Annajiat Alim Rasel • 21266004 - Sadia Afrose Senior Lecturer, • 21366035 - Md. Sabbir Ahmed Department of Computer Science and Engineering 22166003 - Ashik Adnan **BRAC University** 22166009 - Nazifa Khanom Pointwise Output Gate Multiplication **Reset Gate Sigmoid Activation** Function Add Gate Feedforwad Layer Gates Output Vectors Separate Contex Adding Explicit
Context Layer Vector LSTM Weight and Sum Feed Forward Value Vectors LSTM GRUs Negative Attitude towards SRN amd GRU Monitory Softmax Basic Feedforward Add and **Complex Network Architecture** Markov Normalize Assumption **Violation of Privacy** Key/Query Comparisons N Gram Markov Feed Forward Machine Translation **Comparison Beteer** Model network Vanishing Gradient Units Simultaneous Task Problem Performance Normalised by Length Multihead Self Key, Query Value **Text Summarisation** robabilistic Language Attention Vectors Model Autoregressive Generation Perplexity Chain Rule Web Search Query RNN - Loss of Input Vectors relevant (Vi.....Vn) Conditional Language Model information Probabilities to create Language Models Assessment vocabulary distribution Managing Contex in RNNs **Self Attention Network** Deep Learning Architecture for **Sequence Procesing** :: Transformation Potential Harms From Language Models RNN :: Sequence RNNs as Language Training Applications of RNNs RNN Inferences **Bidirectional RNNs** Models Parts of Speech Tagging Training Forward Interface Generation A Simple Recurrent A Feed Forward Stacked RNNs Sentiment Analysis Back Propagation X multiplied by weight W multiplied by Training Set Los Function U (hidden layer at time t-1) and added One - Hot encoded Weight Tying Teacher Forcing Topic Classification Word Embeddings Generation Consists of Multiple Networks Summarisation Vector Back Propagation Single set of Use <s> as first input Cross Entropy Computing Loss representation of Embeddings through line probability End to End Training **Activation Function** Time t+1 Outperforms single Machine Translation distribution Networks Take the word Minimizes error in Bidirectional RNNs SoftMax resulting from predicting next SoftMax distribution Question Answering No. of stacks increase word Backward Interface training cost Softmax Classification Reduces required Generate next word until <s> is in output parameters Backward RNN Forward RNN

Submitted To:

Submitted By: