UNIVERSITY EXAMINATION 2022/2023

EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE IN COMPUTER SCIENCE

SPE 2408: Artificial Neural Networks Year IV Semester II

Date: Wednesday 26th April 2023

Instructions

1. This EXAMINATION has TWO sections.

ii. SECTION A is compulsory. Answer ALL questions in this section.

Answer ANY TWO questions in SECTION B.

Question One:

a) **Differentiation:** i. **Epoch and Batch:** - Epoch: One pass through the entire dataset during training is called an epoch. - Batch: The number of samples processed before the model is updated is called a batch.

ii. **Parameter and Hyperparameter:** - Parameter: Parameters are the variables that the model learns from the training data. They are the coefficients of the model that are optimized during training. - Hyperparameter: Hyperparameters are configuration settings that govern the learning process. They are set before training and control aspects such as model complexity and learning rate.

iii. **Unsupervised Learning and Reinforcement Learning:** - Unsupervised Learning: In unsupervised learning, the model learns patterns from unlabeled data without explicit guidance. The goal is to uncover hidden structures or relationships in the data. - Reinforcement Learning: Reinforcement learning involves an agent learning to make decisions by interacting with an environment. The agent receives feedback in the form of rewards or penalties based on its actions.

iv. **Training Loss and Validation Loss:** - Training Loss: The loss function computed on the training dataset during model training. - Validation Loss: The loss function computed on a

separate validation dataset during model training to evaluate the model's performance on unseen data.

b) **Nonlinear Activation Functions:** i. **ReLU (Rectified Linear Unit):** ReLU is defined as $f(x)=\max(0,x)$. It introduces nonlinearity by outputting the input directly if it is positive, and zero otherwise. ReLU is computationally efficient and helps mitigate the vanishing gradient problem.

ii. **Sigmoid Function:** The sigmoid function is defined as $f(x)=\frac{1}{1+e^{-x}}$. It squashes the input to the range (0, 1), making it suitable for binary classification problems. However, it suffers from the vanishing gradient problem and is not zero-centered.

c) **Single Artificial Neuron:**

- The model of a single artificial neuron consists of:
  - Input values $x_1, x_2, \ldots, x_n$ weighted by corresponding weights $w_1, w_2, \ldots, w_n$.
  - A bias term b.
  - Activation function $f(\cdot)$.
    - The output of the neuron is computed as $\text{output} = f\left(\sum_{i=1}^{n} (w_i \cdot x_i) + b\right)$.

d) **McCulloch-Pitts Neuron:** i. **Requirements:** 1. Binary inputs and outputs. 2. Fixed weights and thresholds. 3. Summation of inputs with weights. 4. Thresholding activation function.

ii. **XOR Function with McCulloch-Pitts Neuron:** - XOR function cannot be implemented with a single McCulloch-Pitts neuron because it's not linearly separable. However, it can be implemented using a multi-layer perceptron with hidden layers.

e) **Neural Network Applications:**

- **Medicine:** Diagnosis of diseases from medical images (MRI, CT scans), drug discovery, personalized treatment plans.
- **Automotive:** Autonomous driving, vehicle diagnostics, predictive maintenance.
- **Banking:** Fraud detection, credit risk assessment, customer relationship management.

f) **Biological Properties Borrowed into ANNs:**

- Neurons: Artificial neurons mimic the functionality of biological neurons.
- Synapses: Connections between neurons in ANNs are analogous to synapses in the brain.
- Learning: ANNs use learning algorithms inspired by synaptic plasticity and neural network dynamics in the brain.
- Layers: The hierarchical organization of neurons into layers in ANNs resembles the layered structure of the brain's cortex.

Let's continue to answer the next set of questions.
Question Two:

a) **Perceptron Implementation:**

```
def perceptron           """ Compute the output of a perceptron. Parameters: W (list):
List of weights. x (list): List of inputs. b (float): Bias term. Returns: int: Output of
the perceptron (1 if Wx + b > 0, else 0). """          sum         for        in zip
        return 1 if          0 else 0
```

b) **Hamming Net Algorithm:**

- The Hamming Net is a type of neural network that uses Hamming distance as a measure of similarity between input patterns.
- It is typically used for pattern recognition and classification tasks.
- The algorithm assigns weights to input patterns based on their similarity to stored patterns and classifies new patterns by computing their Hamming distance from the stored patterns.

c) **Activation Function in Hidden Layer:**

- The choice of activation function in a hidden layer depends on the specific task and the properties of the data.
- ReLU (Rectified Linear Unit) is recommended due to its simplicity, computational efficiency, and ability to mitigate the vanishing gradient problem.
- ReLU is preferred over sigmoid and tanh functions in deep neural networks because it avoids the vanishing gradient problem and allows faster convergence during training.

d) **Forward vs. Backpropagation:**

- **Forward Propagation:** In forward propagation, the input data is passed forward through the network to compute the predicted output. Each layer performs a linear transformation followed by an activation function.
- **Backpropagation:** In backpropagation, the error between the predicted output and the actual output is calculated using a loss function. The error is then propagated backward through the network to update the model's parameters (weights and biases) using gradient descent.

e) **Bird Classifier using Transfer Learning:**

- **Step 1: Data Collection:** Gather a dataset of bird images with corresponding labels.
- **Step 2: Preprocessing:** Preprocess the images (resize, normalize) and split the dataset into training and validation sets.
- **Step 3: Transfer Learning:** Use a pre-trained convolutional neural network (CNN) model (e.g., ResNet, Inception) trained on a large dataset like ImageNet.
- **Step 4: Fine-tuning:** Fine-tune the pre-trained model on the bird dataset by updating the weights of the last few layers while keeping the weights of earlier layers frozen.
- **Step 5: Evaluation:** Evaluate the performance of the trained model on the validation set and fine-tune hyperparameters if necessary.

Let's proceed to answer the next set of questions.

Question Three:

a) **Cross-Entropy Loss:**

- Cross-entropy loss measures the difference between two probability distributions, typically the predicted probability distribution and the true probability distribution.
- It is commonly used as a loss function in classification problems, where the output is a probability distribution over multiple classes.
- For example, in binary classification, the cross-entropy loss is defined as $-\sum\limits_{i}(y_i \log(\hat{y}_i)+(1-y_i)\log(1-\hat{y}_i))$, where y_i is the true label and $\hat{y}_i$ is the predicted probability of the ith class.

b) **Advantages and Challenges of ANN:**

- **Advantages:**
  - Ability to learn complex patterns and relationships from data.

- Versatility in handling various types of data, including structured and unstructured data.
  - **Challenges:**
- Overfitting: ANNs may memorize noise in the training data, leading to poor generalization to unseen data.
- Computational Complexity: Training deep neural networks requires significant computational resources and time.

c) **Steps for Developing an AI Application:**

- **Problem Definition:** Clearly define the problem to be solved and the objectives of the AI application.
- **Data Collection:** Gather relevant data that will be used to train and evaluate the AI model.
- **Data Preprocessing:** Clean, preprocess, and prepare the data for training, including handling missing values, normalization, and feature engineering.
- **Model Selection:** Choose an appropriate machine learning model or neural network architecture based on the problem and data characteristics.
- **Training:** Train the selected model using the prepared data, optimizing the model's parameters to minimize the chosen loss function.
- **Evaluation:** Evaluate the trained model's performance on a separate validation set or through cross-validation techniques.
- **Deployment:** Deploy the trained model into production and integrate it into the target system or application.
- **Monitoring and Maintenance:** Monitor the model's performance in production, retrain periodically with new data, and update as needed.

d) **Self-Organizing Maps (SOMs):**

- SOMs are a type of artificial neural network used for clustering and dimensionality reduction.
- They organize high-dimensional input data into a lower-dimensional grid of neurons, preserving the topological properties of the input space.
- SOMs learn through competitive learning, where neurons compete to represent different regions of the input space.
- During training, the neurons adjust their weights to become similar to the input patterns they are exposed to, resulting in a topological map of the input space.

Let's continue to answer the remaining questions.

Question Four:

a) **Feedback vs. Feedforward in ANN:**

- **Feedback:** Feedback connections in ANNs allow information to flow backward from higher layers to lower layers or from the output to the input. They are commonly used in recurrent neural networks (RNNs) and feedback alignment algorithms.
- **Feedforward:** Feedforward connections in ANNs propagate information forward from the input layer through hidden layers to the output layer. They are prevalent in traditional feedforward neural networks and convolutional neural networks (CNNs).

b) **Differences between Human Brain and Computers:**

- **Processing Mechanism:**

  - Human Brain: Parallel processing with interconnected neurons capable of processing vast amounts of information simultaneously.
  - Computers: Sequential processing with discrete components performing operations one at a time, although modern computers use parallel processing techniques.

- **Learning Capability:**

  - Human Brain: Adaptive learning through synaptic plasticity, enabling continuous learning and adaptation to new environments.
  - Computers: Require explicit programming and training data to learn, lacking the adaptability and generalization capabilities of the human brain.

- **Energy Efficiency:**

  - Human Brain: Highly energy-efficient, consuming relatively low power compared to its processing capacity.
  - Computers: Require significant energy consumption, especially for large-scale computations, leading to concerns about sustainability and environmental impact.

c) **Learning Mechanisms in ANN:** i. **Error-Based Learning:** - In error-based learning, the model adjusts its parameters based on the difference between predicted and true outputs, minimizing the error using optimization algorithms like gradient descent. ii. **Memory-Based Learning:** - Memory-based learning methods store training instances in memory and use them to make predictions or classify new instances. Examples include k-nearest neighbors (KNN) and case-based reasoning (CBR).

Let's proceed to answer the final set of questions.

Question Five:

a) **Explanation of Concepts in ANN:** i. **Kullback-Liebler Divergence:** Also known as relative entropy, Kullback-Liebler Divergence measures the difference between two

probability distributions. It is often used as a loss function in training generative models such as variational autoencoders. ii. **ReLU (Rectified Linear Unit):** ReLU is an activation function commonly used in neural networks. It replaces negative input values with zero, allowing the network to learn nonlinearities and speeding up training by mitigating the vanishing gradient problem. iii. **Stochastic Gradient Descent (SGD):** SGD is an optimization algorithm used to minimize the loss function during training of neural networks. It updates the model parameters using gradients computed on a subset of the training data (mini-batch) rather than the entire dataset, making it computationally efficient. iv. **Convolutional Neural Network (CNN):** CNN is a type of neural network architecture designed for processing grid-like data, such as images. It uses convolutional layers to extract features from input data and is widely used in tasks like image classification, object detection, and image segmentation. v. **Architecture of Perceptron Network:** The architecture of a perceptron network consists of input nodes, each connected to a single output node. The output node computes a weighted sum of the inputs and applies an activation function to produce the output.

b) **Tools Used in Neural Network Programming:**

- TensorFlow: An open-source machine learning framework developed by Google for building and training neural networks.
- PyTorch: A deep learning library built by Facebook's AI Research lab, offering dynamic computational graphs and a Pythonic interface.
- Keras: A high-level neural networks API written in Python, capable of running on top of TensorFlow, Theano, or Microsoft Cognitive Toolkit.
- scikit-learn: A machine learning library in Python that provides simple and efficient tools for data mining and data analysis, including support for various neural network models.

c) **Learning Mechanisms in ANN:** i. **Error-Based Learning:** The model adjusts its parameters based on the error between predicted and true outputs, minimizing the loss function using optimization techniques like gradient descent. ii. **Memory-Based Learning:** The model stores training instances in memory and makes predictions based on similarity measures between new instances and stored instances. It relies on past experiences to make decisions or classifications.

These explanations provide an overview of various concepts and tools used in artificial neural networks and their applications in machine learning tasks. If you have any further questions or need additional clarification, feel free to ask!