make examples to show how feature engineering (e.g. feature selection or change of basis) can be useful.

Question 4. Define the notion of Text (or Document) Classification and provide some example of its application.

2 Early Models: Decision Trees

Question 1. Discuss the class of function that can be learned through decision tree learning.

Question 2. Define the core Decision Tree algorithm and explain its strength and weaknesses.
Question 3. Discuss the role of entropy in the Decision Tree algorithm. Explain its impact through an example.

3 Bayesian Models for Text Classification

Question 1. Describe the process of Text (or Document) Classification and explain why a Bayesian approach can be used for it.

Question 2. Discuss the assumptions of a Naive Bayes approach to Text Classification.

Question 3. Discuss and compare different Naive Bayes approaches to Text Classification.

Question 4. Discuss the role of parameter estimation (and for example the need for smoothing) in Naive Bayes text classification.

4 Bayesian Models: Hidden Markov Models

Question 1. Define the class of models known as Hidden Markov Models and explain their application to sequence labeling tasks.

Question 2. Explain the notion of Hidden Markov Model through an example of sequence labeling task, such as Part-of-Speech (POS) tagging.

Question 3. Discuss the simplifying (Markv) assumptions of Hidden Markov Models in their application to Part-of-Speech (POS) tagging tasks.

Question 4. Discuss the role of parameter estimation (and for example the need for smoothing) in Hidden Markov Models.