

Learning From Differences

Oliver Linger

February 16, 2024

Final-Year Project - B.Sc. in Computer Science

Supervisor: Dr. Derek Bridge

Department of Computer Science

University College Cork

Abstract

Declaration of Originality

In signing this declaration, you are confirming, in writing that the submitted work is entirely your own original work, except where it clearly attributed otherwise, and that it has not been submitted partly or wholly for any other educational award.

1. This body of work is all my own, unless clearly stated otherwise, with the full and proper accreditation;
2. With respect to my work: none of it has been submitted to any educational institution contributing in any way to an educational award.
3. With respect to anyone else's work: all diagrams, text, code, or ideas have been duly attributed to the source in a scholarly manner, whether from books, papers, lecture notes or any other student's work, whether published or unpublished, electronically or print.

Signed

Date: February 16, 2024

0.1 Acknowledgements

Contents

0.1	Acknowledgements	1
1	Introduction	6
2	Literature Review	7
2.1	KNN + ANN for Regression, utilizing (CBR)	8
2.1.1	Learning From Differences Paper	8
2.1.2	Learning Adaptations for Case-based classification	11
3	Design	12
3.1	Linger regression model	12
3.1.1	Objective	12
3.1.2	Architecture Overview	12
3.1.3	Data Collection and Preprocessing	12
3.1.4	Feature Engineering	12
3.1.5	Model Selection	13
3.1.6	Training Procedure	13
3.1.7	Evaluation Metrics	13
3.1.8	Cross-Validation and Testing	13
3.1.9	Hardware and Software Specifications	13

3.1.10	Ethical Considerations	13
3.1.11	Limitations and Future Work	13
4	Implementation	14
5	Evaluation	15
6	Conclusions	16
7	Experiments	17
7.1	Experiment 1: Abalone dataset Linger Regression	17
7.2	Experiment 2: House Prices Dataset Regression	17
7.3	Experiment 3: Glass Dataset Classification	17
7.4	Experiment 4: Raisin Dataset Classification	17
7.5	Experiment 5: Car Safety Dataset Classification	17
8	Results	18

List of Figures

List of Tables

Chapter 1

Introduction

Chapter 2

Literature Review

In recent years, the application of neural networks in regression tasks has gained significant attention within the realm of machine learning and artificial intelligence. Applying neural networks to regression tasks has become common practice within the machine learning community [4]. Traditionally, supervised learning approaches involve presenting datasets described by features alongside their corresponding expected values to the artificial neural network (ANN), facilitating the learning process aimed at predicting these expected values.

Inspired by case-based reasoning, which involves solving new problems by adapting solutions from old ones [5], researchers have developed predictive ensemble models based on the principles of case-based reasoning (CBR). These principles involve retrieving the most similar cases, reusing them to attempt to solve the problem, revising the proposed solution if necessary, and retaining the new solution as part of a new case [6].

Researchers have suggested the potential advantages of deviating from this conventional paradigm by training neural networks on differences between sets of features to predict differences between values. This departure from the traditional approach underscores a shift towards a more nuanced understanding of the learning process, emphasizing the exploration of alternative methodologies to enhance the efficiency and efficacy of neural network training. Additionally, it aims to provide more transparency within the learning process to offer insights into the black box that is a neural network.

2.1 KNN + ANN for Regression, utilizing (CBR)

The exploration of alternative methodologies for enhancing the performance of neural networks in regression tasks has attracted considerable attention within the machine learning community in recent years. One innovative approach, inspired by the principles of case-based reasoning (CBR), involves training neural networks to predict differences between values based on differences between sets of features.

This departure from the conventional supervised learning paradigm aims to leverage the inherent advantages of learning from differences, potentially leading to more interpretable and efficient models. The study under consideration [2] conducts a systematic factorial study to investigate the efficacy of this approach across various datasets and experimental conditions. The findings suggest that neural networks trained on differences demonstrate comparable or even superior performance to those trained using traditional methods, while requiring significantly fewer epochs for convergence. In this section of the literature review, we consider the utilization of case-based reasoning for regression tasks, with a specific focus on the implementation of KNN (K-Nearest Neighbors) and ANN (Artificial Neural Network) frameworks.

2.1.1 Learning From Differences Paper

The paper by [2] introduces the innovative concept of training a neural network based on the disparities between a base case and its closest neighbors. Titled "A Factorial Study of Neural Network Learning from Differences for Regression" [2], the study aims to present novel learning methodologies geared towards enhancing performance and fostering a more interpretable learning process. The research evaluates three model variations: a benchmark neural network, a learning from differences model, and an augmented learning from differences model incorporating the original context. The context here refers to the base case, from which differences with its nearest neighbors are calculated.

According to [2], the study's key findings underscore a significant increase in accuracy when incorporating contextual information in the learning process. Moreover, comparable or superior results were achieved with fewer training epochs compared to the basic neural network. This enhancement is attributed partly to the broader training dataset that encompasses neighbors greater than 1. Notably, the paper advocates for a hybrid approach, combining learning from differences with direct feature-based learning, which yields superior results.

The related work for the paper [2] has addressed the use of neural networks in the case-based reasoning process. These papers have focused on exploiting the base case using the idea that adaptation of information is acquirable from the differences between pairs [3]. Since then, various different approaches have been used. More recently, in relation to using neural networks to predict the differences between problems, a few interesting points have been investigated. These points show the ability of a neural network to correct the solution of the most similar retrieved case. These preliminary works expressed in the paper showed that the use of differences plus context yielded superior results [1]. The previous studies examined in the paper [2] did not examine the impact of different parameters during the (CBR) case-based reasoning experiments. The paper [2] made an effort to explore the effect of epochs in the learning process.

The graphical representations in the paper highlight a distinct trend: rapid attainment of high accuracy levels with the learning from differences method, particularly when contextual information is included. Conversely, the basic neural network requires a substantially larger number of epochs to achieve comparable accuracy, if not worse.

This disparity suggests the efficiency of the learning from differences approach.

In conclusion, the paper suggests a marginal performance enhancement and a substantial reduction in training epochs through the adoption of the differences method and inclusion of contextual information. However, it is crucial to acknowledge the time overhead associated with retrieving and training using similar cases.

The process of learning from differences for regression, as described in the paper by [2], demonstrates the use of the case difference heuristic in the creation of a predictive model. Consider a dataset D of pairs (X_i, Y_i) , where Y_i is a vector of n numerical values corresponding to the result we expect to predict. Within the paper, the method of executing a regression model follows the normal pattern of training and predicting that a standard model would. However, instead of the features themselves being passed into the neural net, the differences between the features and their nearest neighbors are passed in. The differences set is represented by ΔD^{train} . The network, after being fitted with the above, is used again to make predictions. This time, the test set is used, where nearest neighbors in the training set are retrieved using the items in the training set D^{train} . The resultant dataset of test differences is called ΔD^{test} . Predictions are made on this dataset, and the predictions are added to their corresponding differences in D^{train} . The average of the resultant label is taken as the regression prediction.

The paper [2] has an associated library where they have implemented a regression-based machine learning model.

The new model proposed in the paper is tested using the factorial study methodology. This process involved finding a network structure through experimental means for each dataset. It entailed training and testing using a wide range of hyperparameters to determine the optimal settings.

Despite its contributions, the paper falls short in adequately addressing classification and exploring additional variations of the learning from differences model beyond context inclusion. Incorporating original values for the base case in the generated differences dataset provides a valuable framework for understanding the observed disparities. While the model with contextual inclusion demonstrates slight improvements over the base model in learning from differences, a more comprehensive examination of the methodology is warranted for further refinement and enhancement.

The paper tested three models using the same factorial methodology: a base version, which is a basic neural network trained empirically; a differences model employing the (CBH) Case-based Heuristic for predictions; and a differences model with its base context included. The results yielded interesting insights, with similar performance for all three models and a slight improvement in the differences + context model. Notably, the differences and differences + context models required significantly fewer epochs to achieve accuracy compared to the basic neural network, which achieved similar accuracy after more epochs.

Another consideration in the paper [2] was the number of neighbors for both training and prediction. The study revealed no consistent trend across datasets, suggesting dataset-specific requirements. This approach could be seen as a method of data augmentation, as choosing a larger number of neighbors dramatically increases the dataset size.

In conclusion, the learning from differences method achieved equal or higher accuracy compared to the base neural network model with far fewer epochs, potentially leading to faster training times. However, careful consideration must be given to the number of neighbors used, as increasing it can lead to significantly higher time overhead.

An improvement on the study [2] would be to vary the network structure, including but not limited to increasing the number of layers and neurons. Additionally, including datasets that require different types of networks could provide valuable insights. Exploring these parameters could lead to improvements in the existing learning from differences regression model. While the study tested a singular variant through the inclusion of context, exploring more variants of the model could yield further enhancements.

2.1.2 Learning Adaptations for Case-based classification

Chapter 3

Design

3.1 Linger regression model

3.1.1 Objective

The objective of the Linger Regression model is to provide a more dynamic and interpretive model for Regression using differences rather. This model hopes to improve and provide better regression models allow a more interpretable learning process.

3.1.2 Architecture Overview

Provide a concise overview of the architecture of the machine learning device.

3.1.3 Data Collection and Preprocessing

Explain how training and validation datasets were collected and preprocessed.

3.1.4 Feature Engineering

Describe the process of feature extraction or engineering for the machine learning model.

3.1.5 Model Selection

Explain the rationale behind selecting the machine learning model or algorithm.

3.1.6 Training Procedure

Detail the training procedure, including optimization algorithms and hyperparameters.

3.1.7 Evaluation Metrics

Define the evaluation metrics used to assess the performance of the machine learning device.

3.1.8 Cross-Validation and Testing

Describe the methodology used for cross-validation and testing of the machine learning model.

3.1.9 Hardware and Software Specifications

Provide details about the hardware and software infrastructure used in the development and deployment of the machine learning device.

3.1.10 Ethical Considerations

Discuss any ethical considerations associated with the machine learning device.

3.1.11 Limitations and Future Work

Acknowledge limitations and propose future improvements for the machine learning device.

Chapter 4

Implementation

Chapter 5

Evaluation

Chapter 6

Conclusions

Chapter 7

Experiments

7.1 Experiment 1: Abalone dataset Linger Regression

7.2 Experiment 2: House Prices Dataset Regression

7.3 Experiment 3: Glass Dataset Classification

7.4 Experiment 4: Raisin Dataset Classification

7.5 Experiment 5: Car Safety Dataset Classification

Chapter 8

Results

Bibliography

- [1] Susan Craw, Nirmalie Wiratunga, and Ray C Rowe. Learning adaptation knowledge to improve case-based reasoning. *Artificial intelligence*, 170(16-17):1175–1192, 2006.
- [2] Mathieu D’Aquin, Emmanuel Nauer, and Justine Lieber. A factorial study of neural network learning from differences for regression. In Mark T. Keane and Nirmalie Wiratunga, editors, *Case-Based Reasoning Research and Development. ICCBR 2022. Lecture Notes in Computer Science*, volume 13405, page PageNumbersHere. Springer, Cham, 2022.
- [3] Kathleen Hanney and Mark T Keane. Learning adaptation rules from a case-base. In *European Workshop on Advances in Case-Based Reasoning*, pages 179–192. Springer, 1996.
- [4] Leorey Marquez, Tim Hill, Reginald Worthley, and William Remus. Neural network models as an alternative to regression. In *Proceedings of the Twenty-Fourth Annual Hawaii International Conference on System Sciences*, volume 4, pages 129–135. IEEE, 1991.
- [5] Christopher K Riesbeck and Roger C Schank. *Inside case-based reasoning*. Psychology Press, 2013.
- [6] Ian Watson and Farhi Marir. Case-based reasoning: A review. *The knowledge engineering review*, 9(4):327–354, 1994.