
Contents

1. What is Data (Science)?	1
2. Affine Linear, Polynomial and Logistic Regression	7
3. k -Nearest Neighbors	35
4. Clustering	51
5. Graph Clustering	61
6. Best-Fit Subspaces	81
7. Singular Value Decomposition	89
8. Curse and Blessing of High Dimensionality	115
9. Concentration of Measure	127
10. Gaussian Random Vectors in High Dimensions	139
11. Dimensionality Reduction à la Johnson-Lindenstrauss	151
12. Separation and Fitting of High-Dimensional Gaussians	159
13. Perceptron	179
14. Support Vector Machines	189
15. Kernel Method	209
16. Neural Networks	225
17. Gradient Descent for Convex Functions	261
A. Selected Results of Probability Theory	281
Bibliography	289