# CS 181 2018 Midterm 1 Topic List

**The best way to prepare for the midterm is to review homeworks, section notes, the 2017 lecture slides, and the midterm practice questions**. The midterm will be conceptual and analytical, testing ideas and understanding, and does not involve writing pseudocode. You are not expected to memorize formulas such as PDFs, or to memorize complicated matrix cookbook rules, but you should be familiar with methods of probability theory (e.g. Bayes Rule) and the various models we've studied so far in the course. The exam may include material from the max-margin lecture but excludes the SVM lecture. Here is a brief list of topics that you could expect to be asked about. This list emphasizes the main focus areas and is not fully inclusive:

- Linear regression: least squares loss, how to differentiate least squares and solve for weights analytically, be able to work with and interpret alternate (simple) loss functions when given, understand parametric vs. non-parametric regression
- Basis functions [general idea, not specific versions]
- Generative model of linear regression, noise, maximum likelihood estimation
- Bayesian methods: terminology, MAP, posterior predictive, use of conjugate distributions (Beta-Bernoulli, Normal-Normal), interpretation of estimators, Bayesian linear regression [Bayesian model selection is out of scope]
- Binary linear classification. Perceptron (or "hinge") loss. Gradient descent for perceptron loss and the perceptron algorithm [You do not have to memorize the gradient itself]
- Decision boundaries, linear and non-linear separators.
- Generative classification via "class prior" probability and class-conditional distributions (e.g. Gaussian or multinomial Naive Bayes), use of Bayes Rule. Use of MLE estimates for parameters of these distributions [Bayesian Naive Bayes is out of scope]
- Training Logistic Regression with gradient descent, understand why the softmax is useful, understand (but don't memorize) its derivative.
- Understanding of Naive Bayes and Logistic regression as a pair of models for the same task, where NB is generative and LR is discriminative [LR does not model p(x,y)]
- Use of validation for model selection and to avoid over-fitting
- Definition of true/false positives, true/false negatives and AUC.
- Bias-variance trade-off (not full derivation, but understand what role each term has)
- The role and mathematical form of major types of regularization. Particularly when used with linear regression problems
- Neural nets: basic notation for weights in layers and use of sigmoid and ReLU activation functions [you don't need to memorize the functions but you need to understand that they are applied element-wise to vectors to create non-linearities]. Use of neural nets for both classification and regression. Finding gradients on weights using the chain rule [we wouldn't expect you to remember the formula for the sigmoid or its derivative]. High level idea of back-propagation.