Appendix

Data Set		Classes	Samples	Numeric Feat.	Categorical Feat.	Missing Values	Incom. Samples	Minority %
kr-vs-kp	(3)	2	3196	0	37	0	0	47.78
mfeat-factors	(12)	10	2000	216	1	0	0	10.00
mfeat-morph	(18)	10	2000	6	1	0	0	10.00
credit-g	(31)	2	1000	7	14	0	0	30.00
vehicle	(54)	4	846	18	1	0	0	23.52
analcatdata	(458)	4	841	70	1	0	0	6.54
analcatdata	(469)	6	797	0	5	0	0	15.43
ada_agnostic	(1043)	2	4562	48	1	0	0	24.81
kc2	(1063)	2	522	21	1	0	0	20.50
bank-marketin	(1461)	2	45211	7	10	0	0	11.70
blood-transfusi	(1464)	2	748	4	1	0	0	23.80
eeg-eye-state	(1471)	2	14980	14	1	0	0	44.88
nomao	(1486)	2	34465	89	30	0	0	28.56
phoneme	(1489)	2	5404	5	1	0	0	29.35
sa-heart	(1498)	2	462	8	2	0	0	34.63
adult	(1590)	2	48842	6	9	6465	3620	23.93
higgs	(23512)	2	98050	28	1	9	1	47.14
numerai28.6	(23517)	2	96320	21	1	0	0	49.48
connect-4	(40668)	3	67557	0	43	0	0	9.55
Shuttle	(40685)	7	58000	9	1	0	0	0.02
car	(40975)	4	1728	0	7	0	0	3.76
Australian	(40981)	2	690	6	9	0	0	44.49
segment	(40984)	7	2310	19	1	0	0	14.29
jungle_chess	(41027)	3	44819	6	1	0	0	9.67
jasmine	(41143)	2	2984	8	137	0	0	50.00
sylvine	(41146)	2	5124	20	1	0	0	50.00
Jannis	(41168)	4	83733	54	1	0	0	2.01
Helena	(41169)	100	65196	27	1	0	0	0.17

Table 1: List of all tested data sets. Listed are the (abbreviated) name and OpenML id for each data set together with the number of classes, the number of samples, the number of numeric and categorical features per samples, how many values are missing in total (Missing values), how many samples contain at least one missing value (Incomp. Samples) and the percentage of samples belonging to the least frequent class (Minority %).

Meta-Feature
Nr. Instances
Nr. Attributes
Nr. Numerical Attributes
Nr. Categorical Attributes
Nr. Classes
Nr. Missing Values
Pct. Missing Values
Nr. Instances with Missing Values
Pct. Instances with Missing Values
Nr. Attributes with Missing Values
Pct. Attributes with Missing Values
Nr. Outliers
Skewness Mean
Skewness Std.
Kurtosis Mean
Kurtosis Std.
Correlation Mean
Correlation Std.
Covariance Mean
Covariance Std.
Sparsity Mean
Sparsity Std.
Variance Mean
Variance Std.
Class Prob. Mean
Class Prob. Std.
Class Entropy
Attribute Entropy Mean
Attribute Entropy Std.
Mutual Information Mean
Mutual Information Std.
Equal Nr. Attributes
Noisiness Ratio
Decision Tree Nodes
Decision Tree Leaves
Decision Tree Leaves Branch Mean
Decision Tree Leaves Branch Std.
Decision Tree Nodes per Attribute
Decision Tree Leaves per Class Mean
Decision Tree Leaves per Class Std.
Decision Tree Variable Importance Mean
Decision Tree Variable Importance Std.

Table 2: List of all calculated meta-features.

Algorithm	Algorithm Class	# Con.	# Cat.
AdaBoosting	Classifier	2	1
Bernoulli Naive Bayes	Classifier	1	1
Decision Tree	Classifier	3	1
Gradient Boosting	Classifier	7	0
Support Vector Machine	Classifier	5	2
LDA	Classifier	3	1
Multinomial Naive Bayes	Classifier	1	1
Random Forest	Classifier	3	2
SGD	Classifier	6	5
Imputation	Imputation	0	2
KNN Imputation	Imputation	1	2
Max-Abs Scaler	Scaling	0	0
Min-Max Scaler	Scaling	0	0
Normalizer	Scaling	0	0
Quantile Transformer	Scaling	1	1
Robust Scaler	Scaling	2	0
Standard Scaler	Scaling	0	0
Bernoulli RBM	Generation	4	0
Missing Indicator	Generation	0	1
Polynomial Features	Generation	1	2
Random Trees Embedding	Generation	5	1
Factor Analysis	Decomposition	4	1
Fast ICA	Decomposition	1	3
Feature Agglomeration	Decomposition	1	3
Kernel PCA	Decomposition	4	1
PCA	Decomposition	1	1
Truncated SVD	Decomposition	1	0
Binarizer	Discretization	0	0
K-Bins Discretizer	Discretization	1	2
Label Encoder	Encoding	0	0
One-Hot Encoder	Encoding	0	0
Generic Univariate Selection	Selection	1	2
Select k Best	Selection	1	1
Select Percentile	Selection	1	1
Variance Threshold	Filtering	1	0

Table 3: Implemented algorithms with the according algorithm class, number of continuous (# Con.) and categorical (# Cat) hyperparameters.