

max_grad_norm = 5.0, num_layers = 2,
num_steps = 35, hidden_size = 512, max_epoch = 6, max_max_epoch = 39,
keep_prob = 0.5, lr_decay = 0.8, batch_size = 20, vocab_size = 10000,
optimizer = GradDesc, loss_function = sequence_loss_by_example

