

## *MCQ Assessment - Answers Included*

**Question 1: Which of the following neural network architectures has been widely used for sequence modeling and transduction tasks?**

- a: Convolutional Neural Networks
- b: Long Short-Term Memory
- c: Radial Basis Function Networks
- d: Hopfield Networks

Correct Answer: b

**Question 2: What is the main limitation of recurrent models in terms of parallelization?**

- a: Memory constraints
- b: Limited computational resources
- c: Sequential nature of computation
- d: Lack of attention mechanisms

Correct Answer: c

**Question 3: Which researcher proposed replacing recurrent neural networks with self-attention in the Transformer model?**

- a: Jakob
- b: Ashish
- c: Noam
- d: Niki

Correct Answer: a

**Question 4: What is a key advantage of the Transformer model over recurrent models?**

- a: Faster training time
- b: Higher translation quality
- c: Increased parallelization
- d: Improved accuracy on small datasets

Correct Answer: c

**Question 5: Which attention mechanism is used in the Transformer model?**

- a: Additive attention
- b: Dot-product attention
- c: Multi-head attention
- d: Convolutional attention

Correct Answer: c

**Question 6: What is the benefit of using attention mechanisms in sequence models?**

- a: Modeling dependencies without regard to distance
- b: Reducing the number of parameters
- c: Improving computational efficiency
- d: Enhancing interpretability

Correct Answer: a

**Question 7: Which of the following researchers was not involved in the design and implementation of the Transformer models?**

- a: Ashish
- b: Illia

c: Llion

d: Niki

Correct Answer: b

**Question 8: What is the main computational advantage of the Transformer model?**

a: It eliminates sequential computation

b: It reduces memory consumption

c: It parallelizes computation across examples

d: It improves model performance on long sequences

Correct Answer: a