Measurement overview!

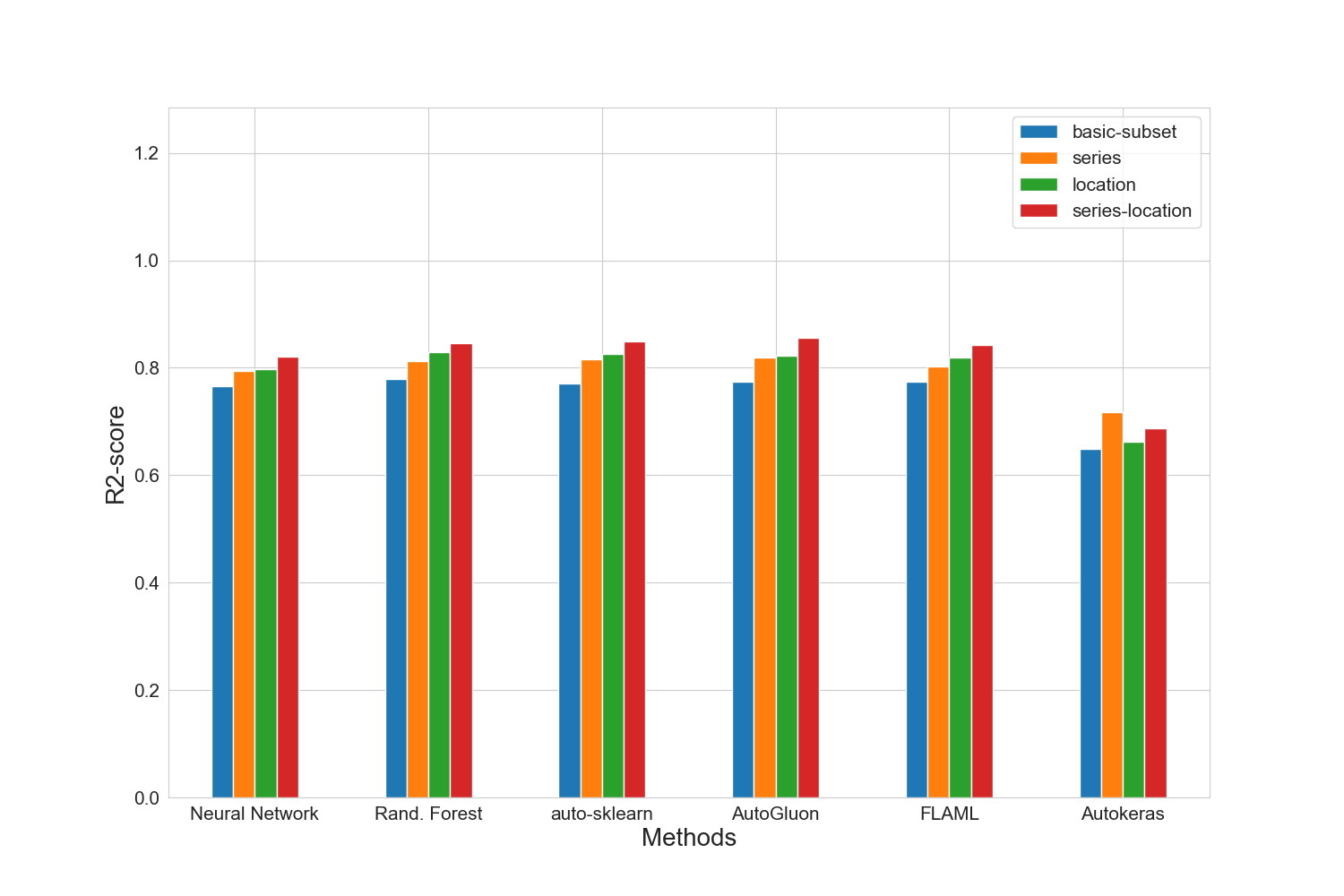
* Measurement date: 2023-12-09
* Random seed: 1
* Number of PCA calculated features: 0
* Outliert detection per bin: True
* Autosklearn runtime per dataset: 1800 seconds
* Autosklearn time limit per algorithm: 180 seconds

# Results for merged-files

* Input data file name: merged-files-final-2022-07-06.csv
* Input data file creation date: 2022-07-06
* Number of processed data points: 2910

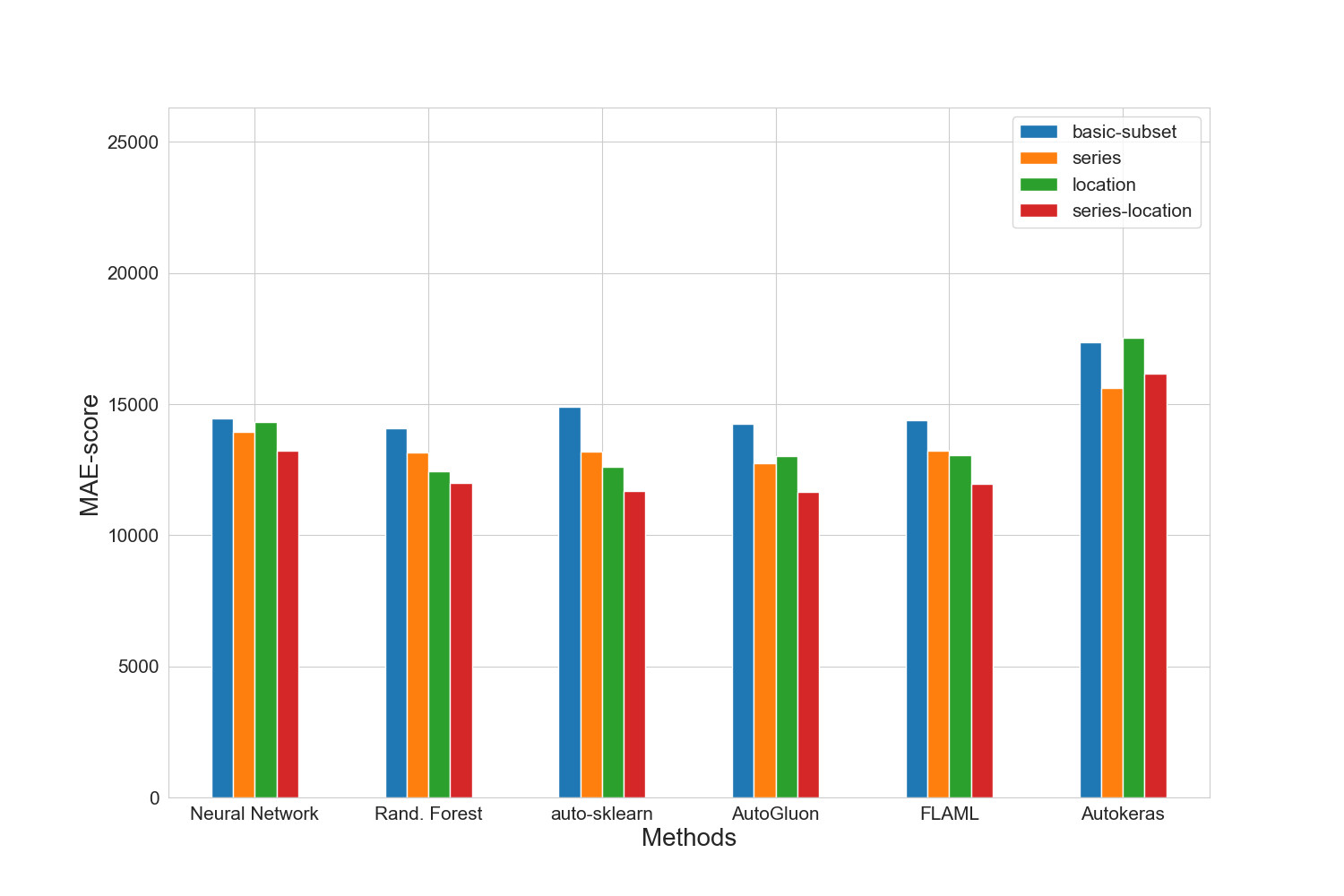
### R2 results for merged-files

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | basic-subset | extension | location | extension-location |
| nn | 0.7658 | 0.7938 | 0.7975 | 0.8209 |
| manual | 0.7789 | 0.8129 | 0.829 | 0.8452 |
| autosklearn | 0.7701 | 0.8156 | 0.8259 | 0.8489 |
| autogluon | 0.774 | 0.8195 | 0.8222 | **0.8564** |
| flaml | 0.7747 | 0.8018 | 0.8184 | 0.8421 |
| autokeras | 0.6496 | 0.7173 | 0.6631 | 0.6871 |



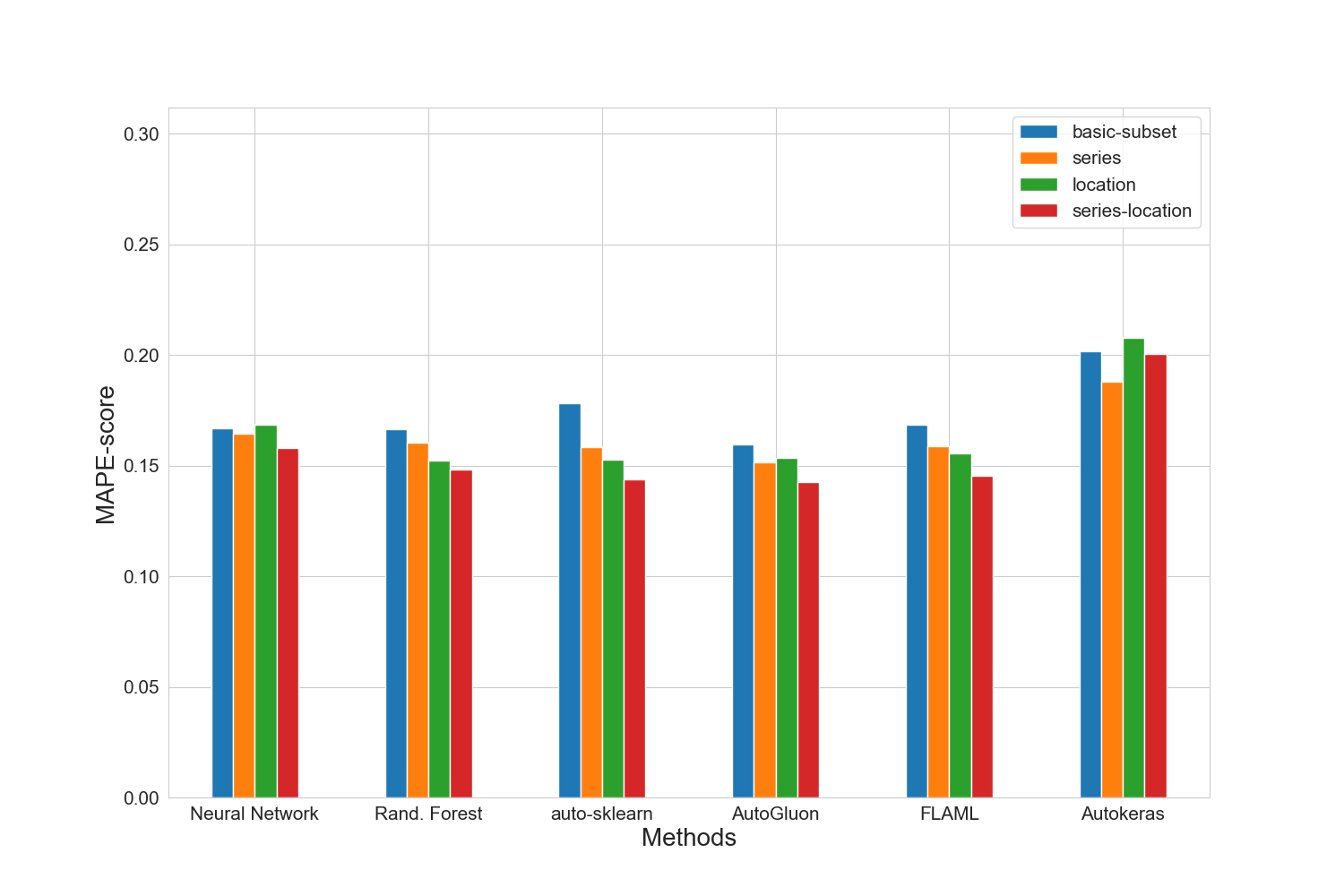
### MAE results for merged-files

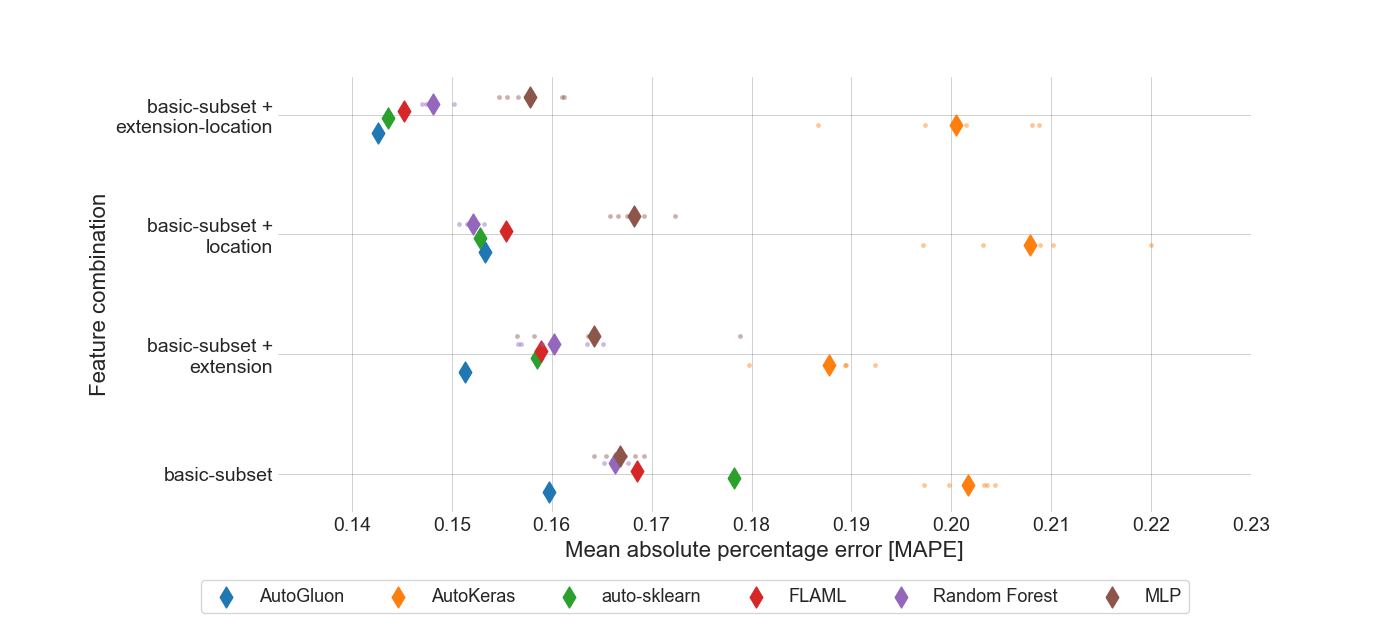
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | basic-subset | extension | location | extension-location |
| nn | 14467.6 | 13929.0 | 14302.8 | 13232.0 |
| manual | 14084.4 | 13163.2 | 12440.4 | 11979.8 |
| autosklearn | 14911.0 | 13204.0 | 12625.0 | 11703.0 |
| autogluon | 14252.0 | 12737.0 | 13013.0 | **11654.0** |
| flaml | 14380.0 | 13232.0 | 13058.0 | 11975.0 |
| autokeras | 17360.6 | 15622.4 | 17543.8 | 16168.0 |



### MAPE results for merged-files

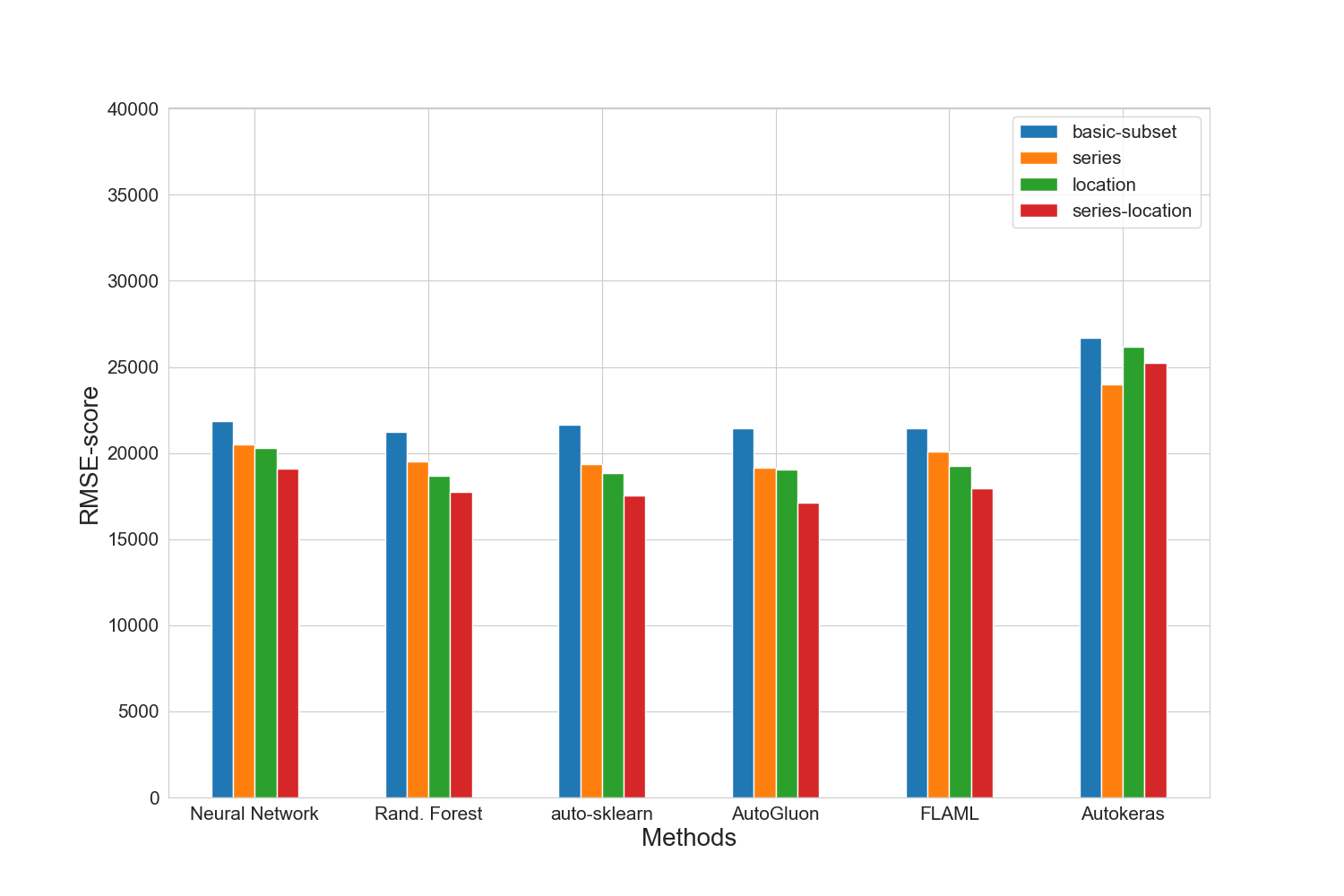
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | basic-subset | extension | location | extension-location |
| nn | 0.1669 | 0.1643 | 0.1683 | 0.1578 |
| manual | 0.1664 | 0.1602 | 0.1521 | 0.1481 |
| autosklearn | 0.1782 | 0.1585 | 0.1528 | 0.1436 |
| autogluon | 0.1597 | 0.1513 | 0.1533 | **0.1426** |
| flaml | 0.1685 | 0.1589 | 0.1554 | 0.1452 |
| autokeras | 0.2017 | 0.1878 | 0.2079 | 0.2005 |





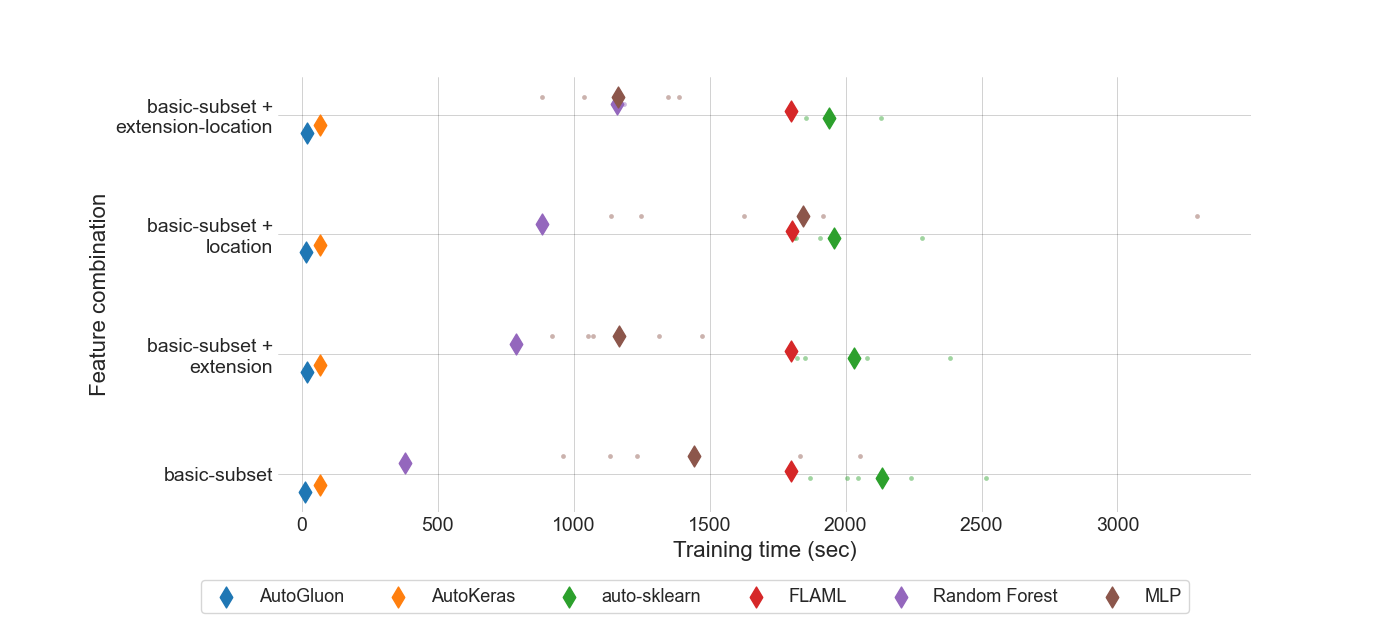
### RMSE results for merged-files

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | basic-subset | extension | location | extension-location |
| nn | 21838.0 | 20488.4 | 20301.4 | 19094.0 |
| manual | 21216.6 | 19519.0 | 18662.8 | 17751.4 |
| autosklearn | 21636.0 | 19379.0 | 18829.0 | 17543.0 |
| autogluon | 21451.0 | 19173.0 | 19028.0 | **17098.0** |
| flaml | 21419.0 | 20090.0 | 19230.0 | 17931.0 |
| autokeras | 26709.0 | 23974.6 | 26186.2 | 25238.4 |



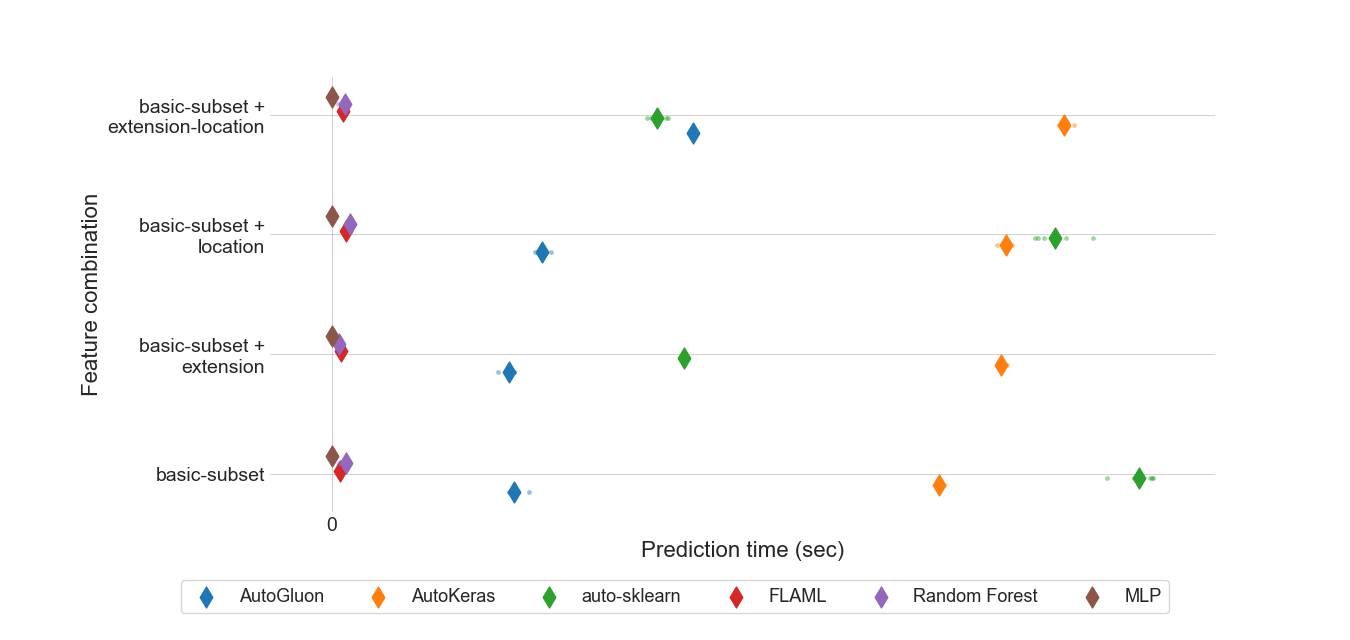
### Training run-time of each method and feature subset in seconds for merged-files

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | basic-subset | extension | location | extension-location |
| nn | 1442.2 | 1166.2 | 1844.2 | 1163.0 |
| manual | 380.5 | 788.3 | 882.3 | 1158.3 |
| autosklearn | 2135.4 | 2030.1 | 1956.4 | 1937.6 |
| autogluon | 12.3 | 17.9 | 16.3 | 18.1 |
| flaml | 1800.2 | 1800.2 | 1801.0 | 1800.4 |
| autokeras | 66.3 | 65.8 | 65.6 | 65.9 |



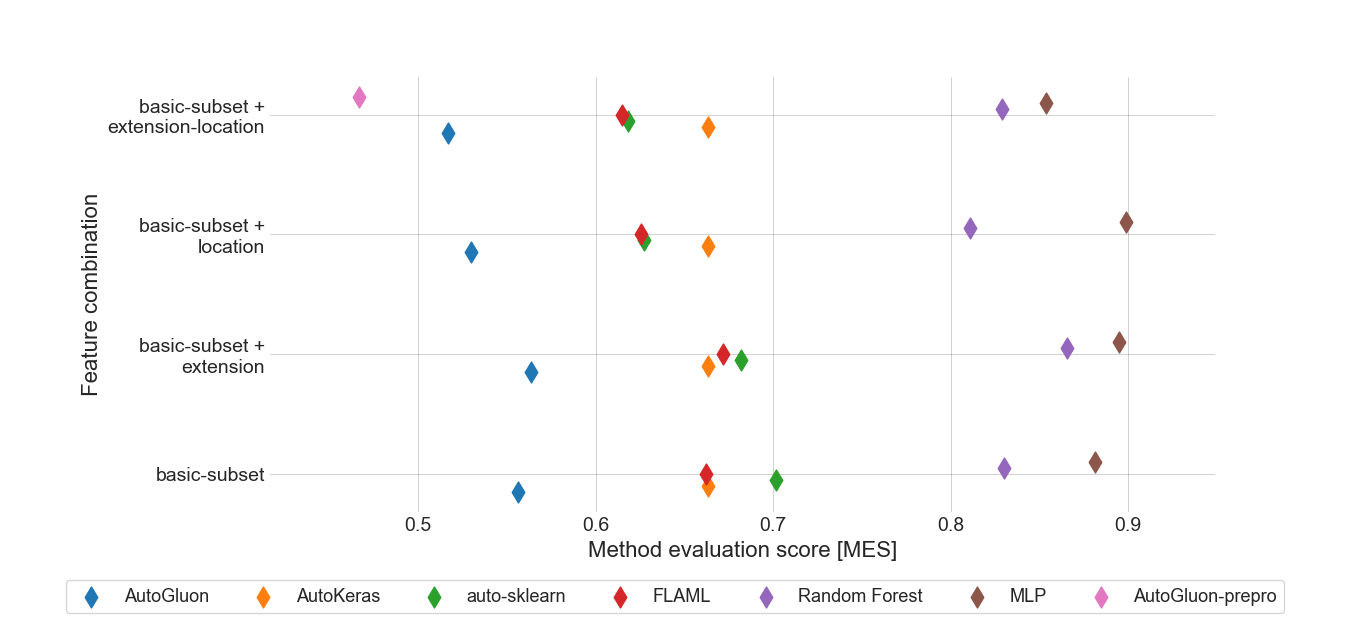
### Prediction time of each method and feature subset in seconds for merged-files

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | basic-subset | extension | location | extension-location |
| nn | 7.506e-07 | 7.01e-07 | 8.26e-07 | 8.354e-07 |
| manual | 3.23658e-05 | 1.64388e-05 | 4.33289e-05 | 3.12581e-05 |
| autosklearn | 0.0019419344 | 0.0008472066 | 0.0017402592 | 0.0007827125 |
| autogluon | 0.0004378285 | 0.0004257194 | 0.0005047946 | 0.0008683983 |
| flaml | 1.90637e-05 | 2.15561e-05 | 3.36132e-05 | 2.68258e-05 |
| autokeras | 0.0014599348 | 0.0016104093 | 0.0016228917 | 0.0017624608 |



### MEV results for all feature sets.

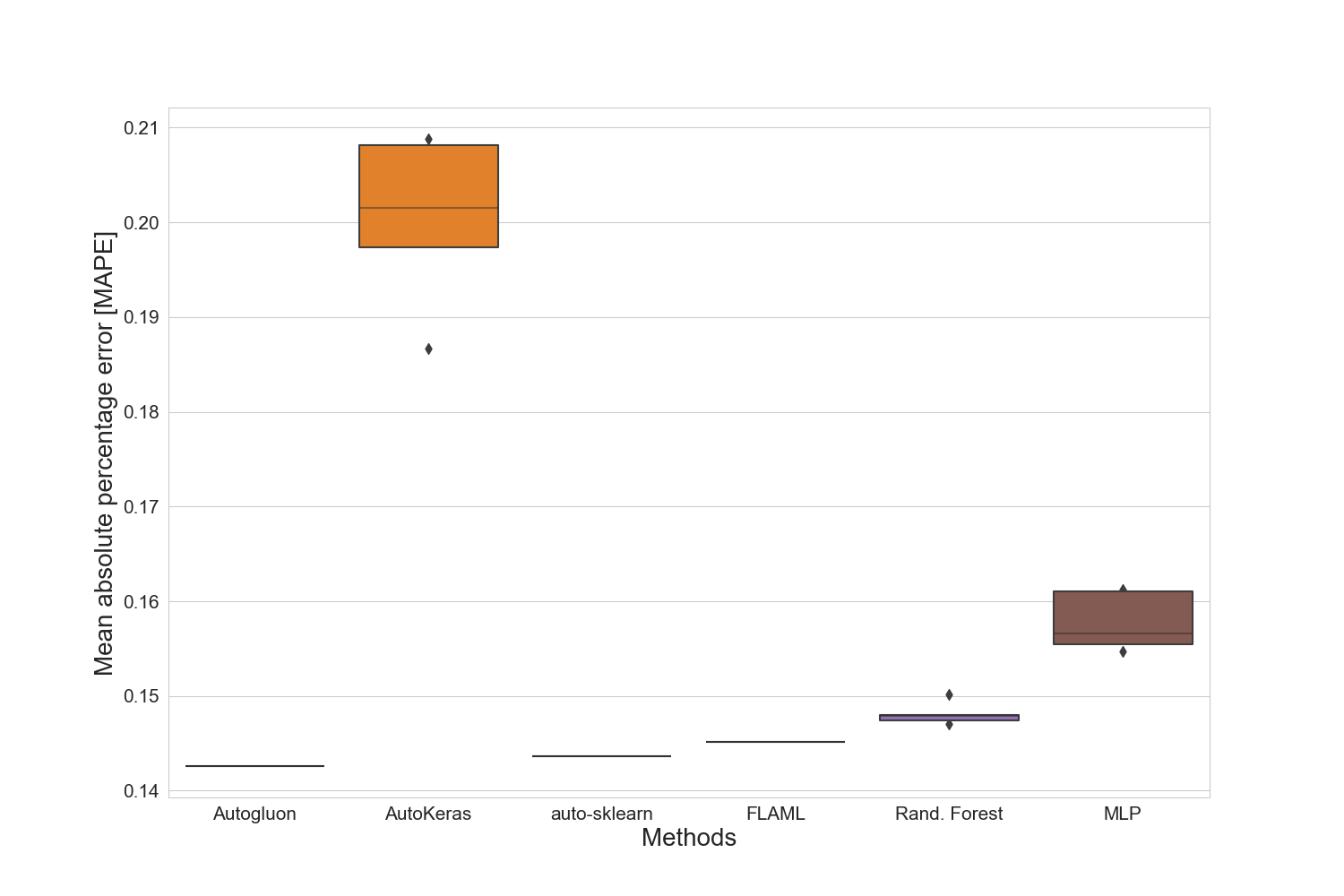
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | basic-subset | extension | location | extension-location |
| nn | 0.881271 | 0.894879 | 0.899027 | 0.853539 |
| manual | 0.830312 | 0.865348 | 0.810899 | 0.829107 |
| autosklearn | 0.701745 | 0.681991 | 0.627484 | 0.618105 |
| autogluon | 0.556461 | 0.563704 | 0.52952 | 0.516545 |
| flaml | 0.662002 | 0.671732 | 0.625794 | 0.615014 |
| autokeras | 0.663105 | 0.663241 | 0.663353 | 0.663401 |
| autogluon-prepro | nan | nan | nan | **0.466477** |



### Standard deviation for MAPE results for extension-location feature set.

|  |  |
| --- | --- |
|  | extension-location |
| autogluon | 0.0 |
| autokeras | 0.00809568 |
| autosklearn | 0.0 |
| flaml | 0.0 |
| manual | 0.00110995 |
| nn | 0.00276188 |

### Boxplot for MAPE results for extension-location feature set.



### Standard deviation for training time for extension-location feature set.

|  |  |
| --- | --- |
|  | extension-location |
| autogluon | 0.17096596 |
| autokeras | 0.18708137 |
| autosklearn | 112.92645682 |
| flaml | 0.15910173 |
| manual | 14.16133225 |
| nn | 189.10804774 |

### Standard deviation for prediction time for extension-location feature set.

|  |  |
| --- | --- |
|  | extension-location |
| autogluon | 7.415449e-06 |
| autokeras | 1.3724808e-05 |
| autosklearn | 2.0234617e-05 |
| flaml | 1.204301e-06 |
| manual | 8.732151e-06 |
| nn | 1.57161e-07 |

### Best manual model for merged-files

|  |
| --- |
| Best manual model |
| RandomForestRegressor(bootstrap=False, max\_features=18, min\_samples\_split=10,  n\_estimators=142) |

### Best NN parameter set for merged-files

|  |  |  |  |
| --- | --- | --- | --- |
| Activation | Hidden-layer-size | Learning rate | Solver |
| relu | (10, 10, 10, 10, 10, 10, 10, 10, 10) | 0.01 | adam |

### Best results for merged-files

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | R2-Score | MAPE | MAE | RMSE | N-RMSE | IQR-RMSE | CV-RMSE |
| Feature-Set | extension-location | extension-location | extension-location | extension-location | extension-location | extension-location | extension-location |
| Method | autogluon | autogluon | autogluon | autogluon | autogluon | autogluon | autogluon |
| Value | 0.8564 | 0.1426 | 11654.0 | 17098.0 | 0.0645 | 0.2771 | 0.1877 |

# Summary of the best results for all data sets

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | # data points | R2-Score | MAPE | MAE | RMSE | N-RMSE | IQR-RMSE | CV-RMSE |
| merged-files | 2910 | **0.8564** | **0.1426** | **11654.0** | **17098.0** | **0.0645** | **0.2771** | **0.1877** |

# Statistics for all data sets

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | # data points | Mean | Median | Std. diviation | Max | Min |
| merged-files | 2910 | 92522 | 80000 | 46927 | 453434 | 1000 |