```
[2]  ✓ 4m 57.5s                                                    Python
```

```
/tmp/ipykernel_522656/1254506153.py:30: DeprecationWarning: `ensemble_size` has been deprecated, please use `ensemble_kwargs = {'ensemble_size': 1}`. Inserting `ensemble_size` into `e
  automl = autosklearn.classification.AutoSklearnClassifier(
Models Trained:
 {8: {'model_id': 8, 'rank': 1, 'cost': 0.14166666666666672, 'ensemble_weight': 1.0, 'data_preprocessor': <autosklearn.pipeline.components.data_preprocessing.DataPreprocessorChoice ob
    gamma=0.028106748647672205, max_iter=-1.0, random_state=42, shrinking=False,
    tol=2.0706197108771777e-05)}}

Best Model Selected:
 (1.0, SimpleClassificationPipeline({'balancing:strategy': 'none', 'classifier:__choice__': 'libsvm_svc', 'data_preprocessor:__choice__': 'feature_type', 'feature_preprocessor:__choic
dataset_properties={
  'task': 2,
  'sparse': False,
  'multilabel': False,
  'multiclass': True,
  'target_type': 'classification',
  'signed': False}))

Best Model Validation Accuracy: 0.8440
```

```
# Comparison the different models tried using auto-sklearn, by default it uses 10-fold cross-validation and it always returns the best model, so we can use the cv_results_ attribut
automl.cv_results_
```

```
[3]  ✓ 0.0s                                                        Python
```

```
  'classifier:libsvm_svc:tol': 2.0706197108771777e-05,
  'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
  'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'standardize',
  'feature_preprocessor:select_rates_classification:alpha': 0.1,
  'feature_preprocessor:select_rates_classification:score_func': 'chi2',
  'classifier:libsvm_svc:coef0': -0.2949567859897677,
  'classifier:libsvm_svc:degree': 5,
  'feature_preprocessor:select_rates_classification:mode': 'fpr'},
```

```
print(f"Best model validation accuracy: {accuracy:.4f}")
```

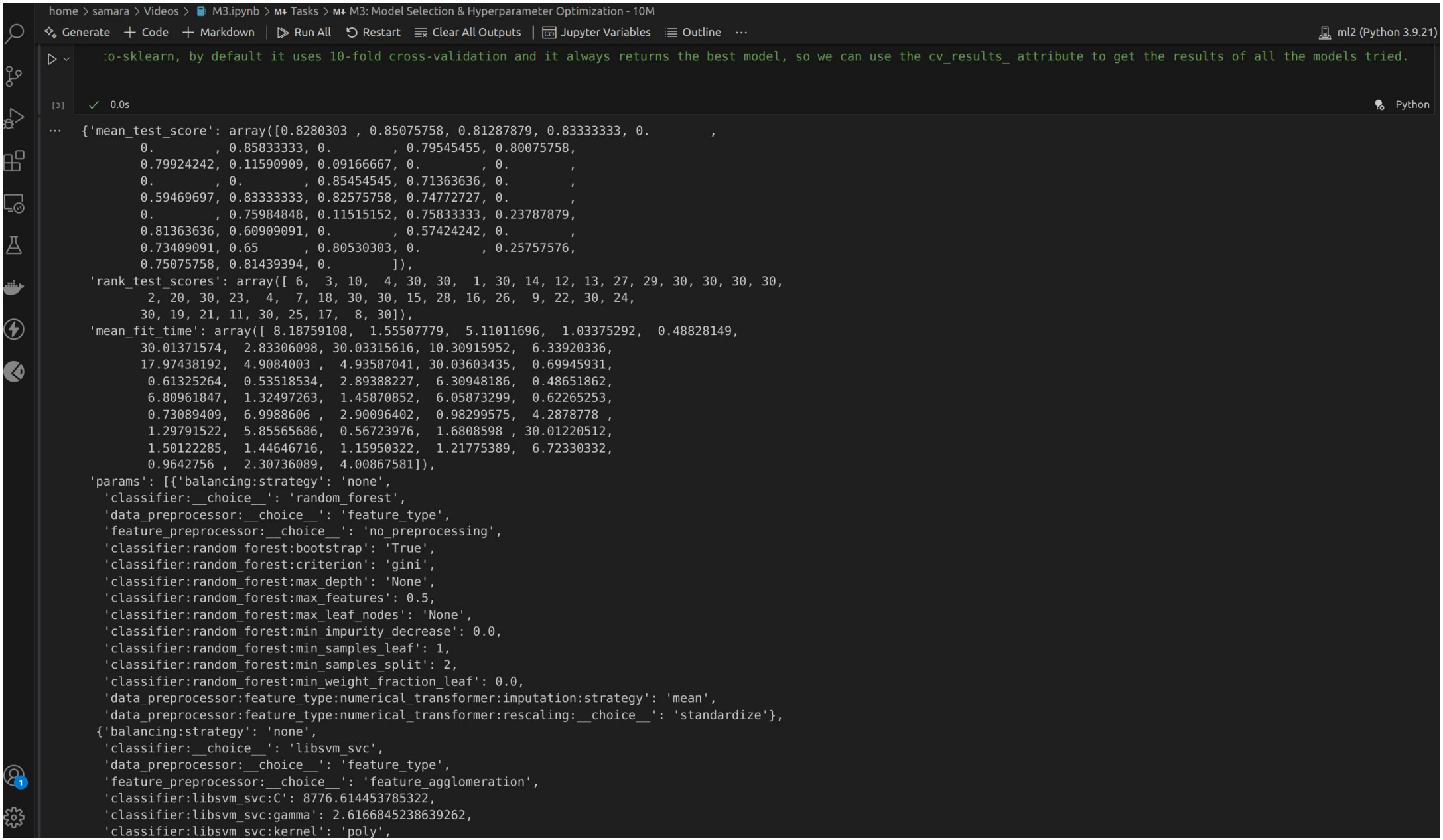
[2] ✓ 4m 57.5s Python

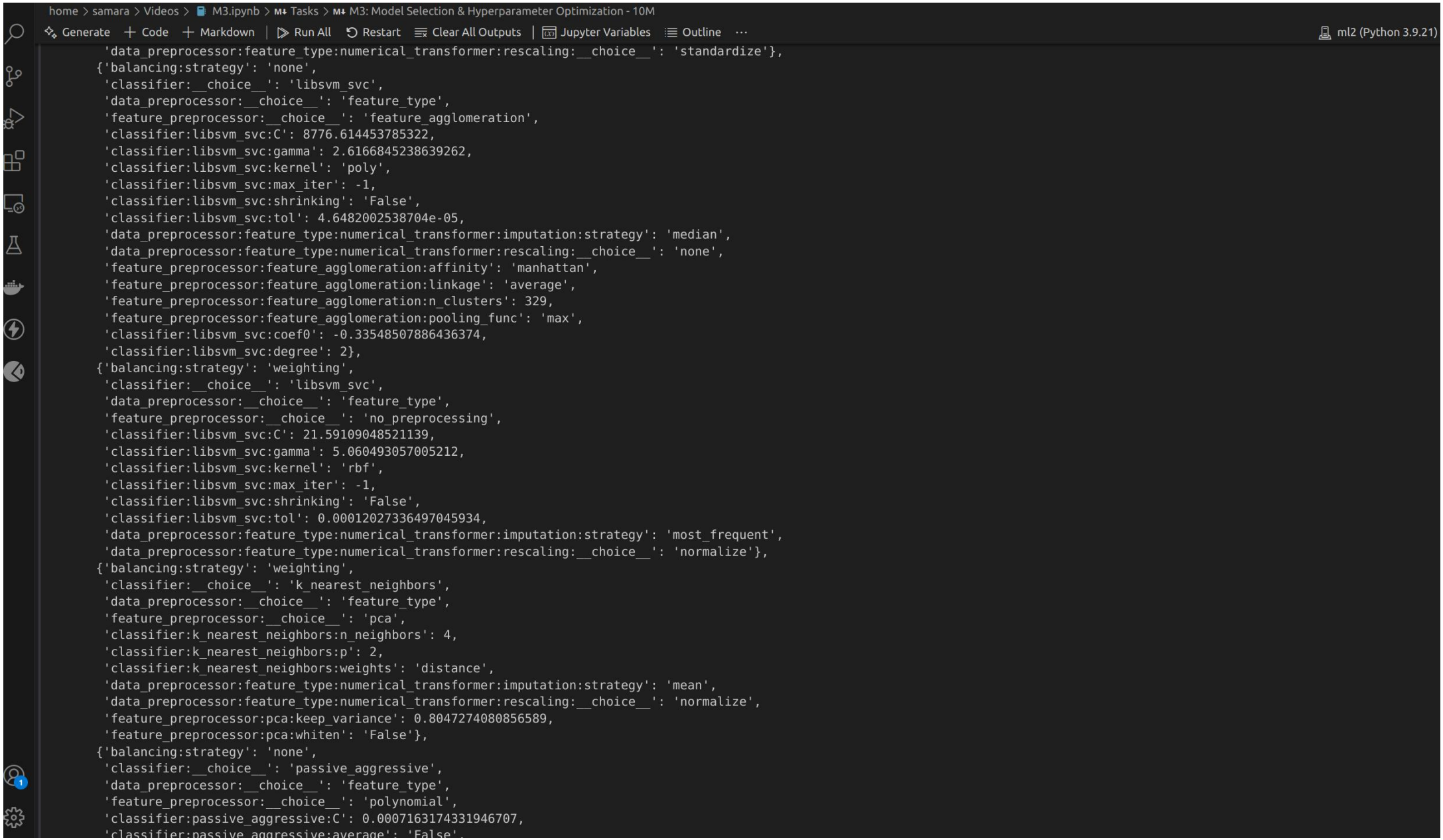
```
le_size` into `ensemble_kwargs` for now. `ensemble_size` will be removed in auto-sklearn 0.16.

ocessorChoice object at 0x704def2de430>, 'balancing': Balancing(random_state=42), 'feature_preprocessor': <autosklearn.pipeline.components.feature_preprocessing.FeaturePreprocessorChoi



rocessor:__choice__': 'no_preprocessing', 'classifier:libsvm_svc:C': 12923.91167672086, 'classifier:libsvm_svc:gamma': 0.028106748647672205, 'classifier:libsvm_svc:kernel': 'rbf', 'cla
```

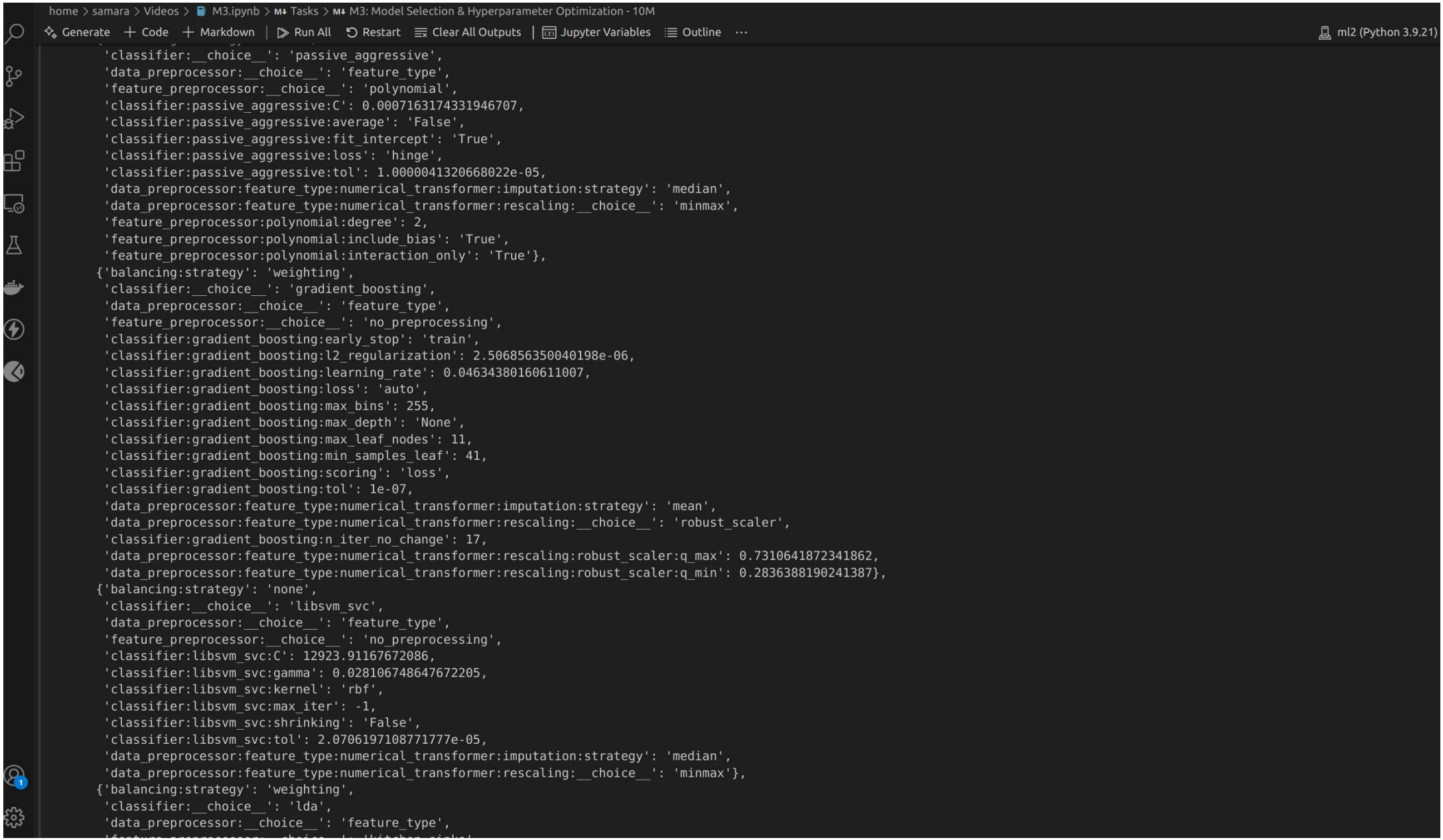
```
# Comparison the different models tried using auto-sklearn, by default it uses 10-fold cross-validation and it always returns the best model, so we can use the cv_results_ attribut
automl.cv_results_
```

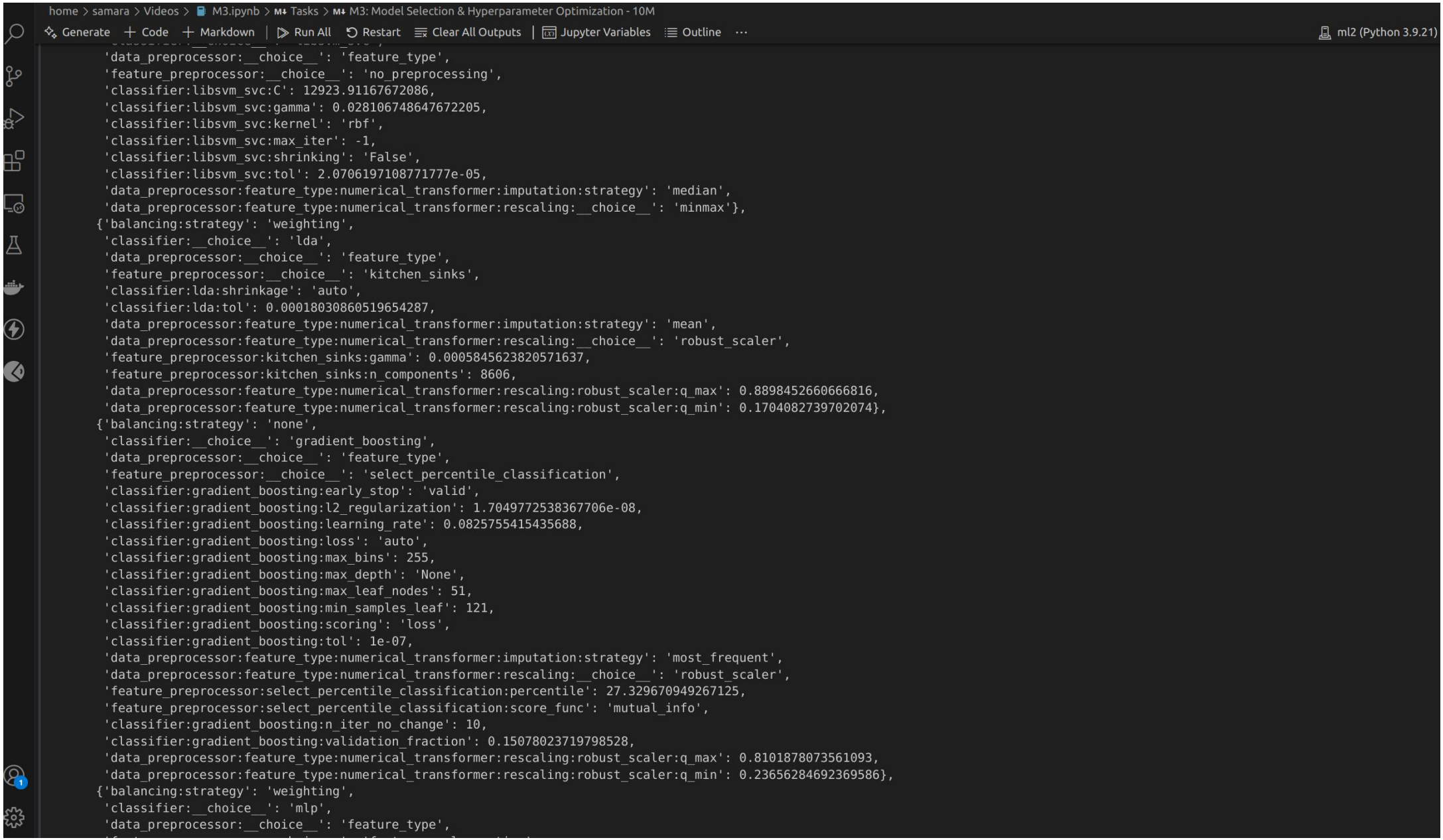
[3] ✓ 0.0s Python

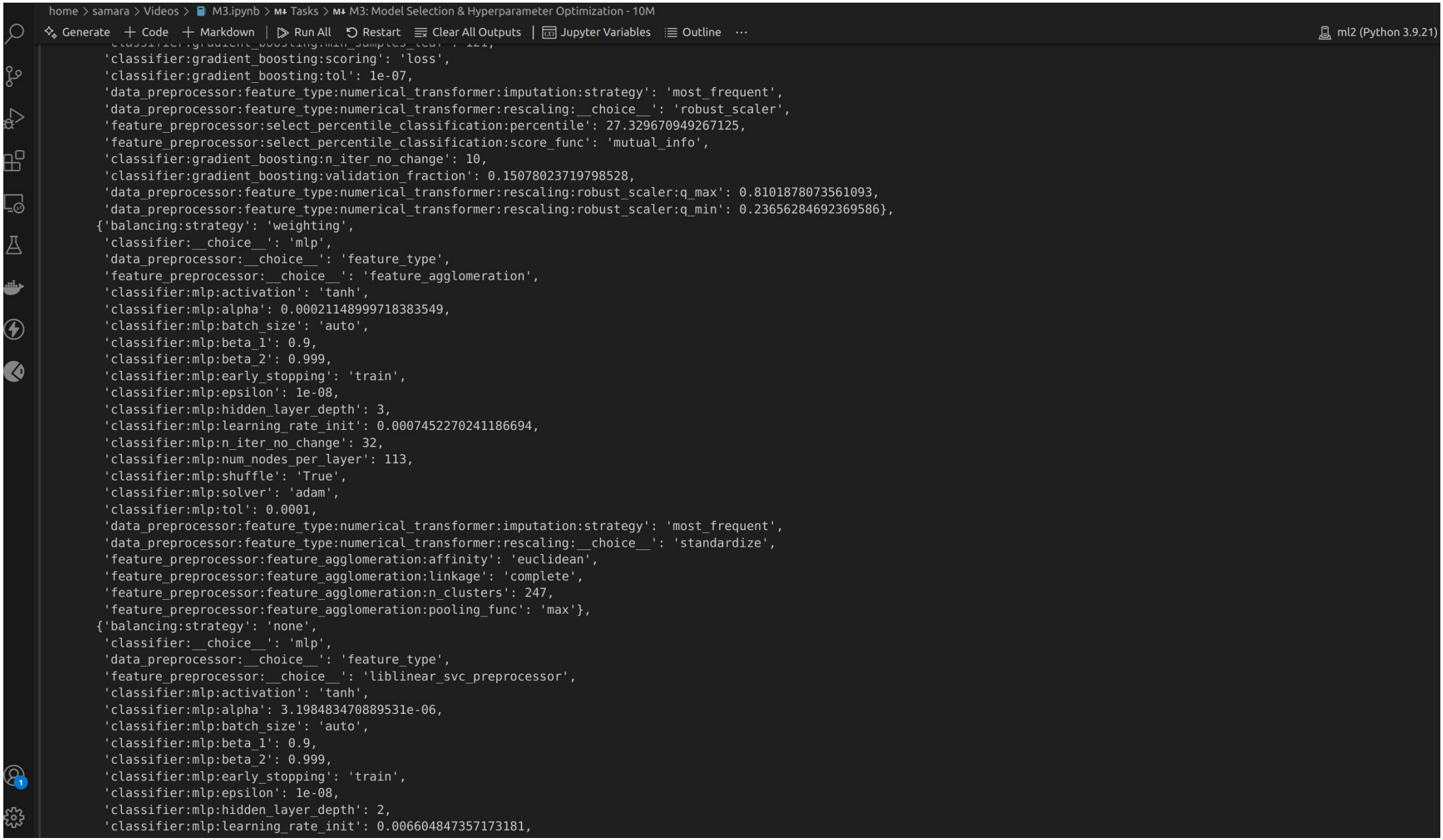
```
'classifier:libsvm_svc:tol': 2.0706197108771777e-05,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'standardize',
'feature_preprocessor:select_rates_classification:alpha': 0.1,
'feature_preprocessor:select_rates_classification:score_func': 'chi2',
'classifier:libsvm_svc:coef0': -0.2949567859897677,
'classifier:libsvm_svc:degree': 5,
'feature_preprocessor:select_rates_classification:mode': 'fpr'},
```

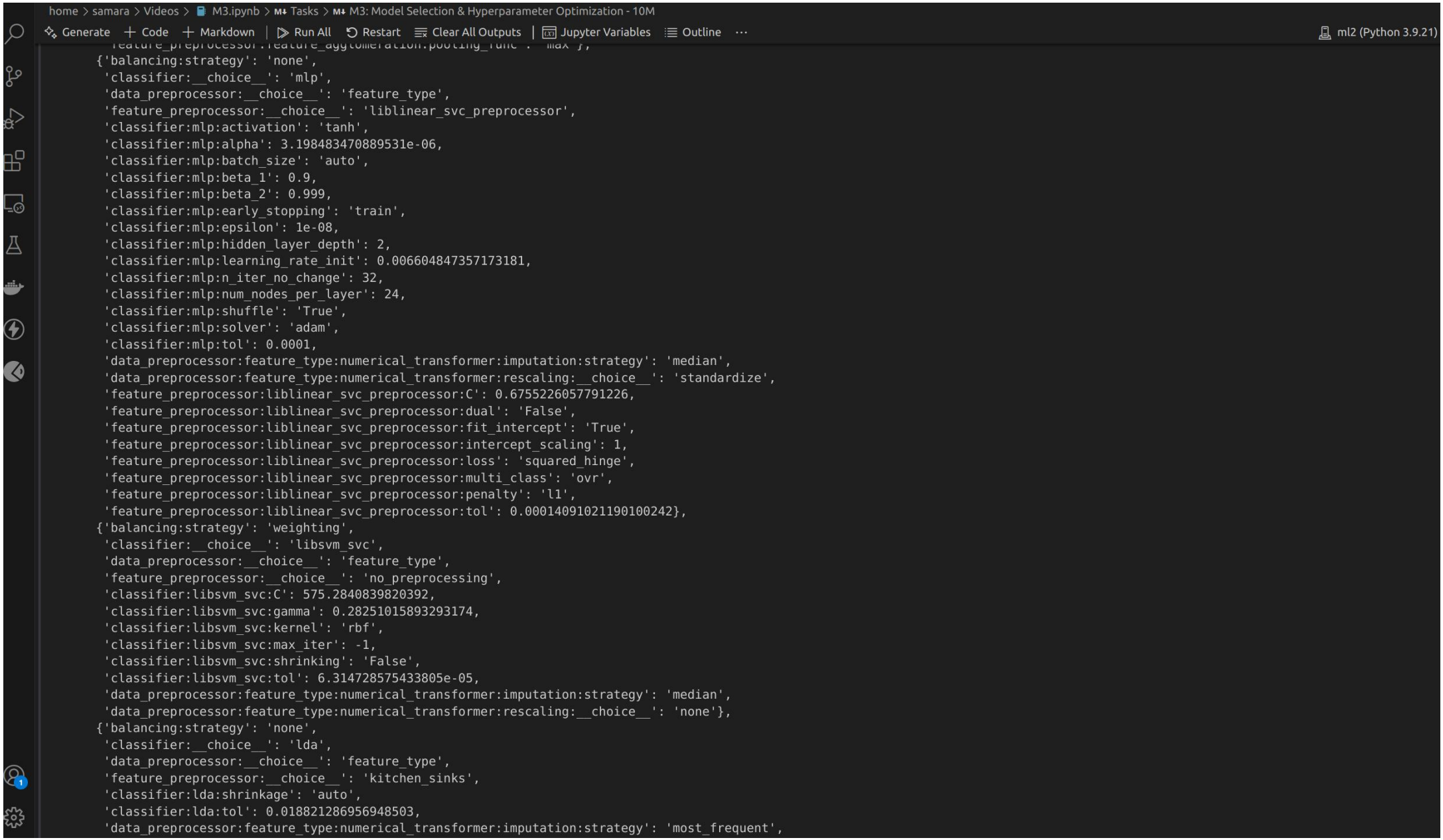

home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
:o-sklearn, by default it uses 10-fold cross-validation and it always returns the best model, so we can use the cv_results_ attribute to get the results of all the models tried.
[3]
0.0s
Python
{'mean_test_score': array([0.8280303 , 0.85075758, 0.81287879, 0.83333333, 0.
0. , 0.85833333, 0. , 0.79545455, 0.80075758,
0.79924242, 0.11590909, 0.09166667, 0. , 0. ,
0. , 0. , 0.85454545, 0.71363636, 0. ,
0.59469697, 0.83333333, 0.82575758, 0.74772727, 0. ,
0. , 0.75984848, 0.11515152, 0.75833333, 0.23787879,
0.81363636, 0.60909091, 0. , 0.57424242, 0. ,
0.73409091, 0.65 , 0.80530303, 0. , 0.25757576,
0.75075758, 0.81439394, 0.]),
'rank_test_scores': array([6, 3, 10, 4, 30, 30, 1, 30, 14, 12, 13, 27, 29, 30, 30, 30, 30,
2, 20, 30, 23, 4, 7, 18, 30, 30, 15, 28, 16, 26, 9, 22, 30, 24,
30, 19, 21, 11, 30, 25, 17, 8, 30]),
'mean_fit_time': array([8.18759108, 1.55507779, 5.11011696, 1.03375292, 0.48828149,
30.01371574, 2.83306098, 30.03315616, 10.30915952, 6.33920336,
17.97438192, 4.9084003 , 4.93587041, 30.03603435, 0.69945931,
0.61325264, 0.53518534, 2.89388227, 6.30948186, 0.48651862,
6.80961847, 1.32497263, 1.45870852, 6.05873299, 0.62265253,
0.73089409, 6.9988606 , 2.90096402, 0.98299575, 4.2878778 ,
1.29791522, 5.85565686, 0.56723976, 1.6808598 , 30.01220512,
1.50122285, 1.44646716, 1.15950322, 1.21775389, 6.72330332,
0.9642756 , 2.30736089, 4.00867581]),
'params': [{'balancing:strategy': 'none',
'classifier:__choice__': 'random_forest',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'no_preprocessing',
'classifier:random_forest:bootstrap': 'True',
'classifier:random_forest:criterion': 'gini',
'classifier:random_forest:max_depth': 'None',
'classifier:random_forest:max_features': 0.5,
'classifier:random_forest:max_leaf_nodes': 'None',
'classifier:random_forest:min_impurity_decrease': 0.0,
'classifier:random_forest:min_samples_leaf': 1,
'classifier:random_forest:min_samples_split': 2,
'classifier:random_forest:min_weight_fraction_leaf': 0.0,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'standardize'},
{'balancing:strategy': 'none',
'classifier:__choice__': 'libsvm_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'feature_agglomeration',
'classifier:libsvm_svc:C': 8776.614453785322,
'classifier:libsvm_svc:gamma': 2.6166845238639262,
'classifier:libsvm_svc:kernel': 'poly',

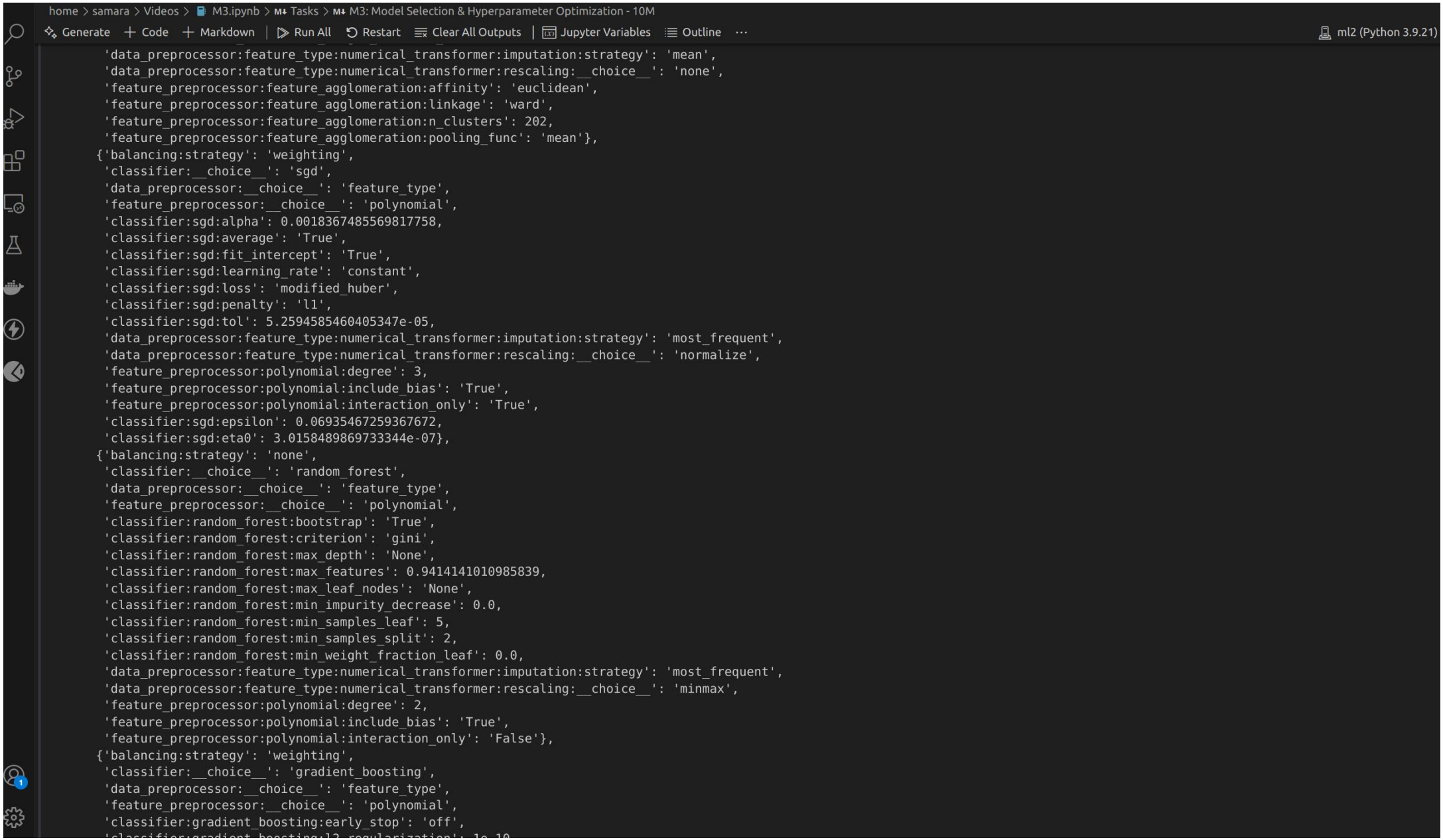

home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'standardize'},
{'balancing:strategy': 'none',
'classifier:__choice__': 'libsvm_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'feature_agglomeration',
'classifier:libsvm_svc:C': 8776.614453785322,
'classifier:libsvm_svc:gamma': 2.6166845238639262,
'classifier:libsvm_svc:kernel': 'poly',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'False',
'classifier:libsvm_svc:tol': 4.6482002538704e-05,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'none',
'feature_preprocessor:feature_agglomeration:affinity': 'manhattan',
'feature_preprocessor:feature_agglomeration:linkage': 'average',
'feature_preprocessor:feature_agglomeration:n_clusters': 329,
'feature_preprocessor:feature_agglomeration:pooling_func': 'max',
'classifier:libsvm_svc:coef0': -0.33548507886436374,
'classifier:libsvm_svc:degree': 2},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'libsvm_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'no_preprocessing',
'classifier:libsvm_svc:C': 21.59109048521139,
'classifier:libsvm_svc:gamma': 5.060493057005212,
'classifier:libsvm_svc:kernel': 'rbf',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'False',
'classifier:libsvm_svc:tol': 0.00012027336497045934,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'normalize'},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'k_nearest_neighbors',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'pca',
'classifier:k_nearest_neighbors:n_neighbors': 4,
'classifier:k_nearest_neighbors:p': 2,
'classifier:k_nearest_neighbors:weights': 'distance',
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'normalize',
'feature_preprocessor:pca:keep_variance': 0.8047274080856589,
'feature_preprocessor:pca:whiten': 'False'},
{'balancing:strategy': 'none',
'classifier:__choice__': 'passive_aggressive',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'polynomial',
'classifier:passive_aggressive:C': 0.0007163174331946707,

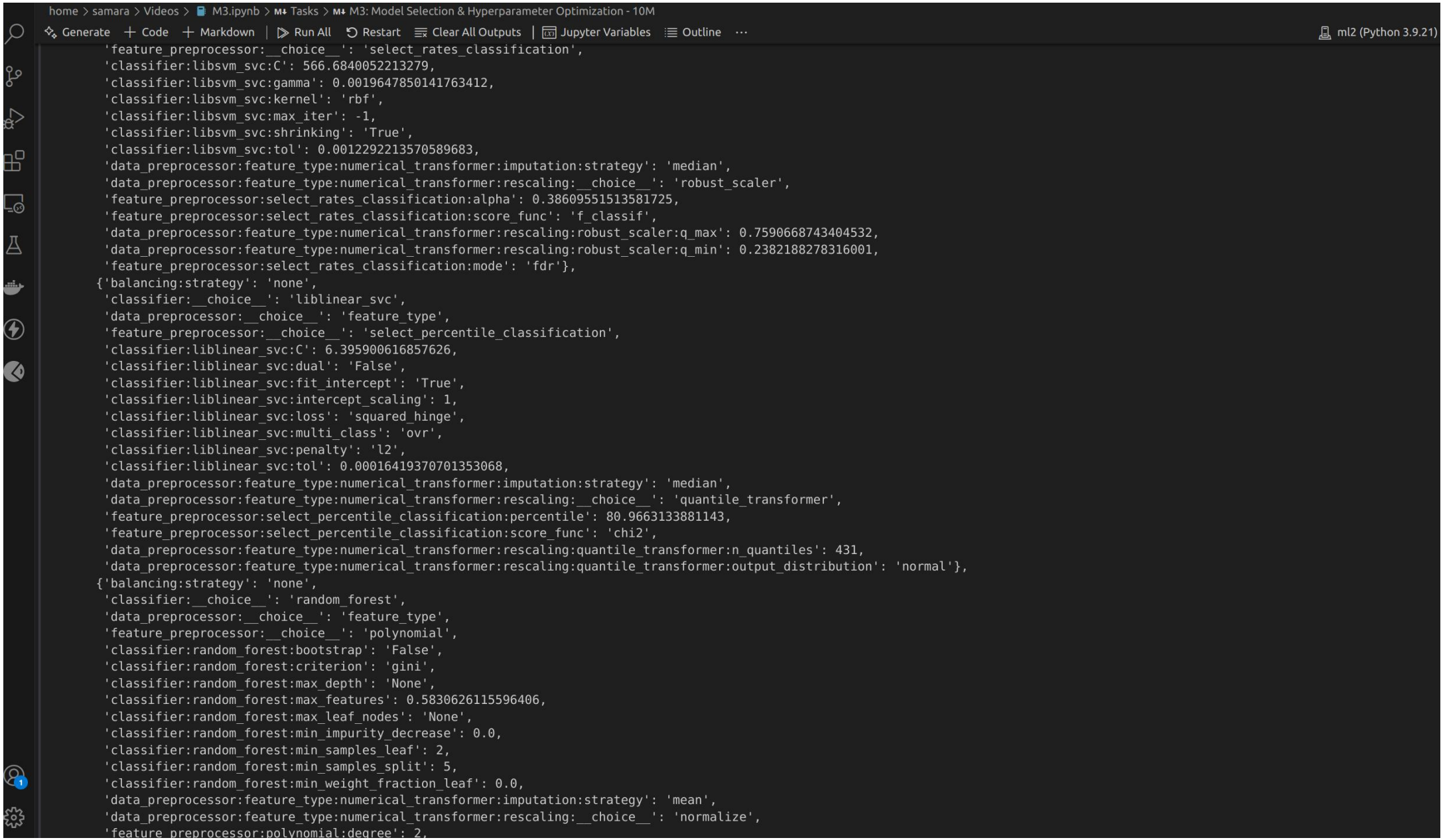

home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
'classifier:__choice__': 'passive_aggressive',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'polynomial',
'classifier:passive_aggressive:C': 0.0007163174331946707,
'classifier:passive_aggressive:average': 'False',
'classifier:passive_aggressive:fit_intercept': 'True',
'classifier:passive_aggressive:loss': 'hinge',
'classifier:passive_aggressive:tol': 1.0000041320668022e-05,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'minmax',
'feature_preprocessor:polynomial:degree': 2,
'feature_preprocessor:polynomial:include_bias': 'True',
'feature_preprocessor:polynomial:interaction_only': 'True'},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'gradient_boosting',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'no_preprocessing',
'classifier:gradient_boosting:early_stop': 'train',
'classifier:gradient_boosting:l2_regularization': 2.506856350040198e-06,
'classifier:gradient_boosting:learning_rate': 0.04634380160611007,
'classifier:gradient_boosting:loss': 'auto',
'classifier:gradient_boosting:max_bins': 255,
'classifier:gradient_boosting:max_depth': 'None',
'classifier:gradient_boosting:max_leaf_nodes': 11,
'classifier:gradient_boosting:min_samples_leaf': 41,
'classifier:gradient_boosting:scoring': 'loss',
'classifier:gradient_boosting:tol': 1e-07,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'robust_scaler',
'classifier:gradient_boosting:n_iter_no_change': 17,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_max': 0.7310641872341862,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_min': 0.2836388190241387},
{'balancing:strategy': 'none',
'classifier:__choice__': 'libsvm_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'no_preprocessing',
'classifier:libsvm_svc:C': 12923.91167672086,
'classifier:libsvm_svc:gamma': 0.028106748647672205,
'classifier:libsvm_svc:kernel': 'rbf',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'False',
'classifier:libsvm_svc:tol': 2.0706197108771777e-05,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'minmax'},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'lda',
'data_preprocessor:__choice__': 'feature_type',

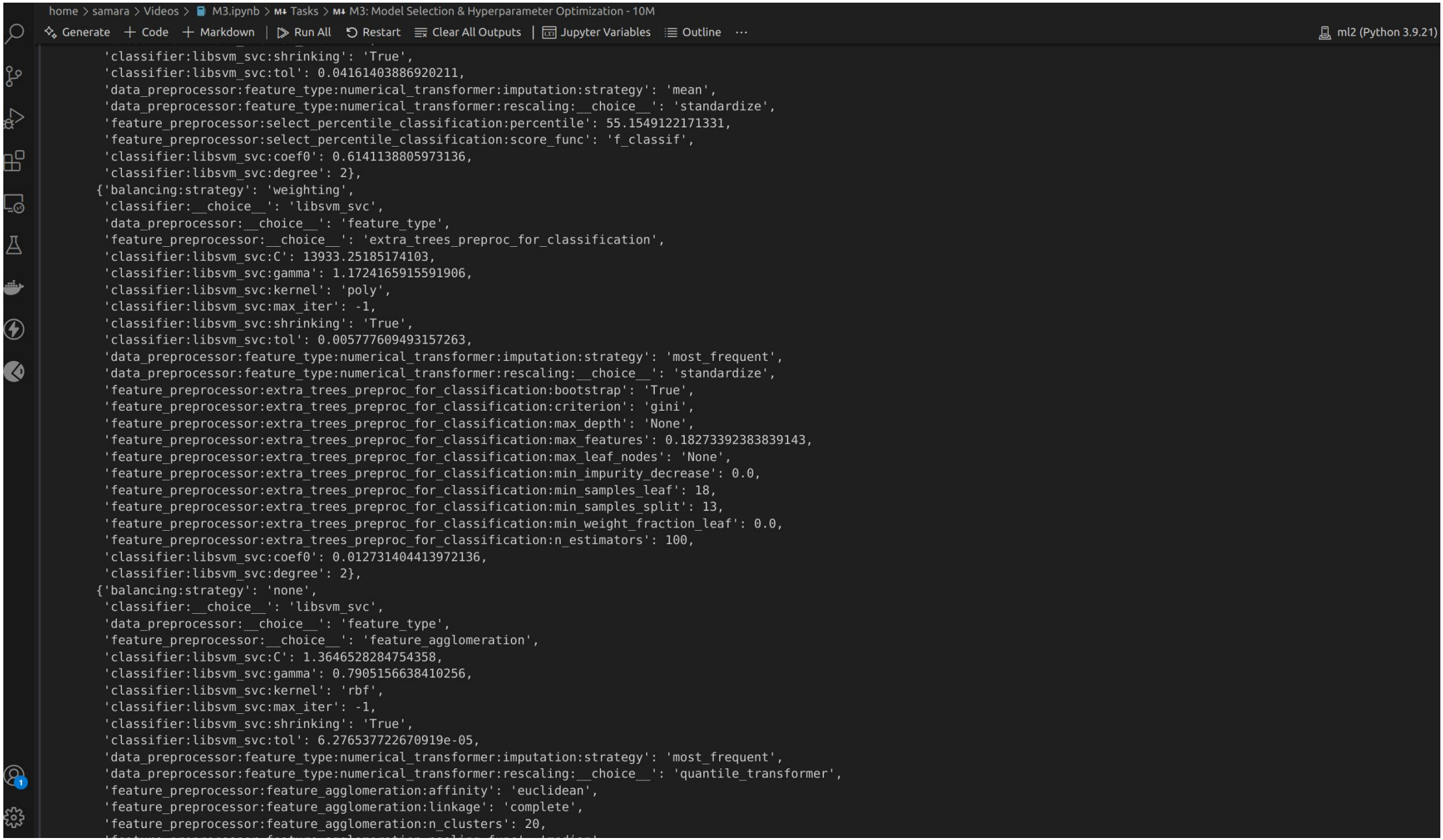

home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'no_preprocessing',
'classifier:libsvm_svc:C': 12923.91167672086,
'classifier:libsvm_svc:gamma': 0.028106748647672205,
'classifier:libsvm_svc:kernel': 'rbf',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'False',
'classifier:libsvm_svc:tol': 2.0706197108771777e-05,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'minmax'},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'lda',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'kitchen_sinks',
'classifier:lda:shrinkage': 'auto',
'classifier:lda:tol': 0.00018030860519654287,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'robust_scaler',
'feature_preprocessor:kitchen_sinks:gamma': 0.0005845623820571637,
'feature_preprocessor:kitchen_sinks:n_components': 8606,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_max': 0.8898452660666816,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_min': 0.1704082739702074},
{'balancing:strategy': 'none',
'classifier:__choice__': 'gradient_boosting',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'select_percentile_classification',
'classifier:gradient_boosting:early_stop': 'valid',
'classifier:gradient_boosting:l2_regularization': 1.7049772538367706e-08,
'classifier:gradient_boosting:learning_rate': 0.0825755415435688,
'classifier:gradient_boosting:loss': 'auto',
'classifier:gradient_boosting:max_bins': 255,
'classifier:gradient_boosting:max_depth': 'None',
'classifier:gradient_boosting:max_leaf_nodes': 51,
'classifier:gradient_boosting:min_samples_leaf': 121,
'classifier:gradient_boosting:scoring': 'loss',
'classifier:gradient_boosting:tol': 1e-07,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'robust_scaler',
'feature_preprocessor:select_percentile_classification:percentile': 27.329670949267125,
'feature_preprocessor:select_percentile_classification:score_func': 'mutual_info',
'classifier:gradient_boosting:n_iter_no_change': 10,
'classifier:gradient_boosting:validation_fraction': 0.15078023719798528,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_max': 0.8101878073561093,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_min': 0.23656284692369586},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'mlp',
'data_preprocessor:__choice__': 'feature_type',

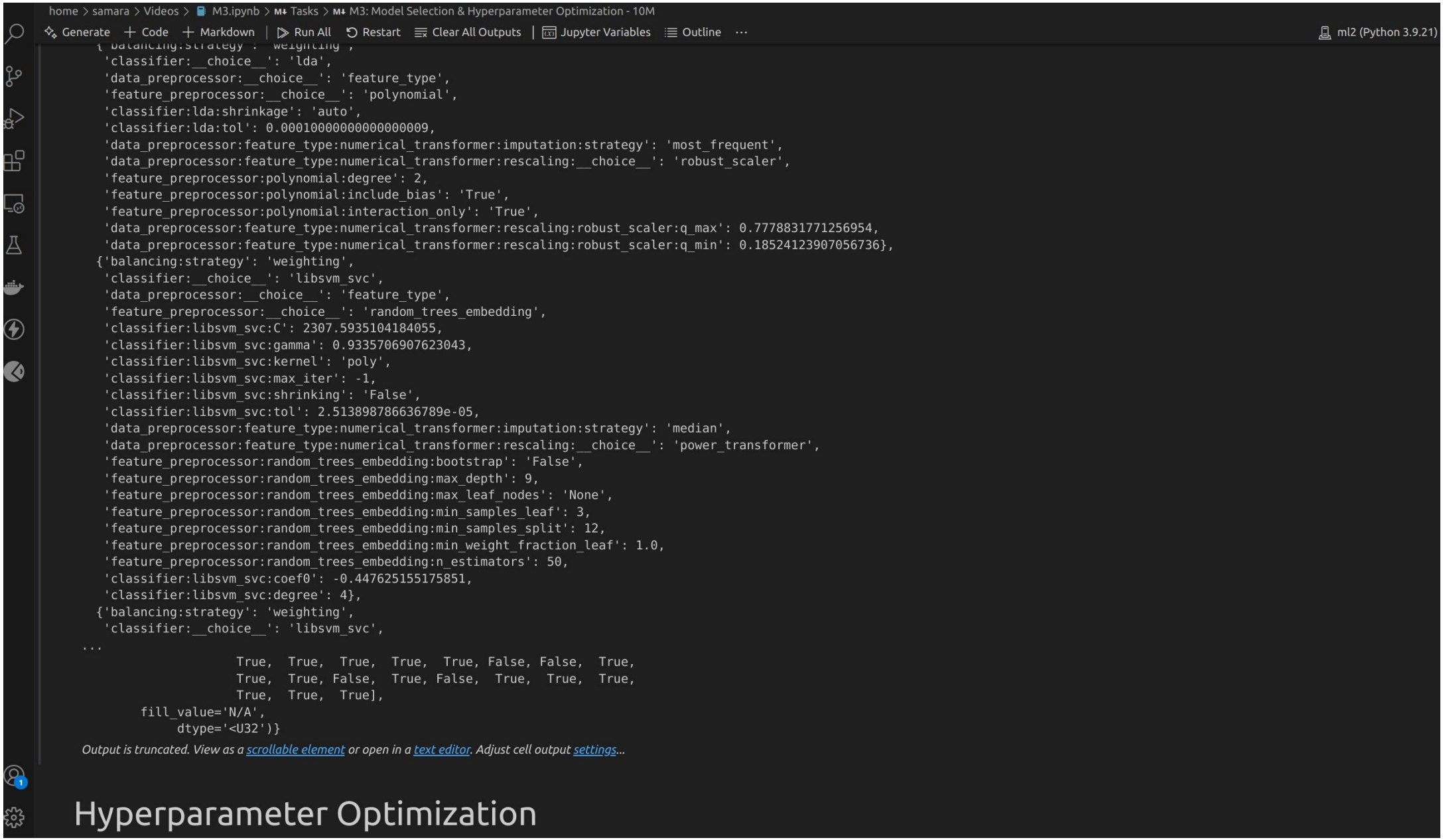

home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
'classifier:gradient_boosting:scoring': 'loss',
'classifier:gradient_boosting:tol': 1e-07,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'robust_scaler',
'feature_preprocessor:select_percentile_classification:percentile': 27.329670949267125,
'feature_preprocessor:select_percentile_classification:score_func': 'mutual_info',
'classifier:gradient_boosting:n_iter_no_change': 10,
'classifier:gradient_boosting:validation_fraction': 0.15078023719798528,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_max': 0.8101878073561093,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_min': 0.23656284692369586},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'mlp',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'feature_agglomeration',
'classifier:mlp:activation': 'tanh',
'classifier:mlp:alpha': 0.00021148999718383549,
'classifier:mlp:batch_size': 'auto',
'classifier:mlp:beta_1': 0.9,
'classifier:mlp:beta_2': 0.999,
'classifier:mlp:early_stopping': 'train',
'classifier:mlp:epsilon': 1e-08,
'classifier:mlp:hidden_layer_depth': 3,
'classifier:mlp:learning_rate_init': 0.0007452270241186694,
'classifier:mlp:n_iter_no_change': 32,
'classifier:mlp:num_nodes_per_layer': 113,
'classifier:mlp:shuffle': 'True',
'classifier:mlp:solver': 'adam',
'classifier:mlp:tol': 0.0001,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'standardize',
'feature_preprocessor:feature_agglomeration:affinity': 'euclidean',
'feature_preprocessor:feature_agglomeration:linkage': 'complete',
'feature_preprocessor:feature_agglomeration:n_clusters': 247,
'feature_preprocessor:feature_agglomeration:pooling_func': 'max'},
{'balancing:strategy': 'none',
'classifier:__choice__': 'mlp',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'liblinear_svc_preprocessor',
'classifier:mlp:activation': 'tanh',
'classifier:mlp:alpha': 3.198483470889531e-06,
'classifier:mlp:batch_size': 'auto',
'classifier:mlp:beta_1': 0.9,
'classifier:mlp:beta_2': 0.999,
'classifier:mlp:early_stopping': 'train',
'classifier:mlp:epsilon': 1e-08,
'classifier:mlp:hidden_layer_depth': 2,
'classifier:mlp:learning_rate_init': 0.006604847357173181,

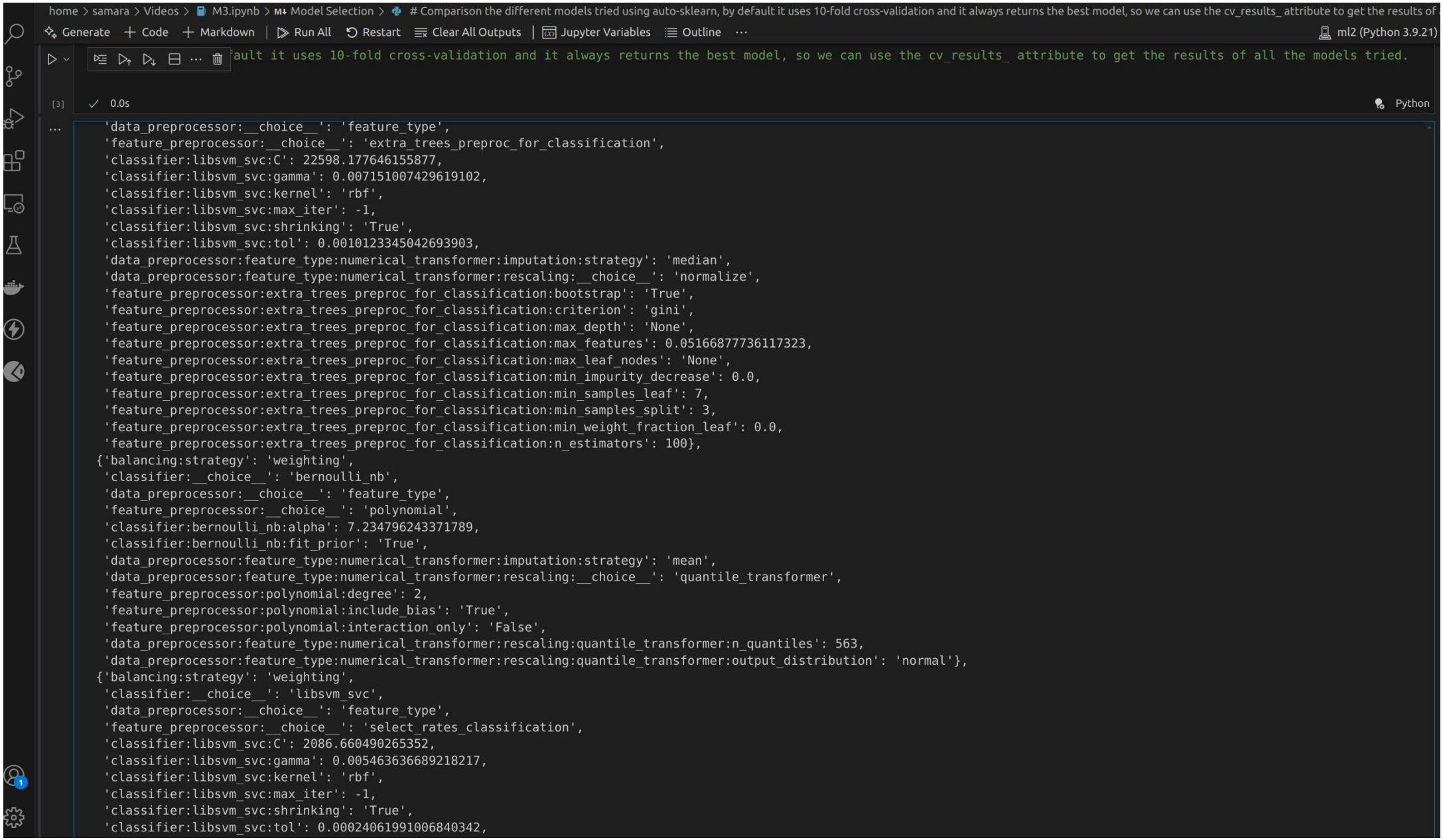

home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
{'balancing:strategy': 'none',
'classifier:__choice__': 'mlp',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'liblinear_svc_preprocessor',
'classifier:mlp:activation': 'tanh',
'classifier:mlp:alpha': 3.198483470889531e-06,
'classifier:mlp:batch_size': 'auto',
'classifier:mlp:beta_1': 0.9,
'classifier:mlp:beta_2': 0.999,
'classifier:mlp:early_stopping': 'train',
'classifier:mlp:epsilon': 1e-08,
'classifier:mlp:hidden_layer_depth': 2,
'classifier:mlp:learning_rate_init': 0.006604847357173181,
'classifier:mlp:n_iter_no_change': 32,
'classifier:mlp:num_nodes_per_layer': 24,
'classifier:mlp:shuffle': 'True',
'classifier:mlp:solver': 'adam',
'classifier:mlp:tol': 0.0001,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'standardize',
'feature_preprocessor:liblinear_svc_preprocessor:C': 0.6755226057791226,
'feature_preprocessor:liblinear_svc_preprocessor:dual': 'False',
'feature_preprocessor:liblinear_svc_preprocessor:fit_intercept': 'True',
'feature_preprocessor:liblinear_svc_preprocessor:intercept_scaling': 1,
'feature_preprocessor:liblinear_svc_preprocessor:loss': 'squared_hinge',
'feature_preprocessor:liblinear_svc_preprocessor:multi_class': 'ovr',
'feature_preprocessor:liblinear_svc_preprocessor:penalty': 'l1',
'feature_preprocessor:liblinear_svc_preprocessor:tol': 0.00014091021190100242},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'libsvm_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'no_preprocessing',
'classifier:libsvm_svc:C': 575.2840839820392,
'classifier:libsvm_svc:gamma': 0.28251015893293174,
'classifier:libsvm_svc:kernel': 'rbf',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'False',
'classifier:libsvm_svc:tol': 6.314728575433805e-05,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'none'},
{'balancing:strategy': 'none',
'classifier:__choice__': 'lda',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'kitchen_sinks',
'classifier:lda:shrinkage': 'auto',
'classifier:lda:tol': 0.018821286956948503,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',


home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'none',
'feature_preprocessor:feature_agglomeration:affinity': 'euclidean',
'feature_preprocessor:feature_agglomeration:linkage': 'ward',
'feature_preprocessor:feature_agglomeration:n_clusters': 202,
'feature_preprocessor:feature_agglomeration:pooling_func': 'mean'},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'sgd',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'polynomial',
'classifier:sgd:alpha': 0.0018367485569817758,
'classifier:sgd:average': 'True',
'classifier:sgd:fit_intercept': 'True',
'classifier:sgd:learning_rate': 'constant',
'classifier:sgd:loss': 'modified_huber',
'classifier:sgd:penalty': 'l1',
'classifier:sgd:tol': 5.2594585460405347e-05,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'normalize',
'feature_preprocessor:polynomial:degree': 3,
'feature_preprocessor:polynomial:include_bias': 'True',
'feature_preprocessor:polynomial:interaction_only': 'True',
'classifier:sgd:epsilon': 0.06935467259367672,
'classifier:sgd:eta0': 3.0158489869733344e-07},
{'balancing:strategy': 'none',
'classifier:__choice__': 'random_forest',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'polynomial',
'classifier:random_forest:bootstrap': 'True',
'classifier:random_forest:criterion': 'gini',
'classifier:random_forest:max_depth': 'None',
'classifier:random_forest:max_features': 0.9414141010985839,
'classifier:random_forest:max_leaf_nodes': 'None',
'classifier:random_forest:min_impurity_decrease': 0.0,
'classifier:random_forest:min_samples_leaf': 5,
'classifier:random_forest:min_samples_split': 2,
'classifier:random_forest:min_weight_fraction_leaf': 0.0,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'minmax',
'feature_preprocessor:polynomial:degree': 2,
'feature_preprocessor:polynomial:include_bias': 'True',
'feature_preprocessor:polynomial:interaction_only': 'False'},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'gradient_boosting',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'polynomial',
'classifier:gradient_boosting:early_stop': 'off',


home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
'feature_preprocessor:__choice__': 'select_rates_classification',
'classifier:libsvm_svc:C': 566.6840052213279,
'classifier:libsvm_svc:gamma': 0.0019647850141763412,
'classifier:libsvm_svc:kernel': 'rbf',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'True',
'classifier:libsvm_svc:tol': 0.0012292213570589683,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'robust_scaler',
'feature_preprocessor:select_rates_classification:alpha': 0.38609551513581725,
'feature_preprocessor:select_rates_classification:score_func': 'f_classif',
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_max': 0.7590668743404532,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_min': 0.2382188278316001,
'feature_preprocessor:select_rates_classification:mode': 'fdr'},
{'balancing:strategy': 'none',
'classifier:__choice__': 'liblinear_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'select_percentile_classification',
'classifier:liblinear_svc:C': 6.395900616857626,
'classifier:liblinear_svc:dual': 'False',
'classifier:liblinear_svc:fit_intercept': 'True',
'classifier:liblinear_svc:intercept_scaling': 1,
'classifier:liblinear_svc:loss': 'squared_hinge',
'classifier:liblinear_svc:multi_class': 'ovr',
'classifier:liblinear_svc:penalty': 'l2',
'classifier:liblinear_svc:tol': 0.00016419370701353068,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'quantile_transformer',
'feature_preprocessor:select_percentile_classification:percentile': 80.9663133881143,
'feature_preprocessor:select_percentile_classification:score_func': 'chi2',
'data_preprocessor:feature_type:numerical_transformer:rescaling:quantile_transformer:n_quantiles': 431,
'data_preprocessor:feature_type:numerical_transformer:rescaling:quantile_transformer:output_distribution': 'normal'},
{'balancing:strategy': 'none',
'classifier:__choice__': 'random_forest',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'polynomial',
'classifier:random_forest:bootstrap': 'False',
'classifier:random_forest:criterion': 'gini',
'classifier:random_forest:max_depth': 'None',
'classifier:random_forest:max_features': 0.5830626115596406,
'classifier:random_forest:max_leaf_nodes': 'None',
'classifier:random_forest:min_impurity_decrease': 0.0,
'classifier:random_forest:min_samples_leaf': 2,
'classifier:random_forest:min_samples_split': 5,
'classifier:random_forest:min_weight_fraction_leaf': 0.0,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'normalize',
'feature_preprocessor:polynomial:degree': 2,


home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
'classifier:libsvm_svc:shrinking': 'True',
'classifier:libsvm_svc:tol': 0.04161403886920211,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'standardize',
'feature_preprocessor:select_percentile_classification:percentile': 55.1549122171331,
'feature_preprocessor:select_percentile_classification:score_func': 'f_classif',
'classifier:libsvm_svc:coef0': 0.6141138805973136,
'classifier:libsvm_svc:degree': 2},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'libsvm_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'extra_trees_preproc_for_classification',
'classifier:libsvm_svc:C': 13933.25185174103,
'classifier:libsvm_svc:gamma': 1.1724165915591906,
'classifier:libsvm_svc:kernel': 'poly',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'True',
'classifier:libsvm_svc:tol': 0.005777609493157263,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'standardize',
'feature_preprocessor:extra_trees_preproc_for_classification:bootstrap': 'True',
'feature_preprocessor:extra_trees_preproc_for_classification:criterion': 'gini',
'feature_preprocessor:extra_trees_preproc_for_classification:max_depth': 'None',
'feature_preprocessor:extra_trees_preproc_for_classification:max_features': 0.18273392383839143,
'feature_preprocessor:extra_trees_preproc_for_classification:max_leaf_nodes': 'None',
'feature_preprocessor:extra_trees_preproc_for_classification:min_impurity_decrease': 0.0,
'feature_preprocessor:extra_trees_preproc_for_classification:min_samples_leaf': 18,
'feature_preprocessor:extra_trees_preproc_for_classification:min_samples_split': 13,
'feature_preprocessor:extra_trees_preproc_for_classification:min_weight_fraction_leaf': 0.0,
'feature_preprocessor:extra_trees_preproc_for_classification:n_estimators': 100,
'classifier:libsvm_svc:coef0': 0.012731404413972136,
'classifier:libsvm_svc:degree': 2},
{'balancing:strategy': 'none',
'classifier:__choice__': 'libsvm_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'feature_agglomeration',
'classifier:libsvm_svc:C': 1.3646528284754358,
'classifier:libsvm_svc:gamma': 0.7905156638410256,
'classifier:libsvm_svc:kernel': 'rbf',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'True',
'classifier:libsvm_svc:tol': 6.276537722670919e-05,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'quantile_transformer',
'feature_preprocessor:feature_agglomeration:affinity': 'euclidean',
'feature_preprocessor:feature_agglomeration:linkage': 'complete',
'feature_preprocessor:feature_agglomeration:n_clusters': 20,


home > samara > Videos > M3.ipynb > Tasks > M3: Model Selection & Hyperparameter Optimization - 10M
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'lda',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'polynomial',
'classifier:lda:shrinkage': 'auto',
'classifier:lda:tol': 0.00010000000000000009,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'most_frequent',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'robust_scaler',
'feature_preprocessor:polynomial:degree': 2,
'feature_preprocessor:polynomial:include_bias': 'True',
'feature_preprocessor:polynomial:interaction_only': 'True',
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_max': 0.7778831771256954,
'data_preprocessor:feature_type:numerical_transformer:rescaling:robust_scaler:q_min': 0.18524123907056736},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'libsvm_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'random_trees_embedding',
'classifier:libsvm_svc:C': 2307.5935104184055,
'classifier:libsvm_svc:gamma': 0.9335706907623043,
'classifier:libsvm_svc:kernel': 'poly',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'False',
'classifier:libsvm_svc:tol': 2.513898786636789e-05,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'power_transformer',
'feature_preprocessor:random_trees_embedding:bootstrap': 'False',
'feature_preprocessor:random_trees_embedding:max_depth': 9,
'feature_preprocessor:random_trees_embedding:max_leaf_nodes': 'None',
'feature_preprocessor:random_trees_embedding:min_samples_leaf': 3,
'feature_preprocessor:random_trees_embedding:min_samples_split': 12,
'feature_preprocessor:random_trees_embedding:min_weight_fraction_leaf': 1.0,
'feature_preprocessor:random_trees_embedding:n_estimators': 50,
'classifier:libsvm_svc:coef0': -0.447625155175851,
'classifier:libsvm_svc:degree': 4},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'libsvm_svc',
...
True, True, True, True, True, False, False, True,
True, True, False, True, False, True, True, True,
True, True, True],
fill_value='N/A',
dtype='<U32')}
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
Hyperparameter Optimization


home > samara > Videos > M3.ipynb > Model Selection > # Comparison the different models tried using auto-sklearn, by default it uses 10-fold cross-validation and it always returns the best model, so we can use the cv_results_ attribute to get the results of
Generate
Code
Markdown
Run All
Restart
Clear All Outputs
Jupyter Variables
Outline
ml2 (Python 3.9.21)
fault it uses 10-fold cross-validation and it always returns the best model, so we can use the cv_results_ attribute to get the results of all the models tried.
[3]
0.0s
Python
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'extra_trees_preproc_for_classification',
'classifier:libsvm_svc:C': 22598.177646155877,
'classifier:libsvm_svc:gamma': 0.007151007429619102,
'classifier:libsvm_svc:kernel': 'rbf',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'True',
'classifier:libsvm_svc:tol': 0.0010123345042693903,
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'median',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'normalize',
'feature_preprocessor:extra_trees_preproc_for_classification:bootstrap': 'True',
'feature_preprocessor:extra_trees_preproc_for_classification:criterion': 'gini',
'feature_preprocessor:extra_trees_preproc_for_classification:max_depth': 'None',
'feature_preprocessor:extra_trees_preproc_for_classification:max_features': 0.05166877736117323,
'feature_preprocessor:extra_trees_preproc_for_classification:max_leaf_nodes': 'None',
'feature_preprocessor:extra_trees_preproc_for_classification:min_impurity_decrease': 0.0,
'feature_preprocessor:extra_trees_preproc_for_classification:min_samples_leaf': 7,
'feature_preprocessor:extra_trees_preproc_for_classification:min_samples_split': 3,
'feature_preprocessor:extra_trees_preproc_for_classification:min_weight_fraction_leaf': 0.0,
'feature_preprocessor:extra_trees_preproc_for_classification:n_estimators': 100},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'bernoulli_nb',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'polynomial',
'classifier:bernoulli_nb:alpha': 7.234796243371789,
'classifier:bernoulli_nb:fit_prior': 'True',
'data_preprocessor:feature_type:numerical_transformer:imputation:strategy': 'mean',
'data_preprocessor:feature_type:numerical_transformer:rescaling:__choice__': 'quantile_transformer',
'feature_preprocessor:polynomial:degree': 2,
'feature_preprocessor:polynomial:include_bias': 'True',
'feature_preprocessor:polynomial:interaction_only': 'False',
'data_preprocessor:feature_type:numerical_transformer:rescaling:quantile_transformer:n_quantiles': 563,
'data_preprocessor:feature_type:numerical_transformer:rescaling:quantile_transformer:output_distribution': 'normal'},
{'balancing:strategy': 'weighting',
'classifier:__choice__': 'libsvm_svc',
'data_preprocessor:__choice__': 'feature_type',
'feature_preprocessor:__choice__': 'select_rates_classification',
'classifier:libsvm_svc:C': 2086.660490265352,
'classifier:libsvm_svc:gamma': 0.005463636689218217,
'classifier:libsvm_svc:kernel': 'rbf',
'classifier:libsvm_svc:max_iter': -1,
'classifier:libsvm_svc:shrinking': 'True',
'classifier:libsvm_svc:tol': 0.00020061991006840342,