Machine Learning in Computational Biology

Assignment 1

First Bonus Question — Using Optuna

Konstantinos Konstantinidis
Student number: 7115152400017

April 6, 2025

Repo: The repository for this assignment can be found here:
`https://github.com/KonsKons26/Assignment-1`

Contents

1 Migration to Optuna

Since the `Regressor` object I had created to perform all regression tasks was highly modularized, with specific methods for each task, moving from `sklearn`'s `GridSearchCV` to `Optuna` was quite simple. I created a new class named `RegressorOptuna` which inherited from `Regressor`, to keep it's basic functionalities the same. Then I modified the private methods handling each model's hyperparameter optimization to allow me to pass the appropriate grid spaces –and in the appropriate format– for `Optuna` to tune the models.

The objective function **minimizes** the RMSE of the test set. The sampler I chose is the `TPESampler`, a tree-structured Parzen estimator, which fits one Gaussian Mixture Model (GMM) (lx) to the set of parameters associated with the best objective values and another GMM ($g(x)$) to the remaining parameter values. It chooses the parameter that minimizes the ration $l(x)/g(x)$.

All files are in the appropriate directories, `src/` and `notebooks/` and the models were saved in `models/bonus1_optuna/`. Since I used the features I selected in the previous task, to keep everything organized, the process includes copying the features files in the `models/bonus1_optuna` directory.

2 Hyperparameter tuning

I set the number of trials for each model tuning to 1000, which seemed appropriate, as it was not such a demanding number for my machine, and since the hyperparameter space has so many dimension, I believe that a small number of trials might not be enough to search the whole space. The complete hyperparameter space, along with the values chosen by `Optuna` are shown in Table 1.

Model	Parameter	Values	Picked value
ElasticNet	α	(0.01, 1.0)	0.11850
	`l1 ratio`	(0.0, 1.0)	0.77838
	`tolerance`	[1e-3, 1e-4, 1e-5, 1e-6, 1e-7]	0.00100
SVR	`kernel`	['rbf', 'linear', 'poly', 'sigmoid']	'rbf'
	`degree`	(2, 5)	4
	γ	['scale', 'auto']	'scale'
	`coef_0`	(0.0, 1)	0.93453
	`tolerance`	[1e-3, 1e-4, 1e-5, 1e-6, 1e-7]	1e-7
	`C`	(0.1, 10)	1.63421
	ϵ	(0.0, 10.0)	0.92788
BayesianRidge	`tolerance`	[1e-3, 1e-4, 1e-5, 1e-6, 1e-7]	1e-7
	α_1	(1e-9, 1e-3)	0.0
	α_2	(1e-9, 1e-3)	0.00065
	λ_1	(1e-3, 1e-9)	0.0009999
	λ_2	(1e-3, 1e-9)	0.0
	`compute_score`	[True, False]	True

Table 1: Hyperparameter spaces for each model.

3 Results