Chapter 6 MCQs

1. What is Transfer Learning in machine learning?
    A. Training a model from scratch for every new task.
    B. Reusing a model developed for one task as a starting point for a second, related task.
    C. Only using small datasets for training.
    D. A method for optimizing model architecture.
2. What is the core idea behind Transfer Learning in image recognition?
    A. To learn only specific features for each new image.
    B. To reuse simple edges and textures learned by early CNN layers across different tasks.
    C. To always train models from scratch.
    D. To avoid using convolutional neural networks.
3. What is a key advantage of Transfer Learning when modeling a second, related task?
    A. It requires a larger dataset.
    B. It allows rapid progress and higher performance with less data.
    C. It increases computational resources.
    D. It makes the model less flexible.
4. In modern image or natural language processing tasks, what is a common practice instead of training a model from scratch?
    A. Using only traditional machine learning models.
    B. Starting from a pre-trained model.
    C. Manually extracting all features.
    D. Avoiding deep learning frameworks.
5. Which of the following is NOT a typical example of models that form the basis of Transfer Learning?
    A. ImageNet
    B. AlexNet
    C. Inception
    D. Decision Tree
6. How does traditional training differ from Transfer Learning regarding initial weights?
    A. Traditional training starts with pre-trained weights.
    B. Transfer learning starts with randomly initialized weights.
    C. Traditional training starts with randomly initialized weights.
    D. Both start with pre-trained weights.
7. What is a requirement for traditional training approaches that is often reduced with Transfer Learning?
    A. Less data
    B. Shorter training time
    C. Large dataset and substantial computational resources
    D. Improved generalization
8. Transfer learning often adapts a pre-trained model to a new task by modifying or fine-tuning which layers?
    A. Only the input layers.
    B. Only the early layers.
    C. Only the final layers.
    D. All layers equally.
9. What is a benefit of Transfer Learning in terms of generalization?
    A. It tends to generalize worse.

B. It builds on features proven effective in a broader context, leading to better generalization.
C. It only works for very specific tasks.
D. It makes models more prone to overfitting.

10. In Computer Vision, Transfer Learning is widely used for tasks like:
    A. Sentiment analysis
    B. Question answering
    C. Image classification and object detection
    D. Time-series forecasting
11. In Natural Language Processing (NLP), what do models like BERT and GPT utilize before being fine-tuned for specific tasks?
    A. Manual feature engineering
    B. Pre-training on vast corpora
    C. Training from scratch on small datasets
    D. Linear regression algorithms
12. What is another term for reusing pre-trained models?
    A. Model compression
    B. Transfer learning
    C. Ensemble learning
    D. Active learning
13. What are pre-trained models?
    A. Models that have not yet been trained.
    B. Models trained on small, specific datasets.
    C. Models that have already been trained on a large dataset to perform a specific task.
    D. Models designed for hardware acceleration only.
14. In computer vision, models like ResNet, VGG, or EfficientNet are commonly pre-trained on which dataset?
    A. MNIST
    B. CIFAR-10
    C. ImageNet
    D. COCO
15. What basic visual features have models pre-trained on ImageNet already learned to detect?
    A. Edges, textures, and shapes
    B. Complex semantic meanings
    C. Human emotions
    D. Audio patterns
16. What is the main advantage of using a pre-trained model?
    A. It increases training time.
    B. It requires more computational resources.
    C. It saves time, resources, and data.
    D. It makes models less accurate.
17. Which of the following is a benefit of using pre-trained weights?
    A. Slower training
    B. Worse performance with less data
    C. Lower computational costs
    D. Decreased generalization
18. Pre-trained weights are particularly helpful when you have what kind of dataset?
    A. A very large dataset

B. A limited dataset
C. A perfectly balanced dataset
D. A dataset with no noise

19. What does using pre-trained weights allow you to skip, making the process more cost-effective?
    A. The final prediction step.
    B. The most resource-intensive parts of the training process.
    C. Data preprocessing.
    D. Model evaluation.
20. Models with pre-trained weights often generalize better to new data because they have learned what?
    A. Specific, low-level features.
    B. Robust, high-level features from diverse datasets.
    C. Only features relevant to the original task.
    D. Biased representations.
21. What is the difference between Feature Extraction and Fine-Tuning in transfer learning?
    A. Feature Extraction trains all layers, while Fine-Tuning freezes them.
    B. Feature Extraction freezes most pre-trained layers, while Fine-Tuning unfreezes some or all.
    C. Feature Extraction is only for text, and Fine-Tuning is only for images.
    D. There is no significant difference.
22. In Feature Extraction, which layers are typically trained on new data?
    A. All pre-trained layers.
    B. Only the newly added classifier layers.
    C. Only the input layer.
    D. A random subset of layers.
23. When is Feature Extraction most effective?
    A. When the new task differs significantly from the original.
    B. When you have a very large dataset.
    C. With limited datasets, especially for tasks closely related to the original.
    D. When computational resources are abundant.
24. What is a characteristic of Fine-Tuning compared to Feature Extraction regarding dataset size?
    A. Requires a smaller dataset.
    B. Requires a larger dataset to avoid overfitting.
    C. Dataset size does not matter.
    D. It works best with unlabeled data.
25. Which approach in transfer learning is generally shorter in training time?
    A. Fine-Tuning
    B. Feature Extraction
    C. Both take the same amount of time.
    D. Neither impacts training time.
26. Which approach has potentially higher accuracy, especially when the new task differs significantly from the original?
    A. Feature Extraction
    B. Fine-Tuning
    C. Both have similar accuracy.
    D. Accuracy is not impacted by the approach.
27. When is Fine-Tuning ideal?

A. When the new task is very similar to the original.
B. When the new task differs significantly from the original, necessitating more adaptation.
C. When data is extremely limited.
D. When you want to keep the model simple.

28. What is one of the compelling advantages of Transfer Learning regarding training time and resources?
A. It increases both.
B. It significantly cuts down on both.
C. It has no impact on either.
D. It only affects training time, not resources.
29. In scenarios where collecting large datasets is challenging, how does Transfer Learning help?
A. It makes model development impossible.
B. It enables robust models with relatively small amounts of data.
C. It requires even more data.
D. It leads to underfitting.
30. What is a benefit of Transfer Learning in terms of model accuracy?
A. It leads to lower accuracy.
B. It has no impact on accuracy.
C. It can lead to higher accuracy in the new task.
D. It only maintains the original model's accuracy.
31. How does Transfer Learning democratize access to advanced machine learning models?
A. By making them more expensive.
B. By limiting access to only large organizations.
C. By allowing organizations with limited resources to utilize sophisticated pre-trained models.
D. By requiring extensive manual coding.
32. What type of models are pre-trained CNN models typically trained on?
A. Small, custom datasets.
B. Large benchmark datasets like ImageNet.
C. Unlabeled data only.
D. Synthetic data.
33. What is a benefit of using pre-trained CNN models?
A. Increased training time.
B. Decreased computational resources.
C. Worse performance with limited data.
D. Reduced generalization.
34. Which pre-trained CNN architecture introduced the concept of "residual learning" to address the vanishing gradient problem?
A. VGG
B. MobileNet
C. EfficientNet
D. ResNet
35. Which pre-trained CNN architecture is known for its simplicity and uniform architecture, utilizing small convolutional filters?
A. ResNet
B. VGG
C. MobileNet

D. EfficientNet

36. Which pre-trained CNN architecture is designed for mobile and embedded vision applications due to its efficiency?
    A. ResNet
    B. VGG
    C. MobileNet
    D. EfficientNet
37. Which pre-trained CNN architecture balances network depth, width, and resolution using a "compound coefficient"?
    A. ResNet
    B. VGG
    C. MobileNet
    D. EfficientNet
38. What was a significant achievement of ResNet in the ILSVRC 2015 competition?
    A. Winning only the object detection task.
    B. Achieving a training error of only 3.57% in classification.
    C. Being the first model to use transfer learning.
    D. Being the smallest model.
39. What does the "deep" in VGG-16 or VGG-19 refer to?
    A. The complexity of the input data.
    B. The number of convolutional layers.
    C. The depth of the feature maps.
    D. The depth of the input image.
40. MobileNet is widely used in real-world applications including:
    A. Speech recognition
    B. Object detection and fine-grained classifications
    C. Time-series forecasting
    D. Generating text
41. What technique does EfficientNet use to scale up models effectively?
    A. Random scaling of width, depth, or resolution.
    B. Compound coefficient scaling.
    C. Manual scaling.
    D. Linear scaling.
42. When implementing pre-trained CNN models, what should you typically exclude when loading the model from frameworks like TensorFlow?
    A. The input layers.
    B. The convolutional layers.
    C. The top (final) layers.
    D. The pooling layers.
43. In TensorFlow, what parameter is set to `False` when loading a pre-trained model like ResNet50 if you want to exclude the top layers?
    A. `weights`
    B. `input_shape`
    C. `include_top`
    D. `trainable`
44. What type of layer is typically added after a Global Average Pooling layer when modifying a pre-trained model for a specific task?
    A. Convolutional layers
    B. Recurrent layers
    C. Dense layers

D. Pooling layers

45. What is the purpose of "freezing" the pre-trained layers during the initial training phase?
    A. To increase their trainability.
    B. To preserve the learned features.
    C. To make the model more complex.
    D. To speed up the forward pass only.
46. When fine-tuning a model, what is typically done with some of the deeper layers of the pre-trained model?
    A. They are removed.
    B. They are frozen permanently.
    C. They are unfrozen to allow further training.
    D. They are replaced with new layers.
47. What is a common optimizer used when compiling a model in TensorFlow/Keras for transfer learning?
    A. SGD
    B. RMSprop
    C. Adam
    D. Adagrad
48. What is a common loss function used for multi-class classification when compiling a model in TensorFlow/Keras for transfer learning?
    A. Mean Squared Error
    B. Binary Cross-Entropy
    C. Categorical Crossentropy
    D. Huber Loss
49. What is a key benefit of using existing architectures and learned features from pre-trained models?
    A. Increased development time.
    B. Significantly reducing development time and computational resources.
    C. Making models less accurate.
    D. Requiring more manual effort.
50. What is the purpose of recompiling the model with a lower learning rate during fine-tuning?
    A. To make the model learn faster.
    B. To prevent any further learning.
    C. To allow for more subtle adjustments to the pre-trained weights.
    D. To increase the loss.