1. **What is Artificial Intelligence (AI)?**

   AI is the simulation of human intelligence in machines to perform tasks like learning, reasoning, and problem-solving.

2. **What are the types of AI?**

   Narrow AI, General AI, and Super AI based on capability.

3. **What is machine learning (ML)?**

   A subset of AI where systems learn patterns from data without explicit programming.

4. **What is supervised learning?**

   ML where models are trained on labeled data with known outputs.

5. **What is unsupervised learning?**

   ML where models identify patterns from unlabeled data.

6. **What is reinforcement learning?**

   ML where agents learn optimal actions through rewards and penalties.

7. **What is a neural network?**

   A computational model inspired by the human brain to recognize patterns.

8. **What are the main components of a neural network?**

   Input layer, hidden layer(s), output layer, and activation functions.

9. **What is an artificial neuron?**

   A unit that processes inputs with weights, adds bias, and applies an activation function.

10. **What is the activation function?**

    It determines the output of a neuron given its input; e.g., sigmoid, ReLU.

11. **What is sigmoid function?**

    A smooth curve mapping input to a range of 0 to 1.

12. **What is ReLU function?**

    Outputs zero for negative input and linear for positive input.

13. **What is softmax function?**

    Converts output vector into probabilities summing to 1.

14. **What is the difference between AI and ML?**

    AI is the broader concept, ML is a technique to achieve AI.

15. **What is overfitting in ML?**

    When a model performs well on training data but poorly on unseen data.

16. **What is underfitting in ML?**

When a model fails to capture patterns, performing poorly on training and test data.

17. **What is backpropagation?**

Algorithm to update neural network weights using gradient descent.

18. **What is gradient descent?**

An optimization method to minimize the loss function by adjusting weights.

19. **What is learning rate?**

A parameter controlling the step size during weight updates.

20. **What is a loss function?**

A measure of how far predictions are from actual values.

21. **What is mean squared error (MSE)?**

Average squared difference between predicted and actual values.

22. **What is cross-entropy loss?**

A loss function measuring the difference between probability distributions.

23. **What is a feedforward neural network?**

Neural network where connections move forward from input to output.

24. **What is a recurrent neural network (RNN)?**

Neural network with cycles allowing sequential data processing.

25. **What is a convolutional neural network (CNN)?**

Network using convolutional layers for image and spatial data processing.

26. **What is pooling in CNN?**

Reduces spatial dimensions to decrease computation and control overfitting.

27. **What is max pooling?**

Selects the maximum value in a region of feature maps.

28. **What is dropout in neural networks?**

Randomly disables neurons during training to prevent overfitting.

29. **What is batch normalization?**

Normalizes layer inputs to stabilize learning and improve speed.

30. **What is the difference between batch and stochastic gradient descent?**

Batch uses all data, stochastic uses one sample per update.

31. **What is a perceptron?**

A simple linear binary classifier in neural networks.

32. **What is multi-layer perceptron (MLP)?**

A feedforward network with one or more hidden layers.

33. **What is weight initialization?**

Setting initial neuron weights to improve convergence.

34. **What is vanishing gradient problem?**

Gradients become too small, slowing or stopping learning in deep networks.

35. **What is exploding gradient problem?**

Gradients become too large, causing unstable updates.

36. **What is a radial basis function (RBF) network?**

Network using radial basis functions as activation functions for hidden neurons.

37. **What is autoencoder?**

A neural network that compresses input into a latent space and reconstructs it.

38. **What is a generative adversarial network (GAN)?**

A network with generator and discriminator competing to create realistic data.

39. **What is reinforcement signal in AI?**

A feedback signal representing reward or penalty for an agent's action.

40. **What is Q-learning?**

A model-free RL algorithm that learns optimal action-value function.

41. **What is policy in reinforcement learning?**

A strategy mapping states to actions for an agent.

42. **What is value function in RL?**

Expected cumulative reward for being in a state under a policy.

43. **What is epsilon-greedy strategy?**

Selects the best-known action most of the time, random action sometimes.

44. **What is supervised vs reinforcement learning?**

Supervised uses labeled data; RL learns from interaction and rewards.

45. **What is feature extraction in AI?**

Transforming raw data into useful inputs for a model.

46. **What is dimensionality reduction?**

Reducing features while preserving information, e.g., PCA.

47. **What is PCA?**

Principal Component Analysis reduces correlated features into uncorrelated components.

48. **What is clustering in ML?**

Grouping similar data points together without labels.

49. **What is K-means clustering?**

An iterative algorithm partitioning data into K clusters by distance.

50. _____

51. Advanced Level (50–98)

52. **What is deep learning?**

Subset of ML using deep neural networks for feature learning and prediction.

53. **What is the difference between shallow and deep networks?**

Shallow has one hidden layer, deep has multiple hidden layers.

54. **What is transfer learning?**

Using a pre-trained model for a new but similar task.

55. **What is fine-tuning in transfer learning?**

Adjusting the weights of a pre-trained model on new data.

56. **What is attention mechanism in neural networks?**

Allows the network to focus on important parts of input sequences.

57. **What is transformer architecture?**

Uses self-attention to process sequences in parallel efficiently.

58. **What is the difference between RNN and LSTM?**

LSTM has memory cells to avoid vanishing gradient in sequential data.

59. **What is GRU?**

Gated Recurrent Unit, a simplified LSTM with fewer parameters.

60. **What is bidirectional RNN?**

Processes sequence in both forward and backward directions.

61. **What is word embedding?**

Maps words to dense vector representations capturing semantics.

62. **What is one-hot encoding?**

Represents categorical data as binary vectors with a single 1.

63. **What is NLP?**

Natural Language Processing deals with human language understanding by machines.

64. **What is sentiment analysis?**

Determining emotions or opinions from text data.

65. **What is sequence-to-sequence model?**

Maps input sequences to output sequences, e.g., translation.

66. **What is attention vs self-attention?**

Attention focuses on input-output relations, self-attention focuses within input.

67. **What is BERT?**

Bidirectional Encoder Representations from Transformers for NLP tasks.

68. **What is GPT?**

Generative Pre-trained Transformer, generates human-like text.

69. **What is reinforcement learning vs supervised learning?**

RL learns via interaction and rewards; supervised learns from labeled data.

70. **What is Markov Decision Process (MDP)?**

Mathematical framework modeling RL problems with states, actions, rewards.

71. **What is reward shaping in RL?**

Modifying rewards to speed up learning without changing optimal policy.

72. **What is exploration vs exploitation in RL?**

Exploration tries new actions, exploitation chooses best-known actions.

73. **What is neural style transfer?**

Combines content of one image with style of another using CNNs.

74. **What is dropout vs batch normalization?**

Dropout prevents overfitting, batch normalization stabilizes learning.

75. **What is gradient clipping?**

Limits gradient values to avoid exploding gradients.

76. **What is adversarial attack in AI?**

Inputs designed to mislead neural networks.

77. **What is explainable AI (XAI)?**

Techniques making AI decisions interpretable to humans.

78. **What is AI bias?**

Systematic error due to skewed or unbalanced data.

79. **What is ethical AI?**

Developing AI respecting fairness, transparency, and accountability.

80. **What is convolution in CNN?**

Sliding filter over input to extract spatial features.

81. **What is stride in convolution?**

Number of pixels the filter moves per step.

82. **What is padding in CNN?**

Adding zeros around input to preserve dimensions.

83. **What is receptive field in CNN?**

Input region affecting a neuron's output.

84. **What is vanishing gradient in deep networks?**

Gradients become too small, halting learning in early layers.

85. **What is residual network (ResNet)?**

Uses skip connections to avoid vanishing gradient in deep networks.

86. **What is GAN mode collapse?**

Generator produces limited outputs failing to capture data diversity.

87. **What is reinforcement learning exploration-exploitation dilemma?**

Choosing between trying new actions or exploiting known rewards.

88. **What is temporal difference learning?**

RL method estimating value functions using bootstrapping.

89. **What is policy gradient method?**

Directly optimizes policy using gradient ascent on expected reward.

90. **What is actor-critic model?**

RL model with separate actor (policy) and critic (value) networks.

91. **What is AI search problem?**

Finding a sequence of actions from initial to goal state.

92. **What is BFS and DFS in AI search?**

Breadth-first explores level by level, depth-first explores deep branches first.

93. **What is heuristic function?**

Estimates cost from current state to goal in informed search.

94. *What is A algorithm?**

Search algorithm using cost + heuristic to find optimal path.

95. **What is constraint satisfaction problem (CSP)?**

Problem where variables must satisfy specific constraints.

96. **What is forward vs backward chaining?**

Forward derives conclusions from facts, backward starts from goal.

97. **What is knowledge representation in AI?**

Encoding information about the world for reasoning.

98. **What is expert system?**

AI system using knowledge base and inference engine to solve problems.

99. **What is fuzzy logic?**

Handles reasoning with degrees of truth rather than binary.

100. **What is reinforcement learning reward function?**

Defines feedback guiding an agent to learn desired behavior.

101. **What is a perceptron in AI?**

A single-layer neural network used for binary classification.

102. **What is multi-class classification?**

Predicting one label from more than two possible categories.

103. **What is a loss landscape in neural networks?**

A visualization of loss function values across different weights.

104. **What is stochastic gradient descent with momentum?**

Enhances SGD by adding previous update velocity to current step.

105. **What is RMSProp optimizer?**

Adjusts learning rate adaptively based on recent gradient magnitudes.

106. **What is Adam optimizer?**

Combines momentum and RMSProp for faster convergence.

107. **What is overparameterization in deep learning?**

Having more parameters than training samples, increasing capacity.

108. **What is early stopping?**

Halting training when validation performance stops improving to prevent overfitting.

109. **What is L1 vs L2 regularization?**

L1 adds absolute weights, L2 adds squared weights to loss.

110. **What is weight decay?**

Another term for L2 regularization applied during training.

111. **What is feature scaling?**

Normalizing features to similar ranges for faster convergence.

112. **What is standardization?**

Subtract mean and divide by standard deviation for normalization.

113. **What is min-max normalization?**

Rescales features to a specific range, typically 0 to 1.

114. **What is one-hot vector representation?**

Binary vector with one "1" and rest zeros representing a class.

115. **What is embedding layer in neural networks?**

Maps categorical data to dense continuous vector space.

116. **What is tokenization in NLP?**

Splitting text into words, subwords, or characters for processing.

117. **What is sequence padding in NLP?**

Adding tokens to make sequences of equal length for batching.

118. **What is dropout rate?**

Fraction of neurons randomly disabled during training.

119. **What is Xavier/Glorot initialization?**

Weights initialized to maintain variance across layers.

120. **What is He initialization?**

Weights initialized for ReLU activations to avoid vanishing gradients.

121. **What is vanishing gradient in RNNs?**

Gradients shrink exponentially, making learning early time steps difficult.

122. **What is exploding gradient in RNNs?**

Gradients grow exponentially, causing unstable weight updates.

123. **What is gradient clipping?**

Restricting gradients to a maximum threshold to stabilize training.

124. **What is attention score?**

Weight assigned to input elements based on relevance to output.

125. **What is self-attention?**

Mechanism computing attention among all elements in a sequence.

126. **What is multi-head attention?**

Using multiple self-attention mechanisms in parallel for richer representation.

127. **What is positional encoding in transformers?**

Adds order information to input tokens since transformers lack recurrence.

128. **What is encoder-decoder architecture?**

Encoder processes input, decoder generates output sequence.

129. **What is teacher forcing in RNN training?**

Using true previous output as next input to speed up training.

130. **What is beam search?**

Heuristic search keeping top-k sequences at each decoding step.

131. **What is greedy decoding?**

Selecting the highest probability token at each step for output.

132. **What is reinforcement learning agent?**

Learner interacting with an environment to maximize reward.

133. **What is Markov property in RL?**

Future state depends only on current state and action.

134. **What is temporal difference error?**

Difference between predicted and actual reward for learning updates.

135. **What is Q-value in Q-learning?**

Expected cumulative reward for a state-action pair.

136. **What is eligibility trace?**

Mechanism in RL assigning credit to previous states/actions.

137. **What is policy iteration?**

Iteratively evaluating and improving policy until convergence.

138. **What is value iteration?**

Updating value function directly to find optimal policy.

139. **What is deep Q-network (DQN)?**

Uses neural networks to approximate Q-values in complex environments.

140. **What is target network in DQN?**

A separate frozen network to stabilize Q-value updates.

141. **What is experience replay?**

Storing and sampling past transitions to break correlation during training.

142. **What is actor-critic in RL?**

Actor selects actions; critic evaluates action quality.

143. **What is advantage function in RL?**

Measures how much better an action is compared to average policy.

144. **What is PPO algorithm?**

Proximal Policy Optimization, an RL method balancing exploration and stability.

145. **What is TRPO algorithm?**

Trust Region Policy Optimization, constrains policy updates for stability.

146. **What is Monte Carlo method in RL?**

Learning value functions from complete episodes of experience.

147. **What is bootstrapping in RL?**

Estimating value using current estimates instead of waiting for final outcome.

148. **What is experience replay buffer?**

Memory storing transitions for random sampling in RL training.

149. **What is deep deterministic policy gradient (DDPG)?**

RL algorithm for continuous action spaces using actor-critic.

150. **What is neural network architecture search (NAS)?**

Automated search for optimal network structure.

151. **What is model compression in deep learning?**

Reducing model size for faster inference on limited hardware.

152. **What is pruning in neural networks?**

Removing insignificant weights or neurons to reduce complexity.

153. **What is quantization in neural networks?**

Reducing precision of weights and activations for efficiency.

154. **What is knowledge distillation?**

Transferring knowledge from a large model to a smaller one.

155. **What is batch size?**

Number of samples processed in one forward/backward pass.

156. **What is epoch in deep learning?**

One complete pass through the entire training dataset.

157. **What is validation set?**

Used to tune model hyperparameters without affecting training.

158. **What is test set?**

Evaluates final model performance on unseen data.

159. **What is confusion matrix?**

Table showing true vs predicted labels in classification tasks.

160. **What is precision?**

Fraction of correctly predicted positive samples among all predicted positives.

161. **What is recall?**

Fraction of correctly predicted positive samples among all actual positives.

162. **What is F1-score?**

Harmonic mean of precision and recall.

163. **What is ROC curve?**

Plots true positive rate vs false positive rate at different thresholds.

164. **What is AUC?**

Area under ROC curve, measures classifier performance.

165. **What is cross-validation?**

Technique to evaluate model generalization by splitting data into folds.

166. **What is k-fold cross-validation?**

Dividing data into k subsets and rotating training/testing.

167. **What is stratified sampling?**

Ensures each class proportion is maintained in splits.

168. **What is bias-variance tradeoff?**

Balancing underfitting (bias) and overfitting (variance).

169. **What is hyperparameter tuning?**

Optimizing parameters like learning rate, batch size, number of layers.

170. **What is grid search?**

Exhaustive search over predefined hyperparameter combinations.

171. **What is random search?**

Randomly sampling hyperparameter combinations for optimization.

172. **What is Bayesian optimization?**

Uses probabilistic model to select promising hyperparameters.

173. **What is recurrent cell in RNN?**

Basic unit maintaining hidden state across time steps.

174. **What is exploding/vanishing gradient solution?**

Use gradient clipping, LSTM/GRU, or proper initialization.

175. **What is deep reinforcement learning?**

Combines deep neural networks with RL for complex tasks.

176. **What is curriculum learning in AI?**

Training model on simpler tasks first, then gradually harder ones.

177. **What is reward shaping?**

Providing intermediate rewards to guide learning in RL.

178. **What is imitation learning?**

Learning by observing expert demonstrations.

179. **What is curiosity-driven learning?**

RL where intrinsic motivation drives exploration.

180. **What is multi-agent RL?**

Multiple agents learning and interacting in a shared environment.

181. **What is hierarchical RL?**

Learning high-level policies to select sub-policies or actions.

182. **What is meta-learning in AI?**

"Learning to learn" by training models to adapt quickly to new tasks.

183. **What is few-shot learning?**

Learning tasks with very limited labeled examples.

184. **What is zero-shot learning?**

Performing tasks without any labeled examples using prior knowledge.

185. **What is continual learning in AI?**

Learning new tasks without forgetting previous ones.

186. **What is catastrophic forgetting?**

When a model forgets previous knowledge while learning new tasks.

187. **What is generative model?**

Models data distribution to generate new samples.

188. **What is discriminative model?**

Models boundary between classes to predict labels.

189. **What is Variational Autoencoder (VAE)?**

Generative model learning latent distribution for data generation.

190. **What is KL-divergence?**

Measures difference between two probability distributions.

191. **What is GAN discriminator?**

Classifier distinguishing real from generated samples.

192. **What is GAN generator?**

Creates samples trying to fool the discriminator.

193. **What is conditional GAN?**

GAN generating samples conditioned on input labels.

194. **What is cycle GAN?**

Transforms data between domains without paired examples.

195. **What is attention visualization?**

Shows which parts of input the model focuses on.

196. **What is explainable AI (XAI) technique?**

LIME or SHAP provide feature importance and model insights.

197. **What is AI ethics challenge?**

Ensuring fairness, transparency, and accountability in AI systems.

198. **What is bias-variance in neural networks?**

Balancing model complexity to avoid underfitting or overfitting.

199. **What is hyperparameter vs parameter in NN?**

Parameter learned during training; hyperparameter set before training.

200. **What is residual connection?**

Shortcut connecting input to output of a layer to ease learning.

201. **What is dense layer in NN?**

Fully connected layer connecting every input to every output neuron.

202. **What is global average pooling?**

Reduces spatial dimensions by averaging each feature map.

203. **What is LSTM forget gate?**

Controls which information from previous state to discard.

204. **What is LSTM input gate?**

Controls which new information to store in memory cell.

205. **What is LSTM output gate?**

Controls which memory to expose as output at each time step.

206. **What is exploding gradient solution in RNN?**

Use gradient clipping or normalized initialization.

207. **What is learning curve in AI?**

Plot of model performance vs training progress.

208. **What is reinforcement learning environment?**

Simulated system where an agent takes actions and receives rewards.

209. **What is policy network in RL?**

Neural network that outputs action probabilities.

210. **What is value network in RL?**

Predicts expected future reward of states or state-action pairs.

211. **What is Monte Carlo Tree Search (MCTS)?**

Search algorithm combining random sampling and tree exploration.

212. **What is AlphaGo architecture?**

Combines deep neural networks with MCTS for board game play.

213. **What is knowledge graph in AI?**

Graph representing entities and their relationships for reasoning.

214. **What is semantic similarity in NLP?**

Degree to which two texts have the same meaning.

215. **What is cosine similarity?**

Measures similarity between vectors using angle between them.

216. **What is BLEU score?**

Metric to evaluate machine-translated text against references.

217. **What is ROUGE score?**

Metric for evaluating text summarization quality.

218. **What is perplexity in language models?**

Measures how well a probability model predicts test data.

219. **What is gradient penalty in GANs?**

Regularization term to stabilize training.

220. **What is mode collapse in GANs?**

Generator produces limited outputs ignoring diversity.

221. **What is Wasserstein GAN?**

GAN variant using Wasserstein distance for stable training.

222. **What is contrastive learning?**

Learning representations by comparing positive and negative pairs.

223. **What is self-supervised learning?**

Learning without labels by creating surrogate tasks.

224. **What is masked language modeling?**

Predicting missing tokens in a sentence, used in BERT.

225. **What is next sentence prediction?**

Task predicting whether two sentences follow each other.

226. **What is sequence labeling?**

Assigning a label to each token in a sequence.

227.   **What is entity recognition?**

Identifying named entities in text like persons, places, and organizations.

228.   **What is word sense disambiguation?**

Determining correct meaning of a word based on context.

229.   **What is beam width in beam search?**

Number of sequences kept at each decoding step.

230.   **What is greedy vs beam search tradeoff?**

Greedy is fast but suboptimal; beam explores more for accuracy.

231.   **What is transformer encoder?**

Processes input sequence and generates context-rich embeddings.

232.   **What is transformer decoder?**

Generates output sequence using encoder context and self-attention.

233.   **What is positional mask in transformers?**

Prevents attention to future tokens during decoding.

234.   **What is masked multi-head attention?**

Attention that masks future tokens for autoregressive modeling.

235.   **What is layer normalization?**

Normalizes inputs across features for each sample.

236.   **What is residual dropout?**

Applies dropout to residual connection for regularization.

237.   **What is knowledge distillation loss?**

Loss comparing teacher model outputs with student predictions.

238.   **What is hierarchical attention?**

Attention applied at multiple levels, e.g., word and sentence.

239.   **What is positional embedding vs positional encoding?**

Embedding is learned; encoding is fixed function of position.

240.   **What is beam search diversity?**

Adding penalty to encourage varied outputs.

241.   **What is catastrophic forgetting solution?**

Replay memory or regularization methods like EWC.

242. **What is attention heatmap?**

Visual showing attention weights over input tokens.

243. **What is sequence length truncation?**

Cutting sequences to a maximum length for batching.

244. **What is multi-label classification?**

Predicting multiple labels per instance simultaneously.

245. **What is label smoothing?**

Adjusting target labels to prevent overconfident predictions.

246. **What is teacher-student training?**

Student learns from soft targets provided by teacher network.

247. **What is curriculum RL?**

Training RL agents from easier tasks to complex tasks sequentially.

248. **What is policy entropy in RL?**

Encourages exploration by maximizing randomness of policy.

249. **What is advantage actor-critic (A2C)?**

Uses advantage function to reduce variance in policy gradient.

250. **What is residual connection in transformers?**

Adds input to output of attention/feedforward layer to ease gradient flow.

251. **What is pre-training in transformers?**

Learning general representations before fine-tuning on tasks.

252. **What is fine-tuning in transformers?**

Adjusting pre-trained model weights for a specific task.

253. **What is multi-task learning?**

Training a single model for multiple related tasks simultaneously.

254. **What is hyperparameter search space?**

Range of possible hyperparameter values to explore.

255. **What is zero-shot classification?**

Predicting classes unseen during training using knowledge transfer.

256. **What is few-shot classification?**

Learning from a very small number of labeled examples.

257. **What is continual learning in neural networks?**

Updating model on new tasks without forgetting old tasks.

258. **What is elastic weight consolidation (EWC)?**

Regularization to reduce forgetting in continual learning.

259. **What is adversarial training?**

Training on adversarial examples to improve robustness.

260. **What is model ensembling?**

Combining predictions of multiple models to improve accuracy.

261. **What is stacked autoencoder?**

Multiple autoencoders stacked to learn hierarchical features.

262. **What is denoising autoencoder?**

Trains to reconstruct input from corrupted versions for robust features.

263. **What is sparse autoencoder?**

Encourages sparse activation to learn meaningful features.

264. **What is neural architecture search (NAS)?**

Automated search for optimal neural network structures.

265. **What is attention visualization in transformers?**

Shows which tokens the model focuses on for predictions.

266. **What is token embedding vs segment embedding?**

Token embedding for word identity; segment embedding for sentence distinction.

267. **What is deep Q-learning vs vanilla Q-learning?**

Deep Q-learning uses neural networks for Q-value approximation.

268. **What is double DQN?**

Uses separate networks to reduce overestimation in Q-learning.

269. **What is prioritized experience replay?**

Samples transitions with high temporal-difference error more frequently.

270. **What is target network soft update?**

Gradually updating target network to stabilize learning.

271. **What is curriculum learning in NLP?**

Training models on easier sequences first before complex ones.

272. **What is masked token prediction?**

Predicting missing tokens in self-supervised NLP tasks.

273. **What is tokenization vs subword tokenization?**

Subword splits rare words into smaller units to handle unknown words.

274. **What is byte-pair encoding (BPE)?**

Algorithm for subword tokenization by merging frequent pairs.

275. **What is sentencepiece tokenizer?**

Language-independent subword tokenizer for NLP.

276. **What is transformer encoder-decoder attention?**

Decoder attends to encoder outputs for sequence generation.

277. **What is cross-attention?**

Attention between different sequences, e.g., decoder attends to encoder.

278. **What is self-supervised pre-training?**

Learning from unlabeled data using surrogate tasks.

279. **What is masked language model vs causal LM?**

Masked predicts missing tokens; causal predicts next token autoregressively.

280. **What is AI model calibration?**

Adjusting model outputs so predicted probabilities reflect true likelihood.

281. **What is temperature scaling?**

Softens model predictions to improve calibration.

282. **What is contrastive predictive coding?**

Learning representations by predicting future latent embeddings.

283. **What is siamese network?**

Network comparing similarity between two inputs.

284. **What is triplet loss?**

Loss encouraging anchor-positive pairs closer than anchor-negative.

285. **What is metric learning?**

Learning a distance function for similarity-based tasks.

286. **What is transformer block?**

Unit with multi-head attention, feedforward network, and residual connections.

287. **What is feedforward network in transformer?**

Applies linear transformations and activation to each position independently.

288. **What is multi-task fine-tuning?**

Fine-tuning model on multiple related tasks at once.

289. **What is unsupervised pre-training?**

Learning representations without labeled data.

290. **What is semi-supervised learning?**

Learning from small labeled dataset plus large unlabeled dataset.

291. **What is curriculum RL vs standard RL?**

Curriculum gradually increases task difficulty to improve learning.

292. **What is knowledge graph embedding?**

Mapping entities and relations into vector space for reasoning.

293. **What is Graph Neural Network (GNN)?**

Neural network operating on graph-structured data.

294. **What is message passing in GNN?**

Nodes exchange information with neighbors to update embeddings.

295. **What is graph attention network (GAT)?**

GNN using attention mechanism to weigh neighbors differently.

296. **What is node classification in GNN?**

Predicting label of nodes based on graph structure.

297. **What is link prediction in GNN?**

Predicting existence of edges between nodes.

298. **What is graph pooling?**

Aggregating node features to form graph-level representation.