

Appendices

A Experimental Results

Below, we present full results for all our experiments in the form of tabular results and learning curves. Tables 1 and 2 enumerate performance metrics for all source, successor, acquisition function combinations after acquiring 10% and 20% of the pool. Figure 1 shows the learning curves for all combinations. We report all average Spearman’s rank correlation coefficients in Table 3.

Text classification										
Successor	pool %	i.i.d.	Acquisition model							
			Uncertainty			QBC			BALD	
			SVM	CNN	LSTM	SVM	CNN	LSTM	CNN	LSTM
Movie reviews										
SVM	10	65.3	65.3	65.8	65.7	64.9	64.9	65.1	64.9	65.2
	20	68.2	69.0	69.4	68.9	68.1	68.4	68.7	68.5	69.0
CNN	10	65.0	65.3	65.5	65.4	64.8	65.1	64.7	65.1	64.9
	20	69.4	69.1	69.5	69.5	68.5	69.1	69.1	68.3	69.1
LSTM	10	63.0	62.0	62.5	63.1	61.9	61.9	62.6	61.7	62.2
	20	67.2	65.1	65.8	67.0	65.4	65.7	66.8	65.6	67.1
Subjectivity										
SVM	10	85.2	85.6	85.3	85.5	85.4	85.0	85.4	85.8	85.4
	20	87.5	87.6	87.4	87.6	87.7	87.0	87.5	87.0	87.6
CNN	10	85.3	85.2	86.3	86.0	85.3	86.0	85.7	86.2	85.7
	20	87.9	87.6	88.4	88.6	88.4	88.5	88.6	88.6	88.3
LSTM	10	82.9	82.7	82.7	84.1	83.3	83.7	84.8	83.1	84.2
	20	86.7	86.3	85.8	87.6	86.9	87.0	87.7	84.7	87.0
TREC										
SVM	10	68.5	68.3	66.8	68.5	68.1	63.1	64.9	68.2	68.3
	20	74.1	74.7	73.2	74.3	73.7	71.6	71.2	74.1	74.1
CNN	10	70.9	70.5	69.0	70.0	67.4	62.8	69.5	71.0	70.5
	20	76.1	77.7	77.3	78.0	76.5	73.7	76.3	79.8	77.7
LSTM	10	65.2	64.5	63.6	63.8	61.7	60.1	64.6	64.1	64.5
	20	71.5	72.7	71.0	73.3	71.4	69.9	71.8	72.9	72.6
Customer reviews										
SVM	10	68.8	70.5	70.3	68.5	70.5	69.5	64.6	70.0	69.2
	20	73.6	74.2	72.9	71.1	73.8	72.6	65.7	73.5	71.7
CNN	10	70.6	70.9	71.7	68.2	71.5	71.4	63.8	72.2	68.4
	20	74.1	74.5	74.8	71.5	74.9	74.9	65.2	75.3	71.3
LSTM	10	66.1	67.2	65.1	65.9	65.0	64.8	64.0	65.2	65.4
	20	68.0	66.6	66.5	66.3	66.3	66.4	65.4	68.3	68.0

Table 1: Text classification accuracy, evaluated for each combination of acquisition and successor models using uncertainty sampling, QBC, and BALD. Accuracies are reported for training sets composed of 10% and 20% of the document pool. Colors indicate performance relative to i.i.d. baselines: Blue implies that a model fared better, red that it performed worse, and black that it performed the same.

Named Entity Recognition							
Successor	pool %	i.i.d.	Acquisition Model				
			Uncertainty		BALD		QBC
			CRF	BiLSTM-CNN	BiLSTM-CNN	CRF	BiLSTM-CNN
CoNLL							
CRF	10	69.2	70.5	70.2	70.3	70.3	70.0
	20	73.6	74.4	74.0	74.1	74.5	74.1
BiLSTM-CNN	10	87.4	87.4	87.8	88.0	87.5	87.7
	20	89.1	89.6	89.6	89.8	89.2	89.5

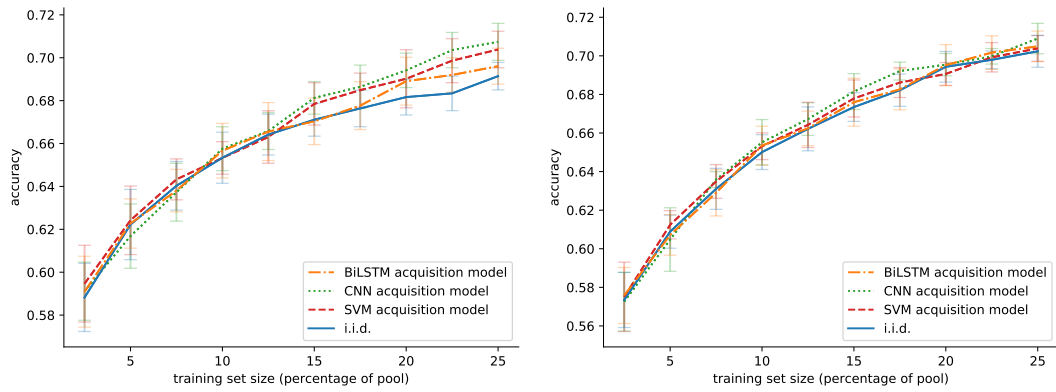
Successor	pool %	i.i.d.	Acquisition Model		
			Uncertainty		BALD
			CRF	BiLSTM-CNN	BiLSTM-CNN
OntoNotes					
CRF	10	73.8	75.5	75.4	75.3
	20	77.6	79.1	78.7	78.7
BiLSTM-CNN	10	82.6	83.1	83.1	83.2
	20	84.6	85.2	84.9	85.1

Table 2: F1 measurements for the NER task, with training sets comprising 10% and 20% of the training pool.

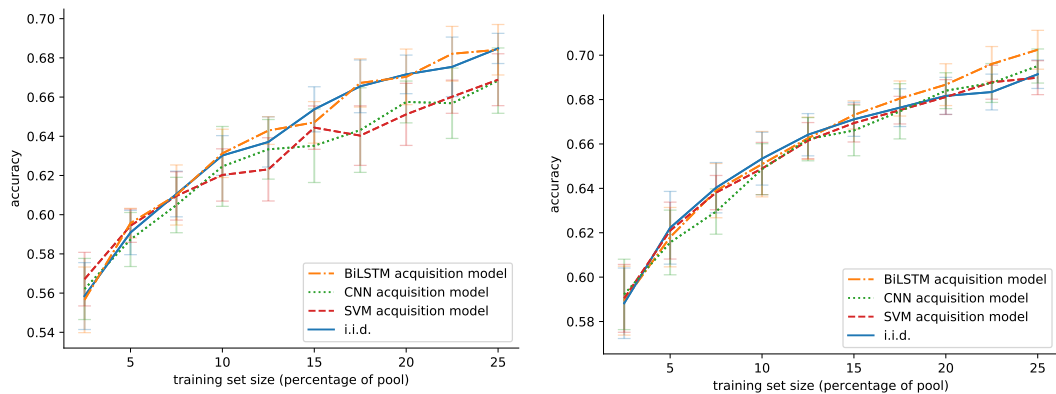
Acquisition Model	Successor							
	Movie Reiews		Subjectivity		TREC		Customer Reviews	
	CNN	LSTM	CNN	LSTM	CNN	LSTM	CNN	LSTM
Uncertainty Sampling								
CNN	–	0.961	–	0.968	–	0.988	–	0.973
LSTM	0.989	–	0.996	–	0.992	–	0.980	–
SVM	0.991	0.961	0.997	0.970	0.990	0.987	0.991	0.974
QBC								
CNN	–	0.956	–	0.970	–	0.985	–	0.972
LSTM	0.989	–	0.996	–	0.990	–	0.988	–
SVM	0.995	0.962	0.997	0.970	0.985	0.986	0.993	0.974
BALD								
CNN	–	0.963	–	0.969	–	0.988	–	0.974
LSTM	0.991	–	0.995	–	0.991	–	0.982	–

Table 3: Average Spearman’s rank correlation coefficients of cosine distances between test set representations learned with native active learning and distances between those learned with transferred actively acquired datasets.

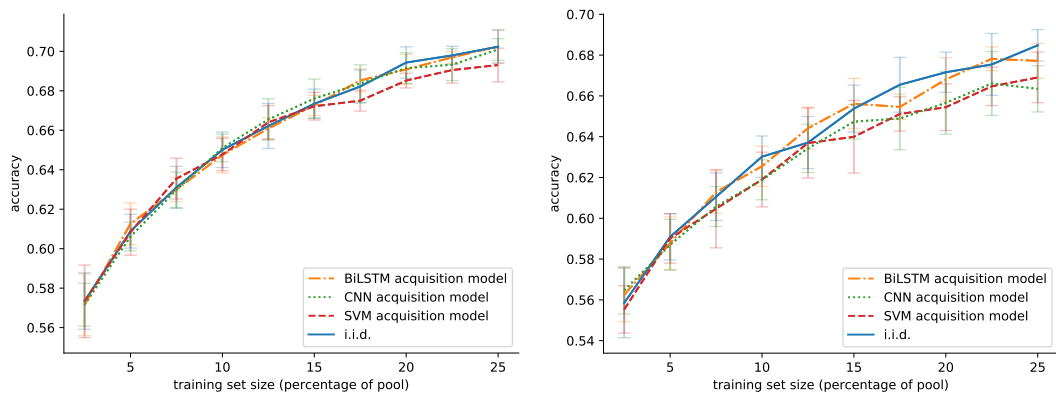
Figure 1: This appendix contains the full set of collected learning curves for the text classification and NER. Error bars represent one standard deviation.



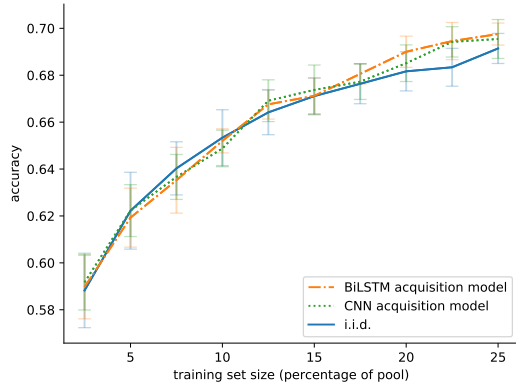
(1) SVM on Movie Reviews dataset using max entropy (2) CNN on Movie Reviews dataset using max entropy



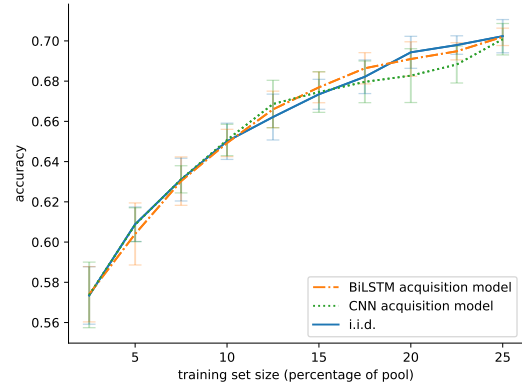
(3) BiLSTM on Movie Reviews dataset using max entropy (4) SVM on Movie Reviews dataset using QBC



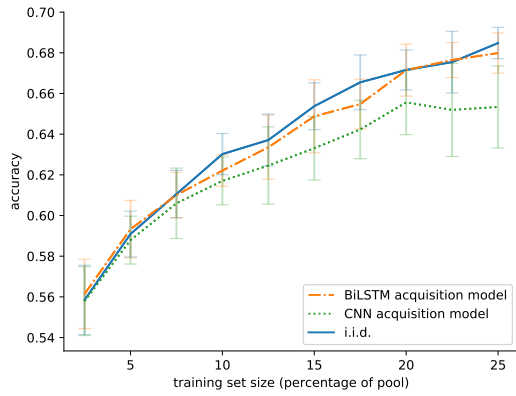
(5) CNN on Movie Reviews dataset using QBC (6) BiLSTM on Movie Reviews dataset using QBC



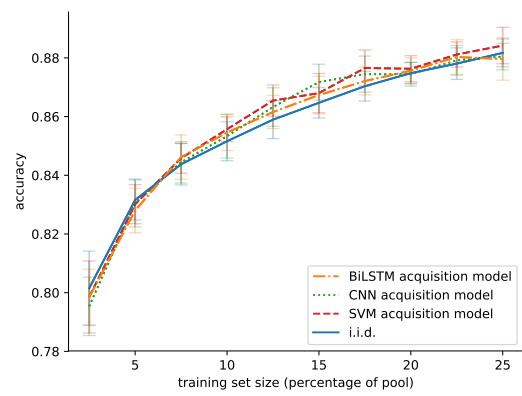
(7) SVM on Movie Reviews dataset using BALD



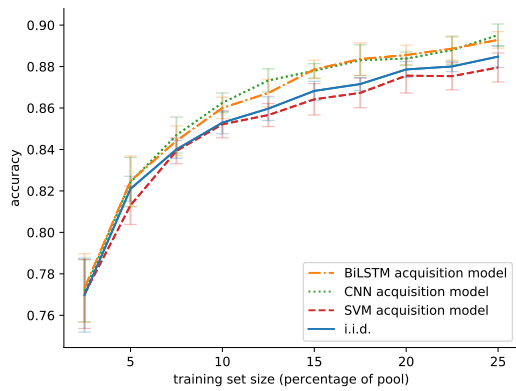
(8) CNN on Movie Reviews dataset using BALD



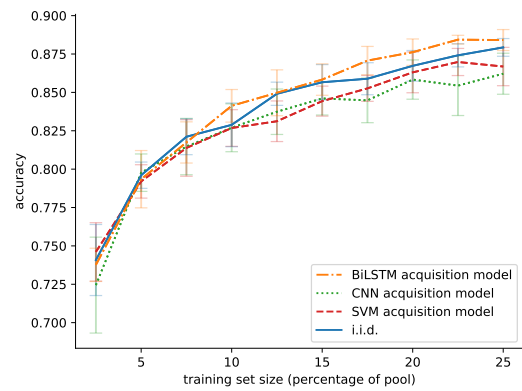
(9) BiLSTM on Movie Reviews dataset using BALD



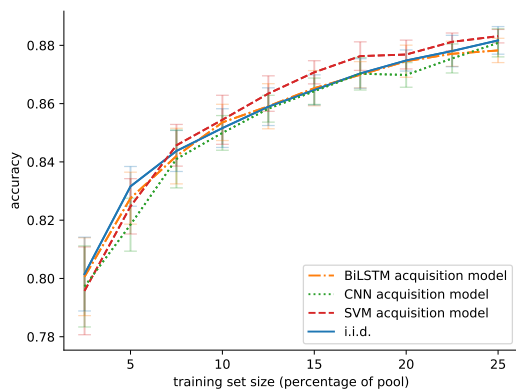
(10) SVM on Subjectivity dataset using max entropy



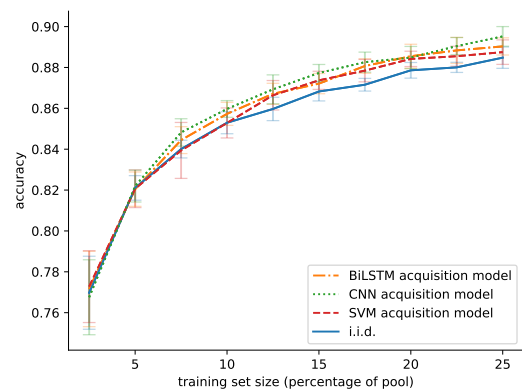
(11) CNN on Subjectivity dataset using max entropy



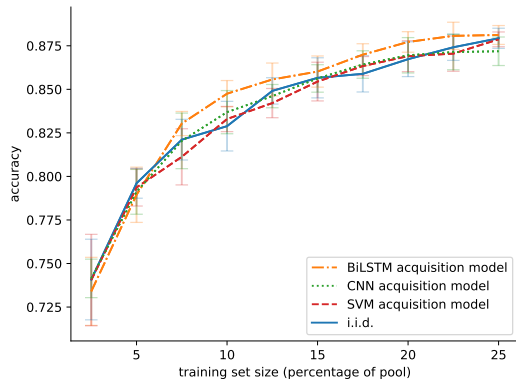
(12) BiLSTM on Subjectivity dataset using max entropy



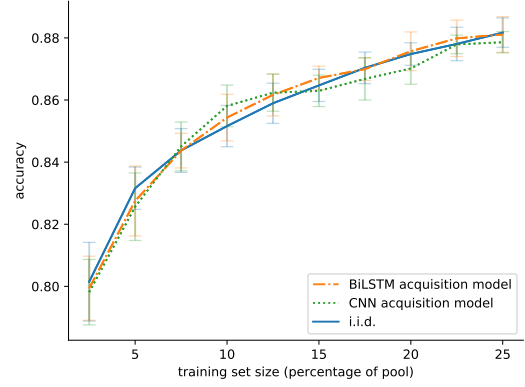
(13) SVM on Subjectivity dataset using QBC



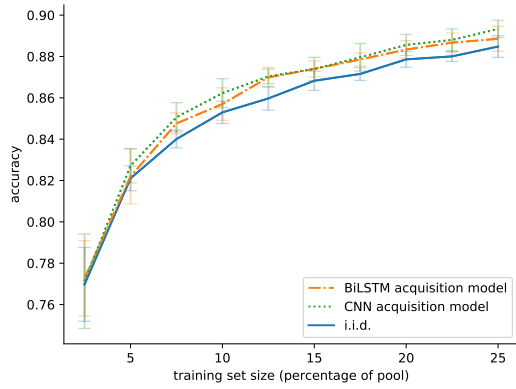
(14) CNN on Subjectivity dataset using QBC



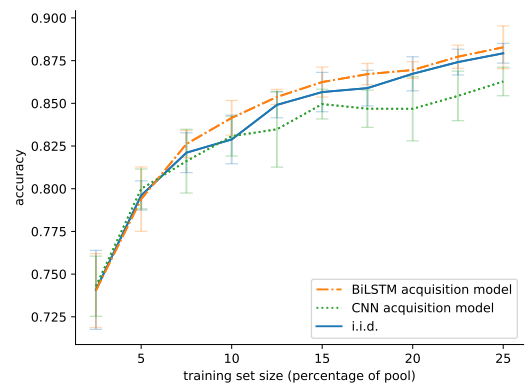
(15) BiLSTM on Subjectivity dataset using QBC



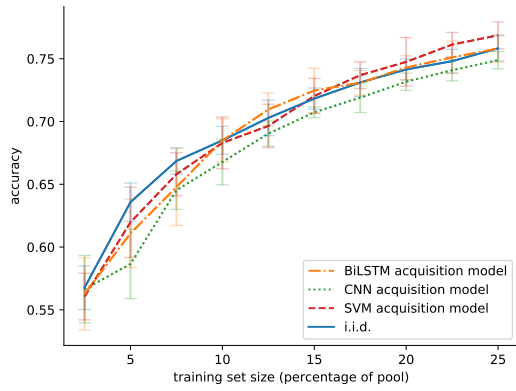
(16) SVM on Subjectivity dataset using BALD



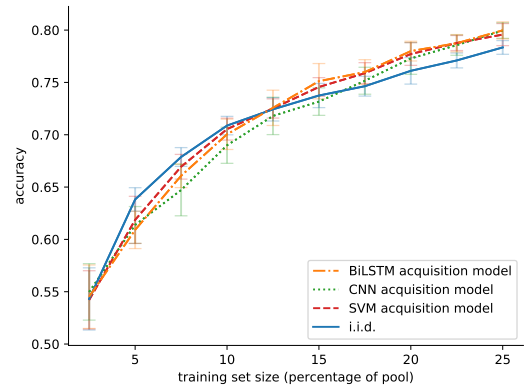
(17) CNN on Subjectivity dataset using BALD



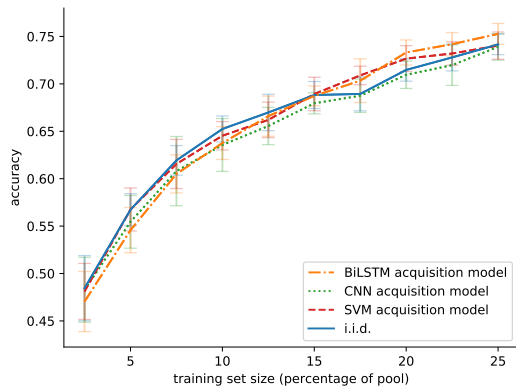
(18) BiLSTM on Subjectivity dataset using BALD



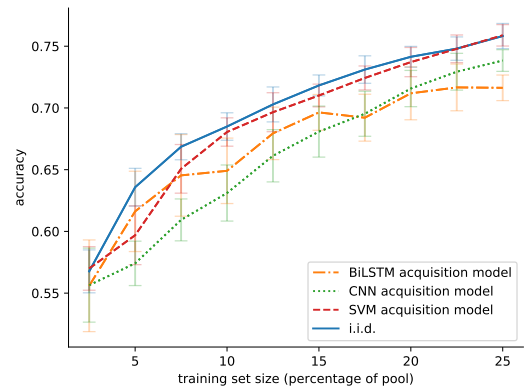
(19) SVM on TREC dataset using max entropy



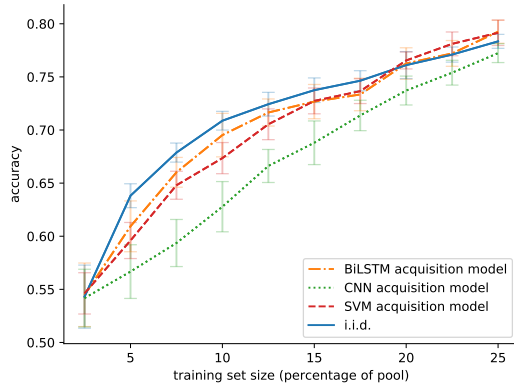
(20) CNN on TREC dataset using max entropy



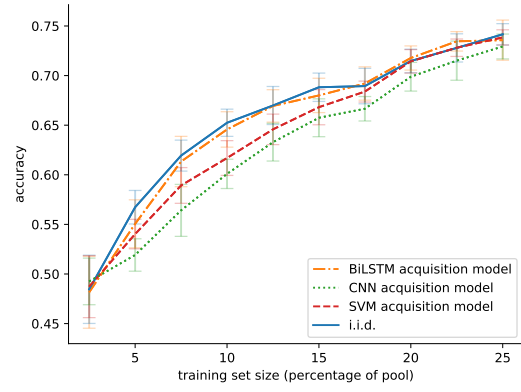
(21) BiLSTM on TREC dataset using max entropy



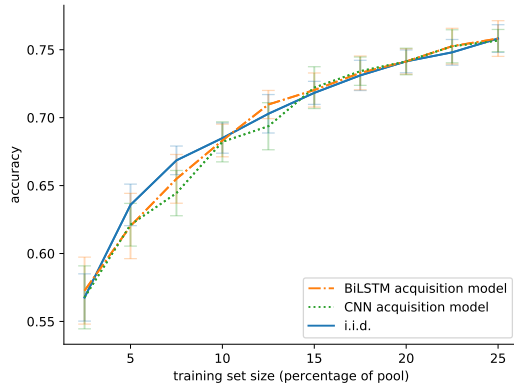
(22) SVM on TREC dataset using QBC



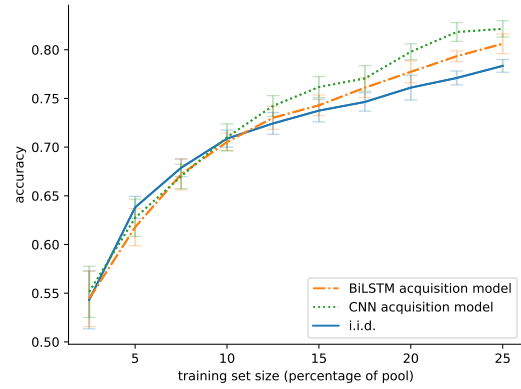
(23) CNN on TREC dataset using QBC



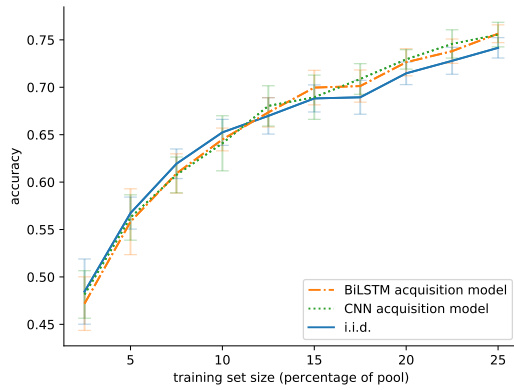
(24) BiLSTM on TREC dataset using QBC



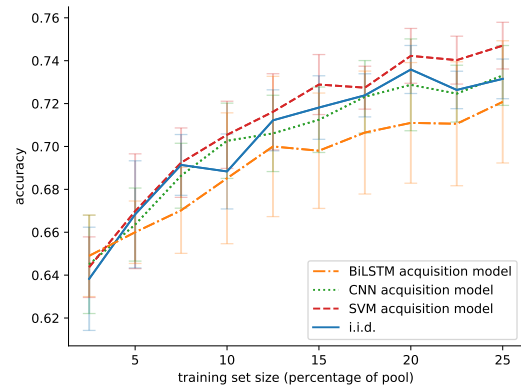
(25) SVM on TREC dataset using BALD



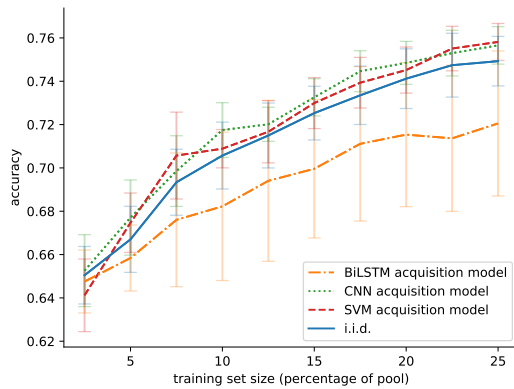
(26) CNN on TREC dataset using BALD



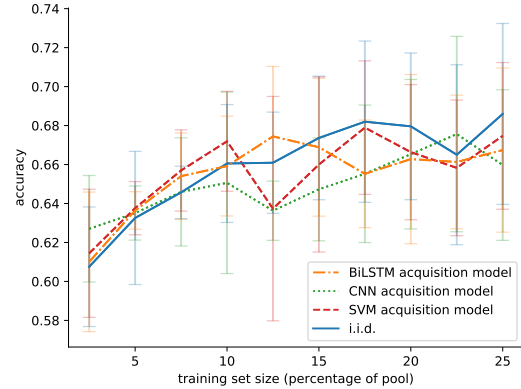
(27) BiLSTM on TREC dataset using BALD



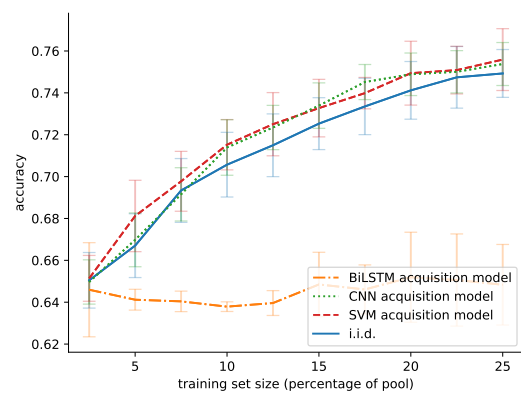
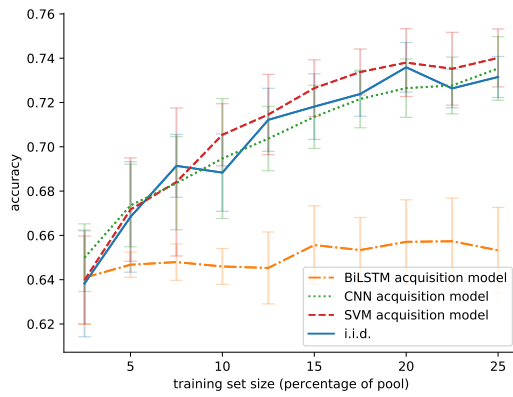
(28) SVM on Customer Review dataset using max entropy



(29) CNN on Customer Review dataset using max entropy

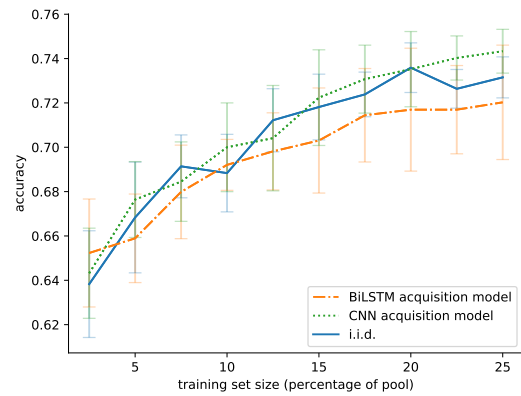
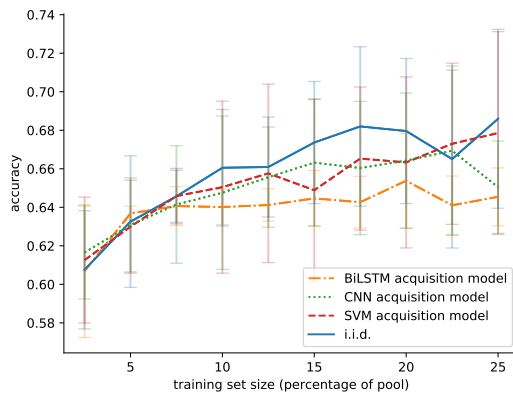


(30) BiLSTM on Customer Review dataset using max entropy



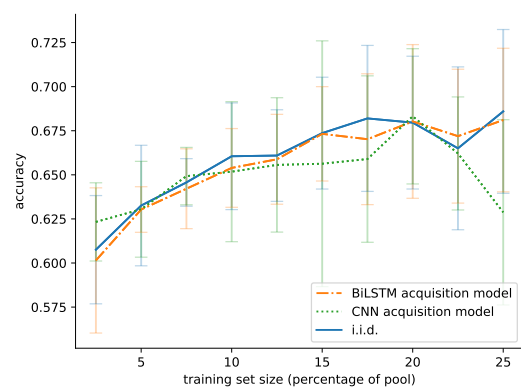
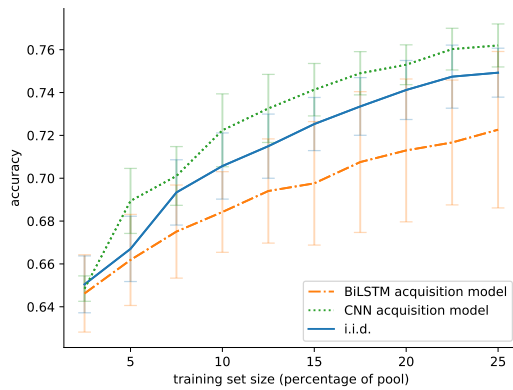
(31) SVM on Customer Review dataset using QBC

(32) CNN on Customer Review dataset using QBC



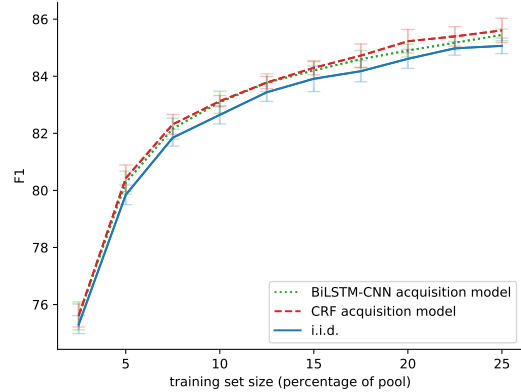
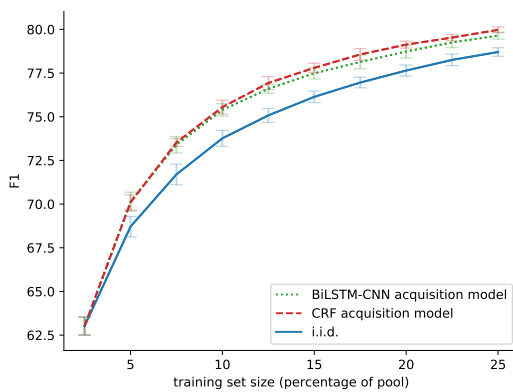
(33) BiLSTM on Customer Review dataset using QBC

(34) SVM on Customer Review dataset using BALD



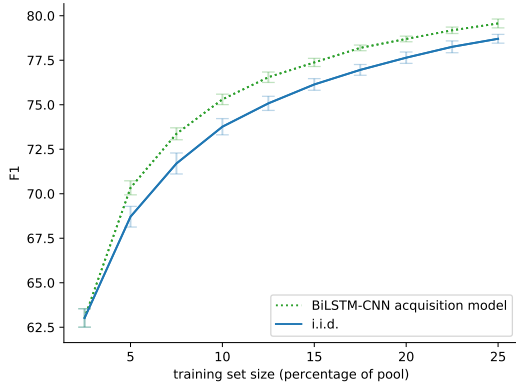
(35) CNN on Customer Review dataset using BALD

(36) BiLSTM on Customer Review dataset using BALD

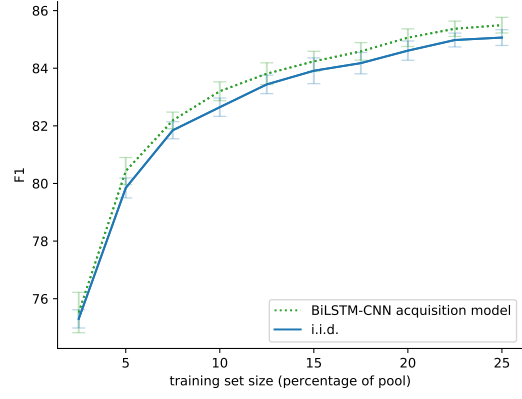


(37) CRF on OntoNotes dataset using max entropy

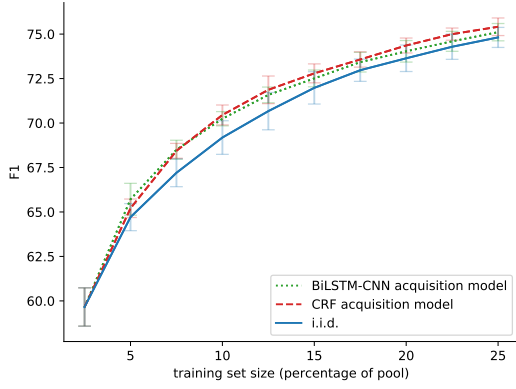
(38) BiLSTM-CNN on OntoNotes dataset using max entropy



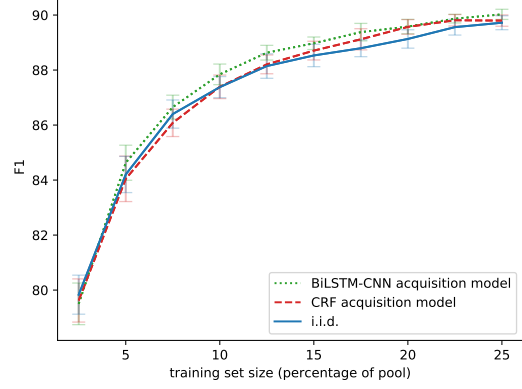
(39) CRF on OntoNotes dataset using BALD



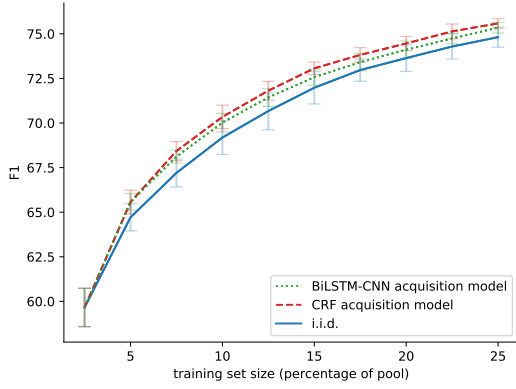
(40) BiLSTM-CNN on OntoNotes dataset using BALD



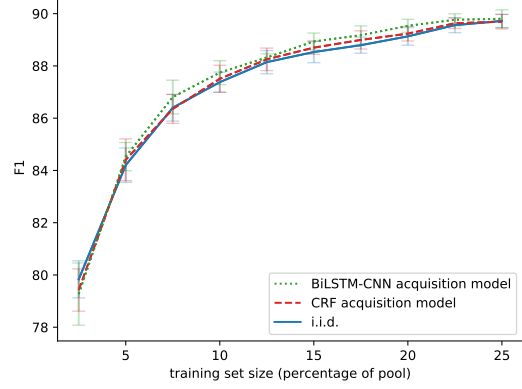
(41) CRF on CoNLL dataset using max entropy



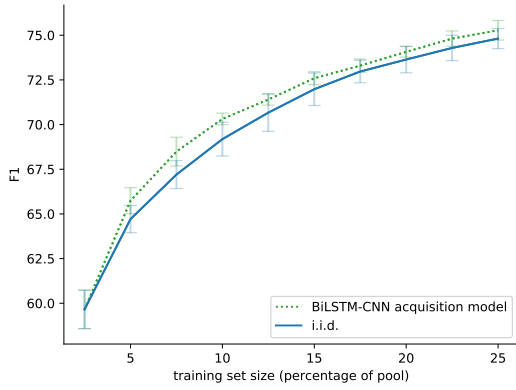
(42) BiLSTM-CNN on CoNLL dataset using max entropy



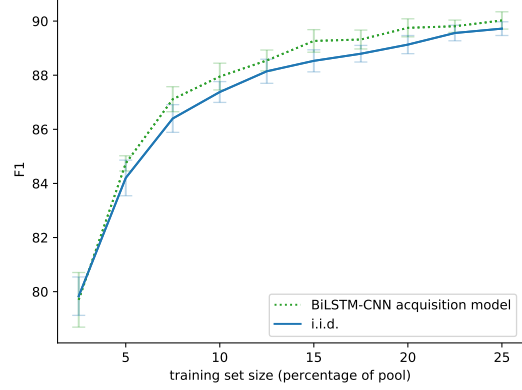
(43) CRF on CoNLL dataset using QBC



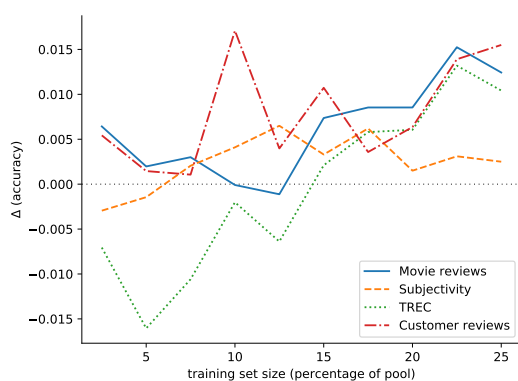
(44) BiLSTM-CNN on CoNLL dataset using QBC



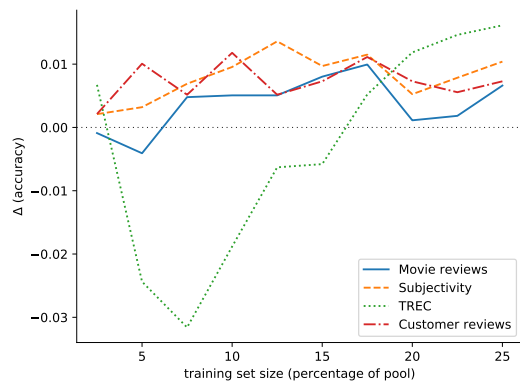
(45) CRF on CoNLL dataset using BALD



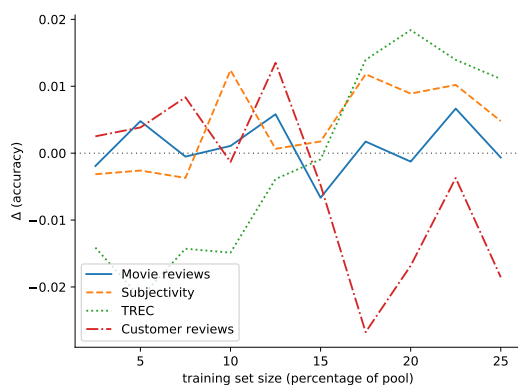
(46) BiLSTM-CNN on CoNLL dataset using BALD



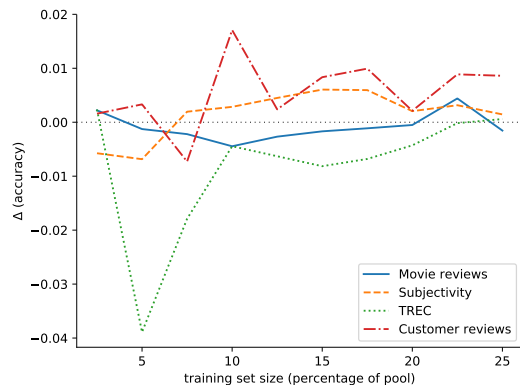
(47) Δ for SVM using max entropy



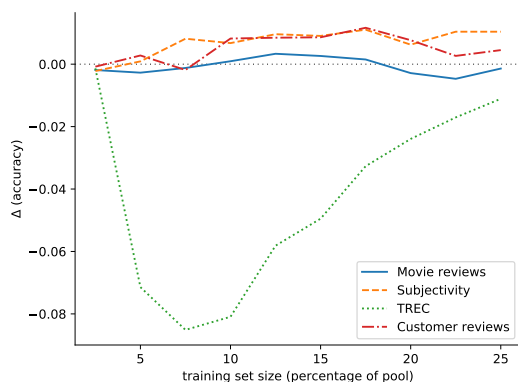
(48) Δ for CNN using max entropy



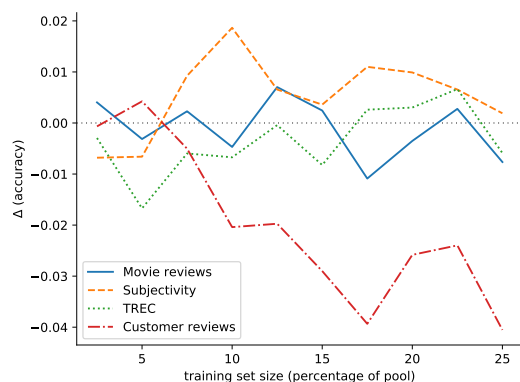
(49) Δ for BiLSTM using max entropy



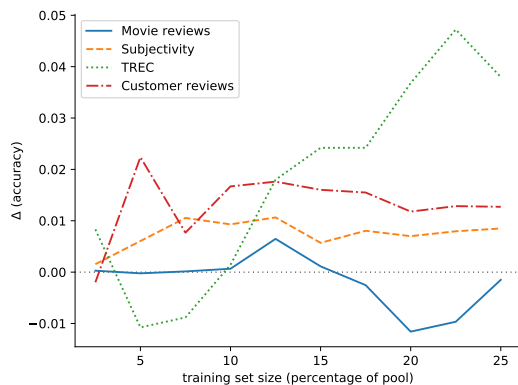
(50) Δ for SVM using QBC



(51) Δ for CNN using QBC



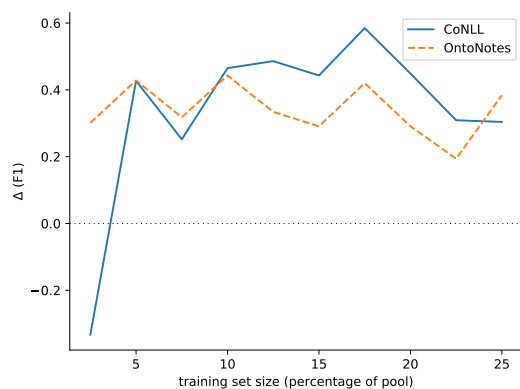
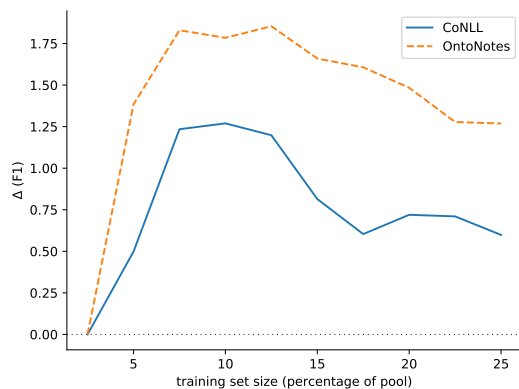
(52) Δ for BiLSTM using QBC



(53) Δ for CNN using BALD

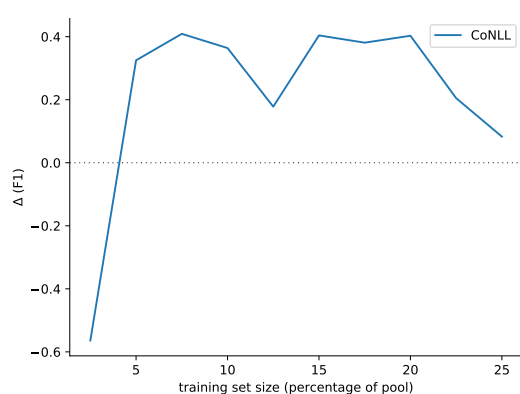
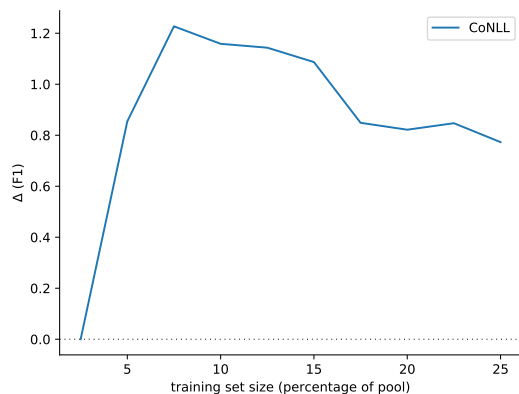


(54) Δ for BiLSTM using BALD



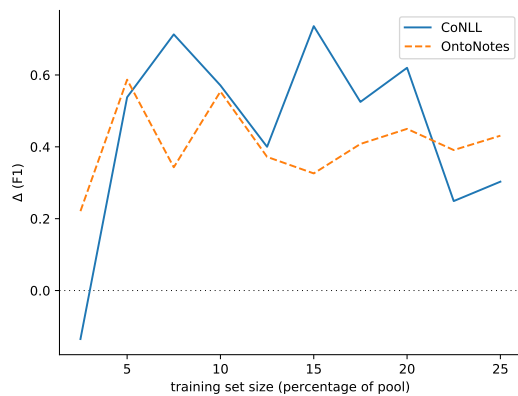
(55) Δ for CRF using max entropy

(56) Δ for BiLSTM-CNN using max entropy



(57) Δ for CRF using QBC

(58) Δ for BiLSTM-CNN using QBC



(59) Δ for BiLSTM-CNN using BALD