

# Machine Learning

4771

Instructors:

Adrian Weller and Ilia Vovsha

# Midterm Review

- Format
- Topics
- Remarks

# Format

- In class, 75 minutes, 5 questions weighted about evenly
- Closed book/notes, but you are allowed to have one sheet of A4 size paper (both sides) with your own notes
- Both conceptual questions and derivations involving math
- Sample midterm available on Thursday.
- Recommended practice before the exam: sample midterm and HW problems

# Topics

COMS4771, Columbia University

1. **Basic statistics:** distributions, properties, Bayesian inference and decision theory, priors, posterior, Maximum Likelihood, MAP, coin tossing
2. **Parametric Approaches:**
  - a. normal distribution, multivariate form, classification task
  - b. function approximation, linear regression, basis functions, overfitting, cross-validation
3. **Perceptron & Neural Networks:** neuron, geometry of linear surface, linear separability, gradient descent, perceptron, online perceptron, convergence proof, neural net definition & graphical representation, expressive power, Lagrange multipliers, back-propagation, activation functions, algorithm deficiencies
4. **Statistical Learning Theory:** general learning model, ERM, convergence modes, consistency, uniform convergence (1/2-sided), indicator functions, Chernoff bounds, entropy conditions, non-falsifiability, relative U.C, bound on achieved risk, structure of capacity concepts, growth function, VC dimension, shattering, large/small sample theory, SRM
5. **Optimal Hyperplane & SVMs:** gap tolerant classifiers, large margin, optimal hyperplane, primal-dual forms & derivations, KKT conditions, support vectors, non-separable data, soft-margin SOCP & QP forms, SVMs, kernels, Mercer's condition

# Remarks

- At most 2 questions exclusively on a particular topic
- Conceptual questions may require you to relate several topics
- If a topic/concept appeared on the HWs, it is definitely more likely to appear on the exam
- Lecture slides are the starting point for possible questions, but you would benefit if you followed class discussion and covered reading material (listed on 'news', see course website)
- If a theorem/result was given without proof, you won't be asked about the proof. However, if the intuition behind the result was discussed (and mentioned in the slides), you might be asked to explain this intuition.