

parameters

Gerbrich Ferdinands

1/14/2020

ASReview takes the following parameters/arguments:

- number of training data (included/excluded)
- a model
- a query strategy
- a balance strategy (fixed)
- feature extraction

configurations	
model	nb, rf, svm, lr
query_strategy	cluster, max, maxcluster, maxuncertainty, maxrandom, clusteruncertainty, cluster*random
feature_extraction	doc2vec, tfidf, sbert, embeddingIdf
training_data	10/10, 5/5, 5/10

This leads to 63 combinations of configurations:

	model	query_strategy	feature_extraction	training_data
1	nb	cluster	doc2vec	10/10
2	nb	max	doc2vec	10/10
3	nb	max*cluster	doc2vec	10/10
4	nb	max*uncertainty	doc2vec	10/10
5	nb	max*random	doc2vec	10/10
6	nb	cluster*uncertainty	doc2vec	10/10
7	nb	cluster*random	doc2vec	10/10
8	nb	cluster	doc2vec	5/5
9	nb	max	doc2vec	5/5
10	nb	max*cluster	doc2vec	5/5
11	nb	max*uncertainty	doc2vec	5/5
12	nb	max*random	doc2vec	5/5
13	nb	cluster*uncertainty	doc2vec	5/5
14	nb	cluster*random	doc2vec	5/5
15	nb	cluster	doc2vec	5/10
16	nb	max	doc2vec	5/10
17	nb	max*cluster	doc2vec	5/10
18	nb	max*uncertainty	doc2vec	5/10
19	nb	max*random	doc2vec	5/10
20	nb	cluster*uncertainty	doc2vec	5/10
21	nb	cluster*random	doc2vec	5/10
43	nb	cluster	sbert	10/10
44	nb	max	sbert	10/10
45	nb	max*cluster	sbert	10/10
46	nb	max*uncertainty	sbert	10/10
47	nb	max*random	sbert	10/10
48	nb	cluster*uncertainty	sbert	10/10

	model	query_strategy	feature_extraction	training_data
49	nb	cluster*random	sbert	10/10
50	nb	cluster	sbert	5/5
51	nb	max	sbert	5/5
52	nb	max*cluster	sbert	5/5
53	nb	max*uncertainty	sbert	5/5
54	nb	max*random	sbert	5/5
55	nb	cluster*uncertainty	sbert	5/5
56	nb	cluster*random	sbert	5/5
57	nb	cluster	sbert	5/10
58	nb	max	sbert	5/10
59	nb	max*cluster	sbert	5/10
60	nb	max*uncertainty	sbert	5/10
61	nb	max*random	sbert	5/10
62	nb	cluster*uncertainty	sbert	5/10
63	nb	cluster*random	sbert	5/10
64	nb	cluster	embeddingIdf	10/10
65	nb	max	embeddingIdf	10/10
66	nb	max*cluster	embeddingIdf	10/10
67	nb	max*uncertainty	embeddingIdf	10/10
68	nb	max*random	embeddingIdf	10/10
69	nb	cluster*uncertainty	embeddingIdf	10/10
70	nb	cluster*random	embeddingIdf	10/10
71	nb	cluster	embeddingIdf	5/5
72	nb	max	embeddingIdf	5/5
73	nb	max*cluster	embeddingIdf	5/5
74	nb	max*uncertainty	embeddingIdf	5/5
75	nb	max*random	embeddingIdf	5/5
76	nb	cluster*uncertainty	embeddingIdf	5/5
77	nb	cluster*random	embeddingIdf	5/5
78	nb	cluster	embeddingIdf	5/10
79	nb	max	embeddingIdf	5/10
80	nb	max*cluster	embeddingIdf	5/10
81	nb	max*uncertainty	embeddingIdf	5/10
82	nb	max*random	embeddingIdf	5/10
83	nb	cluster*uncertainty	embeddingIdf	5/10
84	nb	cluster*random	embeddingIdf	5/10