1. What is an artificial neural network (ANN)?

ANN are computer systems that were inspired by neural networks that are in animal brains. NNs simulate how human brains interpret and process information. They are the foundation of artificial intelligence and allow computers to learn observational data.

Each ANN has an input, output and hidden layer. The hidden layer comprises weights that can transform the input into something the output can use. ANNs can be found solving various fields, such as image and speech recognition, medical diagnosis, financial forecasting, and more by adjusting the weights of the hidden layer.

2. In reference to Neural Networks (NN), what is the use of the Loss functions?

The Loss function has a crucial role when training a NN. It measures the discrepancy between the actual output and the expected output (ground truth) of the network. The major goal of using the loss function is to guide the optimization algorithm in adjusting the network’s weights to minimize the discrepancy.

            When evaluating how poorly the model is performing, the loss function can provide a clear aim for optimizing the process. Such as reducing the loss and improving the model’s predictions. That is why using the loss function is crucial while training an NN in order to accurately perform tasks such as classification, regression, and much more.

3. What is the role of the Activation functions in Neural Networks (NN)?

Activation functions allow NNs to introduce nonlinearity. This allows the function to learn complex patterns. The functions decide if a neuron should activate, which has an influence on the network’s output. This process makes it possible for the network to perform sophisticated tasks and improve learning accuracy through backpropagation.

            By using activation functions it can affect the network’s performance and capability to generalize from the input data.

4. In reference to Neural Networks (NN), what do we mean by Backpropagation?

Backpropagation is a method to use for training the network by adjusting the weights of the connections based on the error rate got in the previous run (forward propagation). Backpropagation will calculate the gradient of the loss function regarding each weight using the chain rule. It moves backwards from the output layer until it hits the input layer. This will optimize the NN’s performance by reducing the error rates and improving the accuracy in the predictions.

5. What is deep learning, and how does it differ from neural networks (NN), if any?

Deep learning is a subset of machine learning that uses NNs with multiple layers. By training and testing deep neural networks, it can learn new information and make intelligent decisions on their own. While all deep learning models are neural networks, not all neural network models are deep. The biggest difference between deep learning and traditional NNs is the size of the NN. Deep learning involves using more layers, allowing the NN to process data with more complexity and abstraction. This makes the NN really good at taking large amounts of unstructured data.

6. In reference to Machine Learning and Deep Learning, explain the impact of the Bias-Variance Trade-off

The bias-variance trade-off is the balance between the model’s simplicity and complexity. If a model has a high bias, it means that the model is too simple. A model with high variance will miss the underlying trends in the data, which will lead to underfitting. However, a model with high variance is too complex. When a model has a high variance, it will capture noise as if it were the real pattern, resulting in overfitting.

The goal is finding the sweet spot between making a model that is too complex or too simple.

7. Describe a machine learning problem where neural networks may not be appropriate?

Since NNs require large datasets, if your data set is small, a machine learning model may not be appropriate. Training on a small dataset will cause overfitting in the model. NN will memorize the training data instead of generalizing from it.

8. Natural language processing (NLP) is a field that focuses on making natural human language usable by computer programs. It developed as a subfield of linguistics, computer science, and artificial intelligence concerned with the interactions between computers and human language.

Think of a situation where NLP might be applied and describe what that situation might be, the type of data you would need, and what you would want to get out of applying NLP to it.

You might apply for an NLP when you want to create a customer service chatbot. I would want this NLP to understand and respond to customer questions, but have it responded in a way the user feels they are talking to another person? To train the NLP, you would need a dataset that has conversations between a customer agent and the customer.

            For example, a healthcare provider or insurance could create a healthcare chatbot company to help their patients with questions about symptoms, booking appointments, understanding insurance coverage or any other question they might have.

9. We discussed Artificial Neural networks (ANN) much in the class, but there are many variations of neural networks. Briefly describe what a Convolutional Neural Network (CNN) is and where they are often applied.

CNN is a type of neural network that is good for working with pictures or videos. By looking at a lot of examples, CNNs learn to recognize unfamiliar patterns, like faces in photos or objects in the video. Some applications use CNNs in your phone or computer to recognize your face, sort photos, and assist doctors in examining X-rays more effectively.

10. We discussed Artificial Neural networks (ANN) a lot in the class, but there are many variations of neural networks. Briefly describe what a Recurrent Neural Network (RNN) is and where they are often applied.

RNNs is a network that can remember things as the past to help decide in the present or in the future. These are great for tasks that you need to rely on previous information/data to make new decisions. For example, using auto fill to fill in the rest of a sentence so a user can type faster. These models are exceptionally useful for translating languages, auto filling text and much more.

*For the following Activation Functions used in Neural Networks, indicate what each function does (i.e. it transfers what type of data to what type of output)*

11. Sigmoid or Logistic Activation Function

* Transforms inputs into values between 0 and 1. Which is ideal for predicting probabilities like if a message is spam.

12. Softmax Activation Function

* Converts scores into probabilities that sum up to 1. This is really good for classifying images into categories.

13 ReLU (Rectified Linear Unit) Activation Function

* Changes negative input to 0, while keeping positives unchanged. This is used in image recognition to help it learn faster.

14. ArgMax Activation Function

* This will picture the highest value making it a 1 and the rest 0. It is useful for choosing the most likely category.

15. Linear or identity Activation Function

* This will output the input as it is received. This function is incredible for predicting a house’s price.

*For the following concepts used in Natural Language Processing, describe each concept and explain why it is needed.*

16. What is tokenizing and why is it needed?

* Tokenization splits words or sentences, making it easier for computers to process and understand language. Completing tasks like counting words or finding specific phrases with large texts requires this.

17. What is stemming and why is it needed?

* Stemming is the act of reducing words to either their base or root form, removing suffixes. The process simplifies the text analysis process by reducing it into different a word to its common base form, which makes it easier to analyze and compare.

18. What is lemmatizing and why is it needed?

* Lemmatizing simplifies words to their base form, but in a more grammatically correct way than if you applied stemming. Lemmatizing considers the words part of speech and uses a dictionary. This process is crucial for understanding the exact meaning of words in text analysis.

19. What is chunking and how is it different than tokenizing?

* Chunking is the act of grouping adjacent tokens into larger pieces based on their syntactic roles, such as noun phrases or verb phrases. It differs from tokenizing, which splits the text into individual words or sentences, chunking combines words and sentences to understand the structure of sentences better.