

	mini-batch units	learning algorithm
(a)	64 sentences	Adam
(b)	32 sentences	Adam
(c)	16 sentences	Adam
(d)	8 sentences	Adam
(e)	2055 or 1742 words	Adam
(f)	64 sentences	SGD