Hyperparameters	Type	Values
Initial learning rate	Uniform	$[10^{-4}, 5 \times 10^{-3}]$
Weight decay	Uniform	$[5 \times 10^{-5}, 10^{-2}]$
Batch size	Categorical	$16, 24$
Optimizer	Categorical	Adam, RMSprop