

Dropout	$d(\mathbf{w})$	$d(\mathbf{h}_{d_i}^j)$	layers	hidden	PER
0.2	64	64	3	128	21.2
	64	32	3	128	21.6
	32	32	3	128	21.4
	64	64	3	250	20.1
	64	32	3	250	20.4
	32	32	3	250	20.6
	64	64	6	250	19.3
	64	32	6	250	20.2
	32	32	6	250	20.2
0.1	64	64	3	128	21.3
	64	64	3	250	20.9
	64	64	6	250	20.4
\times	64	64	6	250	21.9