

SNS COLLEGE OF ENGINEERING

Kurumbapalayam(Po), Coimbatore - 641 107 Accredited by NAAC-UGC with 'A' Grade Approved by AICTE, Recognized by UGC & Affiliated to Anna University, Chennai

Department of AI &DS

Course Name – 19AD602 DEEP LEARNING

III Year / VI Semester

UNIT-4 OPTIMIZATION AND GENERALIZATION

Topic: Recurrent Neural Network Language Models

GULSHAN BANU.A/ AP/AI AND DS / Recurrent Neural Network Language ModelsSNSCE







CASE STUDY:

Enhancing Text Prediction with LSTM-Based Language Modeling

A leading tech company integrated an LSTM-based language model into its chatbot to improve user interactions. By training on vast datasets, the model significantly enhanced sentence generation and reduced perplexity, leading to more natural conversations. As a result, customer engagement and satisfaction increased by 30%, showcasing the effectiveness of recurrent neural networks in NLP applications.





INTRODUCTION:

- Language models predict the probability of word sequences.
- They are used in machine translation, speech recognition, and text generation.
- Traditional models struggle with long-range dependencies.
- RNNs use hidden states to retain information from previous time steps.
- However, they suffer from vanishing/exploding gradients, limiting long-term memory.



Long Short-Term Memory (LSTM)

- LSTM introduces memory cells, input/output/forget gates.
- It allows better retention of long-range dependencies and avoids vanishing gradients.
- Language modeling aims to estimate P(w_t | w_1, ..., w_{t-1}).
- Uses softmax over vocabulary and backpropagation through time (BPTT) for training. lacksquare





Implementation Details

- Model: LSTM with embedding layer, hidden layers, and softmax output.
- Training: Cross-entropy loss, Adam optimizer.
- Evaluation: Perplexity as a metric for model performance.

Perplexity in Language Modeling

- Perplexity measures how well a probability model predicts a sample.
- Lower perplexity indicates a better model.
- Formula: PPL = exp(-sum(log P(w_t | context)) / N).



4/7



Training and Testing the Model

- Training: Feed data in batches, optimize using backpropagation.
- Testing: Evaluate on unseen data, compute perplexity and accuracy. Generating Text with LSTM
- Use the trained model to generate text one word at a time.
- Seed the model with an initial word and predict subsequent words.



CONCLUSION:

- LSTMs improve language modeling over traditional RNNs.
- Further enhancements include Transformer models like GPT and BERT.
- Future work: fine-tuning on specific domains, integrating attention mechanisms.



6/7



THANK YOU

GULSHAN BANU.A/ AP/AI AND DS / Recurrent Neural Network Language Models/SNSCE



7/7