

Evaluation Form – Technical Background Review

Student Name: Adrija Bhattacharya

Project Advisor: Dr. Xiaoli Ma

Team Name: VIASAT-AMEND

Team Members: Chris Rothmann, Mikias Balkew, Shreyas Mhasawade, Tyler Cole

- _____ / 30 Technical Content
- Current state-of-the-art and commercial products
 - Underlying technology
 - Implementation of the technology
 - Overall quality of the technical summary
- _____ / 30 Use of Technical Reference Sources
- Appropriate number of sources (at least six)
 - Sufficient number of source types (at least four)
 - Quality of the sources
 - Appropriate citations in body of text
 - Reference list in proper format
- _____ / 40 Effectiveness of Writing, Organization, and Development of Content
- Introductory paragraph
 - Clear flow of information
 - Organization
 - Grammar, spelling, punctuation
 - Style, readability, audience appropriateness, conformance to standards
- / 100 Total - Technical Review Paper**

VIASAT-AMEND CALIBRATION USING MACHINE LEARNING TOOLBOX

Introduction

In the case of any new technology, modeling and simulations are key to understanding and predicting the behavior of a system that does not exist, or if physical measurement and experimentation is impractical. Simulations can not only help operators attain the desired performance based on its strategic vision but also offer an expected outcome parameter for comparison. This process of comparison of the actual output and the expected output given by a system that uses post-processing as a technique to improve error distribution of a predictive model is called calibration [1].

The main focus of this project is to incorporate machine learning (ML) hooks for automated system tuning in an Autotracking Control System which will keep the Ground Station's parabolic dish pointed at the targeted Low Earth Orbit satellite through tracking and control algorithms. This effort will extend the model to better examine system performance trades and will add scenario modeling to support Machine Learning. An existing Machine Learning tool can be used to examine the feasibility of deriving system calibration factors for ground stations. Through calibration, the existing Simulink model from previous years can be programmed such that the distribution and behavior of the probability predicted is similar to the distribution and behavior of probability observed in training data. This paper focusses on exploring the concept of Machine Learning for calibration systems implemented through an available toolbox software.

Competing Technologies:

A number of software applications for machine learning and calibration are available for consideration. A few of interest are listed below:

- CobWeb 1.0: machine learning toolbox (No price listed): The CobWeb software offers supervised machine learning techniques [2]. The advantage supervised technique offers compared to unsupervised technique is that it is effective in separating non-linear separable data. It also allows the export of data to various formats such as ParaView (.vtk), DSI Studio (.fib) for visualization and animation, and Microsoft® Excel and MATLAB® for numerical calculation and simulations.
- Alteryx (\$5195 per year): Alteryx is the highest peer rated vendor for Data Science & Machine Learning Platforms. It has highly useful functionalities and tools like AutoML which automatically builds a model of the data, Assisted Model Tool that creates a machine-learning pipeline, and Fit

tool that takes input from the dataset and the other machine learning tools and fits the model to the data [3].

- TensorFlow library in Python (Free Opensource): This software offers multiple levels of abstraction so one can choose the right one for their needs. It is an end-to-end open source platform for machine learning [4]. It has a comprehensive, flexible ecosystem of tools, libraries and community resources.
- MATLAB Machine Learning Toolbox (\$1250 Monthly): This software allows the user to draw inferences from data and build predictive models either interactively, using the Classification and Regression Learner apps, or programmatically, using AutoML. It provides supervised, semi-supervised, and unsupervised machine learning algorithms [5].

Some other technologies that offer similar features are: Neuton AutoML (\$2,499.00 Monthly), BigML (\$1000 Monthly), QFlow.ai (\$699 Monthly).

Choosing the best Machine Learning toolbox for this project

Viasat is interested in developing machine learning techniques to automate the derivation of servo calibration factors for new ground stations, and the project over the course of this year draws heavily from previous modeling work done in Simulink. Most of the project will therefore be done primarily in Mathworks products like MATLAB and Simulink, and MATLAB Machine Learning Toolbox would be the best software in terms of compatibility and ease of use with respect to the rest of the project. Although this is a costly toolbox, the availability of Georgia Tech's student license makes it an affordable option for this project.

The "Statistics and Machine Learning Toolbox" in MATLAB provides functions to describe, analyze, and model data [5]. Users can use descriptive statistics and plots to perform exploratory data analysis, fit probability distributions to data, generate random numbers for Monte Carlo simulations, and perform hypothesis tests. It also offers regression and classification algorithms that enables users to draw inferences from data and build predictive models. A challenge with building predictive models for prediction-based machine learning is the time involved in the development of the tools. However, predictive models can be a game changer in the data center industry since it would transform the way data centers are managed in addition to assessing the design strategies prior to their implementation [6]. MATLAB machine learning toolbox provides a simulation interactive environment and a valuable set of tools to quickly develop predictions based on machine learning algorithms, and thus, is best suited for the purpose of this project.

Calibration Methods through Machine Learning

The evaluation of Machine Learning models is a crucial step before deployment, and it is thus essential to go through a behavior analysis of a ML model. In many real-life applications, along with mean error of the model, it is also important to know how this error is distributed and how well probability estimations are made [7].

There are primarily two methods of calibration to understand error distribution [1]:

- Platt's : This technique uses a variation of sigmoid function to fit the distribution of predicted probabilities to the distribution of probability observed in training data.
- Isotonic: This technique uses piecewise constant non-decreasing function instead of sigmoid function. In this we fit like linear regression in the piecewise fashion.

Machine learning algorithms build models by examining many examples in attempts to find a model that minimizes error and loss through an empirical risk minimization [8]. This is primarily done in software. While this project can be completed without physical hardware, solutions may be validated against real systems and / or using real system data upon Viasat's discretion. The specific calibration model data that will be used in the project to determine optimal values of the model parameters formulated by training hidden neurons within a machine learning framework based on available will be discussed later.

Conclusion

Applying optimization algorithms to minimize the deviation between model output and corresponding measurements using gradient based algorithms belongs to the state of the art in calibration technology. The MATLAB Machine Learning Toolbox offers features like supervised algorithms and predictive models which serve this purpose and thus is best suited as the machine learning software for calibration in this project.

References:

- [1] R. Ranjan, "Calibration in Machine Learning," 14-Sep-2019. [Online]. Available: <https://medium.com/analytics-vidhya/calibration-in-machine-learning-e7972ac93555>. [Accessed: 08-Oct-2021].
- [2] S. Chauhan, K. Sell, W. Rühaak, T. Wille and I. Sass, "CobWeb 1.0: machine learning toolbox for tomographic imaging", *Geosci. Model Dev.*, vol. 13, pp. 315-334, Jan. 2020.
- [3] "Gartner 2021 Magic Quadrant for Data Science and Machine Learning Platforms," *Alteryx*. [Online]. Available: https://www.alteryx.com/resources/report/gartner-magic-quadrant-data-science-machine-learning?utm_campaign=NA_Search_Machine_Learning_AO&utm_source=google&utm_medium=cpc&utm_content=&utm_term=machine+learning. [Accessed: 08-Oct-2021].
- [4] "Tensorflow," *TensorFlow*. [Online]. Available: <https://www.tensorflow.org/>. [Accessed: 08-Oct-2021].
- [5] "Statistics and machine learning toolbox," *MATLAB*. [Online]. Available: <https://www.mathworks.com/products/statistics.html>. [Accessed: 08-Oct-2021].
- [6] M. Levy, D. Raviv and J. Baker, "Data Center Predictions using MATLAB Machine Learning Toolbox," 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), 2019, pp. 0458-0464, doi: 10.1109/CCWC.2019.8666624.
- [7] C. Chen et al., "Calibration of Low-Cost Particle Sensors by Using Machine-Learning Method," 2018 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), 2018, pp. 111-114, doi: 10.1109/APCCAS.2018.8605619.
- [8] J. Marpaung, R. Stureborg and J. K. Kimani, "Integrating Machine Learning with MATLAB and Simulink in an Introductory Robotics Class," 2020 Annual Conference Northeast Section (ASEE-NE), 2020, pp. 1-3, doi: 10.1109/ASEENE51624.2020.9292645.