Student 14: Technical but Incomplete

Question 1: Explain the differences between supervised, unsupervised, and reinforcement learning in machine learning. Provide examples of applications for each approach.

Answer:

Supervised learning utilizes labeled training datasets where each input vector x is associated with a corresponding target output y. The learning algorithm attempts to approximate a function $f: X \rightarrow Y$ that minimizes prediction error on the training set while generalizing to unseen data. This paradigm encompasses classification problems with discrete output spaces and regression problems with continuous output spaces.

Classification applications include support vector machines for text categorization, random forests for medical diagnosis, and neural networks for image recognition. Regression applications include linear regression for price prediction, polynomial regression for trend analysis, and deep learning for complex function approximation.

Unsupervised learning operates on datasets containing only input vectors without corresponding target outputs. The objective is to discover latent structures, patterns, or representations within the data distribution. Principal methodologies include clustering algorithms that partition data into homogeneous groups and dimensionality reduction techniques that project high-dimensional data into lower-dimensional spaces while preserving essential characteristics.

Reinforcement learning involves an agent interacting with an environment through a sequence of actions, observations, and rewards. The agent learns an optimal policy π that maximizes expected cumulative reward through exploration and exploitation strategies. The framework is typically modeled as a Markov Decision Process with states, actions, transition probabilities, and reward functions.

Question 2: Describe the architecture and functioning of Convolutional Neural Networks (CNNs) and explain why they are particularly effective for image recognition tasks.

Answer:

CNNs implement spatial convolution operations through learnable filter banks that detect local features across input feature maps. The convolution operation computes dot products between filter weights and local receptive fields, producing activation maps that highlight feature presence.

Key architectural components include:Convolutional layers with shared parameters across spatial dimensions

Non-linear activation functions (typically ReLU)

Pooling layers for spatial downsampling and translation invariance

Fully connected layers for final classification

The hierarchical feature extraction enables learning of increasingly complex representations from simple edge detectors to high-level semantic concepts.

Question 3: Discuss the ethical considerations and potential societal impacts of implementing artificial intelligence systems in critical decision-making processes.

Answer:

Algorithmic bias emerges from training data that contains historical discriminatory patterns or systematic underrepresentation of certain demographic groups. This can result in disparate impact where protected classes experience differential treatment rates.

Explainability challenges arise from the high-dimensional parameter spaces of deep learning models where decision boundaries are not easily interpretable. This creates accountability gaps in high-stakes applications.

Privacy concerns include differential privacy violations, membership inference attacks, and the potential for re-identification through auxiliary information.

Question 4: Explain the concept of transfer learning in deep neural networks and discuss its advantages and limitations.

Answer:

Transfer learning leverages pre-trained neural network weights from source domains to initialize models for target domains. The process typically involves feature extraction or fine-tuning strategies depending on dataset size and domain similarity.

Feature extraction freezes pre-trained convolutional layers and trains only newly added classifier layers. Fine-tuning updates all network parameters with reduced learning rates to preserve learned representations while adapting to new tasks.

Advantages include reduced computational requirements, improved sample efficiency, and better initialization compared to random weight initialization.

Limitations include negative transfer when source and target domains are dissimilar, and potential bias propagation from source domain training data.

Question 5: Describe the principles of natural language processing (NLP) and how transformer-based models like BERT have revolutionized language understanding tasks.

Answer:

Traditional NLP relied on n-gram language models, bag-of-words representations, and hand-crafted features with limited context modeling capabilities.

Transformer architectures implement self-attention mechanisms that compute attention weights between all token pairs in a sequence, enabling parallel processing and long-range dependency modeling.

BERT employs bidirectional training through masked language modeling and next

sentence prediction objectives during pre-training on large text corpora. The self attention mechanism computes contextualized representations where each token's embedding depends on all other tokens in the sequence.
The pre-training and fine-tuning paradigm enables transfer learning across diverse NLP tasks with minimal architectural modifications.