

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Факультет прикладной математики и информатики

Кафедра информационных систем управления

Аннотация к магистерской диссертации

«Обучение глубоких нейронных сетей на основе метода отжига»

Мацкевич Вадим Владимирович

Научный руководитель – доктор технических наук,
профессор Краснопрошин В.В.

Минск, 2020

Реферат

Магистерская диссертация, 50 страниц, 7 рисунков, 3 таблицы, 28 источников, 26 формул.

НЕЙРОННАЯ СЕТЬ, ОГРАНИЧЕННАЯ МАШИНА БОЛЬЦМАНА, ОБУЧЕНИЕ, ВЫБОРКА, ПАРАЛЛЕЛЬНЫЕ ВЫЧИСЛЕНИЯ, МЕТОД ОТЖИГА, ФРЕЙМВОРК.

Объект исследования – система обработки данных на основе нейронных сетей.

Цель работы– изучить проблему эффективности процесса обучения нейронных сетей, разработать архитектуру глубокой доверительной сети и алгоритм ее обучения. На основе теоретических исследований разработать программную систему (в виде фреймворка), ориентированную на использование технологии распараллеливания данных. Проверить эффективность разработанной архитектуры нейронной сети и алгоритма обучения (реализующего метод отжига) на примере задачи сжатия цветных изображений, сравнить эффективность метода отжига с градиентными методами.

Методы исследования– изучение литературы по теме работы, методы обучения нейронных сетей, методы теории параллельных вычислений, вычислительный эксперимент.

Результат – оригинальная архитектура глубокой доверительной сети ориентированная на параллельную обработку данных. Оригинальный алгоритм, реализующий метод отжига, для обучения глубоких доверительных сетей. Дополнительно построена и реализована архитектура оригинальной системы (в виде фреймворка), эффективность которой экспериментально подтверждена на примере решения задачи сжатия цветных изображений.

Область применения – искусственный интеллект.

Abstract

Master's thesis, 50 pages, 7 pictures, 3 tables, 28 references, 26 formulas.

NEURAL NETWORK, RESTRICTED BOLTZMANN MACHINE, TRAINING, DATASET, PARALLEL COMPUTATIONS, ANNEALING METHOD, FRAMEWORK.

Object of research – data processing system based on neural networks.

Research goal– study the problem of efficiency of neural networks training process, develop a deep belief network architecture and its training algorithm. Based on theoretical research, to develop a software system (in the form of framework), focused on the use of data parallelization technology. Check the effectiveness of the developed neural network architecture and the training algorithm (implementing the annealing method) using the example of the color images compression problem, compare the annealing method efficiency with gradient methods.

Research methods – study of literature on the topic of work, methods of training neural networks, methods of the theory of parallel computing, computational experiment.

Result – an original architecture of a deep belief network focused on parallel data processing, an original algorithm implementing the annealing method for training deep belief networks. Additionally, the architecture of the original system (in the form of a framework) was built and implemented, the effectiveness of which was experimentally confirmed by solving the problem of compressing color images as an example.

Application field– artificial intelligence.