

IDF_word2vec Note

window size = 9

LR:

Accuracy: 0.83

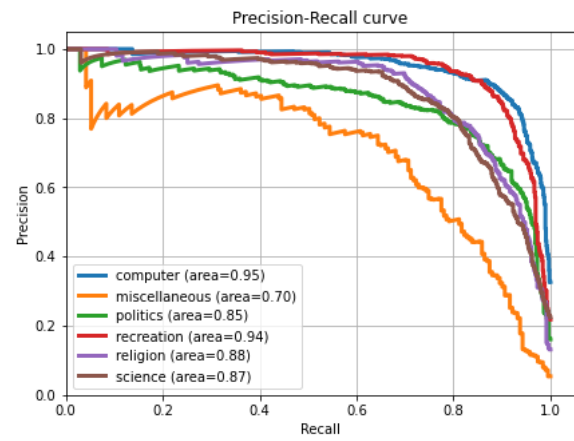
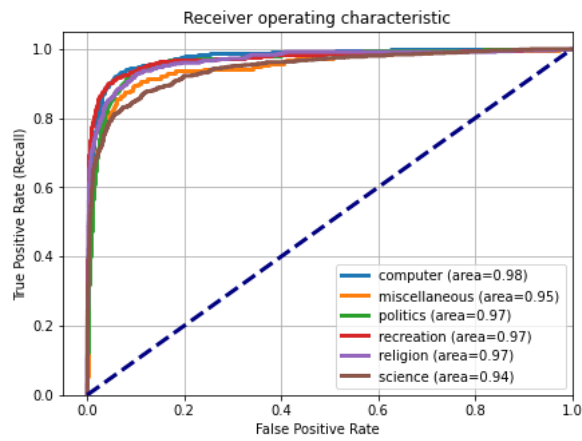
Auc: 0.96

Detail:

	precision	recall	f1-score	support
computer	0.87	0.90	0.89	953
miscellaneous	0.74	0.63	0.68	191
politics	0.77	0.81	0.79	510
recreation	0.87	0.88	0.87	751
religion	0.83	0.78	0.80	478
science	0.80	0.79	0.80	764
accuracy			0.83	3647
macro avg	0.81	0.80	0.80	3647
weighted avg	0.83	0.83	0.83	3647

Confusion matrix

True	computer	855	17	5	13	9	54
	miscellaneous	30	120	3	19	3	16
	politics	4	0	411	26	42	27
	recreation	14	13	21	661	9	33
	religion	15	2	46	19	373	23
	science	60	11	50	22	16	605
	Pred						
	computer	miscellaneous	politics	recreation	religion	science	



SVM:

Accuracy: 0.86

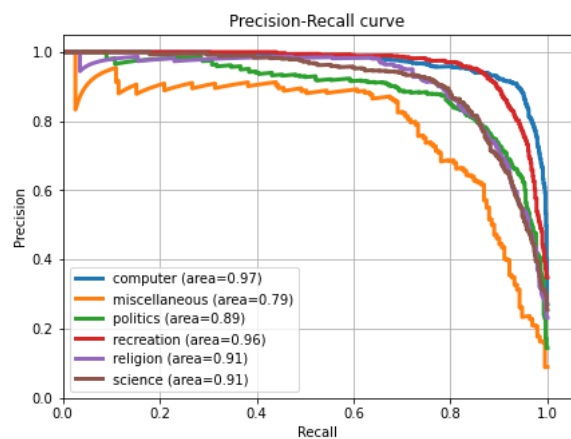
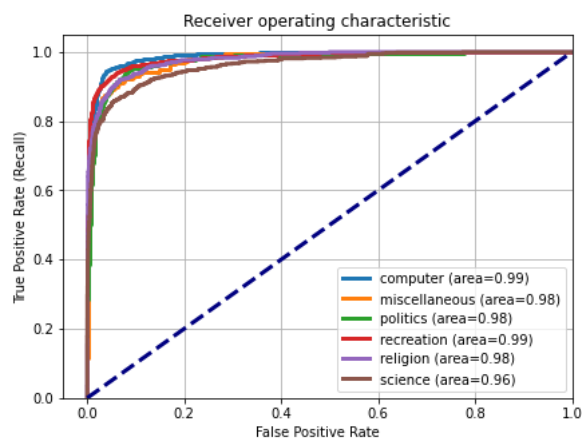
Auc: 0.98

Detail:

	precision	recall	f1-score	support
computer	0.88	0.94	0.91	953
miscellaneous	0.84	0.68	0.75	191
politics	0.77	0.85	0.81	510
recreation	0.89	0.90	0.89	751
religion	0.89	0.79	0.84	478
science	0.85	0.82	0.83	764
accuracy			0.86	3647
macro avg	0.85	0.83	0.84	3647
weighted avg	0.86	0.86	0.86	3647

Confusion matrix

	computer	miscellaneous	politics	recreation	religion	science
True computer	892	11	4	11	5	30
True miscellaneous	28	130	6	15	0	12
True politics	3	1	436	22	27	21
True recreation	14	5	22	673	6	31
True religion	14	0	53	17	379	15
True science	61	8	43	20	8	624
	computer	miscellaneous	politics	recreation	religion	science
Pred						



window size = 38

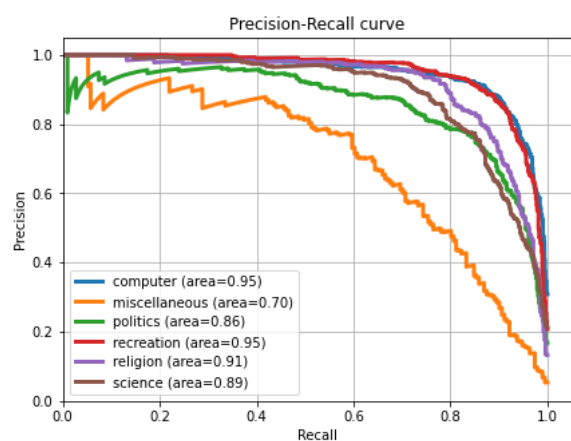
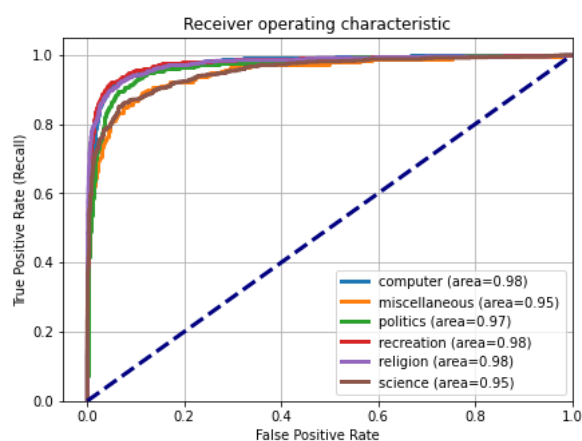
LR:

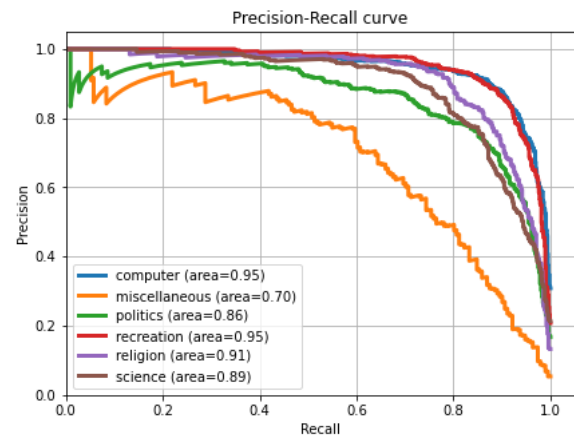
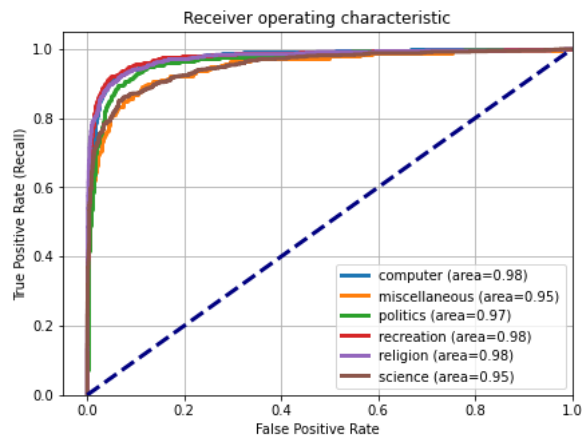
Accuracy: 0.84

Auc: 0.97

Detail:

	precision	recall	f1-score	support
computer	0.88	0.89	0.88	953
miscellaneous	0.67	0.63	0.65	191
politics	0.78	0.81	0.80	510
recreation	0.88	0.88	0.88	751
religion	0.85	0.84	0.84	478
science	0.82	0.81	0.81	764
accuracy			0.84	3647
macro avg	0.81	0.81	0.81	3647
weighted avg	0.84	0.84	0.84	3647





SVM:

Accuracy: 0.87

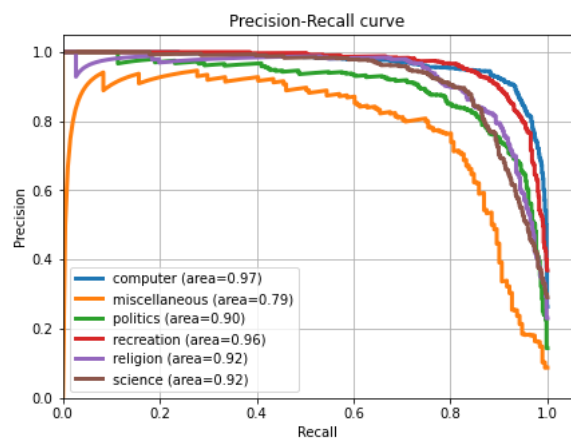
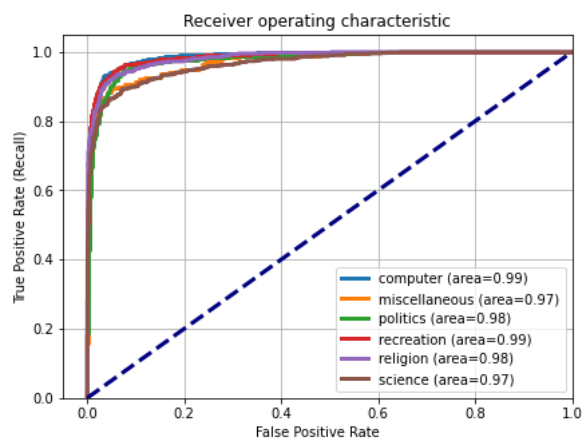
Auc: 0.98

Detail:

	precision	recall	f1-score	support
computer	0.88	0.93	0.91	953
miscellaneous	0.81	0.70	0.75	191
politics	0.79	0.85	0.82	510
recreation	0.91	0.89	0.90	751
religion	0.89	0.83	0.86	478
science	0.87	0.84	0.86	764
accuracy			0.87	3647
macro avg	0.86	0.84	0.85	3647
weighted avg	0.87	0.87	0.87	3647

Confusion matrix

	computer	miscellaneous	politics	recreation	religion	science
True computer	886	15	5	10	3	34
True miscellaneous	28	134	3	15	0	11
True politics	6	1	435	15	36	17
True recreation	19	7	24	672	6	23
True religion	13	0	46	13	396	10
True science	50	9	40	17	5	643
	computer	miscellaneous	politics	recreation	religion	science
	Pred					



window size = 75

LR:

Accuracy: 0.84

Auc: 0.96

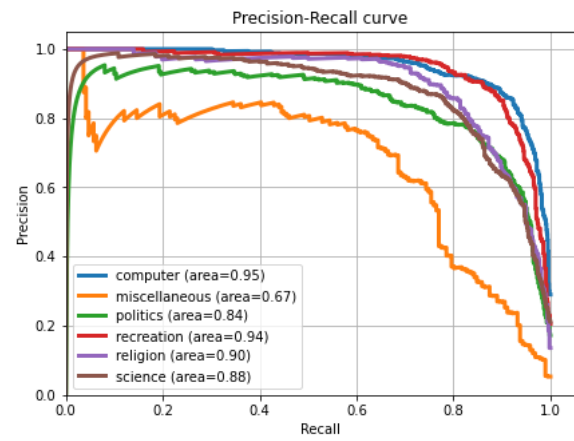
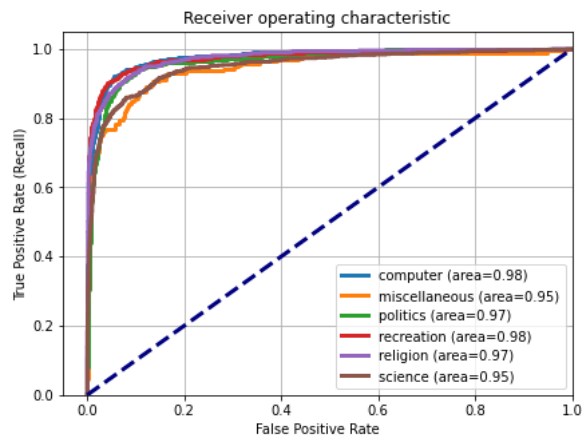
Detail:

	precision	recall	f1-score	support
computer	0.87	0.90	0.88	953
miscellaneous	0.72	0.64	0.68	191
politics	0.77	0.82	0.79	510
recreation	0.88	0.87	0.88	751
religion	0.84	0.81	0.82	478
science	0.81	0.80	0.81	764
accuracy			0.84	3647
macro avg	0.82	0.81	0.81	3647
weighted avg	0.84	0.84	0.84	3647

Confusion matrix

	computer	miscellaneous	politics	recreation	religion	science
True computer	854	24	5	13	2	55
True miscellaneous	29	123	5	16	2	16
True politics	2	2	417	21	41	27
True recreation	18	10	28	657	13	25
True religion	14	0	41	17	387	19
True science	60	11	44	21	16	612
	computer	miscellaneous	politics	recreation	religion	science

Pred



SVM:

Accuracy: 0.87

Auc: 0.98

Detail:

	precision	recall	f1-score	support
computer	0.88	0.93	0.91	953
miscellaneous	0.84	0.70	0.77	191
politics	0.80	0.86	0.83	510
recreation	0.89	0.89	0.89	751
religion	0.89	0.82	0.85	478
science	0.86	0.83	0.85	764
accuracy			0.87	3647
macro avg	0.86	0.84	0.85	3647
weighted avg	0.87	0.87	0.87	3647

Confusion matrix

	computer	miscellaneous	politics	recreation	religion	science
True computer	890	11	3	11	2	36
True miscellaneous	29	134	2	14	0	12
True politics	2	0	437	19	33	19
True recreation	21	4	21	671	8	26
True religion	13	0	46	16	392	11
True science	55	10	37	20	5	637
	computer	miscellaneous	politics	recreation	religion	science
	Pred					

