

Output from the logistic_regression.py code

The logistic_regression.py code shared on (https://github.com/statisticspoland/ecoicop_classification/blob/master/Logistic_Regression/logistic_regression.py) runs ML models on different values of the following parameters: C, fit_intercept, class_weight, solver and multi_class. The parameter max_iter is set at 200.

```
for c in [0.1, 1, 2, 3]:
    for fit_intercept in [True, False]:
        for class_weight in [None, 'balanced']:
            for solver in ["newton-cg", "lbfgs", "liblinear", "sag", "saga"]:
                for multi_class in ["ovr", "multinomial"]:
```

For each combination, the code calculates the accuracy on the model predictions from the validation subset (all predictions combined). It also calculates the accuracy from the training dataset and the F1 score on the validation subset (which is equal to the accuracy because it is calculated on all predictions).

The code outputs the following Excel file: [results_logistic_regression.xlsx](#). Below are the first few records.

	C	fit_intercept	class_weight	solver	multi_class	max_iter	val_accuracy	train_accuracy	f1_score_micro
0	0.1	TRUE		newton-cg	ovr	200	0.8200	0.8566	0.8200
1	0.1	TRUE		newton-cg	multinomial	200	0.8253	0.8667	0.8253
2	0.1	TRUE		lbfgs	ovr	200	0.8200	0.8566	0.8200
3	0.1	TRUE		lbfgs	multinomial	200	0.8253	0.8667	0.8253
4	0.1	TRUE		liblinear	ovr	200	0.8327	0.8679	0.8327
5	0.1	TRUE		sag	ovr	200	0.8200	0.8566	0.8200