HW to chapter 3&4

Non-Programming Assignment

**1. Describe Logistic Regression**

Logistic regression is a supervised learning algorithm used for binary classification tasks. It models the probability that a given input belongs to a particular class. The key idea is to apply the logistic (sigmoid) function to the weighted sum of input features, which constrains the output to be between 0 and 1. The model is trained to find the best set of weights and bias that maximize the likelihood of correctly classifying the input data.

**2. How are Grayscale and Color (RGB) Images Presented as Inputs for the Perceptron?**

* **Grayscale images:** Each pixel in a grayscale image has a single intensity value (ranging from 0 to 255). For perceptron models, these values are typically normalized between 0 and 1 and flattened into a one-dimensional array. For example, a 20x20 grayscale image would be represented as a 400-element input vector.
* **Color (RGB) images:** Each pixel in an RGB image has three intensity values (Red, Green, Blue). To present this as input for a perceptron, the 3D image matrix (height, width, color channels) is flattened. For instance, a 20x20 RGB image would become a 1D array with 1,200 elements (20x20x3).

**3. Is Image Recognition a Logistic Regression Problem? Why?**

Yes, image recognition can be formulated as a logistic regression problem if the task is to classify images into two categories (binary classification). Logistic regression is suitable for simple binary classification tasks. However, for complex tasks like recognizing multiple image classes, logistic regression needs to be extended into multiclass classification models such as softmax regression or neural networks.

**4. Is Home Prices Prediction a Logistic Regression Problem? Why?**

No, predicting home prices is a **regression** problem, not a logistic regression problem. Logistic regression is used for classification, whereas predicting a continuous value like home prices requires linear regression or other regression models.

**5. Is Image Diagnostics a Logistic Regression Problem? Why?**

Yes, image diagnostics, where the task is to classify medical images as belonging to a diseased or healthy class, can be formulated as a logistic regression problem. Logistic regression is suitable for binary classification, which is often the case in medical diagnostics.

**6. How Does Gradient Descent Optimization Work?**

Gradient descent is an optimization algorithm used to minimize the loss function of a model by iteratively adjusting the model parameters (weights). The algorithm calculates the gradient (partial derivatives) of the loss function with respect to the model's parameters and updates the parameters in the opposite direction of the gradient. This process is repeated until the loss converges to a minimum.

**7. How Does Image Recognition Work as a Logistic Regression Classifier?**

When used for image recognition, logistic regression attempts to separate the pixel data into two classes. Each pixel (or group of pixels) contributes to the weighted sum, and the logistic function is applied to produce a probability between 0 and 1. For multiclass image recognition, logistic regression can be extended to models like softmax regression, or more complex neural networks can be used.

**8. Describe the Logistic Regression Loss Function and Explain the Reasons Behind This Choice.**

The loss function used in logistic regression is the **binary cross-entropy (log-loss)** for binary classification. It measures how well the predicted probabilities match the actual labels. The function penalizes confident but incorrect predictions heavily and rewards confident correct predictions, ensuring that the model learns to make accurate predictions with high confidence.

**9. Describe the Sigmoid Activation Function and the Reasons Behind Its Choice.**

The sigmoid activation function outputs values between 0 and 1, making it suitable for logistic regression where the goal is to model probabilities. The sigmoid function is defined as:

σ(x)=11+e−x\sigma(x) = \frac{1}{1 + e^{-x}}σ(x)=1+e−x1​

It compresses input values into the range (0, 1), which is why it is commonly used in binary classification tasks to model the probability that an instance belongs to a class.