VICTORIA UNIVERSITY OF
WELLINGTON
TE HERENGA WAKA

School of Engineering and Computer Science
Te Kura Mātai Pūkaha, Pūrorohiko

Genetic Programming for Antarctic Ice Sheet Modelling

Samuel Mata

Supervisors: Dr. Bach Nguyen, Dr. Bing Xue

Submitted in partial fulfilment of the requirements for
Bachelor of Science with Honors in Artificial Intelligence.

Abstract

With the growing need for accurate long-term modelling of the Antarctic Ice Sheets, the currently used statistical models do not provide the necessary computational efficiency for sufficiently long term predictions. This project aims to investigate and evaluate the application of Genetic Programming *(GP)* and Evolutionary Learning techniques as a potential alternative to traditional methods.

1. Problem Statement

The changing conditions of Antarctica's Ice Sheets are a significant factor in global climate change, particularly with respect to rising sea levels. In this context, the ability to accurately model the long-term behaviour of these ice sheets is of great importance. Current approaches are not computationally efficient enough to provide meaningfully long-term predictions, and so this project aims to investigate the application of Genetic Programming *(GP)* and Evolutionary Learning techniques as a potential alternative. There is some challenge in the application of these techniques, as the models produced must be accurate and efficient enough to meaningfully improve the practicality of long-term predictions. Furthermore, for the results to be trusted the models must be explainable and interpretable.

2. Motivations

Current methods of modelling Antarctic Ice Sheet measurements - such as those being undertaken at Victoria University's Antarctic Research Center *(ARC)* - are typically based on traditional statistical methods, which are computationally intensive and time consuming. This inefficiency limits the practical scope of predictions, excluding the potential of longer-term forecasts. The use of machine learning models shows promise as an approach due to the relatively improved computational efficiency of these predictive systems. Specifically, the use of Genetic Programming *(GP)* and Evolutionary Learning techniques provide the greatest potential due to the increased explainability these methods provide allowing for greater understanding in the results provided. This is a novel application for these techniques, and so the potential for improvement compared to traditional methods is worth investigation.

3. Goals

The primary goal of this project is to investigate the application of Genetic Programming and Evolutionary Learning techniques on the problem of Antarctic Ice Sheet modelling. This will involve several stages:

1. **An initial Exploratory Data Analysis *(EDA)*** of the data provided by the ARC, to greater understand the the nature of the problem and the data. *(2 weeks)*

   This will involve the use of univariate, bivariate, and multi-variate analysis to identify the distribution, shape, and relationships of the features.

   Variables should also be analysed in terms of their spatial and temporal distributions.

   Finally, potential feature engineering opportunities should be explored to identify potential improvements.

2. **Several iterations of model development** and evaluation using Genetic Programming and Evolutionary Learning techniques, to identify the most effective approach. *(8 weeks)*

   Each model should explore implementing a different approach, before analysing and evaluating the results against previous iterations.

3. **Further development of the most effective approach**, targeting the explainability and interpretability of the predicted outputs. *(4 weeks)*

This will involve analysing and visualising the underlying mathematical expressions evolved by the model, and checking these against current scientific understanding of the problem.

4. **A final evaluation of the most effective model**, to determine the potential for improvement over traditional methods. *(4 weeks)*

   Evaluation should consider many measures, particularly the computational efficiency of the model, the accuracy of the predictions, and the interpretability of the results.

Project milestones should be set for each of these stages to ensure that the project remains on track. It is important to note the potential for these goals to change as the direction of the project becomes clearer.

4. Evaluation

The success of the project will be evaluated on three main criteria:

1. **The accuracy of the the model's predictions.**

2. **The computational efficiency of the model.**

3. **The interpretability of the model's results.**

These criteria can be used as metrics to objectively evaluate the performance of the models developed, and to compare them against traditional methods.

5. Resource Requirements

No external or additional resources are required for the project as all tooling is publicly and freely accessible.