# Transformer MCQ

Question 1: What is a key feature of the transformer architecture that differentiates it from RNNs and CNNs in processing sequential data?

A) Sequential data processing

B) Self-attention mechanism

C) Convolutional layers

D) Recurrent layers

Question 2: Which of the following is NOT a typical application of Transformer models outside the field of Natural Language Processing (NLP)?

A) Image classification

B) Music generation

C) Weather forecasting

D) Enhancing physical strength

Question 3: What is the primary application of the Vision Transformer (ViT) model?

A) Automatic speech recognition

B) Object detection in images

C) Image classification

D) Text generation

Question 4: Which Transformer-based model is known for its capabilities in automatic speech recognition?

A) GPT-3

B) BERT

C) Wave2Vec 2.0

D) DALL-E

Question 5: Which model is famous for generating images from textual descriptions?

A) BERT

B) GPT-3

C) DALL-E

D) T5

Question 6: Which model, known for its advanced capabilities in NLP, uses bi-directional context for language understanding?

A) GPT-3

B) BERT

C) XLNet

D) RoBERTa

Question 7: The 'Attention Is All You Need' paper, which introduced the Transformer model, was published in which year?

A) 2015

B) 2017

C) 2019

D) 2021

Question 8: Which Transformer-based model is specifically designed for high-performance on long-sequence time-series forecasting?

A) GPT-3

B) ViT

C) T5

D) Informer

Question 9: What does the 'T' in T5, a model developed by Google, stand for?

A) Transformer

B) Textual

C) Transfer

D) Temporal

Question 10: DALL-E, a model developed by OpenAI for generating images from text, is a variant of which Transformer model?

A) BERT

B) RoBERTa

C) GPT-3

D) XLNet

Mark Scheme: B, D, C, C, D, B, B, D, C, C