

# Appendix: Performance of ML models and feature extraction method

Table I lists the performance metrics of both extraction and classification methods for individual projects in each dataset

Dataset	Project	Method	Fold	F-Score	AUC_ROC
PPI	CPP_IN	RF-BERT	5	0.3876	0.6038
		BERT-BERT	5	0.502	0.618
		RF-TFIDF	5	0.6156	0.7538
	ACCUMULO	RF-BERT	5	0.9564	0.5214
		BERT-BERT	5	0.9668	0.594
		RF-TFIDF	5	0.9668	0.5356
	AEROGEAR	RF-BERT	5	0.418	0.5532
		BERT-BERT	5	0.6054	0.5112
		RF-TFIDF	5	0.514	0.6034
	AMBARI	RF-BERT	5	0.769	0.6034
		BERT-BERT	5	0.7974	0.6316
		RF-TFIDF	5	0.8458	0.5598
	ARROW	RF-BERT	5	0.7274	0.6128
		BERT-BERT	5	0.712	0.6352
		RF-TFIDF	5	0.809	0.5446
	ARTEMIS	RF-BERT	5	0.6232	0.3724
		BERT-BERT	5	0.7456	0.4784
		RF-TFIDF	5	0.6836	0.5302
	BATCH	RF-BERT	5	0.5846	0.646
		BERT-BERT	5	0.6142	0.6864
		RF-TFIDF	5	0.6438	0.6124
	BEAM	RF-BERT	5	0.5192	0.5466
		BERT-BERT	5	0.5188	0.56
		RF-TFIDF	5	0.5222	0.5298
	CALCITE	RF-BERT	5	0.7226	0.5754
		BERT-BERT	5	0.8202	0.5652
		RF-TFIDF	5	0.7848	0.4084
	CARBONDATA	RF-BERT	5	0.5324	0.5896
		BERT-BERT	5	0.6068	0.595
		RF-TFIDF	5	0.5748	0.5664
	EXOJCR	RF-BERT	5	0.2958	0.561
		BERT-BERT	5	0.4452	0.5758
		RF-TFIDF	5	0.3478	0.5604
	GTNPORTAL	RF-BERT	5	0.895	0.5452
		BERT-BERT	5	0.9046	0.5902
		RF-TFIDF	5	0.9252	0.48
	HDDS	RF-BERT	5	0.5564	0.576
		BERT-BERT	5	0.609	0.6132
		RF-TFIDF	5	0.5214	0.5574
	INT	RF-BERT	5	0.7102	0.5448
		BERT-BERT	5	0.844	0.4788
		RF-TFIDF	5	0.8432	0.5784
	JBAS	RF-BERT	5	0.5632	0.6272
		BERT-BERT	5	0.5724	0.6102
		RF-TFIDF	5	0.5524	0.6144
		RF-BERT	2	0.3335	0.5515
		BERT-BERT	2	0.41	0.5355

	RF-TFIDF	2	0.4595	0.5575
	RF-BERT	5	0.83	0.6692
JBEAP	BERT-BERT	5	0.8606	0.6018
	RF-TFIDF	5	0.8078	0.7636
	RF-BERT	5	0.4626	0.5372
JBESB	BERT-BERT	5	0.4602	0.5272
	RF-TFIDF	5	0.4592	0.5264
	RF-BERT	5	0.3288	0.4878
JBFORUMS	BERT-BERT	5	0.5506	0.4866
	RF-TFIDF	5	0.4558	0.439
	RF-BERT	5	0.2518	0.4892
JBLAB	BERT-BERT	5	0.4936	0.5174
	RF-TFIDF	5	0.336	0.4684
	RF-BERT	3	0.369	0.5427
JBPORTAL	BERT-BERT	3	0.499	0.545
	RF-TFIDF	3	0.3693	0.5263
	RF-BERT	5	0.3758	0.5038
JBTM	BERT-BERT	5	0.4762	0.4788
	RF-TFIDF	5	0.4974	0.547
	RF-BERT	5	0.6124	0.5734
METRON	BERT-BERT	5	0.6876	0.5996
	RF-TFIDF	5	0.6198	0.4862
	RF-BERT	2	0.7355	0.511
MNG	BERT-BERT	2	0.806	0.539
	RF-TFIDF	2	0.7925	0.507
	RF-BERT	5	0.5032	0.5544
MXNET	BERT-BERT	5	0.4942	0.5278
	RF-TFIDF	5	0.496	0.5138
	RF-BERT	5	0.7734	0.5156
NETBEANS	BERT-BERT	5	0.8208	0.444
	RF-TFIDF	5	0.7962	0.5426
	RF-BERT	5	0.8916	0.6108
NIFI	BERT-BERT	5	0.9108	0.6096
	RF-TFIDF	5	0.9012	0.5152
	RF-BERT	5	0.4788	0.5996
RF	BERT-BERT	5	0.5554	0.6042
	RF-TFIDF	5	0.549	0.6112
	RF-BERT	5	0.9522	0.4028
SLING	BERT-BERT	5	0.9522	0.6058
	RF-TFIDF	5	0.9522	0.5442
	RF-BERT	5	0.7708	0.5972
SPR	BERT-BERT	5	0.847	0.47
	RF-TFIDF	5	0.829	0.4334
	RF-BERT	5	0.589	0.6214
STDCXX	BERT-BERT	5	0.59	0.6322
	RF-TFIDF	5	0.512	0.551
	RF-BERT	5	0.7174	0.5176
STORM	BERT-BERT	5	0.7678	0.5322
	RF-TFIDF	5	0.802	0.5834
	RF-BERT	5	0.5272	0.532
STS	BERT-BERT	5	0.6868	0.5502
	RF-TFIDF	5	0.575	0.4868
	RF-BERT	5	0.9092	0.525
SWS	BERT-BERT	5	0.9304	0.5416
	RF-TFIDF	5	0.9304	0.3918
	RF-BERT	5	0.5606	0.5034
TS	BERT-BERT	5	0.6922	0.5872
	RF-TFIDF	5	0.6976	0.6124

ZOOKEEPER		RF-BERT	5	0.5592	0.5614
		BERT-BERT	5	0.536	0.5578
		RF-TFIDF	5	0.6096	0.5232
PORRU	MESOS	RF-BERT	5	0.3562	0.5614
		BERT-BERT	5	0.416	0.5432
		RF-TFIDF	5	0.3328	0.5654
	MULE	RF-BERT	5	0.3008	0.5786
		BERT-BERT	5	0.3542	0.5936
		RF-TFIDF	5	0.3036	0.5692
	TIMOB	RF-BERT	2	0.278	0.5175
		BERT-BERT	2	0.4195	0.52
		RF-TFIDF	2	0.3315	0.55
	XD	RF-BERT	5	0.3212	0.5646
		BERT-BERT	5	0.4262	0.6256
		RF-TFIDF	5	0.3772	0.627
DEEP-SE	APSTUD	RF-BERT	5	0.363	0.6176
		BERT-BERT	5	0.4294	0.626
		RF-TFIDF	5	0.4112	0.6226
	BAM	RF-BERT	3	0.4013	0.538
		BERT-BERT	3	0.4923	0.6027
		RF-TFIDF	3	0.3993	0.5343
	CLOV	RF-BERT	5	0.299	0.5754
		BERT-BERT	5	0.4632	0.6416
		RF-TFIDF	5	0.407	0.5914
	STUDIO	RF-BERT	5	0.3004	0.544
		BERT-BERT	5	0.4086	0.538
		RF-TFIDF	5	0.3566	0.553
	TDQ	RF-BERT	4	0.2942	0.6065
		BERT-BERT	4	0.322	0.62
		RF-TFIDF	4	0.2858	0.5842
	TESB	RF-BERT	2	0.4605	0.6555
		BERT-BERT	2	0.5035	0.6865
		RF-TFIDF	2	0.503	0.6355
	TIMOB	RF-BERT	3	0.3127	0.576
		BERT-BERT	3	0.4157	0.6537
		RF-TFIDF	3	0.3833	0.591

TABLE I: Performance of ML models and feature extraction methods. Both F-Score and AUC-ROC are used to measure the performance.