Formalized Generalization Bounds for Perceptron-Like Algorithms

A thesis presented to

the faculty of

the Russ College of Engineering and Technology of Ohio University

In partial fulfillment

of the requirements for the degree

Master of Science

Robin J. Kelby

August 2020

This thesis titled

Formalized Generalization Bounds for Perceptron-Like Algorithms

by

ROBIN J. KELBY

has been approved for

the School of Electrical Engineering and Computer Science

and the Russ College of Engineering and Technology by

Gordon Stewart

Assistant Professor of Electrical Engineering and Technology

Mei Wei

Dean, Russ College of Engineering and Technology

Abstract

KELBY, ROBIN J., M.S., August 2020, Computer Science

Formalized Generalization Bounds for Perceptron-Like Algorithms (15 pp.)

Director of Thesis: Gordon Stewart

Insert your abstract here

Dedication

Dedicated to my Nathan.

Your patience, video games, and good cooking kept me going.

Acknowledgments

Acknowledge later

Table of Contents

List of Tables

List of Figures

1 Introduction

The field of machine learning research has advanced very quickly in the past decade. Machine learning describes the class of computer programs that automatically learn from experience, often employed for classification, recognition, and clustering tasks. One of the classic problems in machine learning is handwritten digit recognition to classify written numbers automatically. Computers have historically struggled to interpret handwritten information because handwriting can vary drastically between writers. While humans can be taught to read as well as learn to read on their own, handwriting recognition can be challenging for computers to accomplish. Several datasets have been created specifically for the problem of handwriting analysis for numerical digits. For example, the MNIST dataset [LBBH98] is one of the primary datasets for computers to learn how to classify handwritten digits into the numbers 0-9. This dataset allows researchers to compare the performance of multiple models, trained and tested on the same data, but using different machine learning algorithms. Some systems have achieved a near-perfect performance on the MNIST dataset for the problem of handwritten digit classification, and ths technology is valuable for processing documents, such as ZIP codes on letters sent through the U.S. Postal Service. Something else

2 BACKGROUND

3 METHODS

4 Results

5 Conclusions

[Ros57] [MGS17] [ABR64] [LTS90] [CCBG07] [DSSS07] [CKS03] [OKC09] [BS19] [BF16] [TD05] [LBBH98] [Var16] [Ler09] [WWP+15] [GSC+16]

References

[ABR64] M. A. Aizerman, E. M. Braverman, and L. I. Rozoner. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control*, 25:821–837, 1964.

[BF16] Adrien Bibal and Benoit Frénay. Interpretability of machine learning models and representations: an introduction. In *ESANN'16 Proceedings*, Bruges, 2016.

[BS19] Alexander Bagnall and Gordon Stewart. Certifying the true error: machine learning in coq with verified generalization guarantees. In *Proceedings of AAAI'19*, pages 2662–2669, Hawaii, 2019.

[CCBG07] Giovanni Cavallanti, Nicolo Cesa-Bianchi, and Claudio Gentile. Tracking the best hyperplane with a simple budget perceptron. *Machine Learning*, 69(23):143–167, 2007.

[CKS03] Koby Crammer, Jaz Kandola, and Yoram Singer. Online classification on a budget. In *Advances in Neural Information Processing Systems 16*. Proceedings of NIPS 2003, 2003.

[DSSS07] Ofer Dekel, Shai Shalev-Shwartz, and Yoram Singer. The forgetron: a kernel-based perceptron on a budget. *SIAM Journal on Computing*, 37(5):1342–1372, 2007.

[GSC+16] Ronghui Gu, Zhong Shao, Hao Chen, Xiongnan (Newman) Wu, Jieung Kim, Vilhelm Sjoberg, and David Costanzo. Certikos: an extensible architecture for building certified concurrent os kernels. In *OSDI'16*, pages 653–669, Savannah, Georgia, 2016.

[LBBH98] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. In *Proceedings of the IEEE*, volume 86(11), pages 2278–2324, 1998.

[Ler09] Xavier Leroy. Formal verification of a realistic compiler. *Communications of the ACM*, 2009.

[LTS90] Esther Levin, Naftali Tishby, and S. A. Solla. A statistical approach to learning and generalization in layered neural networks. In *Proceedings of the IEEE*, volume 78, pages 1568–1574, 1990.

[MGS17] Charlie Murphy, Patrick Gray, and Gordon Stewart. Verified perceptron convergence theorem. In *MAPL'17*, Barcelona, 2017.

[OKC09] Francesco Orabona, Joseph Keshet, and Barbara Caputo. Bounded kernel-based online learning. *Journal of Machine Learning Research*, 10(11):2643–2666, 2009.

[Ros57] Frank Rosenblatt. The perceptron, a perceiving and recognizing automaton. *Report: Cornell Aeronautical Laboratory*, 58(460), 1957.

[TD05] Brian. J. Taylor and Marjorie A. Darrah. Rule extraction as a formal method for the verification and validation of neural networks. In *Proceedings of IEEE International Joint Conference on Neural Networks 2005*, pages 2915–2920, Montreal, 2005.

[Var16] Kush R. Varshney. Engineering safety in machine learning. In *2016 Information Theory and Applications Workshop*, La Jolla, California, 2016.

[WWP+15] James R. Wilcox, Doug Woos, Pavel Panchekha, Zachary Tatlock, Xi Wang, Michael D. Ernst, and Thomas Anderson. Verdi: A framework for implementing and formally verifying distributed systems. In *PLDI'15*, pages 357–368, Portland, Oregon, 2015.