IMPORTANT QUESTION FOR DLNLP EXAM (THEORY)

Unit 1: Introduction to Natural Language Processing

- **Definition:**
  Natural Language Processing (NLP) is a branch of **Artificial Intelligence (AI)** that enables computers to **understand, interpret, and generate human language** in a meaningful way.
  It combines concepts from **linguistics, computer science, and machine learning** to process text or speech data.
- **Example:**
  Voice assistants like **Alexa**, **Google Assistant**, and **ChatGPT** use NLP to understand and respond to human queries.

Importance of NLP:

1. **Human–Computer Interaction:** Enables machines to communicate with humans in natural language.
2. **Automation of Text Tasks:** Used in chatbots, translation, and email classification.
3. **Information Extraction:** Helps extract insights from large text data like reviews or documents.
4. **Sentiment Analysis:** Understands customer emotions and opinions.
5. **Accessibility:** Assists speech recognition and text-to-speech for differently-abled users.

Nlp Process:

The **NLP process** involves several key stages that help machines understand and interpret human language.

1. Text Pre-processing:

- Cleans and prepares raw text for analysis.
- Steps include:
  - **Tokenization:** Splitting sentences into words.
  - **Stop-word Removal:** Removing common words (e.g., “is”, “the”).
  - **Stemming/Lemmatization:** Reducing words to their base form (e.g., *running* → *run*).

2. Lexical (Word Level) Analysis:

- Analyzes structure and meaning of words.
- Identifies **parts of speech** (noun, verb, adjective, etc.) and word patterns.

3. Syntactic Analysis (Parsing):

- Checks **grammar and sentence structure** using parsing trees.
- Example: “He go to school” → syntactically incorrect.

4. Semantic Analysis:

- Determines **meaning of words and sentences**.
- Resolves ambiguity and ensures sentences make logical sense.
  Example: “The car eats grass” → semantically wrong.

5. Pragmatic & Discourse Analysis:

- **Pragmatic Analysis:** Understands **context and intention** behind a sentence.
- **Discourse Analysis:** Understands meaning in a larger text (paragraph or dialogue).

1. **Explain the three waves of NLP: Rationalism, Empiricism, and Deep Learning.**

Ans.

- **First Wave – Rationalism (1950s–1980s):**
  Based on **rule-based systems** and **linguistic knowledge**. Models used grammar rules and logic to understand language.
  Example: ELIZA chatbot, syntax-based parsers.
- **Second Wave – Empiricism (1990s–2010):**
  Relied on **statistical and probabilistic methods** using large datasets. Machine learning algorithms learned patterns from data.
  Example: Hidden Markov Models, Naïve Bayes, and n-grams.
- **Third Wave – Deep Learning (2010–Present):**
  Uses **neural networks** and **representation learning** (like Word2Vec, Transformers) to automatically learn features from raw text.
  Example: ChatGPT, BERT, GPT models.

2. **What makes Natural Language Processing difficult?**

Ans.

1. **Ambiguity:** Words and sentences can have multiple meanings (e.g., “bank” – riverbank or financial bank).
2. **Context Dependence:** Meaning changes based on situation or tone.
3. **Variety in Expression:** Many ways to express the same idea.
4. **Language Diversity:** Grammar and structure differ across languages.
5. **Incomplete Information:** Human communication often relies on common sense or implied meaning, which machines lack.

3. Define and differentiate between syntactic, semantic, and pragmatic analysis.

Ans.

1. **Syntactic Analysis:**
   Deals with **grammar and sentence structure**. It checks if sentences follow correct grammatical rules.
   Example: “He go to school” → Incorrect syntax.
2. **Semantic Analysis:**
   Focuses on **meaning of words and phrases** in context.
   Example: “The apple eats the boy” is syntactically correct but semantically wrong.
3. **Pragmatic Analysis:**
   Considers **real-world context and speaker’s intention**.
   Example: “Can you pass the salt?” – It’s a request, not a question.

4. What is Part-of-Speech (PoS) tagging? Explain its importance in NLP.

Ans.

• **Definition**:
PoS tagging is the process of **assigning a part of speech** (noun, verb, adjective, etc.) to each word in a sentence.
Example: “The dog runs fast.” → dog (noun), runs (verb), fast (adverb).

• **Importance**:

1. Helps in syntactic and semantic analysis.
2. Useful for machine translation, sentiment analysis, and question answering.
3. Provides structure for parsing and word sense disambiguation.

5. Explain the various applications of NLP in real-world systems.

Ans.

1. **Machine Translation:** Google Translate, DeepL.
2. **Chatbots and Virtual Assistants:** ChatGPT, Alexa, Siri.
3. **Sentiment Analysis:** Analyzing customer reviews and feedback.
4. **Text Summarization:** Automatic summarization of long documents.
5. **Speech Recognition:** Converting spoken words to text (e.g., voice typing).
6. **Information Retrieval:** Search engines like Google.

Unit 2: Introduction to Deep Learning

- **Definition:**
  Deep Learning is a subset of **Machine Learning (ML)** that uses **artificial neural networks** with multiple layers to automatically learn patterns and features from large amounts of data.
  It mimics the **human brain's learning process** to understand complex relationships in data such as text, images, or speech.
- **Example:**
  Applications like **ChatGPT, self-driving cars, face recognition**, and **speech assistants** use deep learning models.

Importance of Deep Learning:

1. **Automatic Feature Extraction:** Removes the need for manual feature engineering.
2. **High Accuracy:** Achieves state-of-the-art results in NLP, vision, and speech.
3. **Handles Complex Data:** Works efficiently with large, unstructured data such as images, videos, and natural language.
4. **Continuous Learning:** Improves performance as more data becomes available.
5. **Foundation for Modern AI:** Powers intelligent systems like **Chatbots, Recommendation Systems, and Generative AI models**.

6. **Explain the structure of an Artificial Neural Network (ANN).**

Ans.

- **Definition:**
  An **Artificial Neural Network (ANN)** is a computational model inspired by the structure of the human brain. It consists of layers of interconnected nodes (neurons) that process data.
- **Structure:**
  1. **Input Layer:** Receives raw data or features.
  2. **Hidden Layers:** Perform computations using weighted connections and activation functions.
  3. **Output Layer:** Produces the final prediction or classification.
- **Working:**
  Each neuron receives inputs, multiplies them by **weights**, adds a **bias**, passes the sum through an **activation function**, and sends the output forward.

  Example:
  Used in image recognition, sentiment analysis, etc.

7. **What is an activation function? Explain any three commonly used ones.**

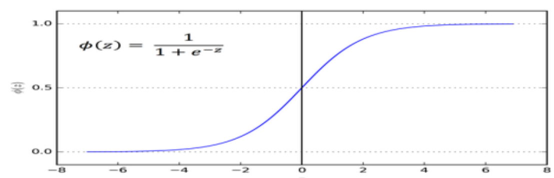
Ans.

• **Definition:**
An **activation function** determines whether a neuron should be activated or not by introducing **non-linearity** into the network, enabling it to learn complex patterns.

• **Common Activation Functions:**

1. **Sigmoid Function:**

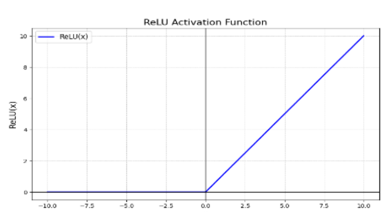
$$f(x) = \frac{1}{1+e^{-x}}$$



   - Output range: (0,1)
   - Used in binary classification models.

2. **ReLU (Rectified Linear Unit):**

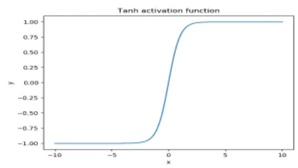
$$f(x) = \max(0, x)$$



   - Fast and simple, avoids vanishing gradient problem.
   - Commonly used in CNNs and deep models.

3. **Tanh (Hyperbolic Tangent):**

$$f(x) = \tanh(x)$$



- Output range: (-1, 1)
- Stronger gradients than sigmoid and centered around zero.

8. **Differentiate between Feedforward Neural Network and Recurrent Neural Network.**

Ans.

Feature	Feedforward Neural Network (FNN)	Recurrent Neural Network (RNN)
Data Flow	Information flows only forward (input → output).	Information flows in loops; previous output feeds next input.
Memory	No memory of past inputs.	Has memory through hidden states.
Use Case	Image classification, static data.	Sequence data like text, speech, time series.
Complexity	Simple architecture.	More complex and harder to train.
Example:	FNN → Image recognition	RNN → Text prediction, speech recognition

9. **Explain the Backpropagation algorithm with a simple example.**

Ans.

- **Definition:**
  Backpropagation is a **supervised learning algorithm** used to **train neural networks** by minimizing the error between predicted and actual outputs.
- **Steps:**
  1. **Forward Pass:** Input is passed through the network to get the output.
  2. **Error Calculation:** Compute loss (difference between predicted and actual output).
  3. **Backward Pass:** Calculate gradients of error w.r.t. weights using **chain rule**.
  4. **Weight Update:** Adjust weights using gradient descent.

Example:
If output is 0.8 and actual value is 1.0, backpropagation reduces the error by updating weights to bring output closer to 1.0.

10. **What is Stochastic Gradient Descent (SGD)? How is it used in training neural networks?**

• **Definition:**
Stochastic Gradient Descent (SGD) is an optimization algorithm used to **update weights** in neural networks by minimizing the **loss function**.

• **Formula:**

$$W_{new} = W_{old} - \eta \times \nabla L(W)$$

Where:
η = learning rate,
∇L(W) = gradient of loss function.

• **Working:**

- Uses **a single training sample** (or mini-batch) at a time instead of the whole dataset.
- Speeds up training and helps escape local minima.

• **Use:**
Essential for training deep networks efficiently and improving model convergence.

Unit 3: Word Vector Representations

11. What is word embedding? How does it improve NLP models?

Ans.

• **Definition:**
Word embedding is a technique used to represent words as **dense numerical vectors** in a continuous vector space, where similar words have similar representations.

• **Improvement in NLP Models:**

1. **Captures Semantic Meaning:** Words with similar meanings are placed closer together (e.g., *king – queen*, *Paris – France*).
2. **Reduces Dimensionality:** More efficient than one-hot encoding.
3. **Improves Model Accuracy:** Enables better understanding of context in NLP tasks like translation, sentiment analysis, and text classification.
4. **Learns Context Relationships:** Embeddings retain both syntactic and semantic word relationships.

12. Explain the architecture of the Skip-Gram model used in Word2Vec.

Ans.

- **Concept:**
  The **Skip-Gram model** predicts **context words** based on a **target (center) word**.
  It learns word embeddings by maximizing the probability of surrounding words given the center word.
- **Architecture:**
  1. **Input Layer:** One-hot encoded vector of the center word.
  2. **Hidden Layer:** Weight matrix that acts as the word embedding representation.
  3. **Output Layer:** Produces probabilities for context words using Softmax.
- **Example:**
  Sentence: *"The cat sits on the mat"*
  Center word = "cat"; Context words = "The", "sits"
  Model learns embedding of "cat" that predicts its surrounding words.

13. Compare Skip-Gram and CBOW models.

Ans.

Feature	Skip-Gram Model	CBOW Model
Objective	Predicts context words from a center word.	Predicts a center word from surrounding context words.
Focus	Works well for rare words.	Works better for frequent words.
Training Speed	Slower due to more predictions per word.	Faster since it averages context vectors.
Example	Input = “cat”, Output = “The”, “sits”	Input = “The”, “sits”, Output = “cat”

14. What is Negative Sampling? Why is it used in Word2Vec?

Ans.

• Definition:
Negative Sampling is a simplified training technique used in Word2Vec to make the **Softmax computation more efficient**.

• Purpose:
Instead of updating all word weights in the vocabulary, the model updates only:

- The **target (positive)** word, and
- A few randomly chosen **negative** (non-context) words.

• Benefits:

1. **Reduces computation time.**
2. **Improves training speed** for large vocabularies.
3. **Enhances word representation** by learning to distinguish correct and incorrect contexts.

15. Explain the concept of GloVe (Global Vectors for Word Representation).

Ans.

• Definition:

GloVe is a **count-based word embedding model** that combines the strengths of global matrix factorization (like LSA) and local context window methods (like Word2Vec).

• **Concept:**
It constructs a **word co-occurrence matrix** and learns word vectors by analyzing how frequently words appear together in a corpus.

• **Formula:**
The model tries to make the ratio of co-occurrence probabilities between words meaningful.

• **Advantages:**

1. Captures both **global statistical information** and **local context**.
2. Pre-trained GloVe vectors (like 50d, 100d, 300d) are widely used in NLP tasks.

16. What are cross-lingual word embedding models? Give examples.

Ans.

• **Definition:**
Cross-Lingual Word Embeddings map words from **different languages** into a **shared vector space**, allowing semantic comparison across languages.

• **Purpose:**
Helps transfer learning from one language to another, improving **machine translation** and **multilingual applications**.

• **Example Models:**

1. **MUSE (Multilingual Unsupervised and Supervised Embeddings)**
2. **fastText Multilingual Embeddings**
3. **LASER by Facebook AI**

• **Example Use Case:**
Words like *"dog" (English)* and *"chien" (French)* have similar vector representations.

Unit 4: Recurrent Neural Network (RNN) for NLP

17. What is the main idea behind recurrence in RNNs?

Ans.

• **Concept:**
The main idea behind **recurrence** in Recurrent Neural Networks (RNNs) is to allow the network to **remember past information** while processing current inputs.

• **How it works:**
Unlike feedforward networks, an RNN takes **input from the current time step** and the **output (hidden state)** from the previous time step.
This creates a **loop (recurrence)** that helps in understanding sequential data such as text, audio, or time series.

- **Mathematical form:**

$$h_t = f(Wx_t + Uh_{t-1} + b)$$

Example:
In a sentence like *"I love India"*, the word *"India"* is understood in context of *"I love"* — recurrence helps retain this context.

18. Explain the meta meaning of the hidden state in an RNN.

Ans.

• **Definition:**
The **hidden state** in an RNN acts as the **memory** of the network. It stores information from **previous time steps** and passes it to future ones.

• **Meta Meaning:**

1. It represents the **context** of the sequence learned so far.
2. It combines both **past knowledge** and **current input** to make predictions.
3. It continuously updates as new inputs arrive.

• **Example:**
In a sentence prediction like "The weather is ___", the hidden state carries context about "The weather is" to predict "sunny" or "cold".

19. What are the limitations of vanilla RNNs and how does LSTM solve them?

Ans.

• **Limitations of Vanilla RNNs:**

1. **Vanishing Gradient Problem:** Gradients become too small during training, making it hard to learn long-term dependencies.
2. **Exploding Gradient Problem:** Gradients become too large, causing instability.
3. **Short Memory:** Can only remember recent inputs.
4. **Difficult Training:** Poor convergence for long sequences.

• **How LSTM Solves These Problems:**

1. Introduces **memory cells** that store information for long durations.
2. Uses **gates (input, forget, output)** to control information flow.
3. Maintains **constant error flow**, preventing vanishing gradients.

20. Describe the structure and functioning of LSTM (Long Short-Term Memory).

Ans.

• **Structure:**

An **LSTM unit** consists of:

1. **Cell State (Ct):** Memory of the network.
2. **Input Gate:** Controls how much new information enters.
3. **Forget Gate:** Decides what information to discard.
4. **Output Gate:** Determines what to output.

• **Functioning:**

1. **Forget Gate:**

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

→ Decides what to forget.

2. **Input Gate:**

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

→ Selects new info to add.

3. **Cell State Update:**
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

4. **Output Gate:**

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

→ Determines final output.

LSTM networks use gates to **store, update, and forget** information, making them ideal for sequential NLP tasks.

21. Explain Sequence-to-Sequence (Seq2Seq) models and their applications.

Ans.

- **Definition:**
  Seq2Seq models are neural network architectures that convert one sequence into another. They consist of:
  1. **Encoder:** Processes input sequence and converts it to a context vector.
  2. **Decoder:** Generates output sequence from that context.
- **Example:**
  - Input: "How are you?"
  - Output: "¿Cómo estás?" (machine translation).
- **Applications:**
  0. Machine Translation
  1. Text Summarization
  2. Chatbots and Conversational AI
  3. Speech Recognition

Seq2Seq models map input sequences to output sequences, widely used in translation and dialogue systems.

22. What are the key differences between standard and advanced Seq2Seq models?

Ans.

Feature	Standard Seq2Seq Model	Advanced Seq2Seq Model
Architecture	Simple Encoder–Decoder without attention.	Uses **Attention Mechanism** or **Transformer** architecture.

Feature	Standard Seq2Seq Model	Advanced Seq2Seq Model
Context Handling	Uses a single fixed-length context vector (information loss).	Dynamically focuses on important parts of the input using attention.
Performance	Struggles with long sequences.	Handles long sequences efficiently.
Examples	Basic RNN-based Seq2Seq.	Attention-based Seq2Seq, Transformer, BERT, GPT.

Advanced Seq2Seq models overcome memory limitations by using **attention and self-attention** mechanisms for better context understanding.

Unit 5: Chatbot Development

23. What is a chatbot? Explain its origin and working mechanism.

Ans.

• **Definition:**
A **chatbot** is a computer program designed to **simulate human conversation** through text or voice interactions. It allows users to communicate with machines in natural language.

• **Origin:**

- The first chatbot, **ELIZA (1966)**, was created by **Joseph Weizenbaum** to mimic a psychotherapist.
- Later, **PARRY** (1972) and **ALICE (1995)** improved on the idea.
- Modern chatbots like **Siri, Alexa, and ChatGPT** use advanced **AI and NLP techniques**.

• **Working Mechanism:**

1. **User Input:** The user types or speaks a query.
2. **Processing:** The chatbot uses **Natural Language Processing (NLP)** to understand intent.
3. **Response Generation:** Based on the query, it retrieves or generates an appropriate response.
4. **Output:** The response is displayed or spoken to the user.

A chatbot interprets user queries using NLP and responds intelligently like a human.

24. Discuss the types of chatbots: rule-based vs AI-based.

Ans.

Feature	Rule-Based Chatbots	AI-Based Chatbots
Working Principle	Uses pre-defined rules or decision trees.	Uses Machine Learning and NLP for learning patterns.
Response Type	Fixed responses; limited flexibility.	Dynamic, context-aware responses.
Learning Ability	Cannot learn from past interactions.	Continuously improves through data and feedback.
Example	Simple FAQ bots, banking query bots.	ChatGPT, Google Assistant, Siri.

Rule-based chatbots follow scripts, while AI-based chatbots **understand, learn, and generate** human-like responses.

25. Explain how text generation works in a chatbot.

Ans.

- **Definition:**
  Text generation in a chatbot refers to creating **natural, contextually appropriate responses** to user input.
- **Working Steps:**
  1. **Input Understanding:**
     The chatbot identifies user intent and key entities using NLP.
  2. **Response Generation:**
     - **Rule-based:** Selects pre-written responses from a database.
     - **AI-based:** Uses deep learning models (like RNN, LSTM, or Transformer) to generate text word-by-word.
  3. **Post-Processing:**
     Adds grammar correction, emotion, and coherence to make responses human-like.
- **Example:**
  User: “Tell me a joke.”
  Chatbot: “Why did the computer show up late? It had a hard drive!” 😄

Chatbots use **language models** to generate meaningful, context-driven text automatically.

26. How has ChatGPT revolutionized conversational AI?

Ans.

- **Definition:**
  ChatGPT (by OpenAI) is an advanced **AI chatbot** based on **Transformer architecture (GPT series)** that can generate human-like text.

- **Revolutionary Features:**
    1. **Context Understanding:** Remembers conversation flow across multiple turns.
    2. **Human-like Responses:** Generates coherent, creative, and emotional text.
    3. **Versatility:** Can write essays, answer questions, code, or chat casually.
    4. **Learning from Data:** Trained on massive datasets using **deep learning** and **reinforcement learning from human feedback (RLHF)**.
    5. **Wide Adoption:** Used in education, customer support, business automation, and personal assistance.

ChatGPT has transformed conversational AI by making interactions **natural, intelligent, and human-like**.

27. Why are chatbots considered a big opportunity in the AI industry?

Ans.

1. 24/7 Availability:
Chatbots provide **round-the-clock customer support**, reducing human workload.

2. Cost Efficiency:
Saves money by automating repetitive customer interactions.

3. Personalization:
Offers customized recommendations and answers using user data.

4. Business Growth:
Improves customer satisfaction, retention, and engagement.

5. Multi-domain Use:
Used in **healthcare, e-commerce, banking, education, and travel**.

6. Integration with AI:
Combined with NLP and deep learning, chatbots can **understand emotions, predict intent**, and give smarter responses.

Chatbots are a massive AI opportunity because they **automate communication**, **save time**, and **enhance user experience** in almost every industry.

Unit 6: Deep Learning in Sentiment Analysis

28. What is Sentiment Analysis? Describe its importance and challenges.

Ans.

- **Definition:**
  Sentiment Analysis, also known as **Opinion Mining**, is a technique in **Natural Language Processing (NLP)** that identifies and classifies opinions expressed in text as **positive, negative, or neutral**. It helps computers understand human emotions and attitudes from written content.
- **Importance:**
  1. **Business Insights:** Helps companies understand customer opinions about products or services.
  2. **Social Media Monitoring:** Tracks public sentiment toward brands or events.
  3. **Customer Experience:** Enhances user satisfaction through emotion-based responses.
  4. **Decision Making:** Supports marketing, politics, and product improvement.
- **Challenges:**
  1. **Sarcasm & Irony:** Hard for models to detect hidden or opposite meanings.
  2. **Context Dependence:** Words can have different meanings depending on context.
  3. **Multilingual and Slang Use:** Informal text is difficult to analyze.
  4. **Domain Variability:** A model trained on movie reviews may fail on tweets or product reviews.

Sentiment Analysis helps interpret emotions from text but struggles with **context, sarcasm, and informal language**.

29. Compare CNN, RNN, and Recursive Neural Networks for sentiment classification.

Ans.

Model Type	Concept	Advantages	Limitations	Use Case
CNN (Convolutional Neural Network)	Uses convolutional filters to extract features from text like local n-grams or phrases.	Fast, efficient, captures local patterns.	Cannot handle long-term dependencies.	Short text or phrase-level sentiment (e.g., tweets).
RNN (Recurrent Neural Network)	Processes input sequentially and keeps memory of past inputs.	Captures temporal or sequential dependencies in text.	Suffers from vanishing gradient, limited long memory.	Sentiment analysis of longer sentences or reviews.
Recursive Neural Network	Builds a tree structure using syntactic parsing to analyze hierarchical sentence meaning.	Understands sentence structure and compositional meaning.	Complex and computationally expensive.	Grammar-based or deep semantic tasks.

• **CNN** → Detects **local features**

• **RNN** → Learns **sequential context**

• **Recursive NN** → Understands **sentence hierarchy**

30. Explain the concept of Aspect-Based Sentiment Analysis and Sarcasm Detection.

Ans.

(a) Aspect-Based Sentiment Analysis (ABSA):

- **Definition:**
  ABSA identifies **specific aspects (features)** of a product or topic and finds the **sentiment** expressed toward each.
- **Example:**
  "The screen is bright but the battery drains fast."
  → *Screen: Positive, Battery: Negative*
- **Importance:**
  Provides detailed insights by identifying **which part** of a product or service users like or dislike.
- **Deep Learning Role:**
  Uses models like **LSTM** or **BERT** to capture context and aspect-specific emotions.

(b) Sarcasm Detection:

- **Definition:**
  Detects sentences where the **literal meaning differs from the intended sentiment** (e.g., "I just love getting stuck in traffic").
- **Challenge:**
  Sarcasm often depends on tone, prior knowledge, or context, making it hard for machines to detect.
- **Techniques Used:**
  Deep learning models like **Bidirectional LSTM** and **Transformers** (e.g., BERT, GPT) learn to recognize emotional cues and contradictions between words and sentiment.

- **ABSA** → Finds sentiment for each **aspect** of text.
- **Sarcasm Detection** → Finds **hidden or opposite emotions** not directly expressed.