

A	B	C	D	E	F	G	Output
DEEP LEARNING MODEL	ACTIVATION FUNCTION	COST FUNCTION	EPOCHS	GRADIENT ESTIMATION	NETWORK ARCHITECTURE	NETWORK INITIALIZATION	Mean BLEU score
Gradient Estimation							
1	ReLU	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.37
2	ReLU	Cross-Entropy	6	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.351
3	ReLU	Cross-Entropy	5	Adagrad	3 layer, 256 nodes, LSTM, vgg16	default	0.404
4	ReLU	Cross-Entropy	5	RMSProp	3 layer, 256 nodes, LSTM, vgg16	default	0.374
5	ReLU	Cross-Entropy	5	Adadelata	3 layer, 256 nodes, LSTM, vgg16	default	0.353
6	ReLU	Cross-Entropy	5	Nadam	3 layer, 256 nodes, LSTM, vgg16	default	0.353
7	ReLU	Cross-Entropy	5	SGD	3 layer, 256 nodes, LSTM, vgg16	default	0.028
Cost Function							
1	ReLU	mean_squared_error	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.215
2	ReLU	hinge	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0
3	ReLU	llback_leibler_divergen	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.373
4	ReLU	cosine_proximity	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0
Network Initialization							
1	ReLU	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	glorot_uniform	0.381
2	ReLU	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	random_uniform	0.388
3	ReLU	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	lecun_uniform	0.367
4	ReLU	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	he_uniform	0.389
5	ReLU	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	glorot_normal	0.398
Activation Function							
1	ReLU	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.374
2	tanh	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.384
3	elu	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.392
4	selu	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.363
5	linear	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.192
6	sigmoid	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.375
7	softsign	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.396
8	softplus	Cross-Entropy	5	Adam	3 layer, 256 nodes, LSTM, vgg16	default	0.381
Epochs							
1	ReLU	Cross-Entropy	3	Adam	3 layers, 256 nodes each	default	0.429
2	ReLU	Cross-Entropy	4	Adam	3 layers, 256 nodes each	default	0.394
3	ReLU	Cross-Entropy	5	Adam	3 layers, 256 nodes each	default	0.408
4	ReLU	Cross-Entropy	6	Adam	3 layers, 256 nodes each	default	0.38
5	ReLU	Cross-Entropy	7	Adam	3 layers, 256 nodes each	default	0.405
Network Architecture							
1	ReLU	Cross-Entropy	5	Adam	3 layers, 256 nodes each	default	0.407
2	ReLU	Cross-Entropy	5	Adam	3 layers, 128 nodes each	default	0.405
3	ReLU	Cross-Entropy	5	Adam	3 layers, 512 nodes each	default	0.394
4	ReLU	Cross-Entropy	5	Adam	4 layers, 256 nodes each	default	0.406
5	ReLU	Cross-Entropy	5	Adam	4 layers, 128 nodes each	default	0.386
6	ReLU	Cross-Entropy	5	Adam	4 layers, 256 nodes each	default	0.421