BIAS - I : Assignment 4

1. Which of the following define bias? (Select all that apply)
    a. Systematic favoritism towards certain groups
    b. Random errors in data collection
    c. Neutral and unbiased judgment
    d. Systematic exclusion of certain data points
    e. Random selection of samples
    f. Systematic deviation from rationality in judgment

2. Which of the following is one of the ways to measure bias
    a. Random sampling
    b. Cross-Validation
    c. Using Benchmark datasets
    d. Chi-Square test

3. Which of the following statements is true?
    a. Demographics do not influence the perception of bias.
    b. Demographics can influence the perception of bias.
    c. Perception of bias is unaffected by demographics.
    d. Demographics and bias perception are unrelated.

4. What does each pair in the CrowS-Pairs dataset consist of?
    a. One stereotype sentence and one neutral sentence
    b. One biased sentence and one unbiased sentence
    c. One stereotype sentence and one opposite stereotype sentence
    d. One stereotype sentence and one less stereotype sentence

5. The statement “Women are bad drivers” is a
    a. Stereotype
    b. Anti-Stereotype
    c. Non-Stereotype
    d. Neutral Statement

6. Which of the following is NOT a common source of bias in data?
    a. Historical inequities reflected in the data
    b. Data collection methods that over-represent certain groups
    c. Balanced representation of all demographic groups
    d. Labeling errors or subjective annotations

7. At which stage of the machine learning process can bias be introduced by over- or under-representing certain groups?
    a. Data Collection
    b. Data Pre-processing
    c. Data Annotation
    d. All of the above

8. Which of the following characteristics can make a statement biased?
    a. Being stereotypical
    b. Being socially aligned
    c. Having inaccurate information
    d. Having emotional language

9. One of the reasons bias exists in models is:
    a. Algorithm complexity
    b. Data
    c. Model architecture
    d. Training duration

10. When is bias a problem?
    a. When it reduces the models performance
    b. When it has negative impact
    c. When it aligns with train data but not test data
    d. When it increases the complexity

11. Which of the following is the correct setting for contrastive learning?
    a. Irrespective of the sentences, minimise the distance between their embeddings
    b. Irrespective of the sentences, maximise the distance between their embeddings
    c. If sentences are similar, minimise the distance between their embeddings
    d. If sentences are different, minimise the distance between their embeddings
    e. If sentences are different, maximise the distance between their embeddings
    f. If sentences are different, minimise the distance between their embeddings