Multi-Objective Bayesian Optimisation

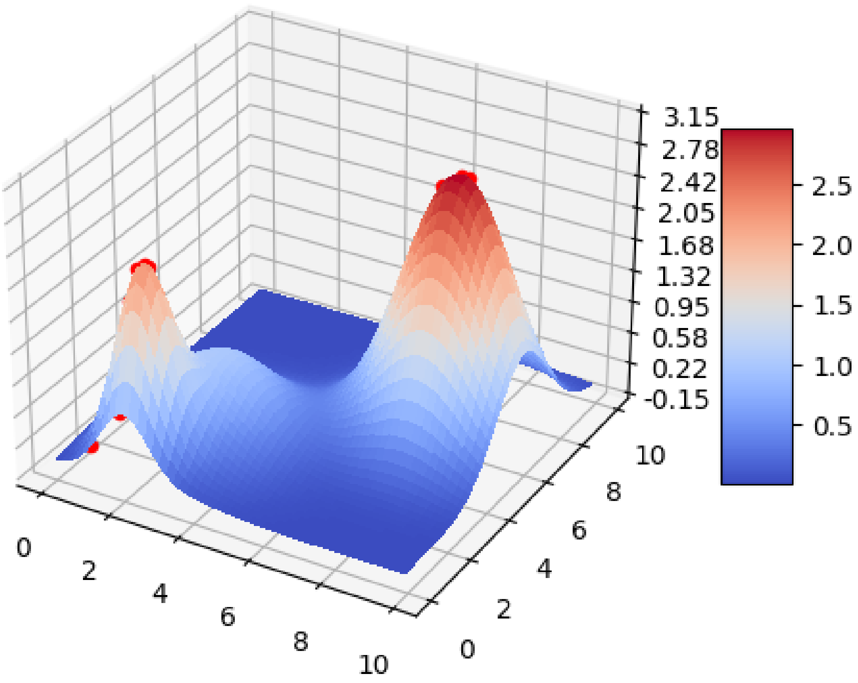
Experimental optimisation using Bayesian Optimisation.

Learning Objectives

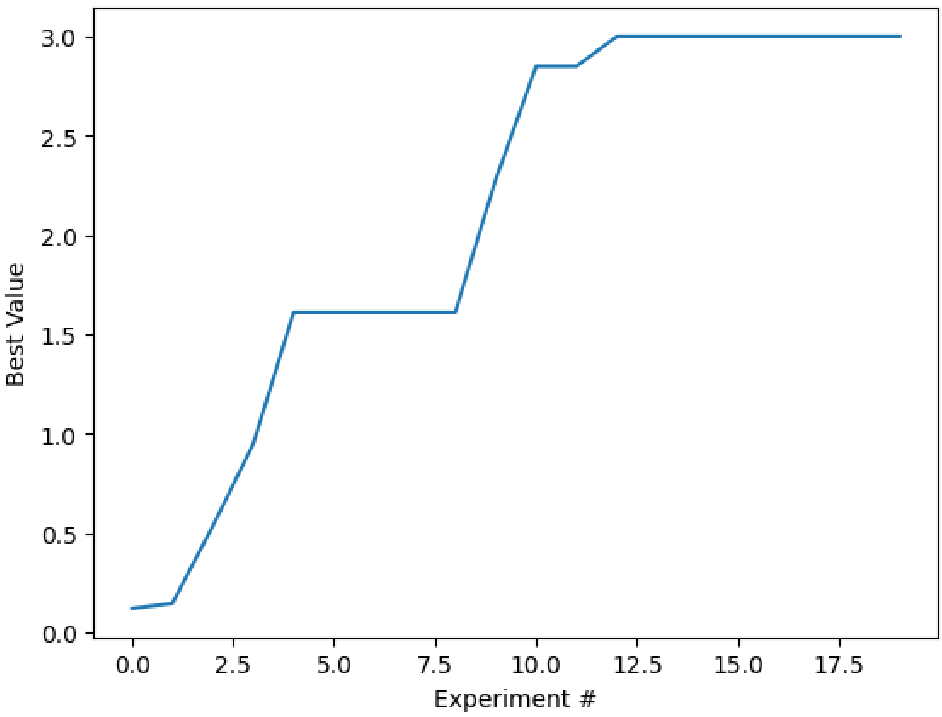
- Bayesian Optimisation
- Multi-Objective Optimisation
- Outcome & Parameter Constraints
- Pareto Frontier
- Parameter Importance
- Data Visualisation
- Hyperparameter Tuning

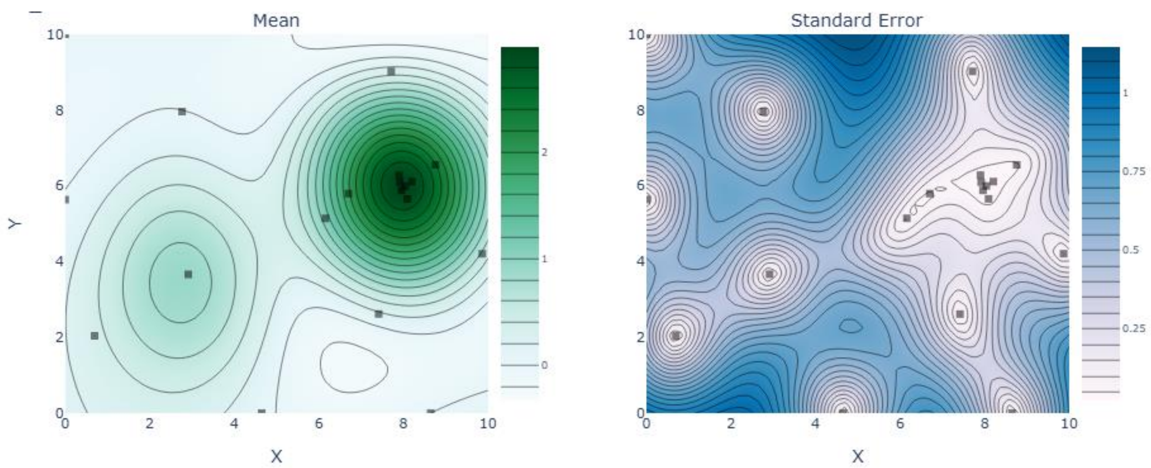
Single-Objective Bayesian Optimisation

2D Gaussian functions were used to create a landscape with varied local maxima. Bayesian optimisation was then used to locate the global maximum. This was achieved after 11 steps. The form of the landscape is shown below. The points sampled by the Bayesian optimisation model have been overlayed in red. The model has sampled points at the edges of the parameter space as well as clusters at the top of the local maxima.



The model first exploited the local maximum it discovered at (1,1) before exploring and finding the global maximum at (8,8).





Optimisation with Constraints

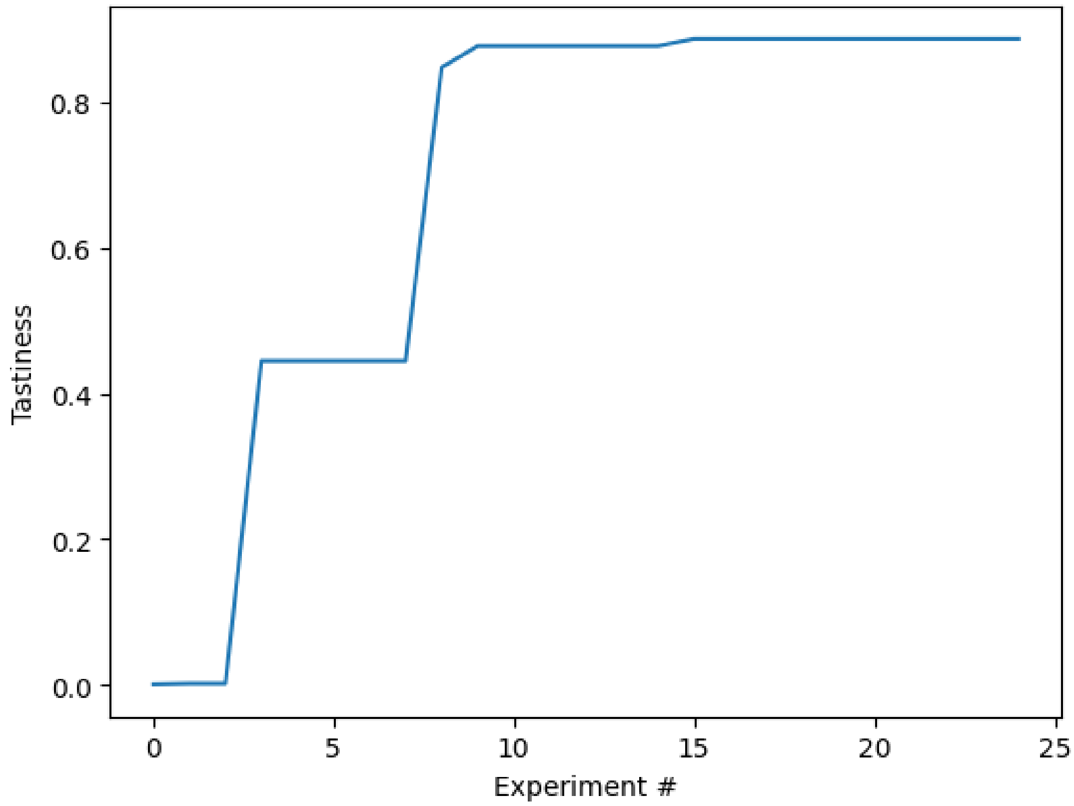
A more complex landscape was modelled around baking cookies using three ingredients: eggs, flour, and sugar. The model took in these ingredients as a ratio, requiring the parameter constraint that the fractional amounts of all the ingredients sum to 1.

The model was tasked with maximising the tastiness of the cookie. This was modelled using a product of Gaussians for each ingredient. The maxima for each ingredient are tabulated below.

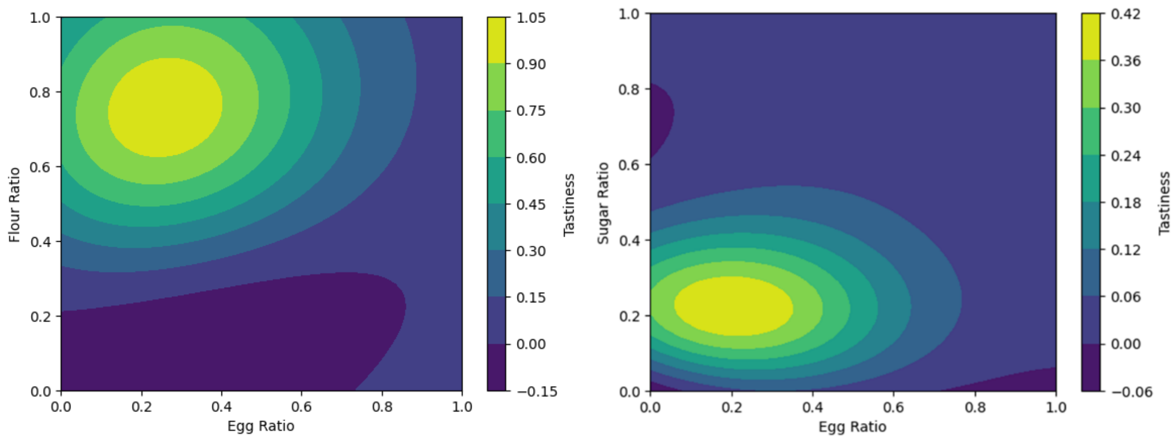
Ingredient	Location of Maximum
Egg Ratio	0.20
Flour Ratio	0.70
Sugar Ratio	0.25

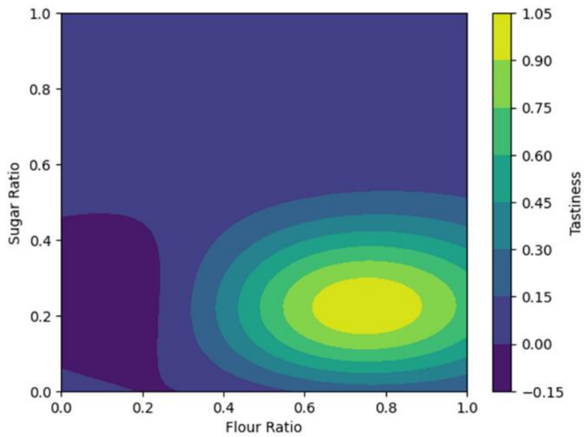
The global maximum thus breaks the parameter constraints and so is not accessible to the model.

A second objective was also defined: crispiness. This was described using a product of two Gaussians centred at Flour Ratio = 0.8 and Sugar Ratio = 0.2. The model was required to achieve a threshold crispiness of at least 0.5, but was not tasked with maximising it.



Contour plots were made from the model's predictions.





The model correctly located the positions of the maxima for each pair of coordinates. However, it predicted tastiness values outside of the mathematical range (between 0 and 1).

The best ratio of ingredients the model identified was 0.148 : 0.655 : 0.197 for eggs : flour : sugar, achieving a tastiness of 0.884 and a crispiness of 0.979 (above the threshold of 0.5).

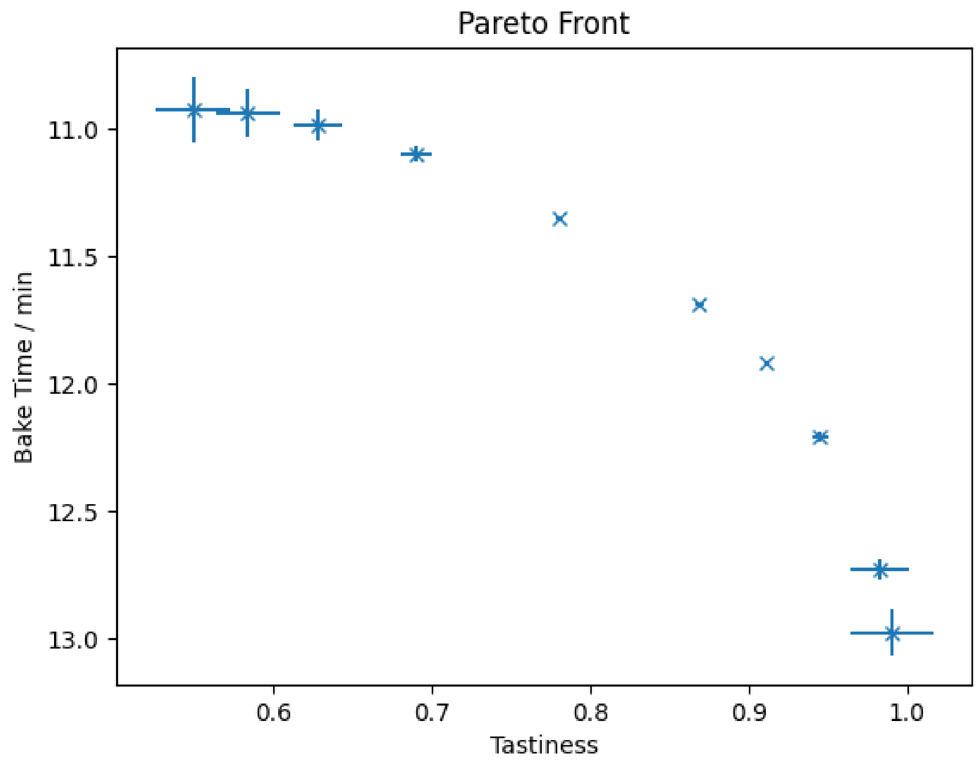
Multi-Objective Bayesian Optimisation

A third objective was added: bake time. The bake time of the cookie (in minutes) was modelled as 10 – 3 x egg ratio + 3 x sugar ratio. The model was then tasked with minimising the bake time, whilst maximising the tastiness. Since tastiness is more important than bake time, it was given 3 times the weighting.

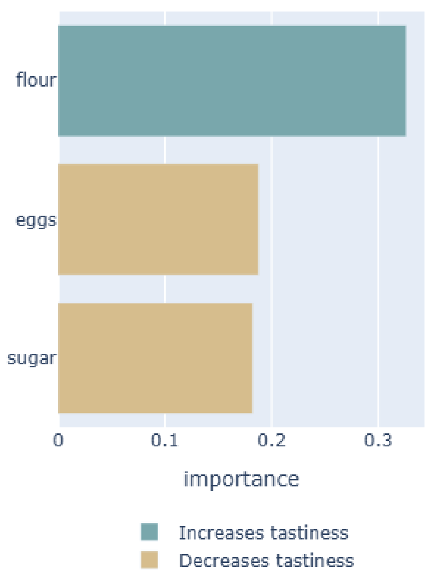
By the way they are defined, tastiness cannot be maximised at the same time as bake time is minimised. Thus, the model will have to make a trade-off, controlled by the relative weights of the objectives.

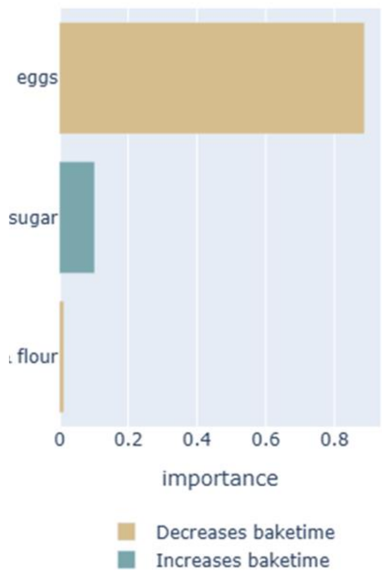
The best ratio of ingredients the model identified was 0.155 : 0.649 : 0.195 for eggs : flour : sugar, achieving a tastiness of 0.886 and a bake time of 12.7 minutes. Clearly, the significantly higher weighting of the tastiness objective led the model to all but disregard the bake time objective.

The Pareto frontier was generated for the bake time and tastiness. For a set of fixed tastiness values, the model minimised the bake time. This allows for visualisation of the trade-off that must be made between the two objectives. Since tastiness is more important than the bake time, parameters should be chosen from the bottom right corner of the Pareto frontier.



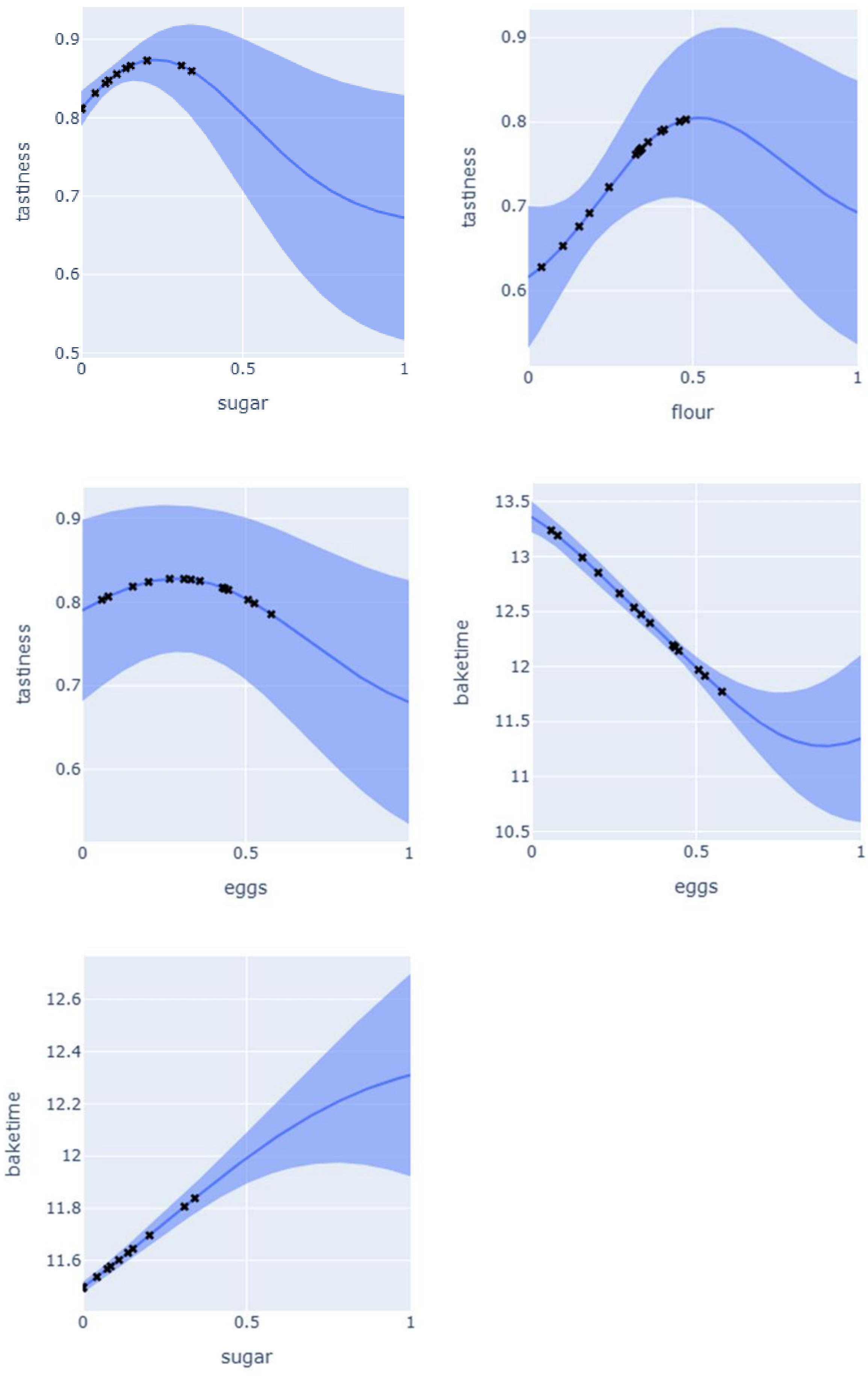
The model was used to extract parameter importance for tastiness and bake time.





The model correctly identified that the flour ratio had little impact on the bake time, but a large flour ratio (~ 0.7) was required for an acceptable tastiness.

Slice plots were also created for each parameter-objective pair. The blue line shows the model's predictions whilst the pale blue region represents the model's uncertainty. The black crosses mark experimental values. It is clear that the model has a higher uncertainty in regions with less experimental data, as would be expected.


tastiness
sugar
0.9
0.8
0.7
0.6
0.5
0
0.5
1
tastiness
flour
0.9
0.8
0.7
0.6
0
0.5
1
tastiness
eggs
0.9
0.8
0.7
0.6
0
0.5
1
baketime
eggs
13.5
13
12.5
12
11.5
11
10.5
0
0.5
1
baketime
sugar
12.6
12.4
12.2
12
11.8
11.6
0
0.5
1

The model expects that the bake time relationships with the egg and sugar ratios are not linear, when in reality they are. This produces the curve seen as the model extrapolates.

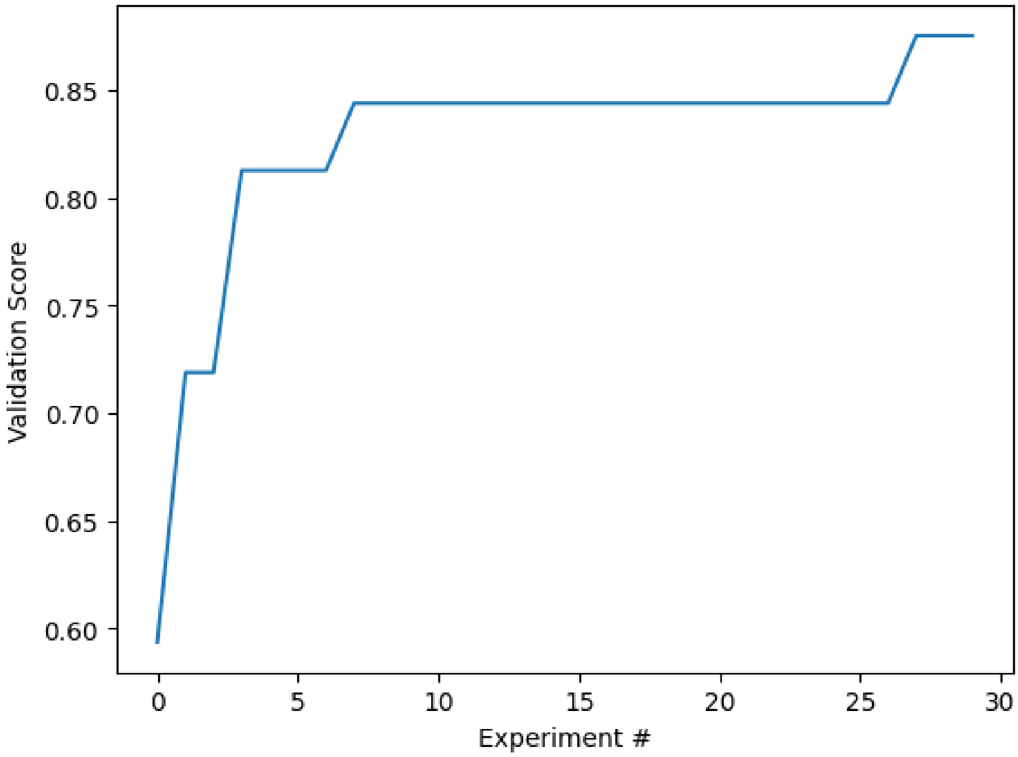
Hyper-Parameter Tuning

A Random Forest Classifier was trained on the fish market dataset. With the default hyperparameters, it achieved a validation score of 0.8125. A grid search with 5-fold cross validation was used to optimise the hyperparameters by testing 320 candidates using the following parameter space:

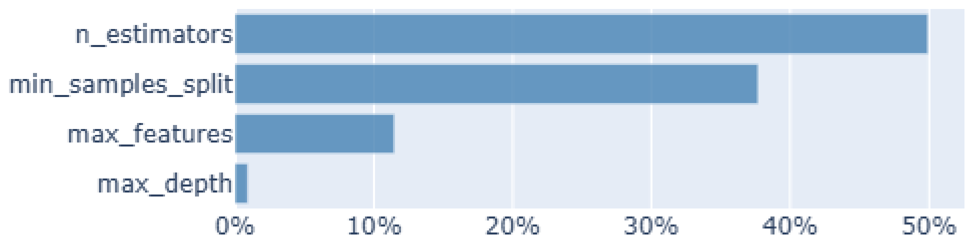
n_estimators: 10, 50, 250, 1000
max_depth: 1, 2, 4, 8, 16
min_samples_split: 2, 4, 8, 16
max_features: 1, 2, 4, 8

The best parameters that were tested were n_estimators: 10, min_samples_split: 4, max_features: 8, and max_depth: 16, achieving a validation score of 0.8438.

Next, Bayesian optimisation was employed to optimise the hyperparameters. The best parameters that the model identified were n_estimators: 213, min_samples_split: 2, max_features: 6, and max_depth: 14, achieving a validation score of 0.875. These parameters were identified in fewer than 30 trials.



Parameter importance was extracted from the Bayesian optimisation model.



It is clear that the max_depth hyperparameter has little impact on the performance of the model. Removing this from the grid search would greatly reduce the number of candidates to test.

Bayesian optimisation performs significantly better at hyperparameter tuning than a full grid search, achieving a higher validation score after a much smaller number of trials. It also has access to a much larger parameter space as it is not restricted to grid points but can instead sample anywhere in a permitted range.

Conclusion

Bayesian optimisation is a powerful tool for single and multi-objective optimisation tasks. It is able to efficiently explore and model the sample space, balancing the exploitation of local optima and the exploration of new regions. Useful insights can also be drawn from the model, such as the trade-offs inherent in multi-objective optimisation and parameter importance for different objectives.