DataInterview.com

80

Deep Learning

Interview Questions 📚

*Seen in **Data Scientist**, **ML Engineer** and **AI Engineer** at FAANGs, startups and consulting firms*

Fundamentals

1. Explain the difference between supervised, unsupervised, and reinforcement learning.
2. Describe the backpropagation algorithm.
3. How does a single-layer perceptron differ from a multi-layer perceptron?
4. What is the purpose of an activation function in a neural network?
5. Explain the difference between weight initialization methods like Xavier and He initialization.
6. Describe the working of the dropout regularization technique.
7. How do pooling layers in CNNs work and why are they important?
8. Explain the concept of "depth" in a neural network.
9. How do LSTMs address the vanishing gradient problem?
10. Describe the difference between batch normalization and layer normalization.
11. What is the skip connection or residual connection in deep networks?
12. Compare and contrast feedforward networks with recurrent networks.
13. Explain the difference between one-hot encoding and word embeddings.
14. How does a max-pooling layer differ from an average pooling layer in a CNN?
15. What are the typical applications of autoencoders?
16. Explain the significance of the bias term in neural networks.
17. What are the potential issues with using a sigmoid activation function in deep networks?
18. How does a self-attention mechanism work in transformers?
19. What challenges arise when training very deep neural networks?
20. Describe the concept of "transfer learning" and its advantages.

Training & Optimization

1. Why is the learning rate considered one of the most important hyperparameters in neural network training?
2. Explain the momentum term in optimization algorithms.
3. What is the batch size in neural network training, and how does it affect convergence?
4. Describe the difference between global and local optima.
5. How do techniques like gradient clipping help in training?
6. What are the benefits of data augmentation in deep learning?
7. How does early stopping prevent overfitting?
8. Explain the role of a validation set in model training.
9. Why might a neural network's training loss decrease while its validation loss increases?
10. Describe the challenges of training a deep network from scratch.
11. How can you handle imbalanced datasets in neural network training?
12. Explain the difference between stochastic gradient descent (SGD) and mini-batch gradient descent.
13. What is the adaptive learning rate, and why is it beneficial?
14. Describe the workings of the Adam optimizer and its advantages.
15. How do learning rate schedulers work, and why are they used?
16. Why do we shuffle the training data after each epoch?
17. What are weight constraints, and how can they benefit training?
18. Explain the significance of the second moment in the Adam optimizer.
19. How does the RMSprop optimizer differ from vanilla SGD?
20. What are common symptoms of overfitting, and how can you diagnose them?

Application Cases

1. How would you approach image segmentation using deep learning?
2. Describe a potential use-case for sequence-to-sequence models.
3. How can CNNs be used in time-series forecasting?
4. Explain how RNNs can be applied to generate text.
5. Describe the primary challenges in applying deep learning to medical image analysis.
6. How might you leverage pre-trained models for a new image classification task?
7. What considerations should be taken into account when using deep learning for real-time applications?
8. Describe how deep learning can be applied in anomaly detection.
9. How can autoencoders be used for dimensionality reduction?
10. Explain the application of deep learning in voice recognition.
11. Describe the role of deep learning in autonomous vehicles.
12. How can transformers be applied to tabular data analysis?
13. Explain the significance of attention mechanisms in machine translation.
14. Describe how reinforcement learning can be used in game playing (e.g., AlphaGo).
15. How might you use deep learning to perform style transfer in images?
16. Explain the potential of GANs in generating artwork or music.
17. How can deep learning assist in sentiment analysis from social media data?
18. Describe a use-case for deep learning in video analysis or video surveillance.
19. How can deep learning be used in recommendation systems?
20. Explain the role of neural networks in predicting stock market movements.

Large Language Models

1. How do GPT-3 and RNNs differ?
2. Describe the transformer architecture.
3. Why is self-attention important in transformers?
4. Define "pre-training" vs. "fine-tuning" in LLMs.
5. Advantages of masked language models like BERT?
6. Why are positional encodings in transformers needed?
7. Ethical concerns with deploying LLMs?
8. Challenges of training GPT-3-sized models?
9. How does knowledge distillation benefit LLMs?
10. What's "few-shot" learning in LLMs?
11. BERT's design tasks and data preparation?
12. Differences between GPT-3, BERT, and T5?
13. Fine-tuning LLMs for question answering?
14. Challenges with LLMs in multilingual tasks?
15. Risks of generating harmful content with LLMs?
16. Role of tokenizers like BPE in LLMs?
17. Importance of "attention heads" in transformers?
18. Handling out-of-vocabulary words in LLMs?
19. LLM applications outside of NLP?
20. Evaluating LLM performance metrics?