Machine Learning Syllabus

- **Prerequisites: (3h)**
  - **Linear Algebra Review**
    - Vectors, Matrices and Tuples
    - Addition, Multiplication, Inverse, Transpose…
  - **Multivariate Calculus**
    - Functions
    - Differential Calculus / Partial Derivatives

- **Introduction to Machine Learning (1j)**
  - What is Machine Learning
  - Applications
  - Supervised vs Unsupervised
  - Introducing ML with Univariate Linear Regression
    - Model Representation
    - Cost Function
    - Optimization
      - Ordinary Least Squares
        - Non-invertibility
      - Gradient Descent

- **Practical Aspects of Machine Learning (1j)**
  - Data Splitting
    - Validation
    - Cross-Validation
  - Bias / Variance diagnosis
  - Bias / Variance correction
  - Learning Curves
    - Learning rate
    - Learning rate decay
  - Regularization
  - Cost Functions & Optimization Techniques
    - Normalization
  - Regression and Classification Evaluation Metrics
    - RMSE, RSquared…
    - Confusion Matrix, Accuracy, Precision, Recall…
    - ROC, AUC
    - Error metrics for skewed classes
  - Hyperparameters tuning
  - Feature Representation & Engineering
    - Numerical Features
    - Categorical Features
    - Images
    - Text
      - BOW
      - TF-IDF
      - N-grams

- **Regression (1j)**
  - Multiple Linear Regression
    - Multiple Features
    - Features and Polynomial regression

        - Gradient descent for multiple variables
    - Lasso Regression (L1 regularization)
    - Ridge Regression (L2 regularization)
    - Decision Trees Regression

- **Classification (1j)**
    - Logistic Regression
        - Decision Boundary
        - Cost Function
    - Multiclass Classification
        - One-vs-all
    - Support Vector Machines
    - Decision Trees
    - Ensemble Methods (Bagging, Boosting, Stacking). Examples:
        - Bagged Trees
            - Random Forest
        - Boosted Trees
            - Gradient Boosting
    - Classification Metrics

- **Unsupervised Learning (1j)**
    - Clustering
        - Partition methods
            - k-means clustering
            - Cluster validity, choosing the number of clusters
        - Distribution models
            - Gaussian Mixture Models & EM
        - Hierarchical methods
        - Density Models
            - DBSCAN
    - Dimensionality Reduction
        - Principal Component Analysis

- **Introduction to deep learning (1j)**
    - Why is deep learning becoming more popular?
    - Perceptron: the artificial neuron
    - Conveying ideas using Binary Classification
        - Model Representation & Notations
        - Cost function
        - Gradient Descent
    - Forward and backward propagation
    - Mini-batch Gradient Descent
    - Activation Functions
    - Optimization setting
        - Weights Initialization
        - Normalization
        - Vanishing / exploding gradients
        - Gradient. checking
    - Regularization
        - Dropout regularization
    - Parameters vs. Hyperparameters
    - Bias/Variance correction strategies
    - Transfer Learning

- **Structuring machine learning projects (1j)**

- Setting goals
  - Single Optimizing metric
  - Optimizing and satisfying metrics
- The ML workflow
- Comparing to human performance
- Manual error analysis
- Mismatched data distributions
- TBD…