

Abstracts Only						
Model	Training Set Accuracy	Independent Test Set Accuracy	Area under the ROC curve	Precision	Recall	Fscore
BIOBERT-nsi + Logistic Regression	0.98324	0.822222	0.980764	0.856296	0.822222	0.814568
TFIDF + Logistic regression	0.9	0.666667	0.975498	0.747654	0.666667	0.664625
TFIDF + Naïve Bayes	0.748603	0.6	0.973499	0.533333	0.6	0.518413
BERTS2L + Logistic Regression	0.994413	0.688889	0.968315	0.679003	0.688889	0.670675
BIOBERT + Logistic Regression	0.944134	0.711111	0.966645	0.707884	0.711111	0.694241
BERT + Logistic Regression	0.98324	0.688889	0.965909	0.684815	0.688889	0.675132
BERT + Logistic Regression	0.972067	0.688889	0.965834	0.702963	0.688889	0.689097
TFIDF + Random Forest	0.994444	0.777778	0.961877	0.765802	0.777778	0.753072
BERTtsb + Logistic Regression	0.955307	0.688889	0.96055	0.708201	0.688889	0.675123
BIOBERT-nsi + Random Forest	0.994413	0.488889	0.95882	0.431111	0.488889	0.421481
BIOBERT-nsi + Naive Bayes	0.955307	0.644444	0.951496	0.562626	0.644444	0.569709
ELECTRA-nsi + Logistic Regression	0.977654	0.644444	0.948568	0.625269	0.644444	0.591834
BERTS2L + Naive Bayes	0.960894	0.577778	0.942748	0.586181	0.577778	0.546848
BERTtsb + Random Forest	0.994413	0.488889	0.940369	0.398413	0.488889	0.420491
ROBERTA + Logistic Regression	0.977654	0.6	0.939774	0.601913	0.6	0.578249
BIOBERT + Random Forest	0.994413	0.6	0.932603	0.626455	0.6	0.55037
BERTCL4 + Naive Bayes	0.96648	0.533333	0.930914	0.50463	0.533333	0.489931
BIOBERT + Naive Bayes	0.949721	0.577778	0.930883	0.496446	0.577778	0.510322
BERTSL4 + Naive Bayes	0.960894	0.533333	0.928548	0.506878	0.533333	0.493051
ELECTRAmed + Logistic Regression	0.932961	0.666667	0.924402	0.730529	0.666667	0.663192
BERTtsb + Naive Bayes	0.905028	0.555556	0.92319	0.504074	0.555556	0.505262
ELECTRA-nsi + Naive Bayes	0.96648	0.511111	0.918362	0.37705	0.511111	0.415873
ELECTRA-nsi + Random Forest	0.994413	0.555556	0.90856	0.611616	0.555556	0.524213
BERT + Random Forest	0.994413	0.422222	0.90767	0.376085	0.422222	0.36037
ROBERTA + Naive Bayes	0.927374	0.466667	0.896585	0.493589	0.466667	0.426781
BERT + Random Forest	0.994413	0.555556	0.893354	0.567513	0.555556	0.506897
ELMO + Logistic Regression	0.972067	0.533333	0.891973	0.579365	0.533333	0.519259
ROBERTA + Random Forest	0.994413	0.422222	0.890633	0.387542	0.422222	0.370297
ELECTRAmed + Naive Bayes	0.988827	0.488889	0.883776	0.435132	0.488889	0.439564
XLNET + Logistic Regression	0.938547	0.511111	0.880796	0.585926	0.511111	0.513698
BERTS2L + Random Forest	0.994413	0.466667	0.877471	0.488084	0.466667	0.43418
ELECTRAmed + Random Forest	0.994413	0.511111	0.870331	0.487037	0.511111	0.449214
GLOVE + Logistic Regression	0.709497	0.444444	0.850406	0.41963	0.444444	0.415564
GLOVE + Naive Bayes	0.899441	0.466667	0.849678	0.360058	0.466667	0.396878
GLOVE + Random Forest	0.994413	0.444444	0.837193	0.395476	0.444444	0.373271
ELMO + Random Forest	1	0.444444	0.815843	0.367901	0.444444	0.380998
XLNET + Random Forest	0.994413	0.244444	0.795012	0.171222	0.244444	0.161195

XLNET + Naive Bayes	0.988827	0.355556	0.780427	0.258451	0.355556	0.278361
ELECTRAmed + Logistic Regression	0.765363	0.333333	0.770364	0.298545	0.333333	0.308107
ELMO + Naive Bayes	0.670391	0.4	0.769745	0.400238	0.4	0.390599
ELECTRA + Logistic Regression	0.832402	0.288889	0.736305	0.311111	0.288889	0.285185
ELECTRAmed + Random Forest	0.994413	0.2	0.696332	0.112963	0.2	0.141019
ELECTRA + Random Forest	0.994413	0.177778	0.652757	0.093276	0.177778	0.1173
ELECTRAmed + Naive Bayes	0.636872	0.222222	0.622479	0.253535	0.222222	0.223757
ELECTRA + Naive Bayes	0.709497	0.177778	0.609068	0.17208	0.177778	0.169024

Methods Only						
Model	Training Set Accuracy	Independent Test Set Accuracy	Area under the ROC curve	Precision	Recall	Fscore
Naive Bayes	0.705556	0.577778	0.935018	0.402183	0.577778	0.460498
Logistic Regression	0.844444	0.688889	0.931979	0.711296	0.688889	0.676572
Random Forest	1	0.577778	0.916867	0.493386	0.577778	0.51675
ELMO + Logistic Regression	0.872222	0.466667	0.842787	0.415741	0.466667	0.404086
BERTtsb + Logistic Regression	0.966667	0.511111	0.835096	0.44037	0.511111	0.469979
XLNET + Logistic Regression	0.927778	0.4	0.834146	0.455185	0.4	0.402264
BERTSL4 + Random Forest	0.994444	0.333333	0.834139	0.196296	0.333333	0.217216
BIOBERT + Logistic Regression	0.972222	0.422222	0.833867	0.373333	0.422222	0.390508
BERTS2L + Logistic Regression	0.994444	0.4	0.833215	0.359259	0.4	0.355778
ROBERTA + Logistic Regression	0.988889	0.377778	0.826928	0.352381	0.377778	0.351534
BERTCL4 + Logistic Regression	0.994444	0.333333	0.825831	0.32928	0.333333	0.31317
BERTSL4 + Logistic Regression	0.994444	0.333333	0.824446	0.32928	0.333333	0.31317
XLNET + Naive Bayes	0.994444	0.444444	0.804493	0.334259	0.444444	0.347677
ROBERTA + Naive Bayes	0.944444	0.355556	0.798399	0.310741	0.355556	0.301747
ROBERTA + Random Forest	0.994444	0.4	0.796613	0.363175	0.4	0.334101
BERTtsb + Random Forest	0.994444	0.311111	0.791666	0.167778	0.311111	0.216904
GLOVE + Random Forest	1	0.4	0.791044	0.274858	0.4	0.319744
ELMO + Random Forest	1	0.377778	0.789852	0.35	0.377778	0.297428
BERTtsb + Naive Bayes	0.916667	0.355556	0.784428	0.279365	0.355556	0.308271
BIOBERT + Naive Bayes	0.961111	0.377778	0.784348	0.252848	0.377778	0.29843
BERTSL4 + Naive Bayes	0.922222	0.355556	0.769885	0.296825	0.355556	0.267901
BERTCL4 + Random Forest	0.994444	0.333333	0.766299	0.230741	0.333333	0.259867
BERTS2L + Random Forest	0.994444	0.288889	0.76291	0.198148	0.288889	0.212243
XLNET + Random Forest	1	0.355556	0.761743	0.271605	0.355556	0.269986
BERTS2L + Naive Bayes	0.927778	0.355556	0.759559	0.244781	0.355556	0.262181
GLOVE + Naive Bayes	0.794444	0.377778	0.752159	0.347593	0.377778	0.346789
GLOVE + Logistic Regression	0.461111	0.266667	0.751481	0.420741	0.266667	0.266772
BERTCL4 + Naive Bayes	0.927778	0.355556	0.750674	0.296825	0.355556	0.267901
BIOBERT + Random Forest	0.994444	0.377778	0.74734	0.274411	0.377778	0.295138
ELECTRAmed + Logistic Regression	0.564246	0.2	0.699266	0.339444	0.2	0.211445
ELECTRAmed + Random Forest	1	0.266667	0.691852	0.145481	0.266667	0.171366
ELMO + Naive Bayes	0.627778	0.244444	0.676678	0.265185	0.244444	0.227327
ELECTRA + Random Forest	0.994444	0.088889	0.611354	0.038095	0.088889	0.053333
ELECTRA + Logistic Regression	0.744444	0.111111	0.603665	0.123016	0.111111	0.09964
ELECTRA + Naive Bayes	0.683333	0.177778	0.590194	0.159259	0.177778	0.163175
ELECTRAmed + Naive Bayes	0.502793	0.177778	0.586923	0.174444	0.177778	0.158879

Abstracts + Methods						
Model	Training Set Accuracy	Independent Test Set Accuracy	Area under the ROC curve	Precision	Recall	Fscore
BERTS2L-AM + Logistic Regression	0.988827	0.733333	0.973282	0.718148	0.733333	0.714638
BERTSL4-AM + Logistic Regression	0.988827	0.711111	0.973146	0.72963	0.711111	0.702011
BERTCL4-AM + Logistic Regression	0.96648	0.777778	0.971598	0.837963	0.777778	0.784138
BIOBERT-AM + Logistic Regression	0.972067	0.733333	0.970227	0.788704	0.733333	0.725801
TFIDF-AM + Logistic Regression	0.860335	0.688889	0.957552	0.729074	0.688889	0.684494
TFIDF-AM + Random Forest	0.994413	0.666667	0.956782	0.603276	0.666667	0.617733
BERT-AM + Logistic Regression	0.949721	0.755556	0.955853	0.802963	0.755556	0.753439
TFIDF-AM + Naive Bayes	0.726257	0.688889	0.953596	0.623369	0.688889	0.611442
BERTS2L-AM + Naive Bayes	0.960894	0.577778	0.94392	0.564488	0.577778	0.543433
BERTCL4-AM + Naive Bayes	0.96648	0.577778	0.93798	0.579012	0.577778	0.524152
ROBERTA-AM + Logistic Regression	0.994413	0.533333	0.934673	0.570696	0.533333	0.510534
BERTSL4-AM + Naive Bayes	0.96648	0.555556	0.927369	0.534858	0.555556	0.513358
BERT-AM + Naive Bayes	0.916201	0.577778	0.921063	0.554656	0.577778	0.54598
BERTS2L-AM + Random Forest	0.994413	0.533333	0.918612	0.542929	0.533333	0.495388
BIOBERT-AM + Random Forest	0.994413	0.555556	0.916674	0.475992	0.555556	0.479043
BERTSL4-AM + Random Forest	0.994413	0.577778	0.916069	0.585079	0.577778	0.546667
BERTCL4-AM + Random Forest	0.994413	0.488889	0.915413	0.470853	0.488889	0.438849
BERT-AM + Random Forest	0.994413	0.466667	0.907761	0.42963	0.466667	0.400808
BIOBERT-AM + Naive Bayes	0.949721	0.511111	0.904857	0.385185	0.511111	0.430963
ELMO-AM + Logistic Regression	0.698324	0.466667	0.88759	0.372646	0.466667	0.392232
ROBERTA-AM + Naive Bayes	0.932961	0.466667	0.885442	0.454435	0.466667	0.417847
ROBERTA-AM + Random Forest	0.994413	0.422222	0.874995	0.424675	0.422222	0.366499
ELMO-AM + Random Forest	1	0.466667	0.867663	0.345036	0.466667	0.37875
XLNET-AM + Logistic Regression	0.994413	0.488889	0.858765	0.505714	0.488889	0.460741
XLNET-AM + Naive Bayes	0.994413	0.422222	0.833238	0.385438	0.422222	0.342424
GLOVE-AM + Naive Bayes	0.73743	0.422222	0.807296	0.422672	0.422222	0.399612
ELECTRAMED-AM + Logistic Regression	0.486034	0.333333	0.79274	0.35037	0.333333	0.325926
GLOVE-AM + Logistic Regression	0.49162	0.288889	0.791933	0.28836	0.288889	0.246384
GLOVE-AM + Random Forest	1	0.333333	0.782005	0.225265	0.333333	0.265085
XLNET-AM + Random Forest	0.994413	0.288889	0.759621	0.216296	0.288889	0.209492
ELECTR-AM + Logistic Regression	0.575419	0.244444	0.747428	0.336905	0.244444	0.23124
ELECTRA-AM + Random Forest	0.994413	0.288889	0.717559	0.158333	0.288889	0.197114
ELECTRAMED-AM + Random Forest	1	0.288889	0.716509	0.215079	0.288889	0.227063
ELMO-AM + Naive Bayes	0.486034	0.244444	0.691356	0.289877	0.244444	0.237773
ELECTRA-AM + Naive Bayes	0.675978	0.155556	0.654665	0.123272	0.155556	0.132999
ELECTRAMED-AM + Naive Bayes	0.52514	0.155556	0.653367	0.124444	0.155556	0.133479

	All AUC-PR scores								
	Abstracts Only			Methods Only			Abstracts + Methods		
	LR	RF	NB	LR	RF	NB	LR	RF	NB
BioBert-hf	*0.85	0.61	0.56	-	-	-	-	-	-
BERTS2L	*0.83	0.47	0.53	*0.44	0.31	0.27	*0.84	0.54	0.53
BERTSL4	*0.81	0.5	0.47	*0.4	0.37	0.27	*0.83	0.52	0.49
BERTCL4	*0.79	0.52	0.49	*0.41	0.33	0.24	*0.82	0.55	0.54
BERT-st	*0.75	0.54	0.53	*0.48	0.35	0.3	*0.73	0.58	0.54
BioBERT-nlu	*0.73	0.63	0.49	*0.46	0.33	0.28	*0.77	0.57	0.48
TFIDF	0.7	*0.8	0.71	0.62	*0.66	0.61	0.73	*0.74	0.71
ELECTRA-hf	*0.64	0.5	0.51	-	-	-	-	-	-
ELECTRAmed-hf	*0.64	0.49	0.41	-	-	-	-	-	-
RoBERTa-st	*0.61	0.5	0.43	*0.38	0.36	0.31	*0.59	0.46	0.4
ELMO	*0.5	0.46	0.31	*0.48	0.38	0.19	*0.51	0.48	0.17
XLNET-nu	*0.46	0.33	0.33	*0.4	0.32	0.34	0.43	0.3	*0.44
GLOVE-nlu	0.33	*0.44	0.36	0.2	*0.36	0.29	0.22	*0.35	0.33
ELECTRAmed-st	*0.24	0.21	0.15	0.17	*0.2	0.13	0.22	*0.27	0.16
ELECTRA-nlu	*0.2	0.17	0.14	0.12	*0.18	0.14	0.2	*0.22	0.14