2. Model selection

1. Multilayer Perceptrons (MLPs):

- **Description:** The most basic form of deep neural network, consisting of feedforward layers that are fully connected. Each neuron is connected to all neurons in the previous layer.
- **Advantages:** Simple in structure, easy to understand, and can be trained with relatively few computational resources.
- **Disadvantages:** Difficult to learn complex patterns and not suitable for structured data such as images or sequences.
- **Suitable Tasks:**
- Tabular data: Data stored in spreadsheets or databases. For example, customer data, financial data, etc.
- Simple classification and regression tasks.
- Basic pattern recognition.
-

2. Convolutional Neural Networks (CNNs):

- **Description:** An architecture specifically designed for image and video processing. It uses convolutional layers to extract spatial features from images and pooling layers to reduce the size of the feature maps.
- **Advantages:** Can effectively learn the spatial features of images and demonstrates excellent performance in various computer vision tasks such as image classification, object detection, and image segmentation.
- **Disadvantages:** More complex in structure than MLPs and requires more computational resources for training.
- **Suitable Tasks:**
- Image classification: Classifying images into specific categories (e.g., cat vs. dog).
- Object detection: Identifying the location and type of objects within an image (e.g., self-driving cars).
- Image segmentation: Classifying each pixel of an image into a specific class (e.g., medical image analysis).
- Video analysis: Tracking objects or detecting events within a video.
-

3. Recurrent Neural Networks (RNNs):

- **Description:** An architecture designed to process sequential data such as text and time series data. It uses recurrent connections to maintain information about past inputs and learn the temporal dependencies of sequences.
- **Advantages:** Can learn the temporal dependencies of sequences and is used in various tasks such as natural language processing, speech recognition, and time series forecasting.
- **Disadvantages:** Can suffer from the vanishing gradient problem when processing long sequences, and training can be relatively difficult.
- **LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit):** Variants of RNNs designed to address the vanishing gradient problem. LSTMs and GRUs use gate mechanisms to selectively retain or delete information, allowing them to better learn long-term dependencies.
- **Suitable Tasks:**
- Natural Language Processing: Text classification, machine translation, text generation, etc.
- Speech Recognition: Converting speech data into text.
- Time Series Forecasting: Stock price prediction, weather forecasting, etc.
- Sentiment Analysis: Analyzing the sentiment (positive, negative, neutral) of text.
-

4. Transformers:

- **Description:** A modern architecture based on self-attention mechanisms that has achieved state-of-the-art performance in natural language processing and computer vision. Self-attention learns the relationships between all positions within a sequence, allowing it to effectively model long-term dependencies.
- **Advantages:** Allows for parallel processing, resulting in faster training, learns long-term dependencies well, and exhibits excellent performance in natural language processing and computer vision.
- **Disadvantages:** More complex in structure than RNNs and requires more computational resources.
- **Suitable Tasks:**
- Natural Language Processing: Machine translation, text summarization, question answering, etc.
- Computer Vision: Image classification, object detection, image generation, etc.
-

5. Autoencoders:

- **Description:** An architecture used for unsupervised learning tasks that aims to encode input data into a low-dimensional latent space and then decode it back to the original input data.
- **Advantages:** Can be used for various tasks such as dimensionality reduction, feature extraction, and anomaly detection.
- **Disadvantages:** Training can be difficult, and performance can vary depending on the characteristics of the data.
- **Suitable Tasks:**
- Dimensionality Reduction: Reducing the dimensionality of data to visualize it or save storage space.
- Feature Extraction: Extracting useful features from data.
- Anomaly Detection: Detecting abnormal data that differs from normal data.
-

6. Generative Adversarial Networks (GANs):

- **Description:** An architecture that trains two neural networks, a generator and a discriminator, competitively to generate new data samples that resemble the training data. The generator creates fake data, and the discriminator tries to distinguish between real and fake data.
- **Advantages:** Can generate various types of data such as realistic images, music, and text.
- **Disadvantages:** Training can be very unstable and difficult, and the quality of the generated data may not be guaranteed.
- **Suitable Tasks:**
- Image Generation: Generating new images.
- Image Editing: Changing the style of an image or adding/removing objects.
- Data Augmentation: Generating new data to increase the amount of training data.
-

Considerations When Choosing a Model Architecture:

- **Complexity of the task:** You should choose an appropriate model architecture depending on how complex the problem you are trying to solve is. More complex problems may require more complex architectures.
- **Size of the dataset:** The size of the dataset has a significant impact on the performance of the model. If the dataset is small, you should choose a simpler model to prevent overfitting.
- **Computational resources:** Training a model can require significant computational resources. You should choose a model architecture considering the available computational resources.