

# CSCI-UA 9472. Machine Learning

## Material for the Midterm

Augustin Cosse

October 2020

### 1 Material covered

1. You must know and understand the [residual sum of squares criterion](#) and derive the [gradient iterations](#) on that function as well as the [closed form solution](#) for the vector of weights (obtained by setting the derivatives to zero).
2. You must be able to list and explain the three main [regularization approaches](#) (Ridge, Lasso, Best Subset Selection)
3. You must be able to explain the [statistical assumptions](#) leading to the residual sum of squares model, the Ridge and the Lasso. You must be able to describe the distributions involved in each regularization approach (Gaussian and Laplace).
4. You must be able to [compare the regularization approaches](#) in terms of their respective efficiency and complexity
5. You must be able to explain how to use the Residual Sum of Squares criterion to [learn a binary classifier](#)
6. You must be able to explain how the [binary classifier can be extended into a multiclass classifier](#) (i.e. one vs rest, one vs one, and one hot encoding)
7. You must be able to explain the [distinction between generative and discriminative classifiers](#) and give an example from each family.
8. You must be able to explain and derive the expression for the [logistic regression classifier](#)
9. You must be able to explain and derive the expression for the [Linear/Gaussian Discriminant Analysis Classifier](#)
10. You must be able to discuss the use of Kernels and motivate this use through gradient descent on large feature vectors.
11. You must be able to define the notion of [Mercer Kernel](#) (i.e. list the two properties that a matrix has to satisfy to be derived from such a Kernel)
12. You must be able to discuss the [kernel trick](#) and apply it to the Residual sum of squares criterion to derive a formulation that only depends on the similarities and not on feature vectors.
13. You must be able to explain the notion of [Maximal Margin Classifier](#). You must be able to derive the [optimization problem](#) that one has to solve to learn this classifier from the distance of a point to a plane.

14. You must be able to give the [final expression of the Maximal Margin Classifier/SVM](#). In particular, you must be able to use this expression to [illustrate the notion of support vectors](#).
15. You must be able to explain the [perceptron model](#) as well as the [perceptron learning rule](#) and the associated [convergence theorem](#)
16. You must be able to understand and explain [how the perceptron can be extended into a neural network](#) to learn non linearly separable datasets
17. You must be able to provide the [general expression for a neural network](#) and draw the [corresponding diagram](#)
18. You must be able to explain the [backpropagation algorithm](#) (list the main steps) and provide the associated equations.