Ques 1
A neuron is a single unit in live creatures that processes and transmits information, whereas a neural network is a computer model made of several artificial neurons organized in layers to perform complicated tasks like pattern recognition, decision-making, and prediction. The form and function of biological neurons inspire neural networks, which are built for computational reasons.

Ques 2
1. The cell body (soma) is the core component of the neuron and includes the nucleus, which contains the genetic material (DNA). It is in charge of maintaining the neuron's general function and is critical in integrating information received from other neurons.
2. Dendrites are branching structures that protrude from the cell body. These neurons are the major recipients of messages from other neurons. Dendrites are coated with synapses, which are specialized connections where neurotransmitters (chemical messengers) from other neurons are released to interact with the cell.
3. The axon is a long, thin, cable-like protrusion extending from the cell body. It transports the electrical signal created in the cell body (by signal integration) away from the neuron and towards other neurons or effector cells (e.g., muscle cells or glands).
4. Myelin Sheath: The axon of certain neurons is coated in a myelin sheath generated by glial cells. The myelin coating works as an insulator, reducing electrical signal leakage and allowing for quicker signal transmission down the axon.
5. Axon Terminals (Synaptic Terminals): Axon terminals are tiny structures at the end of the axon. Synaptic vesicles containing neurotransmitters are found at axon terminals.
6. Synapse: A synapse is a tiny gap or connection between one neuron's axon terminal and the dendrite (or, in rare cases, cell body) of another neuron. Neurotransmitters produced from the first neuron's axon terminal connect to receptors on the dendrite of the second neuron, causing an electrical signal to be sent to the receiving cell.

Ques 3
Architecture
1. Input layer = this is responsible for receiving of the input data.
2. Weights = each input layer in node is associated with a weight. Weights represents the strength of each input
3. Activation function = The output of the weighted sum of inputs is passed through an activation function

Functioning
1. Input and weights = the perceptron receives input data along with the crresponding weights
2. Weighted sum = the weighted sum of the inputs is calculated
3. Activation function = the weighted sum obtained is passed through the activation function thus calculating the output.
4. Decision = the perceptron makes decision on the output from the previous step

Ques 4

The primary distinction between a perceptron and a multilayer perceptron is one of construction and capability. The perceptron is a single-layer neural network that can only handle linearly separable issues, but the multilayer perceptron contains many hidden layers and non-linear activation functions that allow it to handle non-linearly separable problems and do more difficult tasks

Ques 5

.forward propagation is the core process through which data flows through a neural network, undergoing weighted sums and activations, ultimately leading to the generation of predictions or outputs.

Ques 6

Backpropagation is an important technique in neural network training because it allows weights to be adjusted to minimise the loss function, allowing the model to learn and generalise from the training data to make correct predictions on new data.

QUES 7

The chain rule enables backpropagation to alter the weights and enhance the model's performance via training by quickly computing the gradients of the loss function with respect to the weights in a neural network.

QUES 8

Loss functions are important in neural networks because they assess model performance, guide training by giving gradients for weight updates, and assist optimise the model to make correct predictions on fresh data. The proper loss function is determined by the job at hand and the type of the target output that the neural network is attempting to anticipate.

Ques 9

1. Classificatio loss func
2. Regression loss func
3. Sequence gen func

Ques 10

Optimizers in neural networks are in charge of changing the model's weights during training based on backpropagation gradients. Their goal is to minimise the loss function, allowing the model to learn and predict accurately on fresh data. Different optimizers employ different algorithms to effectively update the weights, and their selection and tuning are critical for the success of the neural network training process.

Ques 11
The exploding gradient problem arises during deep neural network training, particularly in the context of gradient-based optimisation methods such as stochastic gradient descent (SGD) and its derivatives. When the gradients of the model's parameters grow exceptionally huge during the backpropagation process, the weights update by a substantial amount. As a result, the network may experience inconsistent and divergent training, making it difficult for it to learn properly.

1. Gradient Clipping:
2. Weight initialization
3. batch norm
4. Learning rate scheduling
5. Gradient regu
6. Reducing model complexity

Ques 12
Another issue that develops during deep neural network training is the vanishing gradient problem, which is especially problematic in multi-layer designs. During the backpropagation process, the gradients of the model's parameters become incredibly tiny as they propagate backward from the output layer to the input layer. As a result, the weights of the older layers are changed seldom or not at all, resulting in these layers essentially "lagging behind" in learning when compared to the later layers.

1. Slow conv
2. Stalling  learning
3. Difficulty capturing long term dependencies
4. Model performance suffers

Ques 13
Regularisation is an effective strategy for preventing overfitting in neural networks and other machine learning models. When a model grows too complicated, it begins to memorise the training data rather than learning general patterns. As a result, it does well on training data but badly on fresh, previously unknown data (validation or test data). Regularisation adds restrictions or penalties to the model during training to keep it from fitting the noise in the data and encourages it to acquire more robust and generalizable representations.

Ques 14
Normalisation techniques not only increase convergence during training but also add to the neural network's overall stability and generalisation. They make it easy for optimisation algorithms to discover the best answer and keep the model from being sensitive to the size of input characteristics. Furthermore, normalisation can improve performance by allowing the network to acquire meaningful representations in the data, which is important for tasks like as image recognition, natural language processing, and other complicated data domains.

Ques 15

1. Relu
2. Leakay relu
3. Elu
4. Sigmoid
5. Tan h
6. Softmax

Qus 16

Batch normalisation is a technique for normalising the activations of each layer in a mini-batch during neural network training. It was developed to overcome two major deep learning challenges: internal covariate shift and the vanishing/exploding gradient problem.

Advan

1. Faster conv
2. Higher learning rate
3. Reduced overfitting
4. independence

Ques 17

Weight initialization is a fundamental step in neural network training. It refers to the process of initialising the model's weights and biases before to training. Proper weight initialization is critical since it has a substantial influence on the neural network's convergence speed, performance, and stability throughout training.

1. Vanishin ggradients
2. Slow conv
3. Stagnation

Ques 18

1. Faster cov
2. Stablization
3. Overcoming space gradients
4. Smoothing trajectories

Ques 19

Both L1 and L2 regularisation strategies are useful for reducing overfitting in neural networks, and the decision relies on the specific situation and desired model properties.

Ques 20

Early halting may be used as a simple yet efficient regularisation approach in neural networks to prevent overfitting. It helps generate models that generalise better to fresh data by terminating the training process before overfitting begins.

Ques 21
Dropout regularisation is a straightforward yet extremely successful approach. During each training cycle, a certain number of neurons in a layer (usually 20% to 50%) are "dropped out" or deactivated at random. This implies that their outputs are set to zero, and they contribute nothing to the forward or backward pass (gradient computations). The dropout is used stochastically, which means that it is unique for each training sample or mini-batch.

1. Neural network train
2. Regi]ularization
3. Flexibility
4. Tuning dropout rate

qUes 22
It is critical to choose an adequate learning rate while training neural networks. It has an immediate effect on the model's convergence, stability, and generalisation performance. It is a hyperparameter that should be carefully set and frequently necessitates experimentation to get the best value for a certain job and architecture. Techniques like as learning rate decay, adaptive learning rate algorithms, and employing learning rate schedules can help in determining an appropriate learning rate and improving the training process.

Ques 23
Vanishing
Exploding
Overrfitting
Large dataset requirement
Computational demands
Data augmentation
long training time

Ques 24
The primary distinction between a CNN and a conventional NN is seen in their design and connection. CNNs are especially intended to handle grid-like data, such as photographs, by taking advantage of local patterns and lowering the number of parameters through weight sharing. This architecture makes CNNs very effective for a variety of computer vision tasks and has been critical in delivering cutting-edge performance in a variety of image-related applications.

Ques 25

Pooling layers in CNNs serve to minimise computational complexity, boost translation invariance, and encourage feature selection by downsampling feature maps. They are vital in obtaining the most significant information from feature maps while rejecting less important elements. Pooling layers, along with convolutional layers, are critical components of CNN designs and have contributed to CNN success in a variety of computer vision applications.

Ques 26
A recurrent neural network (RNN) is a sort of artificial neural network that is meant to handle sequential input by retaining hidden states that contain information from prior time steps. RNNs, as opposed to feedforward neural networks, feature loops that allow information to persist over time, making them well-suited for tasks requiring sequences and time series data.

1. Nlp
2. Speech reco
3. Gesture reco
4. Time series pre
5. Music gene
6. Video anal
7. Healthcare

Ques 27
LSTM networks have transformed sequence modelling and contributed significantly to the success of several natural language processing and sequential data jobs. Because of their capacity to capture long-term relationships and alleviate vanishing gradient concerns, they are an essential tool in current deep learning applications.

Ques 28
Ian Goodfellow and his colleagues proposed Generative Adversarial Networks (GANs) in 2014 as a class of deep learning models. GANs are made up of two neural networks, the generator and the discriminator, that are competitively trained together. GANs are generally used for generative jobs such creating realistic pictures, audio, video, or other forms of data.

1. The generator (G) accepts random noise as input and creates synthetic data samples. In picture generation, for example, the generator takes a random vector (typically chosen from a standard normal distribution) and generates an image that is ideally similar to the real images from the training dataset.
2. Discriminator (D): A discriminator is a binary classifier that attempts to discriminate between genuine data samples from the training dataset and produced samples from the generator. Its purpose is to accurately recognise actual samples as real (a label of 1) and manufactured samples as false (a label of 0).

Ques 29
Autoencoder neural networks are a form of unsupervised learning model that uses compression and reconstruction to learn efficient data representations. Autoencoders are mostly used for dimensionality reduction and feature learning. They are frequently used for data compression, denoising, and anomaly detection.

Functioning

1. Architect
2. Training
3. Latent space represen

Purpose
1. Dimensioonalty
2. Feature learn
3. Data compression
4. Image generation

Ques 30
Self-Organizing Maps are useful for visualising and comprehending large datasets, as well as for tackling a variety of unsupervised learning challenges. Their capacity to organise data based on topological links qualifies them for a wide range of applications in many disciplines.

Ques 31
1. Data prepara
2. Model archi
3. Input norm
4. Loss func
5. Output layer
6. Model train
7. Hyperparameter tuning
8. Evaluation
9. Prediction
10. monitor

Ques 32
Computational resources
Memory constraints
Long training time
Overfitting

Tuning
Augmentation
Data imbalance
Transfer learning

Ques 33
Transfer learning is a machine learning approach that uses knowledge obtained from training a model on one task or domain to enhance performance on another but related task or area. Transfer learning in neural networks is using pre-trained models as a starting point for a new task rather than developing a model from scratch.

Benefit
1. Less train time
2. Less dependency
3. Improved generali
4. Handling domains
5.

Ques 34
During the training phase of neural networks for anomaly detection, it is critical to have a well-labeled dataset with both normal and anomalous examples. In other circumstances, the training data may contain only normal occurrences, and the model is fine-tuned on a small sample of labelled anomalies or semi-supervised approaches to efficiently detect anomalies in unseen data.

Ques 35
Model interpretability in neural networks is a hotly debated topic. Researchers are working on creative approaches to increase interpretability while maintaining performance. Techniques like as integrated gradients, occlusion sensitivity, and adversarial example analysis are constantly improving in order to give improved insights into model behaviour.

Ques 36
Deep learning is a strong and adaptable technique to machine learning, especially when dealing with complicated, unstructured data. Its capacity to learn key traits and adapt to varied domains has resulted in significant achievements in a variety of industries. Deep learning, on the other hand, presents obstacles in terms of data needs, computing resources, interpretability, and the potential of overfitting. Traditional machine learning techniques, on the other hand, may be better suited for smaller datasets, interpretable models, and scenarios with more organised and well-defined data. The decision between deep learning and classical machine learning is determined by the job at hand, the availability of data, the criteria for interpretability, and the available resources.

Ques 37
The use of neural networks in ensemble learning is an effective strategy for improving model performance and generalisation. It can be used to generate more accurate and resilient models in a variety of methods, but it comes with higher processing costs and issues in model selection and interpretation. The efficiency of ensemble learning, like any other machine learning technique, is dependent on the individual issue as well as the quality and quantity of data available for training.

Ques 38
1 word embeddings
Text classification
Machine translation.
Summarization
sentiment analysis
Question ans
Speech reco
Clustering

Ques 39
The primary principle behind self-supervised learning is to use the data's inherent structure or relationships to generate relevant learning signals without the usage of external annotations. This is accomplished by generating fictitious "targets" or "pseudo-labels" from the input data and training the model to predict these labels. Once trained on this pretext task, the representations obtained may be used for a variety of downstream tasks including as classification, object identification, semantic segmentation, and more.

Applica
1. Representation learning
2. Nlp
3. Cv
4. Speech and audio processing
5. Robotics
6. Drug discovery
7. Medical imaging

Ques 40
1. Biased learning
2. Class bouadry confusion
3. Poor generalization
4. Rare class detection

5. Loss function imbalance overfitting
6. Sampling bias

QUES 41
With tiny changes in input data, adversarial assaults on neural networks influence models. They take advantage of model flaws and can result in inaccurate forecasts. Adversarial training, gradient masking, and ensemble approaches are examples of mitigation strategies. Obtaining perfect adversarial robustness, on the other hand, remains a difficulty. The discipline is continually evolving in order to increase model security and dependability.

Ques 42
A important feature of model training is the trade-off between model complexity and generalisation performance in neural networks. Regularisation methods and careful model selection are essential for striking the correct balance between allowing the model to generalise successfully to new, previously unknown data and avoiding overfitting or underfitting.

Ques 43
Removing missing data
Mean mode imputation
Multiple imputation
Autoencoders
Masking techniques
Data augmentation
Knn ummputauion

Ques 44
1. Shap
   Concept
   SHAP values are based on cooperative game theory and aim to fairly distribute the "credit" or "importance" of each feature in a prediction across all possible combinations of features.

   Benefits
   1. Global interprity
   2. Consistency and fairness
   3. Model debugging and improvements
   4. Feature selection
2. Shap
- Local interpretability
- Trustworthiness
- Debugging
- Sensitivity analysis

Shap values provide global insights into feature importance , highlights  trends and patterns across the entire dataset.
Lime provides local interrpretability, explaining individual predictions and enabling users to identify cases where the model may not be reliable.

Ques 45

1. Model optimization
2. Hardware acceleration
3. Model optimization
4. edge cloud collaboration
5. efficient modle architecture
6. Model running
7. On device catching
8. Multi model optimization
9. Task offroading
10. Latency aware design
11. energy efficiency

Ques 46

1. Data parallelism vs model parallelism
2. Communication overhead
3. Syncronization
4. Coordination
5. Fault tolerance
6. Scalability and load balancing
7. Resource management
8. Batch size and learning rate consistency
9. Hybrid training
10. Algo challenges
11. Cost and infra

Ques 47

1. Bias and fairness
2. Transparency
3. Privacy concerns
4. Adversarial attacks
5. Accountability and liability
6. Human autonomy and control
7. Distribution of benefits and harms
8. Job displacement

Ques 48
Reinforcement learning (RL) is a machine learning paradigm in which an agent learns to make decisions through interaction with its surroundings. Based on its activities, the agent receives feedback in the form of rewards or penalties, and its purpose is to maximise the cumulative reward over time. The primary concept is to develop an optimum policy—a mapping from states to actions—that directs the agent to adopt activities that result in the greatest long-term payoff.

Applications
1. Game playing
2. Robotics
3. Autonomous vehicles
4. Recommendation system
5. Resource allocation
6. Nlp
7. finance

Ques 49
1. Computational efficiency - large batch size can affect memory requirements and takes soo much time during training.
2. Gradient estimation - The gradient of the loss function is computed using a subset (mini-batch) of the training data in stochastic gradient descent (SGD), the most frequent optimisation approach for training neural networks.
3. Generalisation-The batch size can have an effect on the trained model's generalisation performance. Larger batch sizes can occasionally result in more generalizable models because they average out noise in gradient updates and explore a more consistent path in the weight space.
4. Convergence and learning rate - The learning rate can also be influenced by the batch size. Smaller batch sizes may necessitate a lower learning rate in order to prevent overshooting the ideal weights as a result of more frequent and noisy updates.
5. Parallelization - Larger batch sizes can take better advantage of parallel processing capabilities provided by GPUs or TPUs, leading to faster training times on appropriate hardware.
6. Batch noramalization

Ques 50
1. Data requirement
2. Computational resources
3. Generalization in small data settiiings
4. Interpretability
5. Adversial vulnerability
6. Bias and fairness

7. Transfer learning to new domains