Hyper-Parameter	Value
Activation function	Relu
Dropout probability	0.25
Learning rate	10^{-6}
Loss function	negative log-likelihood
Optimization method	Adam
Minibatch	10
Epochs	20