

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ
БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Механико-математический факультет
Кафедра теории функций

ДУБЯГА
ВЛАДИСЛАВ МИХАЙЛОВИЧ

Аннотация к дипломной работе

Методы оптимизации в машинном обучении

Научный руководитель:
кандидат физико-математических наук
доцент Бондарев Сергей Александрович

Минск
2023

Реферат

Дубяга В.М. Методы оптимизации в машинном обучении (дипломная работа). — Минск: БГУ, 2023. — 51 с.

Дипломная работа содержит: 51 страниц, 32 использованных источников.

Ключевые слова: МАШИННОЕ ОБУЧЕНИЕ, МЕТОДЫ ОПТИМИЗАЦИИ, ГРАДИЕНТНЫЙ СПУСК, СТОХАСТИЧЕСКИЙ ГРАДИЕНТНЫЙ СПУСК, УСКОРЕННЫЕ МЕТОДЫ, МЕТОД НЬЮТОНА, КВАЗИНЬЮТОНОВСКИЕ МЕТОДЫ, НЕЙРОННЫЕ СЕТИ, ЛИНЕЙНАЯ РЕГРЕССИЯ, ЛОГИСТИЧЕСКАЯ РЕГРЕССИЯ.

В дипломной работе изучаются методы оптимизации, активно используемые в машинном обучении. Проведен анализ и рассмотрены теоретические оценки сходимости градиентного спуска с постоянным шагом для выпуклых функций и сильно выпуклых функций с липшицевым градиентом. Рассмотрена линейка методов градиентного спуска:

1. градиентный спуск с постоянным шагом
2. стохастический градиентный спуск,
3. метод тяжелого шарика,
4. метод Нестерова.

Также рассмотрены методы второго порядка и квазиньютоновские методы. Сделан вывод о том, какие методы лучше использовать для задач линейной и логистической регрессии.

Завершается дипломная работа рассмотрением нейронных сетей. Дается краткий обзор методов оптимизации, используемых для обучения нейронных сетей — RMSprop, Adam, AdaGrad, AMSGrad.

Abstract

Dubjaga V.M. Optimization Methods in Machine Learning (Thesis).
— Minsk: BSU, 2023. — 51 p.

The thesis contains: 51 pages, 32 sources used.

Keywords: MACHINE LEARNING, OPTIMIZATION METHODS, GRADIENT DESCENT, STOCHASTIC GRADIENT DESCENT, ACCELERATED METHODS, NEWTON'S METHOD, QUASI-NEWTON METHODS, NEURAL NETWORKS, LINEAR REGRESSION, LOGISTIC REGRESSION.

The thesis explores optimization methods actively used in machine learning. An analysis is conducted and theoretical convergence assessments of gradient descent with a constant step for convex functions and strongly convex functions with Lipschitz gradient are considered. A range of gradient descent methods is examined:

1. gradient descent with a constant step,
2. stochastic gradient descent,
3. heavy ball method,
4. Nesterov's method.

Second-order methods and quasi-Newton methods are also reviewed. A conclusion is drawn about which methods are best used for linear and logistic regression tasks.

The thesis concludes with a discussion of neural networks. A brief overview of the optimization methods used for training neural networks — RMSprop, Adam, AdaGrad, AMSGrad — is provided.

Рэферат

Дубяга В.М. Метады аптымізацыі ў машынным навучанні (дыпломная праца). — Мінск: БДУ, 2023. — 51 ст.

Дыпломная праца ўключае ў сябе: 51 старонак, 32 выкарыстаных крыніц.

Ключавыя словы: МАШЫННАЕ НАВУЧАННЕ, МЕТАДЫ АПТЫМІЗАЦЫІ, ГРАДЫЕНТНЫ СПУСК, СТОХАСТЫЧНЫ ГРАДЫЕНТНЫ СПУСК, УСКОРАНЫЯ МЕТАДЫ, МЕТАД НЬЮТОНА, КВАЗІНЬЮТОНАЎСКІЯ МЕТАДЫ, НЕЙРАННЫЯ СЕТКІ, ЛІНЕЙНАЯ РЭГРЕСІЯ, ЛАГІСТЫЧНАЯ РЭГРЕСІЯ.

У дыпломнай працы вывучаюцца метады аптымізацыі, якія актыўна выкарыстоўваюцца ў машынным навучанні. Праведзены аналіз і разгледжаны тэарэтычныя ацэнкі збегнасці градыентнага спуску з пастаянным крокам для выпуклых функцый і моцна выпуклых функцый з ліпшыцавым градыентам. Разгледжана шэраг метадаў градыентнага спуску:

1. градыентны спуск з пастаянным крокам
2. стохастычны градыентны спуск,
3. метада цяжкага шарыка,
4. метада Несцерава.

Таксама разгледжаны метады другога парадка і квазіньютонаўскія метады. Зроблены вынік аб тым, якія метады лепш выкарыстоўваць для задач лінейнай і лагістычнай рэгресіі.

Дыпломная праца заканчваецца разглядам нейранных сетак. Даецца кароткі абзор метадаў аптымізацыі, якія выкарыстоўваюцца для навучання нейранных сетак — RMSprop, Adam, AdaGrad, AMSGrad.