

# Elements of Machine Learning and Data Science

Part II: Machine Learning — Exam Notes (Living Document)

Emir Pisirici

January 29, 2026

Exam likelihood: High (overall Machine Learning part)

This document is structured to match the lecture topics exactly and is designed for adding **exam-style notes, common traps, and visual summaries**.

## Contents

<b>1</b>	<b>Introduction to ML</b>	<b>2</b>
1.1	Motivation . . . . .	2
1.2	Forms of Learning . . . . .	2
1.3	Terms, Concepts, and Notation . . . . .	2
1.4	Bayes Decision Theory . . . . .	2
<b>2</b>	<b>Probability Density Estimation</b>	<b>3</b>
2.1	Probability Distributions . . . . .	3
2.2	Parametric Methods . . . . .	3
2.3	Nonparametric Methods . . . . .	3
2.4	Mixture Models . . . . .	3
2.5	Bayes Classifier . . . . .	3
2.6	K-NN Classifier . . . . .	3
<b>3</b>	<b>Linear Discriminants</b>	<b>4</b>
3.1	Motivation: Discriminant Functions . . . . .	4
3.2	Linear Discriminant Functions . . . . .	4
3.3	Least-Squares Classification . . . . .	4
3.4	Generalized Linear Discriminants . . . . .	4
3.5	Basis Functions . . . . .	4
<b>4</b>	<b>Linear Regression</b>	<b>5</b>
4.1	Motivation . . . . .	5
4.2	Least-Squares Regression . . . . .	5
4.3	Regularization . . . . .	5
4.4	Ridge Regression . . . . .	5
4.5	The Bias-Variance Tradeoff . . . . .	5
<b>5</b>	<b>Logistic Regression</b>	<b>6</b>
5.1	Logistic Regression Formulation . . . . .	6
5.2	Motivation and Background . . . . .	6
5.3	Iterative Optimization . . . . .	6
5.4	First-Order Gradient Descent . . . . .	6
5.5	Second-Order Gradient Descent . . . . .	6
5.6	Error Function Analysis . . . . .	6

<b>6</b>	<b>Support Vector Machines</b>	<b>7</b>
6.1	Maximum Margin Classification . . . . .	7
6.2	Primal Formulation . . . . .	7
6.3	Dual Formulation . . . . .	7
6.4	Soft-Margin SVMs . . . . .	7
6.5	Non-linear SVMs . . . . .	7
6.6	Error Function Analysis . . . . .	7
<b>7</b>	<b>Neural Network Basics</b>	<b>8</b>
7.1	Perceptrons . . . . .	8
7.2	Multi-Layer Perceptrons . . . . .	8
7.3	Loss Functions . . . . .	8
7.4	Backpropagation . . . . .	8
7.5	Computational Graphs . . . . .	8
7.6	Stochastic Gradient Descent . . . . .	8

## 1 Introduction to ML

- 1.1 Motivation
- 1.2 Forms of Learning
- 1.3 Terms, Concepts, and Notation
- 1.4 Bayes Decision Theory

## **2 Probability Density Estimation**

- 2.1 Probability Distributions**
- 2.2 Parametric Methods**
- 2.3 Nonparametric Methods**
- 2.4 Mixture Models**
- 2.5 Bayes Classifier**
- 2.6 K-NN Classifier**

### **3 Linear Discriminants**

- 3.1 Motivation: Discriminant Functions**
- 3.2 Linear Discriminant Functions**
- 3.3 Least-Squares Classification**
- 3.4 Generalized Linear Discriminants**
- 3.5 Basis Functions**

## **4 Linear Regression**

**4.1 Motivation**

**4.2 Least-Squares Regression**

**4.3 Regularization**

**4.4 Ridge Regression**

**4.5 The Bias-Variance Tradeoff**

## **5 Logistic Regression**

- 5.1 Logistic Regression Formulation**
- 5.2 Motivation and Background**
- 5.3 Iterative Optimization**
- 5.4 First-Order Gradient Descent**
- 5.5 Second-Order Gradient Descent**
- 5.6 Error Function Analysis**

## **6 Support Vector Machines**

- 6.1 Maximum Margin Classification**
- 6.2 Primal Formulation**
- 6.3 Dual Formulation**
- 6.4 Soft-Margin SVMs**
- 6.5 Non-linear SVMs**
- 6.6 Error Function Analysis**

## 7 Neural Network Basics

- 7.1 Perceptrons
- 7.2 Multi-Layer Perceptrons
- 7.3 Loss Functions
- 7.4 Backpropagation
- 7.5 Computational Graphs
- 7.6 Stochastic Gradient Descent