SNS COLLEGE OF ENGINEERING

Kurumbapalayam (Po), Coimbatore – 641 107

AN AUTONOMOUS INSTITUTION

Approved by AICTE, New Delhi and Affiliated to Anna University, Chennai

INTERNAL ASSESSMENT EXAMINATION – II ANSWER KEY

VI Semester

B.Tech-Artificial Intelligence and Data Science

19AD602 -DEEP LEARNING

PART-A

1. Define generalization in the context of neural networks and explain why it is important.
The ability of a neural network to perform well on unseen data, showing it has learned meaningful patterns rather than memorizing training data.

2. How does a Recurrent Neural Network Language Model predict the next word in a sequence?
Uses word embeddings and hidden states from previous time steps to predict the next word in a sequence, enabling context-aware generation.

3. What is the purpose of a spatial transformer module in a convolutional neural network?
A neural network component that learns to apply spatial transformations, like scaling or rotation, to focus on important features in input data.

4. Compare object detection and image classification in terms of functionality and output.

Object detection locates and labels multiple objects within an image, while classification assigns a single label to the entire image.

5. List common challenges encountered in face recognition systems.

Performance is affected by pose, lighting, occlusion, and facial expression, making accurate recognition across conditions difficult.

PART-B

6.a)Describe the working of WaveNet and its applications in speech synthesis. How does it differ from traditional audio generation models?

1. Autoregressive model using dilated causal convolutions for raw audio generation.
2. Captures long-range temporal dependencies effectively.
3. Trained on actual audio waveforms, not spectrograms.
4. Generates high-fidelity human-like speech.
5. Applications include voice assistants and TTS (e.g., Google Assistant).
6. Differs from traditional models by avoiding predefined features.
7. Uses deep CNNs instead of RNNs, enabling parallel processing during training.

b)Explain the architecture and impact of ImageNet in deep learning. How has it contributed to advancements in computer vision?

1. Large-scale visual database with 14M+ images across 20K categories.
2. AlexNet's success on ImageNet sparked the deep learning revolution.
3. Used for benchmarking computer vision models.
4. Introduced ReLU, dropout, and data augmentation in CNNs.
5. ImageNet Challenge (ILSVRC) pushed model performance limits.
6. Enabled pretraining and transfer learning.
7. Boosted applications in medical imaging, robotics, and surveillance.

7.a)(i) Analyze how regularization techniques improve generalization in content classification models

(i) Regularization Techniques in Content Classification

1. L1/L2 regularization penalizes large weights.
2. Dropout prevents co-adaptation of neurons.
3. Early stopping avoids overfitting on training data.
4. Data augmentation enhances robustness.
5. Improves model generalization on unseen content.
6. Reduces model complexity effectively.
7. Encourages smoother decision boundaries.

(ii) How can neural networks learn to apply spatial transformations to inputs or feature maps in differentiable manner?

Differentiable Spatial Transformations

1. Spatial Transformer Networks (STNs) introduce learnable modules.
2. They perform affine transformations (scale, rotate, translate).
3. Use localization network and grid generator.
4. Implemented in a differentiable manner via bilinear sampling.
5. Enable dynamic focus on regions of interest.
6. Integrated into standard CNN architectures.
7. Improve performance in tasks with spatial variation.

b)Explain Word2Vec in detail, including its two main models (CBOW and Skip-gram). How does it help in vector representation of words?

1. Learns vector representations of words from context.
2. CBOW predicts a word from surrounding context words.
3. Skip-gram predicts surrounding words from the current word.
4. Uses negative sampling or hierarchical softmax.
5. Captures semantic and syntactic relationships.
6. Helps clustering similar words in vector space.
7. Enables transfer learning in NLP tasks.

PART-C

8.a)How does the use of Long Short-Term Memory (LSTM) networks improve the performance of speech recognition systems compared to traditional models like HMMs or feedforward neural networks? Discuss with reference to the challenges of temporal dependencies in speech and provide an example of an LSTM-based architecture used in a real-world speech recognition application.

1. LSTMs handle long-term dependencies using memory cells and gates.
2. Outperform HMMs by avoiding fixed state transitions.
3. Can model variable-length audio input sequences.
4. Reduce vanishing gradient issues seen in traditional RNNs.
5. Enable context-aware decoding in speech.
6. Used in Google's voice recognition systems.
7. Bidirectional LSTMs further improve accuracy.

b)How can deep learning models be applied to predict protein structures or gene expression patterns in bioinformatics? Analyze a specific deep learning approach (e.g., CNN, RNN, Transformer, or AlphaFold) used in solving a complex biological problem, and discuss the advantages and limitations of using such models in real-world biological data analysis.

1. CNNs and transformers used to model protein structures.
2. AlphaFold predicts 3D structure from amino acid sequence.
3. Trained on protein databases like PDB.
4. Uses attention mechanisms to capture residue interactions.
5. Outperforms traditional simulation-based methods.
6. Open-sourced by DeepMind, democratizing access.
7. Applications: drug discovery, genomics.

SUBJECT IN-CHARGE HoD