Title: Automated Feature and Hyperparameter Selection for Artificial Neural Networks

Name: Nathan Lamplough

Abstract

In recent times artificial neural networks have surged in popularity due to their plethora of uses. A major problem with them is that they can not be used off the shelf due to multiple hyperparameters that need to be inputted and tuned by the user. To solve this problem a python library called "AutoFH" has been created that will automatically select the optimal hyperparameters and features for artificial neural networks.

Motivation

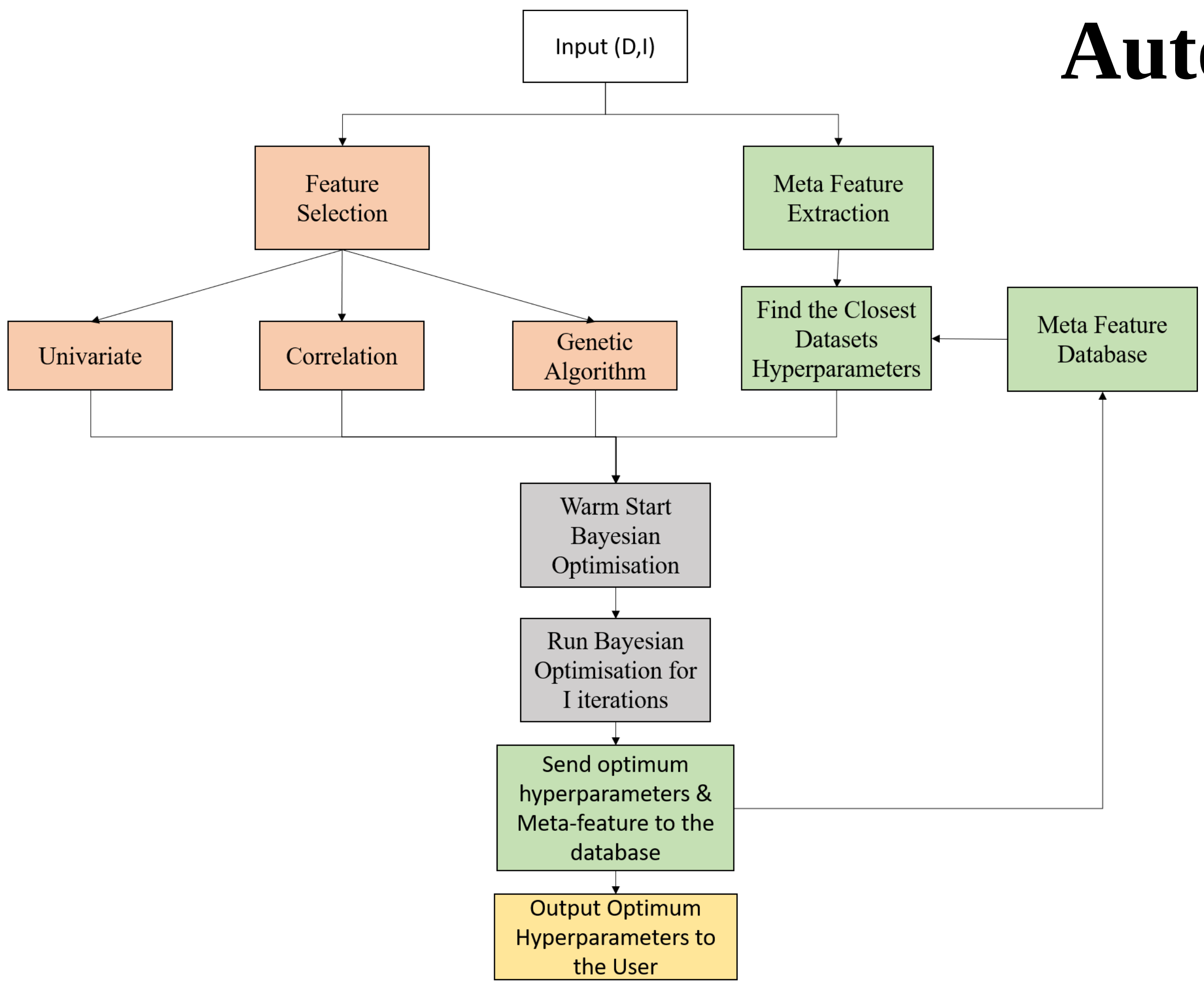
- AutoFH will allow Machine Learning practitioners to save time and focus on other important tasks.
- AutoFH will be a great asset to novices of Machine Learning as it reduces the skill gap needed for using neural networks.

Primary Goals

Aim: Create a python library (AutoFH) that outperforms Bayesian Optimisation & Random Search in finding optimum hyperparameters

Objective: Conduct experiments that illustrate the performance of AutoFH against Bayesian Optimisation and Random Search. Then analyse and draw conclusions from the results.

AutoFH Algorithm



1. A dataset **D** is inputted by the user along with iterations **I** for Bayesian Optimisation runs.
2. Feature selection is performed using 3 different techniques, the feature subsets are stored.
3. Meta features are extracted from **D** and used to find the most similar dataset from the database using Euclidean distance. The closest datasets optimum hyperparameters are then retrieved.
4. Bayesian Optimisation is then warm started using the hyperparameters from step 3. (The feature selection subsets are selected as hyperparameter).
5. Bayesian Optimisation is run for **I** iterations and outputs the optimum hyperparameters. These hyperparameters are stored with **D**s meta-features in the database, with the user's permission.
6. The optimum hyperparameters are returned to the user.

The Hyperparameters

- Epochs
- Hidden Layers
- Nodes
- Learning Rate
- Momentum
- Dropout
- Regularisation
- Regularisation Parameter

Testing & Evaluation

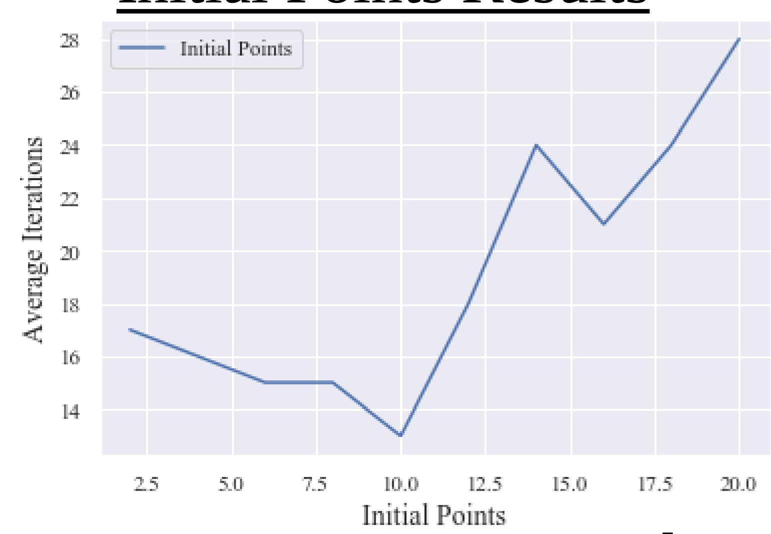
After the first prototype was built experiments/tests were done to optimise the results of AutoFH. The following were tested:

- Bayesian Optimisation Initial Points
- Distance Measure
- Meta Feature Sets
- Method for Choosing the Feature Selection Set

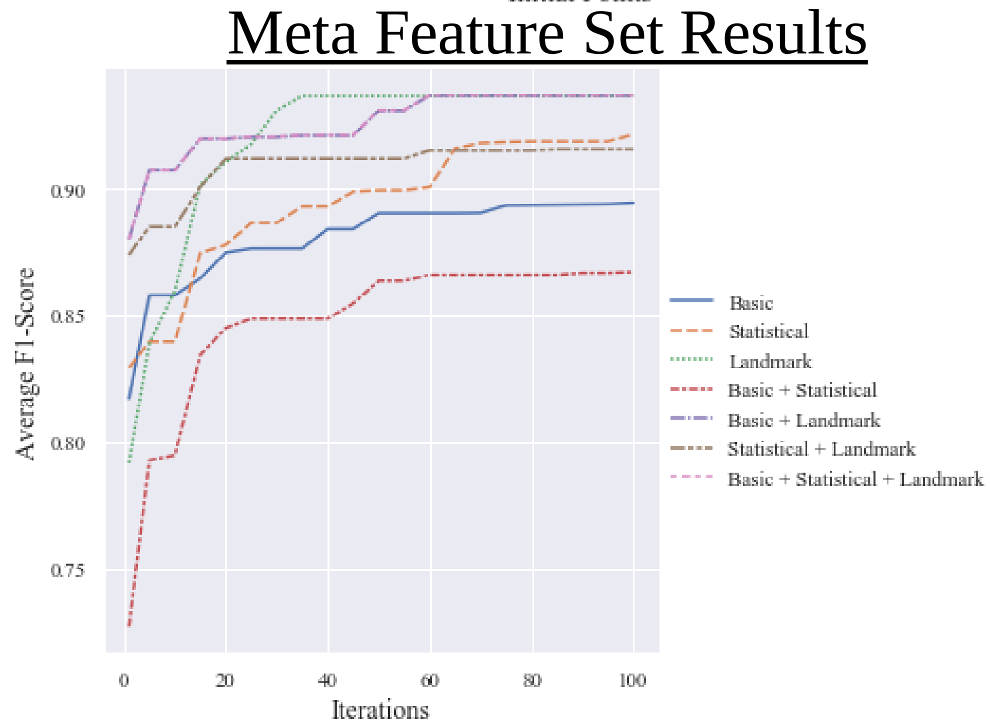
In addition to optimising AutoFH unit and performance testing was done.

- Unit testing was completed throughout development to make debugging and revisions easier.
- Performance testing was completed on a range of different specification computers to ensure AutoFH could be used by a wide range of users.

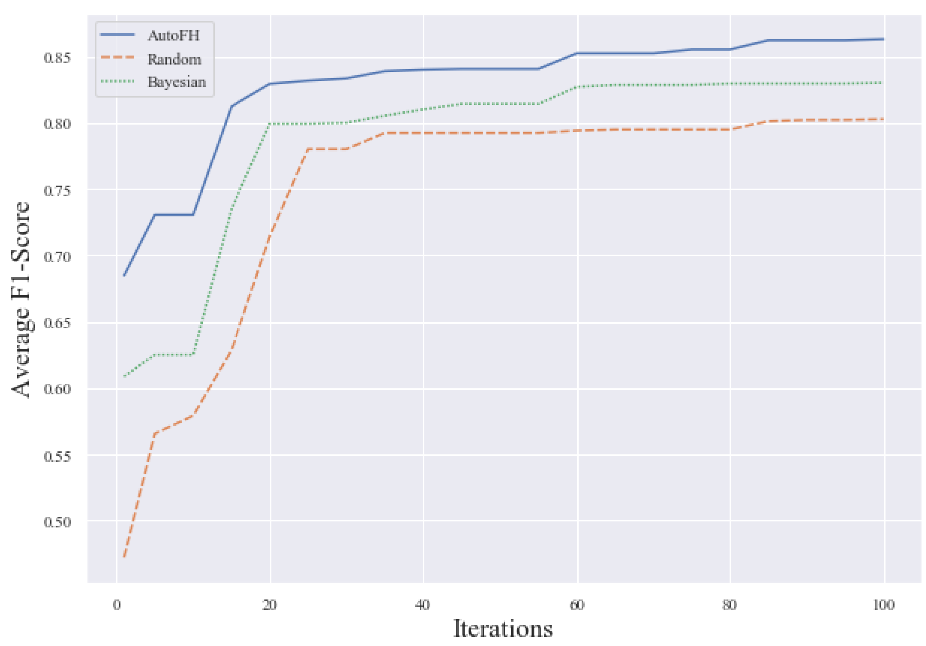
Initial Points Results



Meta Feature Set Results



Final Results & Conclusions



The final evaluation of AutoFH involved comparing its performance against the most popular current methods across 100 iterations and 5 datasets.

The project can be deemed a success as the main aim was achieved. AutoFH performed better than Bayesian Optimisation in final average F1-Score by 0.03 and Random Search by 0.06. AutoFH was also able to reach higher F1-Scores in fewer iterations than the other two methods and find more optimum hyperparameters after 1 iteration due to meta-learning.

Future Work

- User testing
- Add support for different types of neural networks
- Investigate Parallel Random Search