Titel der Seminarausarbeitung

Proseminar Data Mining

Vorname Name
Fakultät für Informatik
Technische Universität München
Email: name@in.tum.de

Abstract— In this paper we introduce the reader to neural networks–a beautiful, biology-inspired machine learning paradigm.

Index Terms— Schlüsselworte

I. Einleitung

blabla

II. Kapitel

blabla

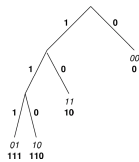### A. Unterkapitel

blabla mit drei Quellenangaben [1]–[3]

Fig. 1. Baum

TABLE I
BEISPIELTABELLE

Spalte1	Spalte2
0	1

III. Gradient Descent

A. Introduction

Gradient Descent is an algorithm used to iteratively minimize functions $f : \mathbb{R}^n \to \mathbb{R}$ of multiple values.

B. Directional derivatives

Since f is a function of multiple values, it does not suffice to.

From the definition of the directional derivative it follows that it evaluates to $\nabla f \cdot u$. A rigorous proof can be found in [1], but as an intuition, the change of $f(x)$ in direction u can be thought of as u_1 times the change in x_1 plus u_2 times the change in x_2 plus ... which results in $\sum_{i=0}^{n} \frac{\partial f}{\partial x_i} u_i = \nabla f \cdot u$.

Following Goodfellow et al. [2], we can find the direction in which f decreases fastest using the directional derivative:

$$\min_u \nabla f \cdot u$$
$$= \min_u \|u\|_2 \|\nabla f\|_2 \cos \theta$$

...

Our goal is to choose a Δv that minimizes $\Delta C \approx \nabla C \cdot \Delta v$. The Cauchy–Schwarz inequality tells us that $|\nabla C \cdot \Delta v|$ is constrained by $\|v\|\|\nabla C\|$ where $|\nabla C \cdot \Delta v| = \|v\|\|\nabla C\|$ if and only if $\Delta v = \eta \nabla C$. Since $\nabla C \cdot \eta \nabla C = \eta \|\nabla C\|^2 > 0$ we can choose $\Delta v = -\eta \nabla C$ to minimize ΔC.

> Beweis mit Cauchy Schwarz oder directional derivatives?

IV. Zusammenfassung und Ausblick

blabla

Todo list

References

[1] B. Claise, "IPFIX protocol specifications," Internet-Draft, draft-ietf-ipfix-protocol-07, December 2004.

[2] A. C. Snoeren, C. Partridge, L. A. Sanchez, C. E. Jones, F. Tchakountio, S. T. Kent, and W. T. Strayer, "Hash-based IP traceback," in *ACM SIGCOMM 2001 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, 2001.

[3] A. Belenky and N. Ansari, "IP traceback with deterministic packet marking," *IEEE Communications Letters*, vol. 7, no. 4, pp. 162–164, 2003.