

Comparative Analysis of Gradient Descent Methods: Stochastic, Mini-batch, and Batch

Hamza Rashid¹

¹Mathematics and Statistics, McGill University, Montreal, Canada

April 7, 2024

Abstract

The abstract should briefly summarize the aim of the research, the methods used, the key findings, and the main conclusions.

1 Introduction

In the field of machine learning, Gradient descent is the standard method to minimize the cost function and optimize model parameters. Among the variations of gradient descent, Stochastic Gradient Descent (SGD), Mini-batch Gradient Descent, and Batch Gradient Descent each have distinct characteristics that affect their performance in different computational and data scenarios. In addition, there are various optimization techniques built on top of Gradient descent that improve convergence speed, reduce oscillation, and adapting learning rates. This research aims to compare these methods systematically, focusing on their convergence rates, computational efficiency, and applicability to various machine learning models and data sizes. Through this analysis, we seek to provide insights and guidelines for selecting the most appropriate gradient descent technique in practice.

1.1 Background Information

Gradient Descent is a first-order iterative method for finding a minimum value of a function.

1.2 Problem Statement

Discuss the variations of gradient descent methods and their relevance.

1.3 Research Objectives

Outline the main objective: to compare stochastic, mini-batch, and batch gradient descent methods in various aspects.

2 Literature Review

3 Methodology

4 Results

5 Discussion

6 Conclusion

References