

Overview of n-gram models with different phrase lengths. Each column name shows the minimum and maximum width of considered phrases. Phrase-length 1 is equivalent to a bag-of-words model.

		1	1-2	1-3	1-4	1-5
Adv. Cancer	<i>P</i>	44	41	41	41	35
	<i>R</i>	77	55	55	55	48
	<i>F1</i>	56	47	47	47	41
	<i>AUC</i>	90	88	87	86	85
Adv. Heart Disease	<i>P</i>	70	73	73	78	78
	<i>R</i>	32	27	32	36	42
	<i>F1</i>	44	40	45	49	55
	<i>AUC</i>	85	86	86	86	85
Adv. Lung Disease	<i>P</i>	21	27	27	23	21
	<i>R</i>	29	39	32	36	32
	<i>F1</i>	24	32	30	28	25
	<i>AUC</i>	76	79	78	78	77
Chronic Neuro	<i>P</i>	47	48	49	40	42
	<i>R</i>	46	53	54	63	62
	<i>F1</i>	46	50	51	49	50
	<i>AUC</i>	72	71	71	71	70
Chronic Pain	<i>P</i>	33	37	42	39	42
	<i>R</i>	54	39	38	46	46
	<i>F1</i>	41	38	40	43	44
	<i>AUC</i>	68	67	67	67	67
Alcohol Abuse	<i>P</i>	100	55	57	53	52
	<i>R</i>	50	64	57	57	57
	<i>F1</i>	67	59	57	55	54
	<i>AUC</i>	89	88	87	86	85
Substance Abuse	<i>P</i>	62	75	71	77	83
	<i>R</i>	50	30	33	33	33
	<i>F1</i>	56	43	45	47	48
	<i>AUC</i>	90	89	88	87	86
Obesity	<i>P</i>	27	16	40	42	44
	<i>R</i>	35	45	20	20	20
	<i>F1</i>	30	24	27	27	28
	<i>AUC</i>	72	72	72	72	71
Psychiatric Disorders	<i>P</i>	47	53	55	55	55
	<i>R</i>	53	39	31	29	31
	<i>F1</i>	50	45	39	38	39
	<i>AUC</i>	77	76	75	74	73
Depression	<i>P</i>	51	47	51	51	52
	<i>R</i>	67	78	73	73	71
	<i>F1</i>	58	59	60	60	60
	<i>AUC</i>	77	78	78	78	78