Introduction to Large Language Models
Assignment- 1

Number of questions: 10   **Total marks: 10 X 1 = 10**

QUESTION 1:
Based on Distributional Semantics, which of the following statements is/are true?
(i) The meaning of a word is defined by its relationship to other words.
(ii) The meaning of a word does not rely on its surrounding context.
a. Both (i) and (ii) are correct
b. Only (ii) is correct
c. Only (i) is correct
d. Neither (i) nor (ii) is correct

Correct Answer: d
Solution: According to Distributional Semantics, the meanings of words can be derived from their distributional properties in a large corpora of text. It relies on the context in which words appear.

QUESTION 2:

Which of the following words have multiple senses?
a. light
b. order
c. letter
d. buffalo

Correct Answer: a, b, c, d
Solution: Each of these words has multiple meanings.
- Light can mean illumination, something not heavy, or a pale color.
- Order can refer to a sequence, a command, or a request for goods.
- Letter can mean an alphabet character or a written message.
- Buffalo can also have different senses. As an attributive noun (acting as an adjective), it refers to a specific place named Buffalo, such as the city of Buffalo, New York; as the verb *to buffalo*, means "to bully, harass, or intimidate" or "to baffle"; and as a noun it refers to the animal (either the true buffalo or the bison).

QUESTION 3:

Consider the following sentences:
Sentence 1: Amit forgot to set an alarm last night.
Sentence 2: Amit woke up late today.

Does Sentence 1 entail Sentence 2?
a. True
b. False

Correct Answer: b
Solution: Amit forgetting the alarm does not necessarily mean he woke up late.

QUESTION 4:

What issues can be observed in the following text?

On a much-needed #workcation in beautiful Goa. Workin & chillin by d waves!

a. Idioms
b. Non-standard English
c. Tricky Entity Names
d. Neologisms

Correct Answer: b, d
Solution: We can observe usage of non-standard English such as "Workin" and "chillin", neologisms such as "workcation" in the text.

QUESTION 5:

Consider the following sentences:
Sentence 1: The bats flew out of the cave at sunset.
Sentence 2: Rohan bought a new bat to practice cricket.

Question: Does the word "bat" have the same meaning in both sentences?
a. Yes
b. No

Correct Answer: b
Solution: In Sentence 1, "bat" refers to the flying mammal, while in Sentence 2, "bat" refers to a piece of sports equipment used in cricket.

QUESTION 6:

Which of the following statements is/are true?
a. Apple is a hypernym of fruit.
b. Leaf is a meronym of tree.
c. Flower is a holonym of petal.
d. Parrot is a hyponym of bird.

Correct Answer: b, c, d

Solution: Hyponymy (subset relation), Hypernymy (superset relation), Meronymy (part-of relation), Holonymy (has-a relation)

QUESTION 7:

_____ deals with word formation and internal structure of words.
a. Pragmatics
b. Discourse
c. Semantics
d. Morphology

Correct Answer: d
Solution: Morphology is the study of words, how they are formed, and their relationship to other words in the same language. It analyzes the structure of words and parts of words, such as stems, root words, prefixes, and suffixes.

QUESTION 8:

Consider the following sentences:
Sentence 1: Priya told Meera that she had completed the report on time.
Sentence 2: Meera was impressed by her dedication.

Which of the following statements is/are true?
a. In Sentence 1, "she" refers to Meera.
b. In Sentence 1, "she" refers to Priya.
c. In Sentence 2, "her" refers to Priya.
d. In Sentence 2, "her" refers to Meera.

Correct Answer: b, c
Solution: This is an example of co-reference resolution in natural language processing. In Sentence 1, "she" refers to Priya as per the structure of the sentence where Priya is the subject, and the report was completed by her. "her" in the second sentence also refers to Priya.

QUESTION 9:

In semantic role labelling, we determine the semantic role of each argument with respect to the _____ of the sentence.
a. noun phrase
b. subject
c. predicate
d. adjunct

Correct Answer: c

Solution: In semantic role labelling, we determine the semantic role of each argument with respect to the predicate of the sentence.

QUESTION 10:

Which of the following statements is/are true?
(i) Artificial Intelligence (AI) is a sub-field of Machine Learning.
(ii) LLMs are deep neural networks for processing text.
(iii) Generative AI (GenAI) involves only Large Language Models (LLMs).

a. Only (i) and (ii) are correct
b. Only (ii) is correct
c. Only (ii) and (iii) are correct
d. All of (i), (ii), and (iii) are correct
e. Neither (i), (ii), or (iii) is correct

Correct Answer: b
Solution: Let's analyze each statement.
(i) Artificial Intelligence (AI) is a sub-field of Machine Learning.
Incorrect: AI is a broad field encompassing various techniques for making machines perform intelligent tasks. Machine Learning (ML) is a sub-field of AI that focuses on learning patterns from data. Since AI includes rule-based systems, expert systems, and other non-ML approaches, it is incorrect to say that AI is a sub-field of ML. Instead, ML is a sub-field of AI.

(ii) LLMs are deep neural networks for processing text.
Correct: Large Language Models (LLMs) are a class of deep neural networks specifically designed for text processing tasks such as language generation, translation, summarization, etc. "Processing" in this context refers to how LLMs transform input text into meaningful representations before generating an output. Note that not all classes of transformer-based language models are trained to generate text. For example, encoder-only models are particularly useful for creating contextual encoding of sentences/sequences (e.g., SentenceBERT is a popularly used model for getting sentence encoding). Autoregressive LMs are scaled to huge sizes nowadays to "generate text", but several other text processing tasks are also performed by LLMs, and transformer-based LMs in general.

(iii) Generative AI (GenAI) involves only Large Language Models (LLMs).
Incorrect: While LLMs are a major part of Generative AI, they are not the only models used in this domain. Generative AI includes models for various data modalities, such as image generation (e.g., DALL-E, Stable Diffusion), audio generation (e.g., WaveNet), and video generation. GenAI encompasses a broader range of generative models beyond just LLMs.