

Student 17 – Answer Sheet

Question 1:

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

Answer:

Supervised learning is a machine learning approach where models are trained using datasets that include both inputs and their corresponding outputs. The model learns to associate inputs with the correct outputs, making it suitable for tasks like classifying emails as spam or not, predicting house prices, or diagnosing diseases from patient data. This method is most effective when there is a wealth of labeled historical data available.

Unsupervised learning, in contrast, deals with data that lacks explicit output labels. Algorithms in this category aim to uncover hidden patterns or groupings within the data. Techniques such as clustering and dimensionality reduction are commonly used, with applications including grouping customers by purchasing behavior, detecting unusual activity in cybersecurity, or extracting topics from large collections of documents.

Reinforcement learning involves an agent that interacts with an environment, learning to make a series of decisions to maximize cumulative rewards. The agent receives feedback in the form of rewards or penalties based on its actions, gradually improving its strategy through trial and error. Notable applications include training AI to play complex games like Go, controlling robots, or optimizing resource allocation in large-scale systems¹.

Question 2:

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

Answer:

Convolutional Neural Networks are specialized neural networks designed to process data with a grid-like topology, such as images. Their architecture includes convolutional layers that use filters to scan across the input image, extracting features like edges or textures. Activation functions (such as ReLU) introduce non-linearity, while pooling layers reduce the spatial size of the data, making the computation more efficient and helping the network focus on the most relevant features. The final layers are typically fully connected and responsible for making predictions or classifications.

CNNs are particularly powerful for image recognition because their local connectivity and shared weights reflect the spatial structure of images, allowing them to efficiently capture hierarchical patterns—from simple edges in early layers to complex objects in deeper layers. This structure not only reduces the number of parameters compared to traditional networks but also provides robustness to variations in position, scale, and orientation of objects within images, making CNNs the dominant choice for computer vision tasks such as image classification and object detection¹.

Question 3:

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

Answer:

The integration of AI into critical decision-making brings forth several ethical challenges. One major concern is algorithmic bias, where models trained on historical data may reinforce or even amplify existing societal prejudices, potentially leading to unfair outcomes in areas like hiring, lending, or law enforcement. This raises difficult questions about what constitutes fairness and how to ensure it.

Another issue is transparency. Many advanced AI models, especially deep learning systems, operate as "black boxes," making their decision processes difficult to interpret. This lack of explainability is problematic in high-stakes fields such as healthcare, where trust and understanding are essential.

Privacy is also a significant concern, as AI systems often require access to large amounts of sensitive personal data, raising issues around consent and data security. Furthermore, the question of accountability remains unresolved—when AI-driven decisions cause harm, it is often unclear who should be held responsible.

Finally, widespread AI adoption may disrupt labor markets, potentially increasing economic inequality if the benefits of automation are not widely shared. Addressing these issues requires a combination of technical solutions, policy frameworks, and ongoing stakeholder engagement to ensure AI serves the public good¹.

Question 4:

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

Answer:

Transfer learning is a technique where a neural network trained on one task is repurposed for a different but related task. Typically, a model is first trained on a large dataset (such as ImageNet for images), and then its learned features are adapted to a new task with a smaller dataset by fine-tuning some or all of its layers.

The main advantage of transfer learning is that it allows models to leverage knowledge gained from large, diverse datasets, which is especially useful when the target task has limited labeled data. This approach reduces the need for extensive data and computational resources and often leads to better performance by providing a strong starting point for learning.

However, transfer learning works best when the source and target tasks are similar; transferring knowledge across unrelated domains can be ineffective or even detrimental. Additionally, pre-trained models may bring over biases from their original training data, and their architectures may not always be optimal for the new task¹.

Question 5:

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

Answer:

Natural Language Processing (NLP) involves developing computational methods to analyze and generate human language. Traditional NLP relied on rule-based systems and statistical models that often struggled to capture the complexities of language context.

Transformer-based models, such as BERT, have transformed NLP by introducing self-attention mechanisms that allow models to consider all words in a sentence simultaneously, capturing long-range dependencies and contextual relationships more effectively. BERT, in particular, uses bidirectional context, meaning it looks at both the left and right context of each word, leading to richer language representations.

These models are pre-trained on massive text corpora and then fine-tuned for specific tasks, which enables them to achieve state-of-the-art performance on a wide range of applications, from sentiment analysis to question answering. Despite challenges such as high computational demands and potential biases, transformer architectures now form the backbone of modern NLP systems, enabling more nuanced and effective language understanding