DATA AND DOMAIN UNCERTAINTY IN MACHINE LEARNING

by

Artin Majdi

A Dissertation Submitted to the Faculty of the

Department of Electrical and Computer Engineering

In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy

In the Graduate College

The University of Arizona

June 27, 2023

Get the official approval page
from the Graduate College
before your final defense.

Statement by Author

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Signed: _____

Approval by Dissertation Director

This dissertation has been approved on the date shown below:

_____   _____
Jeffrey J. Rodriguez   Date
Associate Professor

Acknowledgments

As I reflect upon the completion of this invigorating journey, my heart brims with profound gratitude towards those who journeyed with me, imparting their wisdom, support, and guidance at every step.

First and foremost, my deepest appreciation is directed towards my mentor, Dr. Jeffrey J. Rodriguez, whose steadfast guidance and support throughout the project has been unparalleled. His expertise, dedication, and continual encouragement have been priceless to me, for which I am eternally thankful.

My sincere gratitude goes out to Mr. Nirav Merchant, Ms. Maliaca Oxnam, and Drs. Carlos Alsua, Manoj Saranathan, and Mahesh Keerthivasan, whose unceasing support has been invaluable throughout this endeavor. Additionally, I wish to convey my gratitude to Drs. Ali Bilgin and Abhijit Mahalanobis for their insightful contributions and feedback which significantly enhanced the depth and quality of my work. Your invaluable support and mentorship have played a crucial role in my personal and academic evolution.

An acknowledgment wouldn't be complete without a heartfelt tribute to my family. Your unwavering faith in my capabilities, your patience during the tough times, and your consistent encouragement have served as my strongest pillars of support.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Brief Background on Machine Learning and Uncertainty Management

The world is currently experiencing an era of rapid development in the field of machine learning. This exciting progression has triggered massive transformations across numerous sectors, including healthcare, automotive industries, and crowdsourcing platforms. In healthcare, machine learning algorithms help in diagnosing diseases, predicting patient outcomes, and personalizing treatment plans. In the automotive industry, the emergence of self-driving vehicles is revolutionizing transportation and mobility. In crowdsourcing, machine learning is used to aggregate information sourced from a large number of individuals to solve complex problems or provide insights. However, as we leverage these technologies to drive innovation and efficiency, we are also confronted with a critical challenge: *management of uncertainty and domain fluctuations in machine learning models*. Real-world applications of machine learning are riddled with situations where the provided data can contain missing, incorrect, or uncertain values. This often results in the unpredictable behavior of the machine learning models, leading to fluctuating performance across different domains. The precision of model predictions is crucial in high-stakes decision-making situations, such as medical diagnosis or autonomous driving. Given the inherent variability of input data and the susceptibility of models to error, accuracy alone is no longer sufficient. It is crucial to also account for uncertainty in the

model's predictions. With sufficient knowledge of model and data uncertainty, we can make informed decisions regarding whether to trust the model's predictions or seek additional information before making a final decision. This capability is particularly important in critical situations where inaccurate predictions could have severe repercussions.

1.1.1   Uncertainty in Machine Learning Models

Uncertainty is an integral part of any predictive model, and machine learning models are no exception. Recognizing, quantifying, and managing this uncertainty is a critical aspect of developing robust and reliable machine learning systems. Uncertainties in the field of machine learning are typically divided into two categories: aleatoric and epistemic. Aleatoric uncertainty, also referred to as statistical uncertainty, is linked to the data's inherent noise or variability. This type of uncertainty is often irreducible as it stems from factors such as measurement errors or inherent randomness in the system being modeled. Epistemic uncertainty, on the other hand, is related to our lack of knowledge about the system. This form of uncertainty, also known as model uncertainty, arises from our inability to perfectly specify the parameters of our model or from using a model that does not perfectly capture the true underlying process. The sources of uncertainty in machine learning models are complex and can be attributed to various factors. Measurement errors, missing data, and inherent noise are examples of data-related sources of uncertainty. In contrast, model-related sources of uncertainty include the choice of model structure, or model parameters, and the use of approximations in model computations. Uncertainty can significantly impact the performance and behavior of machine learning models. Unmanaged uncertainty can lead to reduced model performance

as the model struggles to make accurate predictions in the face of noisy data or imperfect model specifications. Uncertainty can also lead to biased predictions, as models may overfit to noisy data or become too confident in their predictions, ignoring the inherent uncertainty in the process. Overconfidence is a particularly dangerous effect of unmanaged uncertainty. When a model is overconfident, it may produce very certain predictions even when they are inaccurate. This can lead to poor decision-making, as users of the model may place too much trust in its predictions. In high-stakes domains such as healthcare or autonomous driving, overconfidence can lead to severe negative outcomes. Therefore, understanding and managing uncertainty is a critical task in machine learning. By providing uncertainty estimates along with predictions, machine learning models can become more reliable and trustworthy, facilitating their use in real-world decision-making tasks. Recent methods tackle the problem of missing and uncertain data, but they assume the data is drawn from well-known distributions defined by a few parameters. Some studies approach the issue of lack of accurate labels by pre-training a network on a large dataset with noisy labels and fine-tuning it for a smaller target dataset [1]. Semi-supervised techniques have also been proposed where the data with noisy labels is discarded [2]. However, they suffer from model complexity and cannot be applied to large-scale datasets. Even loss functions that are perceived as noise-robust are not completely robust to label noise [3]. With the rise and influence of ML in medical applications and the need to translate newly developed techniques into clinical practice, questions about the safety and uncertainty of models have gained more importance. Prior research mostly focused on assessing the correctness of individual decisions and modeling the behavior of individual labelers (human experts or non-experts who assign labels to data) [4]. Label bias becomes more crucial when we move from using manual delineation as our gold standard to

using existing software (e.g., Free Surfer [5] for subcortical segmentation tasks).

1.2   Objectives and Scope

The primary objective of this dissertation is to advance the understanding of uncertainty management and, to a lesser extent, domain fluctuation in machine learning models, with the ultimate aim of enhancing models' performance, reliability, and applicability in a variety of real-world settings. We have developed, and evaluated novel methods for estimating and mitigating the impact of uncertainty on machine learning outcomes. By doing so, we hope to facilitate the more widespread adoption of these technologies across various sectors and real-world applications and, in the process, harness their full potential.

A significant part of this objective is devoted to investigating ways to incorporate uncertainty information into existing techniques with the goal of improving the accuracy of the prediction models as well as providing an additional confidence score that can enhance the models' interpretability and applicability in real-life settings. The challenge facing the machine learning community is to establish a comprehensive understanding of these issues, laying the foundation for the design of new techniques and strategies that manage uncertainty effectively in both supervised and unsupervised settings.

Addressing this challenge, we present several approaches to estimate and mitigate the effect of uncertainty in model outcomes. We propose strategies to manage uncertainty, aiming to make machine learning models more robust, reliable, and trustworthy. We focus primarily on healthcare, autonomous driving, and crowdsourcing, as these are areas where either the

stakes are high or their applicability is widespread, and thus accurate uncertainty estimates can significantly enhance decision-making and outcomes.

In healthcare, we aim to improve the performance of machine learning models used for tasks such as disease diagnosis and organ segmentation. In the realm of autonomous driving, we introduce a novel transfer learning approach for driver distraction detection. In the field of crowdsourcing and ensemble learning, we have proposed new label aggregation techniques that take into account the consistency and accuracy of annotators (models in the case of ensemble learning). Further, we investigate the impact of domain fluctuation on model accuracy when segmenting thalamic nuclei and provide a novel technique that facilitates the utilization of low-contrast imaging sequences by developing a model that, albeit trained on advanced imaging technologies, can readily be used on low-contrast yet more widely available imaging sequences without sacrificing too much accuracy. By keeping a broad scope, we have developed various methods that are not only effective in handling uncertainty and domain fluctuations but are also versatile and adaptable across different domains and applications, and thus have enhanced the utility of these techniques across different domains.

1.3 Dissertation's Organization

This dissertation contributes to the ongoing research in uncertainty management and domain fluctuation in machine learning models. The methods proposed in this dissertation take into account the inherent uncertainty and variability in the data, leading to improved model performance. These methods provide a framework that can be extended to other machine learn-

ing applications. Chapter 2 presents a novel method, "crowd-certain", that provides a more accurate and reliable label aggregation technique, leading to improved overall performance in both crowdsourcing and ensemble learning scenarios. This method takes into account the consistency and accuracy of the annotators as a measure of their reliability, which allows us to obtain a weight that closely follows the annotator's degree of reliability. Chapter 3 proposes a novel hierarchical multilabel classification technique that utilizes the taxonomic relationship between different classes to improve classification accuracy. Further, to reduce the effect of domain fluctuation and improve the generalizability of the model to images obtained from other sources, the proposed technique provides one model trained on multiple large publicly available chest X-ray datasets (CheXpert [6], NIH [7], and PADCHEST [8]). Chapter 4 presents a fast and accurate convolutional neural network (CNN) for segmentation of thalamic nuclei that is optimized for various diseases, magnetic field strengths, and image modalities. It demonstrates the potential of the proposed method for improving our understanding of the thalamic nuclei's involvement in neurological diseases. Finally, Chapters 5 and 6 provide a transfer learning approach for the detection of driver distraction and primary cilia cells, respectively. The proposed technique uses a combination of a CNN and a random decision forest to improve classification accuracy.

References

[1] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and Transferring Mid-Level Image Representations Using Convolutional Neural Networks," in *2014 IEEE Conf. Comput. Vis. Pattern Recognit.*, (Columbus, OH, USA), pp. 1717–1724, IEEE, June 2014.

[2] X. Zhu and Z. Ghahramani, "Learning From Labeled and Unlabeled Data With Label Propagation," technical Report CMU-CALD-02–107, School of Computer Science, Carnegie Mellon University, 2002.

[3] P. L. Bartlett, M. I. Jordan, J. D. McAuliffe, M. I. J. Peter L. BARTLETT, J. D. MCAULIFFE, P. L. Bartlett, M. I. Jordan, and J. D. McAuliffe, "Convexity, Classification, and Risk Bounds," *J. Am. Stat. Assoc.*, 2006.

[4] V. C. Raykar, S. Yu, L. H. Zhao, A. Jerebko, C. Florin, G. H. Valadez, L. Bogoni, and L. Moy, "Supervised Learning From Multiple Experts: Whom to Trust When Everyone Lies a Bit," in *Proc. 26th Annu. Int. Conf. Mach. Learn. - ICML 09*, (Montreal, Quebec, Canada), pp. 1–8, ACM Press, 2009.

[5] B. Fischl, "FreeSurfer," *NeuroImage*, vol. 62, pp. 774–781, Aug. 2012.

[6] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Ilcus, C. Chute, H. Marklund, B. Haghgoo, R. Ball, K. Shpanskaya, J. Seekins, D. A. Mong, S. S. Halabi, J. K. Sandberg, R. Jones, D. B. Larson, C. P. Langlotz, B. N. Patel, M. P. Lungren, and A. Y. Ng, "CheXpert: A Large Chest Radiograph Dataset With Uncertainty Labels and Expert Comparison," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, pp. 590–597, July 2019.

[7] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," in *2017 IEEE Conf. Comput. Vis. Pattern Recognit. CVPR*, (Honolulu, HI), pp. 3462–3471, IEEE, July 2017.

[8] A. Bustos, A. Pertusa, J.-M. Salinas, and M. de la Iglesia-Vayá, "Padchest: A Large Chest X-Ray Image Dataset With Multi-Label Annotated Reports," *Medical Image Analysis*, vol. 66, p. 101797, Dec. 2020.