Data set	Features $(d)$	Instances $(N)$
news20	1,355,191	19,954
url	3,231,961	2,396,130
webspam	16,609,143	350,000
kdd2010	29,890,095	19,264,097