Module 15: Week 17: Required Assignment

### **Task 1 - Deep Learning vs. Traditional Machine Learning**

Explain the differences between deep learning and traditional machine learning models. Use the hints below:

* Deep learning excels at handling unstructured data like images and text.
* Traditional ML techniques often require more manual data preprocessing and feature extraction.
* Mention the higher computational cost of deep learning models.

**Key Differences:**  
While both aim to teach machines to learn from data, deep learning (DL) and traditional machine learning (ML) take very different approaches:

| **Aspect** | **Traditional ML** | **Deep Learning** |
| --- | --- | --- |
| **Data Type** | Structured data (e.g. Excel spreadsheets). | Unstructured data (images, audio, raw text). |
| **Feature Extraction** | Needs humans to manually select features (domain expertise required). | Figures out features automatically (due to its deep neural networks). |
| **Model Complexity** | Uses simpler models (e.g., Decision Trees, SVM). | Builds brain-like networks with many layers (millions of parameters). |
| **Data Hunger** | Works fine with smaller datasets. | Demands massive data to avoid overfitting. |
| **Compute Needs** | Less CPU needs | Craves GPUs/TPUs |

**When to Use Which?**

* **Pick Traditional ML** if data is structured, labeled, and don’t have massive compute resources.
* **Go Deep Learning** for complex, unstructured data (like recognizing dog photos) if have enough data and computing power.

### **Task 2 – LSTM Networks in NLP (2 Marks)**

Explain how Long Short-Term Memory (LSTM) networks are used in NLP to capture long-range dependencies in sequence data. Describe how LSTMs overcome the limitations of traditional RNNs.

**Long Short-Term Memory (LSTM)** networks are a type of **Recurrent Neural Network (RNN)** designed to learn long-term dependencies in sequential data.

### **Why LSTMs in NLP?**

* Natural Language Processing (NLP) tasks — such as text classification, translation, and speech recognition — require understanding **context** from earlier in a sequence.
* Traditional RNNs often forget information when the gap between relevant inputs and outputs is large, due to the **vanishing gradient problem**.
* LSTMs solve this by introducing **gates** that regulate the flow of information.

### **How LSTMs Work**

* **Forget Gate:** Decides what information from the previous cell state to discard.
* **Input Gate:** Decides what new information to store in the cell state.
* **Output Gate:** Decides what part of the cell state to output as the hidden state.
* This gating mechanism allows LSTMs to **remember important context** for long periods and ignore irrelevant details.

**Example in NLP**  
  
"I didn’t like the movie, but the ending was amazing."

* A **basic RNN** might give too much weight to the early negative phrase (“didn’t like”) and overlook the later positive sentiment.
* An **LSTM**, by maintaining longer-term dependencies, can capture both sentiments and determine that the overall tone is **mixed** or even **positive**, depending on how the model weights each part.

**Summary:**  
LSTMs are crucial in NLP because they can retain relevant information across long sequences, overcoming the short-term memory limitations of standard RNNs.