

Global Hyperparameter	Range
Learning Rate	[0.0001, 0.1]
Momentum	[0.68, 0.99]
Shared Embedding Size	[128, 512]
Embedding Dropout	[0, 0.7]
LSTM Recurrent Dropout	{True, False}
Nesterov Momentum	{True, False}
Weight Initialization	{Glorot normal, He normal}
Node Hyperparameter	Range
Layer Type	{Dense, LSTM}
Merge Method	{Sum, Concat}
Layer Size	{128, 256}
Layer Activation	{ReLU, Linear}
Layer Dropout	[0, 0.7]