TABLE OF CONTENTS

1. Introduction 1
2. Basic Statistics Review 2
2.1. Basic Statistical Parameters 2
2.2. Distributions 2
2.3. P-value 2
2.4. T-test 2
3. Data Science Project Steps 3
3.1. Project Management 3
3.2. Data Collection 3
3.3. Data Visualization and Discovery 3
3.4. Data Processing 3
3.4.1. Handling Outliers and Missing Values 3
3.5. Model Selection and Evaluation 3
3.5.1. Cross Validation 3
3.5.2. ROC and AUC 3
4. Regression Analysis 4
4.1. Gradient Descent 4
4.1.1. Batch Gradient Descent 4
4.1.2. Stochastic Gradient Descent 7
4.2. Linear Regression 9
4.2.1. Least-Squares Estimation 10
4.2.2. Least Mean Squares Estimation 11
4.3. Logistic Regression 12
4.3.1. Maximum Likelihood Estimation 12
4.3.2. Cost Function and Update Rule 12
5. Tree-based Models 15
6. Support Vector Machines 17
7. Deep Learning 19
7.1. Perceptron 19

7.2. Multilayer Perceptron Model 21
7.2.1. Activation Function 23
7.2.2. Backpropagation Algorithm 24
7.2.3. Derivation 26
7.2.4. Convergence 29
7.3. Regularization 32
7.3.1. Ridge and Lasso Regression 32
7.4. Convolutional Neural Networks 32
8. Dimensionality Reduction 33
8.1. PCA 33
9. Clustering 34
9.1. K-Means 34
REFERENCES 41