

# **Applying Temperature Scaling to Reduce DNN Miscalibration**

K.M. Kalani Hasantha

199471X

Degree of Master of Science in Artificial Intelligence

Department of Computational Mathematics

Faculty of Information Technology

University of Moratuwa Sri Lanka

March 2022

# **Applying Temperature Scaling to Reduce DNN Miscalibration**

K.M. Kalani Hasantha

199471X

Thesis submitted in partial fulfillment of the requirements for the degree of Master  
of Science in Artificial Intelligence

Department of Computational Mathematics

Faculty of Information Technology

University of Moratuwa Sri Lanka

March 2022

## **Declaration**

I declare that this is my work and this thesis does not incorporate without acknowledgment any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Also, I hereby grant to the University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic, or another medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name of Student:

Kalani Hasantha

Signature of Student:

Date:

The above candidate has carried out research for the Master's Dissertation under my supervision.

Name of Supervisor:

Dr. Subha Fernando

Signature of Supervisor:

Date:

## **Acknowledgments**

I would like to express my sincere gratitude to my supervisor Dr. Subha Fernando for her immense support and guidance throughout the research. This research would not have been completed without her extensive expertise and continuous support. Further, I would like to thank all academic staff of the Department of Computational Mathematics for providing sufficient knowledge to complete this research.

Finally, I wish to express my sincere thanks to my family and colleagues for their support and encouragement.

## Abstract

Deep Neural Networks (DNNs) exhibit state of art performance in a wide variety of fields. However, with the modern architectural trends and increased network capacity, DNN tends to be poorly calibrated. Only well-calibrated models can make probabilistic predictions that reflect real-world probabilities. This issue limits the usage of DNN base models in critical decision scenarios. For achieving a reliable probability score for pre-trained models, several techniques have been introduced including histogram binning, Bayesian Binning into Quantiles, isotonic regression, and Platt scaling. Temperature scaling (TS) was recently introduced, which achieves superior performance compared to the other techniques. But TS does not work properly with small validation datasets, highly accurate networks, and non-highly accurate ones. Therefore, DNN calibration remains a research challenge.

This research focuses on improving the performance of DNN calibration by introducing changes to the state of art post-hoc calibration technique, Temperature Scaling. To overcome the limitations of the Temperature Scaling technique, the novel Class-wise Temperature Scaling technique is proposed. The proposed technique calculates the optimum temperature for each class by binning the predictions on class and the confidence values are rescaled using class-wise temperature values. Class-wise Temperature Scaling can further minimize calibration error in classification models, up to 10% than the standard Temperature scaling technique.

# Table of Content

Declaration.....	i	
Acknowledgments.....	ii	
Abstract.....	iii	
Table of Content .....	iv	
List of Figures .....	vii	
List of Tables .....	viii	
Abbreviation .....	ix	
Chapter 1	Introduction..... 1	
1.1 Prolegomena .....	1	
1.2 Background and motivation.....	1	
1.3 Aims and objectives.....	2	
1.4 Problem definition .....	2	
1.5 Class-wise Temperature Scaling .....	3	
1.6 Resource requirements.....	3	
1.7 Structure of the thesis.....	3	
1.8 Summary .....	3	
Chapter 2	Uncertainty Scaling .....	4
2.1 Introduction.....	4	
2.2 Early approaches of uncertainty scaling .....	4	
2.3 Breakthrough in DNN calibration techniques.....	6	
2.4 Challenges in DNN calibration .....	9	
2.5 Problem definition .....	11	
2.6 Summary .....	11	
Chapter 3	Definitions and Technologies.....	12
3.1 Introduction.....	12	
3.2 Reliability Diagram.....	12	
3.3 Expected Sample Accuracy .....	13	
3.4 Average Sample Confidence.....	13	
3.5 Measurements of Model Calibration.....	14	
3.5.2 Maximum Calibration Error (MCE) .....	14	

3.5.1 Expected Calibration Error (ECE) .....	14	
3.5.3 Negative Log-Likelihood (NLL) .....	15	
3.6 Neural Network Architectures .....	15	
3.6.1 LeNet-5 .....	15	
3.6.2 VGGNet .....	16	
3.6.3 ResNet.....	17	
3.6.4 DenseNet.....	17	
3.7 Summary .....	18	
Chapter 4	Class-wise Temperature Scaling Approach.....	19
4.1 Introduction.....	19	
4.2 Hypothesis.....	19	
4.3 Input .....	19	
4.4 Output .....	19	
4.5 Process .....	20	
4.5.1 Analyzing Miscalibration in DNN .....	20	
4.5.2 Identifying drawbacks of Temperature Scaling .....	21	
4.5.3 Identifying Approaches to Prevent drawbacks of Temperature Scaling.....	22	
4.5.4 Class-wise Temperature Scaling System .....	23	
4.6 Users .....	23	
4.7 Features .....	23	
4.8 Summary .....	24	
Chapter 5	Design of Class-wise Temperature Scaling.....	25
5. 1 Introduction.....	25	
5.2 Class-wise Temperature Calculation Module .....	25	
5.3 Confidence Scaling Module.....	26	
5.4 Summary .....	26	
Chapter 6	Implementation .....	27
6.1 Introduction.....	27	
6.2 Data Sets .....	27	
6.3 Model Training .....	27	
6.4 Class-wise Temperature Scaling Algorithm .....	29	
6.5 Class-wise Temperature Calculator Implementation .....	29	
6.6 Confidence Scaler Module Implementation.....	30	
6.7 Summary .....	30	

Chapter 7	Evaluation .....	31
7.1 Introduction.....		31
7.2 Model Calibration Evaluation.....		31
7.3 Summary .....		34
Chapter 8	Conclusion & Further Work .....	35
8.1 Introduction.....		35
8.2 Conclusion .....		35
8.2.1 Achievement of Project Objectives.....		35
8.2.2 Overall Conclusion .....		36
8.3 Limitations and further work .....		36
8.4 Summary .....		37
References.....		38
Appendix.....		42
Appendix I: Standard Temperature Calculation .....		42
Appendix II: Binning on Confidence Range.....		42
Appendix III: Standard ECE and MCE Calculation .....		43
Appendix IV: Reliability Diagram.....		43
Appendix V: Confidence Calibration using Temperature Scaling .....		44
Appendix VI: Class-wise temperature calculation.....		44
Appendix VII: Confidence Calibration using Temperature Scaling.....		44
Appendix VIII: Binning on Predicted Class .....		45

# List of Figures

Figure 2. 1 Overview of DNN Calibration Techniques.....	4
Figure 3.2. 2 Reliability diagram (bar chart) .....	13
Figure 3.2. 2 Reliability diagram (line chart) .....	13
Figure 3.6.1 LeNet5 architecture diagram.....	16
Figure 3.6.2 VGG16 architecture diagram.....	16
Figure 3.6.3 ResNet50 architecture diagram.....	17
Figure 3.6.4 DenseNet201 architecture diagram.....	18
Figure 4.2 Approach.....	19
Figure 4.5.1 Reliability diagrams for uncalibrated models trained with Cifar10.....	21
Figure 4.5.2.1 Reliability diagrams after calibrating with TS (models trained with Cifar10) .....	21
Figure 4.5.2.2 Data distribution between Bins in probability range binning approach.....	22
Figure 4.5.3 Data distribution between Bins in class wise binning approach.....	22
Figure 5.1 High-level design of class-wise Temperature Scaling approach.....	25
Figure 7.2.1 Reliability diagrams after calibrating with Class-wise TS.....	31
Figure 7.2.2 ECE, MCE evaluation of the model with CIFAR10 and CIFAR100 datasets.....	32

## List of Tables

Table 2.4.1 summary of benefits and limitations of post-hoc calibration techniques.....	10
Table 4.5.1 DNN architecture depth and parameters comparison.....	20
Table 6.3 Implemented models and the architectures.....	28
Table 7.2 ECE, MCE comparison before and after applying calibration techniques.....	31

# Abbreviation

<b>Abbreviation</b>	<b>Definition</b>
DNN	Deep Neural Network
CNN	Convolutional Neural Networks
ECE	Expected Calibration Error
MCE	Maximum Calibration Error
NLL	Negative Log Likelihood
TS	Temperature Scaling
ATS	Attended temperature scaling
BTS	Bin-wise temperature scaling
UTS	Unsupervised temperature scaling
PTS	Parameterized temperature scaling