Machine Learning Project

“Calories Estimation Problem”

2022/2023

Zaza Maria Elena

Perna Grazia

Index

1. Introduction
   1. Motivation
2. Dataset
3. Pre-processing
   1. StandardScaler
   2. MinMaxScaler
   3. RobustScaler
4. Model
   1. Ensemble
   2. Extra Trees Regressor
5. Workflow
6. Results
   1. Metrics
7. Conclusion
8. Bibliography

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| MODEL | SCALER | MSE | R2 | MAE |
| Random Forest | Standard Scaler | 9.888113665273435 | 0.9999306410590445 | 0.807497810918827 |
| Random Forest | Robust Scaler | 10.18623634899898 | 0.9999285499146343 | 0.827860155792307 |
| Random Forest | MinMax Scaler | 11.009184409701504 | 0.999922777448026 | 0.8177112057257595 |
| AdaBoost Regressor | Standard Scaler | 10656.164396994798 | 0.9252536629085176 | 88.69877624899232 |
| AdaBoost Regressor | Robust Scaler | 10518.728963153746 | 0.9262176866307017 | 88.67405675478258 |
| AdaBoost Regressor | MinMax Scaler | 10086.200386862425 | 0.9292516044233259 | 88.05302717313046 |
| Isolation Forest | Standard Scaler | 4667681.827301569 | -31.74087244728156 | 2127.2349117842878 |
| Isolation Forest | Robust Scaler | 4667726.0914022345 | -31.74118293229907 | 2127.2349117842878 |
| Isolation Forest | MinMax Scaler | 4667952.591891982 | -31.742771691755785 | 2127.290837217044 |
| Extra Trees Regressor | Standard Scaler | 6.229307878644889 | 0.99995630529624 | 0.73653881025302 |
| Extra Trees Regressor | Robust Scaler | 7.753100697036747 | 0.9999456168414248 | 0.8062014480693446 |
| Extra Trees Regressor | MinMax Scaler | 6.689025658728041 | 0.999953080663166 | 0.781645593209158 |
| Gradient Boosting Regressor | Standard Scaler | 420.13475489540315 | 0.9970530171229232 | 15.606831269704083 |
| Gradient Boosting Regressor | Robust Scaler | 419.62835611565333 | 0.9970565691940516 | 15.600160260634157 |
| Gradient Boosting Regressor | MinMax Scaler | 420.10608800028814 | 0.9970532182032863 | 15.607906665917982 |
| Bagging Regressor  (estimator=ExtraTreesRegressor) | MinMax Scaler | 7.7496598063654005 | 0.9999456409771236 | 0.9191475494674172 |

Migliori iperparametri: {'max\_depth': 35, 'min\_samples\_leaf': 1, 'min\_samples\_split': 2, 'n\_estimators': 720} CV=10

Migliori iperparametri: {'max\_depth': 35, 'min\_samples\_leaf': 1, 'min\_samples\_split': 2, 'n\_estimators': 500}

CV=5

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| MODEL | MSE | MAE | R2 | CV |
| Extra Trees Regressor | 4.3975269234300844 | 0.6118285538220723 | 0.9999719484688838 | 10 |
| Extra Trees Regressor | 4.724539953145269 | 0.6418285538220723 | 0.9999698624745652 | 5 |