

Survey Statistics	
Viewed	297
Started	54
Completed	54
Completion Rate	100%
Drop Outs (After Starting)	0
» Average time taken to complete survey : 12 minute(s)	

Have you ever used Active Learning in order to speed up annotation/labeling work of any linguistic data?

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Yes	11	20.37%					
2.	No	43	79.63%					
	Total	54	100%					

Key Analytics

Mean	1.796
Confidence Interval @ 95%	[1.688 - 1.905] n = 54
Standard Deviation	0.407
Standard Error	0.055

Which of the following is the annotation project (where you used AL) you are reporting on? (Select one alternative)

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Most recent one	3	27.27%	<input type="text"/>				
2.	Most successful one	2	18.18%	<input type="text"/>				
3.	Largest one in terms of labeled units	2	18.18%	<input type="text"/>				
4.	Largest one in terms of spent annotation time	1	9.09%	<input type="text"/>				
5.	Most interesting one	3	27.27%	<input type="text"/>				
6.	Other	0	0.00%	<input type="text"/>				
	Total	11	100%					

Key Analytics

Mean	2.909	<p>Key Facts</p> <p>➔ 54.55% chose the following options :</p> <ul style="list-style-type: none"> ➤ Most recent one ➤ Most interesting one
Confidence Interval @ 95%	[1.940 - 3.879] n = 11	
Standard Deviation	1.640	
Standard Error	0.495	

What kind of annotation had to be done? (Multiple answers allowed)

Frequency Analysis				
	Answer	Count	Percent	
				20% 40% 60% 80% 100%
1.	Document classification	3	17.65%	<input type="text"/>
2.	Part-of-speech	1	5.88%	<input type="text"/>
3.	Chunk	1	5.88%	<input type="text"/>
4.	Parse	0	0.00%	<input type="text"/>
5.	Word senses	2	11.76%	<input type="text"/>
6.	Sentiment	0	0.00%	<input type="text"/>
7.	Named entity	5	29.41%	<input type="text"/>
8.	Relations/Events	4	23.53%	<input type="text"/>
9.	Other	1	5.88%	<input type="text"/>
	Total	17	100%	

Key Analytics		
Mean	5.529	Key Facts <input type="checkbox"/> 52.94% chose the following options : <ul style="list-style-type: none"> ➤ Named entity ➤ Relations/Events
Confidence Interval @ 95%	[4.184 - 6.875] n = 17	
Standard Deviation	2.831	
Standard Error	0.687	

From which domain(s) where the annotated instances? (Multiple answers allowed)

Frequency Analysis				
	Answer	Count	Percent	20% 40% 60% 80% 100%
1.	Newswire	3	25.00%	<input type="text"/>
2.	Bio-medical	4	33.33%	<input type="text"/>
3.	Other	5	41.67%	<input type="text"/>
	Total	12	100%	

Key Analytics		
Mean	2.167	Key Facts <input type="checkbox"/> 75% chose the following options : > Other > Bio-medical <input type="checkbox"/> Least chosen option 25% : > Newswire
Confidence Interval @ 95%	[1.694 - 2.639] n = 12	
Standard Deviation	0.835	
Standard Error	0.241	

What was the principal language? (Select one alternative)

Frequency Analysis				
	Answer	Count	Percent	
				20% 40% 60% 80% 100%
1.	English	11	100.00%	<input type="text"/>
2.	Hindi	0	0.00%	<input type="text"/>
3.	Mandarin	0	0.00%	<input type="text"/>
4.	Arabic	0	0.00%	<input type="text"/>
5.	Spanish	0	0.00%	<input type="text"/>
6.	French	0	0.00%	<input type="text"/>
7.	German	0	0.00%	<input type="text"/>
8.	Other	0	0.00%	<input type="text"/>
	Total	11	100%	

Key Analytics	
Mean	1.000
Confidence Interval @ 95%	[1.000 - 1.000] n = 11
Standard Deviation	0.000
Standard Error	0.000

How did you assemble the corpus of unlabeled annotation instances (i.e., the pool from which your AL mechanism selected)? (Select one alternative)

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	We took all annotation instances available to us	3	27.27%	<input type="text"/>				
2.	Random subsample of all annotation available instances	3	27.27%	<input type="text"/>				
3.	Subsample of all annotation available instances based on introspection	0	0.00%	<input type="text"/>				
4.	Subsample of all annotation available instances based on key words	3	27.27%	<input type="text"/>				
5.	Other	2	18.18%	<input type="text"/>				
	Total	11	100%					

Key Analytics

Mean	2.818	Key Facts <input type="checkbox"/> 54.55% chose the following options : <ul style="list-style-type: none"> ➤ We took all annotation instances available to us ➤ Random subsample of all annotation available instances
Confidence Interval @ 95%	[1.872 - 3.764] n = 11	
Standard Deviation	1.601	
Standard Error	0.483	

What kind(s) of active learning approach have you used? (Multiple alternatives allowed)

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Uncertainty sampling	2	18.18%	<input type="text"/>				
2.	Query-by-Committee (including variants such as Query-by-Bagging, Query-by-Boosting, Co-testing etc)	2	18.18%	<input type="text"/>				
3.	Variance Reduction	0	0.00%	<input type="text"/>				
4.	Expected error reduction	2	18.18%	<input type="text"/>				
5.	Other	5	45.45%	<input type="text"/>				
	Total	11	100%					

Key Analytics

Mean	3.545	<p>Key Facts</p> <p>➤ 63.64% chose the following options :</p> <ul style="list-style-type: none"> ➤ Other ➤ Uncertainty sampling
Confidence Interval @ 95%	[2.544 - 4.547] n = 11	
Standard Deviation	1.695	
Standard Error	0.511	

How did you decide when to stop the annotation process? (Select one alternative)

Frequency Analysis				
	Answer	Count	Percent	20% 40% 60% 80% 100%
1.	Evaluation of model learned on annotations with a held-out gold standard	4	36.36%	<input type="text"/>
2.	Money and/or time available for annotation used up	4	36.36%	<input type="text"/>
3.	All relevant documents annotated	1	9.09%	<input type="text"/>
4.	Expected gains in model performance compared to additional annotation costs below critical threshold	1	9.09%	<input type="text"/>
5.	Other	1	9.09%	<input type="text"/>
	Total	11	100%	

Key Analytics		
Mean	2.182	Key Facts ➤ 72.73% chose the following options : ➤ Evaluation of model learned on annotations with a held-out gold standard ➤ Money and/or time available for annotation used up ➤ Least chosen option 9.09% : ➤ All relevant documents annotated
Confidence Interval @ 95%	[1.397 - 2.967] n = 11	
Standard Deviation	1.328	
Standard Error	0.400	

What was your primary motivation to use AL? (Select one alternative)

Frequency Analysis								
	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Minimize labeling cost	8	72.73%	<input type="text"/>				
2.	Gather experience with/test AL	0	0.00%	<input type="text"/>				
3.	Test your annotation guidelines against hard cases	2	18.18%	<input type="text"/>				
4.	Other	1	9.09%	<input type="text"/>				
	Total	11	100%					

Key Analytics			Key Facts + 90.91% chose the following options : <ul style="list-style-type: none"> ➤ Minimize labeling cost ➤ Test your annotation guidelines against hard cases
Mean		1.636	
Confidence Interval @ 95%		[0.974 - 2.298] n = 11	
Standard Deviation		1.120	
Standard Error		0.338	

Have your expectations with respect to AL been met? (Select one alternative)

Frequency Analysis				
	Answer	Count	Percent	20% 40% 60% 80% 100%
1.	Yes	4	36.36%	<input type="text"/>
2.	No	1	9.09%	<input type="text"/>
3.	Partially	6	54.55%	<input type="text"/>
	Total	11	100%	

Key Analytics		
Mean	2.182	Key Facts <input checked="" type="checkbox"/> 90.91% chose the following options : > Partially > Yes <input checked="" type="checkbox"/> Least chosen option 9.09% : > No
Confidence Interval @ 95%	[1.602 - 2.762] n = 11	
Standard Deviation	0.982	
Standard Error	0.296	

Would you use AL again in an annotation project of this type and scope? (Select one alternative)

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Yes	11	100.00%	<input type="text"/>				
2.	No, because	0	0.00%	<input type="text"/>				
	Total	11	100%					

Key Analytics

Mean	1.000
Confidence Interval @ 95%	[1.000 - 1.000] n = 11
Standard Deviation	0.000
Standard Error	0.000

Which of the following is the annotation project you are reporting on? (Select one alternative)

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Most recent one	28	65.12%	<input type="text"/>				
2.	Most successful one	5	11.63%	<input type="text"/>				
3.	Largest one in terms of labeled units	7	16.28%	<input type="text"/>				
4.	Largest one in terms of spent annotation time	1	2.33%	<input type="text"/>				
5.	Most interesting one	1	2.33%	<input type="text"/>				
6.	Other	1	2.33%	<input type="text"/>				
	Total	43	100%					

Key Analytics

Mean	1.721	<p>Key Facts</p> <ul style="list-style-type: none"> ➤ 81.4% chose the following options : <ul style="list-style-type: none"> ➤ Most recent one ➤ Largest one in terms of labeled units ➤ Least chosen option 2.33% : <ul style="list-style-type: none"> ➤ Largest one in terms of spent annotation time
Confidence Interval @ 95%	[1.362 - 2.080] n = 43	
Standard Deviation	1.202	
Standard Error	0.183	

What kind of annotation had to be done? (Multiple alternatives allowed)

Frequency Analysis								
	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Document classification	4	5.26%	<input type="checkbox"/>				
2.	Part-of-speech	10	13.16%	<input type="checkbox"/>				
3.	Chunk	4	5.26%	<input type="checkbox"/>				
4.	Parse	7	9.21%	<input type="checkbox"/>				
5.	Word senses	4	5.26%	<input type="checkbox"/>				
6.	Sentiment	2	2.63%	<input type="checkbox"/>				
7.	Named entity	12	15.79%	<input type="checkbox"/>				
8.	Relations/Events	13	17.11%	<input type="checkbox"/>				
9.	Other	20	26.32%	<input type="checkbox"/>				
	Total	76	100%					

Key Analytics		
Mean	6.105	Key Facts <input type="checkbox"/> 43.42% chose the following options : <ul style="list-style-type: none"> ➤ Other ➤ Relations/Events <input type="checkbox"/> Least chosen option 2.63% : <ul style="list-style-type: none"> ➤ Sentiment
Confidence Interval @ 95%	[5.485 - 6.726] n = 76	
Standard Deviation	2.760	
Standard Error	0.317	

From which domain(s) where the annotated instances? (Multiple alternatives allowed)

Frequency Analysis				
	Answer	Count	Percent	20% 40% 60% 80% 100%
1.	Newswire	9	18.37%	<input type="text"/>
2.	Bio-medical	13	26.53%	<input type="text"/>
3.	Other	27	55.10%	<input type="text"/>
	Total	49	100%	

Key Analytics		
Mean	2.367	Key Facts <input type="checkbox"/> 81.63% chose the following options : <ul style="list-style-type: none"> ➤ Other ➤ Bio-medical <input type="checkbox"/> Least chosen option 18.37% : <ul style="list-style-type: none"> ➤ Newswire
Confidence Interval @ 95%	[2.148 - 2.586] n = 49	
Standard Deviation	0.782	
Standard Error	0.112	

What was the principal language of the annotated texts? (Select one alternative)

Frequency Analysis				
	Answer	Count	Percent	
				20% 40% 60% 80% 100%
1.	English	27	62.79%	<input type="text"/>
2.	Hindi	1	2.33%	<input type="text"/>
3.	Mandarin	0	0.00%	<input type="text"/>
4.	Arabic	0	0.00%	<input type="text"/>
5.	Spanish	4	9.30%	<input type="text"/>
6.	French	3	6.98%	<input type="text"/>
7.	German	1	2.33%	<input type="text"/>
8.	Other	7	16.28%	<input type="text"/>
	Total	43	100%	

Key Analytics		
Mean	3.023	Key Facts + 79.07% chose the following options : > English > Other
Confidence Interval @ 95%	[2.169 - 3.877] n = 43	
Standard Deviation	2.858	
Standard Error	0.436	

How did you assemble the corpus of annotation instances to be labeled? (Select one alternative)

Frequency Analysis								
	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	We took all annotation instances available to us	17	39.53%	<input type="text"/>				
2.	Random subsample of all available annotation instances	13	30.23%	<input type="text"/>				
3.	Subsample of all available annotation instances based on introspection	3	6.98%	<input type="text"/>				
4.	Subsample of all available annotation instances based on key words	4	9.30%	<input type="text"/>				
5.	Other	6	13.95%	<input type="text"/>				
	Total	43	100%					

Key Analytics		
Mean	2.279	Key Facts ➤ 69.77% chose the following options : ➤ We took all annotation instances available to us ➤ Random subsample of all available annotation instances ➤ Least chosen option 6.98% : ➤ Subsample of all available annotation instances based on introspection
Confidence Interval @ 95%	[1.850 - 2.708] n = 43	
Standard Deviation	1.436	
Standard Error	0.219	

How did you decide when to stop the annotation process? (Select one alternative)

Frequency Analysis								
	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Evaluation of model learned on annotations with a held-out gold standard	1	2.33%	<input type="checkbox"/>				
2.	Cross-validation of model on annotations	1	2.33%	<input type="checkbox"/>				
3.	Money and/or time available for annotation used up	19	44.19%	<input type="text" value="44.19%"/>				
4.	All relevant documents annotated	15	34.88%	<input type="text" value="34.88%"/>				
5.	Expected gains in model performance compared to additional annotation costs below critical threshold	2	4.65%	<input type="checkbox"/>				
6.	Other	5	11.63%	<input type="text" value="11.63%"/>				
	Total	43	100%					

Key Analytics		
Mean	3.721	Key Facts <input type="checkbox"/> 79.07% chose the following options : <ul style="list-style-type: none"> ➤ Money and/or time available for annotation used up ➤ All relevant documents annotated <input type="checkbox"/> Least chosen option 2.33% : <ul style="list-style-type: none"> ➤ Evaluation of model learned on annotations with a held-out gold standard
Confidence Interval @ 95%	[3.393 - 4.049] n = 43	
Standard Deviation	1.098	
Standard Error	0.167	

Are there specific reasons to why you did not apply Active Learning in this project?
(Multiple alternatives allowed)

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Never heard of Active Learning	8	11.11%	<input type="text"/>				
2.	Insufficient expertise/knowledge	18	25.00%	<input type="text"/>				
3.	Did not want to spend overhead in implementing Active Learning-based annotation environment	12	16.67%	<input type="text"/>				
4.	Did not meet my projects specific requirements	10	13.89%	<input type="text"/>				
5.	Wanted a corpus with a different instance distribution	2	2.78%	<input type="text"/>				
6.	Was not convinced that Active Learning would reduce labeling cost	7	9.72%	<input type="text"/>				
7.	Doubts that Active Learning would work well in my scenario	8	11.11%	<input type="text"/>				
8.	Other	7	9.72%	<input type="text"/>				
	Total	72	100%					

Key Analytics

Mean	3.944	<p>Key Facts</p> <ul style="list-style-type: none"> ➤ 41.67% chose the following options : <ul style="list-style-type: none"> ➤ Insufficient expertise/knowledge ➤ Did not want to spend overhead in implementing Active Learning-based annotation environment ➤ Least chosen option 2.78% : <ul style="list-style-type: none"> ➤ Wanted a corpus with a different instance distribution
Confidence Interval @ 95%	[3.417 - 4.472] n = 72	
Standard Deviation	2.282	
Standard Error	0.269	

Would you consider applying active learning in future annotation projects? (Select one alternative)

Frequency Analysis				
	Answer	Count	Percent	
				20% 40% 60% 80% 100%
1.	Yes	17	39.53%	<input type="text"/>
2.	No	0	0.00%	<input type="text"/>
3.	Maybe	24	55.81%	<input type="text"/>
4.	Only under specific circumstances:	2	4.65%	<input type="text"/>
	Total	43	100%	

Key Analytics	
Mean	2.256
Confidence Interval @ 95%	[1.942 - 2.569] n = 43
Standard Deviation	1.049
Standard Error	0.160

Key Facts
 95.35% chose the following options :
 > Maybe
 > Yes

Where are you currently working? (Select one alternative)

Frequency Analysis				
	Answer	Count	Percent	
				20% 40% 60% 80% 100%
1.	Academia	46	85.19%	<input type="text"/>
2.	Industry	3	5.56%	<input type="checkbox"/>
3.	Governmental organization	3	5.56%	<input type="checkbox"/>
4.	Other	2	3.70%	<input type="checkbox"/>
	Total	54	100%	

Key Analytics		
Mean	1.278	Key Facts <input type="checkbox"/> 90.74% chose the following options : > Academia > Industry <input type="checkbox"/> Least chosen option 3.7% : > Other
Confidence Interval @ 95%	[1.081 - 1.475] n = 54	
Standard Deviation	0.738	
Standard Error	0.100	

What is your educational background? (Select one alternative)

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Linguistics	12	22.22%	<input type="text"/>				
2.	Computational linguistics	25	46.30%	<input type="text"/>				
3.	Computer Science/Informatics	12	22.22%	<input type="text"/>				
4.	Mathematics	0	0.00%	<input type="text"/>				
5.	Other	5	9.26%	<input type="text"/>				
	Total	54	100%					

Key Analytics

Mean	2.278	Key Facts <input type="checkbox"/> 68.52% chose the following options : <ul style="list-style-type: none"> ➤ Computational linguistics ➤ Linguistics
Confidence Interval @ 95%	[1.983 - 2.573] n = 54	
Standard Deviation	1.106	
Standard Error	0.151	

Which field are you currently working in? (Select one alternative)

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	General Text Mining	15	27.78%	<input type="text"/>				
2.	Bio Text Mining	11	20.37%	<input type="text"/>				
3.	Machine Learning	10	18.52%	<input type="text"/>				
4.	Other	18	33.33%	<input type="text"/>				
	Total	54	100%					

Key Analytics

Mean	2.574	Key Facts <input type="checkbox"/> 61.11% chose the following options : <ul style="list-style-type: none"> ➤ Other ➤ General Text Mining <input type="checkbox"/> Least chosen option 18.52% : <ul style="list-style-type: none"> ➤ Machine Learning
Confidence Interval @ 95%	[2.248 - 2.900] n = 54	
Standard Deviation	1.222	
Standard Error	0.166	

May we contact you with follow-up questions? (requires that you have specified you email address above)

Frequency Analysis

	Answer	Count	Percent	20%	40%	60%	80%	100%
1.	Yes	30	55.56%					
2.	No	24	44.44%					
	Total	54	100%					

Key Analytics

Mean	1.444
Confidence Interval @ 95%	[1.311 - 1.578] n = 54
Standard Deviation	0.502
Standard Error	0.068

Analytics Powered by [QuestionPro](#)