NLP & CV MCQ Bank: 25 Questions per Section

🔤 Section 1: Language Basics, Ambiguity, and Linguistics

1. **Language is considered a _____ system shared by a group of people to express thoughts.**

- **A. Computational**
- **B. Rule-governed**
- **C. Semantic**
- **D. Arbitrary**
- **Answer: B**

2. **Lexical ambiguity arises when:**

- **A. Sentence structure is unclear**
- **B. The same word has multiple meanings**
- **C. Multiple clauses confuse meaning**
- **D. Emotions affect sentence delivery**
- **Answer: B**

3. **In the sentence "I saw the boy with the telescope", ambiguity is:**

- **A. Lexical**
- **B. Structural**
- **C. Morphological**
- **D. Phonological**
- **Answer: B**

4. **Pragmatic ambiguity depends on:**

- **A. Grammar**

- **B. Syntax**
- **C. Context and speaker intention**
- **D. Word length**
- **Answer: C**

5. **Which type of relationship involves substitution of words in the same category?**

- **A. Syntagmatic**
- **B. Semantic**
- **C. Paradigmatic**
- **D. Schematic**
- **Answer: C**

6. **Which branch of linguistics studies sentence structure?**

- **A. Morphology**
- **B. Syntax**
- **C. Semantics**
- **D. Phonetics**
- **Answer: B**

7. **Structural ambiguity arises due to:**

- **A. Cultural references**
- **B. Multiple possible sentence parses**
- **C. Misspelled words**
- **D. Accent variations**
- **Answer: B**

8. **What does the term 'natural language' refer to?**

- **A. Computer programming languages**
- **B. Formally structured codes**
- **C. Human languages used for communication**
- **D. Translated languages**
- **Answer: C**

9. The form-function-meaning framework includes:

- **A. Syntax, morphology, semantics**
- **B. Physical, emotional, social levels**
- **C. Structure, purpose, conveyed content**
- **D. Sound, script, content**
- **Answer: C**

10. 'Wanted' in the sentence "Outlaws are wanted" demonstrates:

- **A. Structural ambiguity**
- **B. Semantic neutrality**
- **C. Lexical ambiguity**
- **D. Morphological analysis**
- **Answer: C**

11. Which is not a type of ambiguity?

- **A. Lexical**
- **B. Structural**
- **C. Morphological**
- **D. Predictive**
- **Answer: D**

12. "Break a leg!" is an example of:

- **A. Idiomatic ambiguity**
- **B. Pragmatic meaning**
- **C. Morphology**
- **D. Direct speech**
- **Answer: B**

13. Morphological ambiguity involves:

- **A. Unclear word order**
- **B. Multiple meanings for a word root**
- **C. Phoneme misinterpretation**
- **D. Accent shift**
- **Answer: B**

14. The word 'run' in 'She will run a campaign' vs 'She runs fast' shows:

- **A. Structural ambiguity**
- **B. Lexical ambiguity**
- **C. POS ambiguity**
- **D. Context-free parsing**
- **Answer: B**

15. Which field is focused on speech sounds?

- **A. Semantics**
- **B. Pragmatics**
- **C. Phonetics**
- **D. Morphology**
- **Answer: C**

16. The syntagmatic relationship involves:

- **A. Substitution**
- **B. Context-independent meaning**
- **C. Co-occurrence of words**
- **D. Root analysis**
- **Answer: C**

17. Lexicology studies:

- **A. Sentence structures**
- **B. Sound patterns**
- **C. Word meanings and relations**
- **D. Sentence chunking**
- **Answer: C**

18. 'Unbelievable' consists of how many morphemes?

- **A. 1**
- **B. 2**
- **C. 3**
- **D. 4**
- **Answer: C**

19. 'Teacher' in "She is a teacher" has:

- **A. Implied meaning**
- **B. Cultural meaning**
- **C. Literal meaning**
- **D. Context-dependent**
- **Answer: C**

20. 'Wow, you are early!" — implied meaning is an example of:

- **A. Sarcasm**
- **B. Phonology**
- **C. Pragmatics**
- **D. Idioms**
- **Answer: A**

21. Which level of analysis focuses on the meaning in larger discourse?

- **A. Morphology**
- **B. Syntax**
- **C. Discourse**
- **D. Semantics**
- **Answer: C**

22. A morpheme is:

- **A. A sentence chunk**
- **B. Smallest unit of syntax**
- **C. Smallest unit of meaning**
- **D. A sound segment**
- **Answer: C**

23. Prefix 'un-' in 'unhappy' is a:

- **A. Free morpheme**
- **B. Bound morpheme**
- **C. Root**
- **D. Suffix**
- **Answer: B**

24. Free morphemes can:

- **A. Only exist in compounds**
- **B. Not form complete words**
- **C. Stand alone as words**
- **D. Be used as suffixes**
- **Answer: C**

25. Linguistic nativism refers to:

- **A. Learning language only via school**
- **B. Language being innate to humans**
- **C. Artificial language evolution**
- **D. Morphological decoding**
- **Answer: B**

⚙️ Section 2: NLP Pipeline and Preprocessing

1. **What is the first step in an NLP pipeline?**

- **A. POS tagging**
- **B. Parsing**
- **C. Tokenization**
- **D. Lemmatization**
- **Answer: C**

2. **Which technique converts text to lowercase and removes punctuation?**

- **A. Vectorization**
- **B. Stemming**
- **C. Normalization**

- **D. Parsing**
- **Answer: C**

3. **Which of the following is a stop word?**

- **A. Python**
- **B. Apple**
- **C. The**
- **D. Run**
- **Answer: C**

4. **Lemmatization returns:**

- **A. Shortened stems**
- **B. N-grams**
- **C. Root dictionary word**
- **D. Tokens**
- **Answer: C**

5. **What is the key difference between stemming and lemmatization?**

- **A. Lemmatization is faster**
- **B. Lemmatization gives valid words**
- **C. Stemming is more accurate**
- **D. Stemming requires context**
- **Answer: B**

6. **Which NLP library is rule-based?**

- **A. PyTorch**
- **B. spaCy**
- **C. NLTK**

- **D. Keras**
- **Answer: C**

7. **The output of tokenization is:**

- **A. Sentence vectors**
- **B. POS tags**
- **C. Tokens**
- **D. Named entities**
- **Answer: C**

8. **Which of the following is not a token?**

- **A. 'apple'**
- **B. '?'**
- **C. 'run'**
- **D. 'noun'**
- **Answer: D**

9. **Which component removes commonly used but insignificant words?**

- **A. Lemmatizer**
- **B. POS tagger**
- **C. Stop word filter**
- **D. Chunker**
- **Answer: C**

10. **What is the output of chunking?**

- **A. Character tokens**
- **B. Parse tree**
- **C. POS-tagged sentences**

- **D. Phrase-level units**
- **Answer: D**

11. POS tagging assigns:

- **A. Lemmas**
- **B. Sentence chunks**
- **C. Grammatical category**
- **D. Embeddings**
- **Answer: C**

12. Which method uses regex to identify patterns in text?

- **A. Stemming**
- **B. Chunking**
- **C. Parsing**
- **D. NER**
- **Answer: B**

13. Shallow parsing is also called:

- **A. Dependency parsing**
- **B. Light parsing**
- **C. POS tagging**
- **D. Word segmentation**
- **Answer: B**

14. The process of converting words to numbers is called:

- **A. Lemmatization**
- **B. Text vectorization**
- **C. Disambiguation**

- **D. Parsing**
- **Answer: B**

15. Bag-of-Words model ignores:

- **A. Word counts**
- **B. Document size**
- **C. Word order**
- **D. Term uniqueness**
- **Answer: C**

16. What is a document-term matrix?

- **A. Tree of phrases**
- **B. Summary of entities**
- **C. Table of word frequencies per document**
- **D. Table of sentence structures**
- **Answer: C**

17. Which model gives weightage to rare but important words?

- **A. BoW**
- **B. Word2Vec**
- **C. TF-IDF**
- **D. CNN**
- **Answer: C**

18. A unigram model considers:

- **A. One word at a time**
- **B. Two words**
- **C. Trigrams**

- **D. Characters**
- **Answer: A**

19. The term 'normalization' in NLP refers to:

- **A. Setting word frequency to 1**
- **B. Equalizing sentence length**
- **C. Standardizing text format**
- **D. Removing nouns**
- **Answer: C**

20. Which technique is more accurate but slower?

- **A. Lemmatization**
- **B. Stemming**
- **C. Tokenization**
- **D. Chunking**
- **Answer: A**

21. Which part of a word does stemming retain?

- **A. Prefix**
- **B. Suffix**
- **C. Stem/root**
- **D. Lemma**
- **Answer: C**

22. In NLP, ‘chunking’ groups words into:

- **A. Syntax trees**
- **B. Noun/verb phrases**
- **C. Random clusters**

- **D. Characters**
- **Answer: B**

23. POS tagging is usually applied after:

- **A. Parsing**
- **B. Chunking**
- **C. Tokenization**
- **D. Entity linking**
- **Answer: C**

24. What is a limitation of rule-based NLP?

- **A. Overfitting**
- **B. Poor performance on known data**
- **C. Inflexibility and poor generalization**
- **D. High memory usage**
- **Answer: C**

25. TF in TF-IDF represents:

- **A. Total frequency**
- **B. Term formula**
- **C. Term frequency**
- **D. Text feature**
- **Answer: C**

🧱 Section 3: Syntax, Parsing, and Structure

1. **Which component determines grammatical structure?**

- **A. Lemmatizer**
- **B. Syntax parser**
- **C. NER**
- **D. Stop word remover**
- **Answer: B**

2. **Shallow parsing results in:**

- **A. Named entity extraction**
- **B. Phrase chunks**
- **C. Syntax trees**
- **D. Token segmentation**
- **Answer: B**

3. **Deep parsing builds:**

- **A. Word clouds**
- **B. Dependency or constituency trees**
- **C. Document embeddings**
- **D. Grammar rules**
- **Answer: B**

4. **Which parser type uses probabilities?**

- **A. Chart parser**
- **B. Recursive descent parser**
- **C. CFG parser**
- **D. Probabilistic context-free grammar parser**
- **Answer: D**

5. **POS tagging uses which kind of model?**

- **A. Supervised learning**
- **B. Unsupervised learning**
- **C. Reinforcement learning**
- **D. Rule-based only**
- **Answer: A**

6. **The tag ‘VBZ’ in POS tagging stands for:**

- **A. Verb in base form**
- **B. Verb, third-person singular present**
- **C. Verb, past tense**
- **D. Verb phrase**
- **Answer: B**

7. **Which of the following identifies phrase relationships?**

- **A. Lemmatizer**
- **B. Named Entity Recognizer**
- **C. Constituency parser**
- **D. Tokenizer**
- **Answer: C**

8. **Dependency parsing focuses on:**

- **A. Chunking**
- **B. Semantic matching**
- **C. Word-to-word relations**
- **D. Embeddings**
- **Answer: C**

9. **The root of a dependency tree is usually:**

- **A. An adjective**
- **B. A determiner**
- **C. The main verb**
- **D. A noun phrase**
- **Answer: C**

10. Shift-reduce parsing is used in:

- **A. NER**
- **B. spaCy**
- **C. Keras**
- **D. TF-IDF**
- **Answer: B**

11. Which method produces a parse tree from rules?

- **A. Chart parser**
- **B. Neural net parser**
- **C. Word2Vec**
- **D. N-gram model**
- **Answer: A**

12. What is ambiguity in parsing?

- **A. When tokenization fails**
- **B. When multiple parse trees are possible**
- **C. When chunking fails**
- **D. When vectors overlap**
- **Answer: B**

13. Top-down parsers work by:

- **A. Reading tokens line by line**
- **B. Predicting rules from root to leaf**
- **C. Evaluating word frequency**
- **D. Compressing phrases**
- **Answer: B**

14. Which of these can handle ambiguity with backtracking?

- **A. Recursive descent parser**
- **B. TF-IDF model**
- **C. Unigram tagger**
- **D. Stemming**
- **Answer: A**

15. Which is not a syntactic category?

- **A. NP**
- **B. VP**
- **C. DET**
- **D. TF**
- **Answer: D**

16. A constituent is:

- **A. A token**
- **B. A valid phrase in a parse tree**
- **C. A synonym group**
- **D. A vectorized word**
- **Answer: B**

17. Sentence segmentation is part of:

- **A. Syntax**
- **B. Morphology**
- **C. POS tagging**
- **D. Parsing**
- **Answer: D**

18. The output of a constituency parser is:

- **A. Tokens**
- **B. Parse tree with nested phrases**
- **C. Word embeddings**
- **D. Root verbs**
- **Answer: B**

19. Bracketing notation represents:

- **A. Word vectors**
- **B. Syntax trees**
- **C. Parsing failures**
- **D. NER confidence**
- **Answer: B**

20. POS tags are critical for:

- **A. Syntax parsing**
- **B. Tokenization**
- **C. Sentence generation**
- **D. TF-IDF**
- **Answer: A**

21. Parsing helps in:

- **A. Calculating vector distances**
- **B. Structuring grammatical relationships**
- **C. Lemmatizing words**
- **D. Predicting document labels**
- **Answer: B**

22. Recursive structure in parsing means:

- **A. Every sentence has the same pattern**
- **B. Sentences include repeated elements or embedded clauses**
- **C. Words are duplicated**
- **D. Word order is fixed**
- **Answer: B**

23. A sentence with more than one parse tree is:

- **A. Morphologically ambiguous**
- **B. Structurally ambiguous**
- **C. Semantically irrelevant**
- **D. Lexically invalid**
- **Answer: B**

24. The output of POS tagger ‘DT NN VBZ’ is:

- **A. Verb phrase**
- **B. Prepositional clause**
- **C. Noun phrase**
- **D. Determiner, noun, verb**
- **Answer: D**

25. In NLP, parsing is essential for:

- **A. Finding word vectors**
- **B. Understanding syntactic structure**
- **C. Topic modeling**
- **D. TF-IDF indexing**
- **Answer: B**

📊 Section 4: Vectorization and Feature Extraction

1. **Which model represents text as a vector of word counts?**

- **A. Word2Vec**
- **B. TF-IDF**
- **C. Bag-of-Words**
- **D. GloVe**
- **Answer: C**

2. **TF in TF-IDF stands for:**

- **A. Term frequency**
- **B. Token factor**
- **C. Text frequency**
- **D. Term formula**
- **Answer: A**

3. **Which model considers the rarity of terms across documents?**

- **A. Word2Vec**
- **B. CountVectorizer**
- **C. TF-IDF**

- **D. CBOW**
- **Answer: C**

4. **Which is a disadvantage of the Bag-of-Words model?**

- **A. Simple computation**
- **B. Preserves context**
- **C. Ignores word order**
- **D. Requires parsing**
- **Answer: C**

5. **High TF-IDF score indicates:**

- **A. Very common word**
- **B. Rare and informative word**
- **C. Stop word**
- **D. Word from title**
- **Answer: B**

6. **Which vectorization method produces sparse matrices?**

- **A. TF-IDF**
- **B. Word2Vec**
- **C. GloVe**
- **D. Doc2Vec**
- **Answer: A**

7. **What is the size of the BoW feature vector determined by?**

- **A. Number of sentences**
- **B. Number of unique characters**
- **C. Vocabulary size**

- **D. Corpus length**
- **Answer: C**

8. **Which of the following handles semantics best?**

- **A. BoW**
- **B. TF-IDF**
- **C. Word2Vec**
- **D. N-grams**
- **Answer: C**

9. **CBOW in Word2Vec predicts:**

- **A. Context words**
- **B. Target word**
- **C. Sentence**
- **D. Phrase chunk**
- **Answer: B**

10. **Which embedding uses matrix factorization techniques?**

- **A. GloVe**
- **B. CBOW**
- **C. Skip-gram**
- **D. TF-IDF**
- **Answer: A**

11. **What kind of vector does Word2Vec produce?**

- **A. Binary**
- **B. One-hot**
- **C. Dense**

- **D. Sparse**
- **Answer: C**

12. N-grams model helps capture:

- **A. Word frequency only**
- **B. Word co-occurrence**
- **C. Semantic relations**
- **D. POS tagging**
- **Answer: B**

13. Which vector representation preserves context?

- **A. One-hot**
- **B. TF-IDF**
- **C. BERT embeddings**
- **D. Bag-of-Words**
- **Answer: C**

14. The dimensionality of one-hot encoding is:

- **A. Fixed to 100**
- **B. Equal to corpus length**
- **C. Equal to vocabulary size**
- **D. Variable by word count**
- **Answer: C**

15. What does ‘inverse’ in IDF penalize?

- **A. Rare words**
- **B. Unique tokens**
- **C. Common words**

- **D. Verb phrases**
- **Answer: C**

16. Sentence embedding differs from word embedding in:

- **A. Encoding only nouns**
- **B. Context-free structure**
- **C. Representing full sentence meaning**
- **D. Using fewer tokens**
- **Answer: C**

17. Doc2Vec is useful for:

- **A. Named Entity Recognition**
- **B. Document similarity**
- **C. Syntax parsing**
- **D. Entity linking**
- **Answer: B**

18. The output of TF-IDF vectorizer is usually:

- **A. A list**
- **B. Dense matrix**
- **C. Sparse matrix**
- **D. Tensor**
- **Answer: C**

19. Contextual embeddings change:

- **A. With document length**
- **B. For each token instance**
- **C. Based on corpus size**

- **D. With vocabulary**
- **Answer: B**

20. Cosine similarity is often used to:

- **A. Find entity types**
- **B. Compare syntactic rules**
- **C. Measure vector similarity**
- **D. Normalize documents**
- **Answer: C**

21. What is the problem with high-dimensional sparse vectors?

- **A. Slower search**
- **B. Overfitting**
- **C. Dimensionality curse**
- **D. All of the above**
- **Answer: D**

22. Which technique reduces vector dimensionality?

- **A. Tokenization**
- **B. TF-IDF**
- **C. PCA**
- **D. POS tagging**
- **Answer: C**

23. Word2Vec embeddings are trained using:

- **A. LSTMs**
- **B. CNNs**
- **C. Neural nets with context prediction**

- **D. Rule-based models**
- **Answer: C**

24. Which is better for large corpora and capturing analogies?

- **A. Word2Vec**
- **B. TF-IDF**
- **C. BoW**
- **D. POS tags**
- **Answer: A**

25. BERT embeddings are:

- **A. Static**
- **B. Rule-based**
- **C. Contextual and dynamic**
- **D. One-hot encoded**
- **Answer: C**

🧠 Section 5: NLU, NLG, and NLI

1. **NLU focuses on:**

- **A. Parsing**
- **B. Understanding meaning and intent**
- **C. Sentence generation**
- **D. Tokenization**
- **Answer: B**

2. **NLG is used to:**

- **A. Translate intent into natural language**
- **B. Parse sentence structure**
- **C. Tokenize text**
- **D. Generate embeddings**
- **Answer: A**

3. **The task of inferring logical conclusions is part of:**

- **A. NLG**
- **B. NLI**
- **C. POS tagging**
- **D. NER**
- **Answer: B**

4. **Intent detection is crucial for:**

- **A. Entity recognition**
- **B. Summarization**
- **C. Dialogue systems**
- **D. Machine translation**
- **Answer: C**

5. **Which is a common NLU application?**

- **A. Data visualization**
- **B. Sentiment analysis**
- **C. Image processing**
- **D. Data scraping**
- **Answer: B**

6. **In the sentence “Book me a flight,” ‘flight’ is a:**

- **A. Token**
- **B. Named Entity**
- **C. Slot value**
- **D. Parsing error**
- **Answer: C**

7. **Slot filling is part of:**

- **A. Intent classification**
- **B. Tokenization**
- **C. Named Entity Linking**
- **D. Dialogue management**
- **Answer: A**

8. **Which task involves resolving contradictions in text?**

- **A. Parsing**
- **B. NLI**
- **C. Lemmatization**
- **D. Embedding**
- **Answer: B**

9. **The phrase “Can you pass the salt?” demonstrates:**

- **A. Syntax parsing**
- **B. Lexical analysis**
- **C. Indirect intent**
- **D. Semantic similarity**
- **Answer: C**

10. **Which is NOT part of the NLU pipeline?**

- **A. Intent classification**
- **B. Entity recognition**
- **C. Text summarization**
- **D. Slot filling**
- **Answer: C**

11. NLI often evaluates:

- **A. Contradiction, entailment, neutral**
- **B. Sentiment levels**
- **C. Frequency of tokens**
- **D. POS structure**
- **Answer: A**

12. NLG model converts:

- **A. Structured data to text**
- **B. Text to embeddings**
- **C. Parsing tree to tokens**
- **D. Numbers to vectors**
- **Answer: A**

13. A chatbot interpreting “weather in Delhi?” as a query is an example of:

- **A. Tokenization**
- **B. Parsing**
- **C. Intent detection**
- **D. Embedding**
- **Answer: C**

14. Coreference resolution helps in:

- **A. Sentiment classification**
- **B. Matching pronouns to entities**
- **C. Identifying sentence boundaries**
- **D. Extracting topics**
- **Answer: B**

15. An utterance like “Turn it off!” requires:

- **A. Summarization**
- **B. Contextual intent parsing**
- **C. Morphological analysis**
- **D. Word segmentation**
- **Answer: B**

16. Which models are used in NLU classification?

- **A. CNNs**
- **B. Naive Bayes, SVM, LSTM**
- **C. Decision trees**
- **D. GANS**
- **Answer: B**

17. “Apple” as a fruit or company is disambiguated in:

- **A. NLI**
- **B. Intent detection**
- **C. Word Sense Disambiguation**
- **D. NLG**
- **Answer: C**

18. Semantic parsing converts:

- **A. Text into vector**
- **B. Sentence into logical form**
- **C. Document into summary**
- **D. Chat into conversation**
- **Answer: B**

19. Which task ensures correct pronoun linkage?

- **A. NLG**
- **B. NLI**
- **C. Coreference resolution**
- **D. Lemmatization**
- **Answer: C**

20. Named Entity Recognition helps with:

- **A. Word generation**
- **B. Entity classification (like PERSON, DATE)**
- **C. Syntax analysis**
- **D. Vector construction**
- **Answer: B**

21. In NLI, “All dogs bark” and “My dog barks” is:

- **A. Entailment**
- **B. Contradiction**
- **C. Neutral**
- **D. Irrelevant**
- **Answer: C**

22. Which tool is used for NLG?

- **A. spaCy**
- **B. GPT**
- **C. NLTK**
- **D. NumPy**
- **Answer: B**

23. Slot filling extracts:

- **A. Topic keywords**
- **B. Named entities relevant to intent**
- **C. Embeddings**
- **D. POS tags**
- **Answer: B**

24. Sentiment analysis is part of:

- **A. NLU**
- **B. POS tagging**
- **C. Parsing**
- **D. Chunking**
- **Answer: A**

25. Sequence-to-sequence models are often used in:

- **A. POS tagging**
- **B. Text normalization**
- **C. Machine translation (NLG)**
- **D. Parsing**
- **Answer: C**

🏷️ Section 6: Named Entity Recognition and Information Extraction

1. **Named Entity Recognition (NER) is used to:**

- **A. Identify parts of speech**
- **B. Extract named entities like persons, locations**
- **C. Parse sentence structure**
- **D. Tokenize words**
- **Answer: B**

2. **A named entity example is:**

- **A. run**
- **B. quickly**
- **C. Paris**
- **D. if**
- **Answer: C**

3. **Information extraction refers to:**

- **A. Parsing grammar**
- **B. Summarizing entire documents**
- **C. Pulling structured data from text**
- **D. Predicting next word**
- **Answer: C**

4. **Which is not a common named entity category?**

- **A. PERSON**
- **B. LOCATION**
- **C. VERB**
- **D. ORGANIZATION**

- **Answer: C**

5. **Rule-based NER uses:**

- **A. Neural networks**
- **B. Regular expressions and heuristics**
- **C. BoW model**
- **D. TF-IDF**
- **Answer: B**

6. **Statistical NER models are often based on:**

- **A. Regex**
- **B. Decision trees**
- **C. CRFs and HMMs**
- **D. Sentiment classifiers**
- **Answer: C**

7. **Which method is best for recognizing new unseen entities?**

- **A. Rule-based system**
- **B. POS tagger**
- **C. Deep learning models**
- **D. Regex parser**
- **Answer: C**

8. **NER helps in:**

- **A. Improving grammar**
- **B. Document vectorization**
- **C. Knowledge graph construction**
- **D. Token normalization**

- **Answer: C**

9. The phrase "Elon Musk founded Tesla" contains how many named entities?

- **A. 1**
- **B. 2**
- **C. 3**
- **D. 0**
- **Answer: B**

10. Entity linking differs from NER in that it:

- **A. Links words to topics**
- **B. Assigns semantic roles**
- **C. Maps entities to real-world concepts (like Wikipedia)**
- **D. Does POS tagging**
- **Answer: C**

11. Entity recognition in spaCy is performed using:

- **A. TF-IDF**
- **B. Neural models and transition-based parsing**
- **C. One-hot vectors**
- **D. Rule templates only**
- **Answer: B**

12. Which library provides pre-trained NER models?

- **A. NumPy**
- **B. Pandas**
- **C. spaCy**
- **D. Matplotlib**

- **Answer: C**

13. The BIO format used in NER stands for:

- **A. Begin-In-Out**
- **B. Binary Integer Output**
- **C. Biased Input Output**
- **D. Base Indexed Observation**
- **Answer: A**

14. Which is a subtask of information extraction?

- **A. Tokenization**
- **B. POS tagging**
- **C. Coreference resolution**
- **D. Lowercasing**
- **Answer: C**

15. Which task groups entities by relation?

- **A. Clustering**
- **B. Entity normalization**
- **C. Relation extraction**
- **D. Parsing**
- **Answer: C**

16. Temporal expressions are extracted as:

- **A. POS tags**
- **B. DATE entities**
- **C. Modifiers**
- **D. Features**

- **Answer: B**

17. The F1-score in NER balances:

- **A. Precision and recall**
- **B. Accuracy and runtime**
- **C. Recall and specificity**
- **D. Confidence and size**
- **Answer: A**

18. Nested entity recognition is:

- **A. Identifying duplicate entities**
- **B. Finding multi-level or embedded entities**
- **C. Normalizing tokens**
- **D. Chunking sequences**
- **Answer: B**

19. A challenge in NER is:

- **A. Context-insensitive words**
- **B. Training on images**
- **C. Predicting punctuation**
- **D. Lemmatizing entities**
- **Answer: A**

20. Ontologies in IE help with:

- **A. Syntax analysis**
- **B. Defining structured relationships among entities**
- **C. Word segmentation**
- **D. Parsing logic**

- **Answer: B**

21. A named entity disambiguation task maps:

- **A. Tokens to embeddings**
- **B. Ambiguous entities to correct identities**
- **C. Words to syntax**
- **D. Chunks to vectors**
- **Answer: B**

22. Co-reference resolution helps IE by:

- **A. Counting entities**
- **B. Aligning embeddings**
- **C. Linking pronouns to entities**
- **D. POS conversion**
- **Answer: C**

23. Named entities improve:

- **A. Grammar**
- **B. Visualizations**
- **C. Retrieval accuracy in QA systems**
- **D. Punctuation prediction**
- **Answer: C**

24. What is meant by entity drift?

- **A. Shifting locations of tokens**
- **B. Changing relevance of entity types over time**
- **C. Vector changes**
- **D. Parse tree evolution**

- **Answer: B**

25. The use of NER in chatbots is primarily to:

- **A. Understand grammatical intent**
- **B. Fill dialogue slots with relevant real-world data**
- **C. Create syntax trees**
- **D. Adjust token boundaries**
- **Answer: B**

🧬 Section 7: Word Embeddings & Semantic Understanding

1. **Word embeddings convert words into:**

- **A. Strings**
- **B. Numbers**
- **C. Syntax trees**
- **D. Vector representations**
- **Answer: D**

2. **Which embedding model is based on local context prediction?**

- **A. GloVe**
- **B. BERT**
- **C. Word2Vec**
- **D. TF-IDF**
- **Answer: C**

3. **GloVe captures:**

- **A. Syntax rules**

- **B. Global word-word co-occurrence**
- **C. Local character encoding**
- **D. POS tags**
- **Answer: B**

4. **Which model learns word representations via context prediction?**

- **A. TF-IDF**
- **B. Word2Vec**
- **C. SVM**
- **D. Regex**
- **Answer: B**

5. **Embeddings help with:**

- **A. Syntax analysis**
- **B. Understanding semantic similarity**
- **C. Sentence segmentation**
- **D. Parsing**
- **Answer: B**

6. **Context-free embeddings give:**

- **A. Same vector for every occurrence**
- **B. Unique vector per context**
- **C. Sentence-level info**
- **D. TF counts**
- **Answer: A**

7. **Contextual embeddings are generated by:**

- **A. Static models**

- **B. Deep language models like BERT**
- **C. Rule-based engines**
- **D. Regex matchers**
- **Answer: B**

8. **Analogy tasks test embeddings by:**

- **A. Word frequency**
- **B. Vector arithmetic (king - man + woman = ?)**
- **C. Syntax errors**
- **D. POS tagging**
- **Answer: B**

9. **Skip-gram in Word2Vec tries to:**

- **A. Predict context from target word**
- **B. Predict target word from context**
- **C. Skip low-frequency tokens**
- **D. Model sentence grammar**
- **Answer: A**

10. **CBOW in Word2Vec tries to:**

- **A. Predict context from sentence**
- **B. Predict sentence type**
- **C. Predict target word from context**
- **D. Predict grammar pattern**
- **Answer: C**

11. **Embedding dimension size affects:**

- **A. Word count**

- **B. Vector length and representation capacity**
- **C. POS tag set**
- **D. Sentence length**
- **Answer: B**

12. Word2Vec is trained using:

- **A. Backpropagation and gradient descent**
- **B. Decision tree**
- **C. POS rules**
- **D. Frequency tables**
- **Answer: A**

13. GloVe uses:

- **A. Contextual prediction**
- **B. Word co-occurrence matrix factorization**
- **C. Rule-based tagging**
- **D. Token counts**
- **Answer: B**

14. Word similarity in embeddings is measured by:

- **A. Cosine similarity**
- **B. Euclidean distance only**
- **C. TF counts**
- **D. Entity linking**
- **Answer: A**

15. Which vector similarity is scale-invariant?

- **A. Jaccard**

- **B. Cosine**
- **C. Manhattan**
- **D. L1**
- **Answer: B**

16. Pre-trained embeddings are useful when:

- **A. We lack training data**
- **B. We parse documents**
- **C. We need rule sets**
- **D. We cluster tokens**
- **Answer: A**

17. Subword embeddings (e.g. FastText) help with:

- **A. Only frequent words**
- **B. Rare and OOV words**
- **C. Regular expressions**
- **D. Chunking**
- **Answer: B**

18. FastText represents words as:

- **A. Integer indices**
- **B. Bag-of-characters**
- **C. Sums of n-gram embeddings**
- **D. Tokens only**
- **Answer: C**

19. Sentence embeddings encode:

- **A. Individual word meanings**

- **B. Whole sentence semantics**
- **C. Entity classes**
- **D. Token counts**
- **Answer: B**

20. Which model gives different vectors for the same word in different contexts?

- **A. Word2Vec**
- **B. TF-IDF**
- **C. BERT**
- **D. CountVectorizer**
- **Answer: C**

21. Embedding space geometry allows:

- **A. Grammar checking**
- **B. Word clustering and analogy reasoning**
- **C. POS extraction**
- **D. Stop word removal**
- **Answer: B**

22. Fine-tuning embedding layers allows:

- **A. Manual vector editing**
- **B. Learning domain-specific representations**
- **C. Limiting vocabulary size**
- **D. Removing stop words**
- **Answer: B**

23. GloVe is trained on:

- **A. Context windows**

- **B. Word co-occurrence statistics**
- **C. Sentence length**
- **D. Lemmas**
- **Answer: B**

24. Cosine similarity of 1 indicates:

- **A. Orthogonal vectors**
- **B. Identical orientation**
- **C. Zero similarity**
- **D. No relation**
- **Answer: B**

25. OOV stands for:

- **A. Output Optimized Vector**
- **B. Out-Of-Vocabulary**
- **C. Online Vector Value**
- **D. Overfitted Output Vector**
- **Answer: B**

🔁 Section 8: Transformers & Contextual Embeddings

1. **Transformers primarily rely on:**

- **A. Convolutional filters**
- **B. Recurrence**
- **C. Attention mechanisms**
- **D. Word co-occurrence**

- **Answer: C**

2. **What does the 'self' in self-attention mean?**

- **A. Attention is given only to other tokens**
- **B. Each word attends to all tokens in the input, including itself**
- **C. Attention is shared across sentences**
- **D. Only verbs get attention**
- **Answer: B**

3. **Positional encoding is required because:**

- **A. Transformers are recurrent**
- **B. Input tokens are unordered in transformers**
- **C. It adds stop words**
- **D. It replaces embeddings**
- **Answer: B**

4. **Which architecture is used by BERT?**

- **A. Encoder-decoder**
- **B. Encoder only**
- **C. Decoder only**
- **D. Convolutional blocks**
- **Answer: B**

5. **GPT model uses:**

- **A. Decoder-only transformer**
- **B. Encoder-only transformer**
- **C. RNN**
- **D. Attention-less transformer**

- **Answer: A**

6. **The term 'contextual embedding' implies:**

- **A. Fixed word vectors**
- **B. Same vector for each word**
- **C. Different vectors for a word in different contexts**
- **D. Vectors based on grammar rules**
- **Answer: C**

7. **The term 'multi-head attention' allows:**

- **A. Multiple models to run in parallel**
- **B. Focus on different parts of the sentence simultaneously**
- **C. Multiple embeddings per sentence**
- **D. Different tokenizers**
- **Answer: B**

8. **Which transformer component aggregates word importance?**

- **A. Feed-forward layer**
- **B. Positional encoder**
- **C. Attention layer**
- **D. Input embedding**
- **Answer: C**

9. **BERT stands for:**

- **A. Binary Encoding for Recurrent Tokens**
- **B. Bidirectional Encoder Representations from Transformers**
- **C. Backward Embedding for Recurrent Transformers**
- **D. Bi-layer Enhanced Recursive Transformer**

- **Answer: B**

10. Transformer training requires:

- **A. Backpropagation through recurrence**
- **B. Training one token at a time**
- **C. Parallelizable training on full sequences**
- **D. Grammar trees**
- **Answer: C**

11. In transformers, attention score is computed using:

- **A. Dot product of Query and Key vectors**
- **B. Difference between word indices**
- **C. Sum of token positions**
- **D. LSTM memory state**
- **Answer: A**

12. The mask in transformers is used to:

- **A. Prevent dropout**
- **B. Hide future tokens during training**
- **C. Improve syntax parsing**
- **D. Reduce vocabulary**
- **Answer: B**

13. Fine-tuning a transformer refers to:

- **A. Retraining only the embedding layer**
- **B. Training on a new corpus from scratch**
- **C. Adapting a pre-trained model to a downstream task**
- **D. Changing model architecture**

- **Answer: C**

14. RoBERTa improves on BERT by:

- **A. Using fewer layers**
- **B. Removing dropout**
- **C. Training longer on more data**
- **D. Replacing attention with CNN**
- **Answer: C**

15. Which is NOT a benefit of transformer models?

- **A. Capturing long-range dependencies**
- **B. Parallelizable training**
- **C. Fixed-size vocabulary**
- **D. Transfer learning**
- **Answer: C**

16. The key innovation of transformer compared to RNN is:

- **A. Embeddings**
- **B. Attention mechanism and parallelism**
- **C. Bag-of-Words encoding**
- **D. Syntax tree parsing**
- **Answer: B**

17. In BERT, the [CLS] token is used for:

- **A. Padding**
- **B. Segment separation**
- **C. Classification tasks**
- **D. Ignoring irrelevant text**

- **Answer: C**

18. A transformer block consists of:

- **A. RNN + CNN**
- **B. Self-attention + Feed-forward layers**
- **C. GRU + Memory network**
- **D. Embedding only**
- **Answer: B**

19. Token embeddings + positional encodings =

- **A. Input to attention layer**
- **B. Final output vector**
- **C. Masked sequence**
- **D. N-gram probability**
- **Answer: A**

20. Transformers solve vanishing gradient problem by:

- **A. ReLU activations**
- **B. Using GRUs**
- **C. Removing recurrence and using attention**
- **D. Bag-of-words trick**
- **Answer: C**

21. Pre-trained transformers are typically trained on:

- **A. Specific application data**
- **B. Small labeled corpora**
- **C. Large unlabeled corpora using self-supervised tasks**
- **D. Only Wikipedia**

- **Answer: C**

22. The output dimension of transformer layers depends on:

- **A. Number of input tokens**
- **B. Vocabulary size**
- **C. Hidden size parameter**
- **D. Embedding function**
- **Answer: C**

23. Which model generalizes better in multilingual NLP?

- **A. BERT**
- **B. GPT**
- **C. mBERT or XLM-R**
- **D. Naive Bayes**
- **Answer: C**

24. Self-attention helps by:

- **A. Keeping only subject and object**
- **B. Allowing model to attend to every token at each layer**
- **C. Skipping irrelevant parts of text**
- **D. Reducing matrix dimensions**
- **Answer: B**

25. Sentence-transformers are used for:

- **A. Named entity recognition**
- **B. Sentence-level semantic similarity**
- **C. Syntax tagging**
- **D. Word segmentation**

- **Answer: B**

❤️ Section 9: Sentiment Analysis & Classification

1. **Sentiment analysis helps determine:**

- **A. Topic category**
- **B. Grammatical errors**
- **C. Emotional tone or opinion in text**
- **D. Part-of-speech**
- **Answer: C**

2. **Which machine learning algorithm is often used for sentiment classification?**

- **A. Linear regression**
- **B. K-means**
- **C. Naive Bayes**
- **D. DBSCAN**
- **Answer: C**

3. **A lexicon-based approach to sentiment analysis uses:**

- **A. Predefined word lists with sentiment scores**
- **B. Deep neural networks**
- **C. TF-IDF weights only**
- **D. Clustering techniques**
- **Answer: A**

4. **Which of the following is a polarity label in binary sentiment analysis?**

- **A. Neutral**

- **B. Mixed**
- **C. Positive**
- **D. Subjective**
- **Answer: C**

5. **The VADER sentiment tool is best suited for:**

- **A. Long legal documents**
- **B. Tweets and social media text**
- **C. Technical manuals**
- **D. News summarization**
- **Answer: B**

6. **Sentiment scores typically range between:**

- **A. 0–100**
- **B. -1 to +1**
- **C. 1–10**
- **D. 0–1**
- **Answer: B**

7. **Which of these is a limitation of lexicon-based sentiment analysis?**

- **A. Cannot process tweets**
- **B. Ignores sentence structure/context**
- **C. Requires labeled data**
- **D. Overfits quickly**
- **Answer: B**

8. **Supervised sentiment analysis requires:**

- **A. Only raw text**

- **B. Rule-based grammar**
- **C. Labeled training data**
- **D. Parsing trees**
- **Answer: C**

9. Word embeddings help in sentiment analysis by:

- **A. Compressing sentences**
- **B. Capturing contextual meaning of words**
- **C. Vectorizing syntax trees**
- **D. Removing punctuation**
- **Answer: B**

10. Which of these is a common dataset for sentiment analysis?

- **A. IMDB reviews**
- **B. COCO dataset**
- **C. SQuAD**
- **D. MNIST**
- **Answer: A**

11. Which deep learning model is used for sequential sentiment analysis?

- **A. CNN**
- **B. RNN / LSTM**
- **C. Decision tree**
- **D. Random Forest**
- **Answer: B**

12. What does the softmax function do in classification?

- **A. Reduces dimensionality**

- **B. Extracts features**
- **C. Converts logits into probabilities**
- **D. Tokenizes inputs**
- **Answer: C**

13. A sentence with both positive and negative phrases may be:

- **A. Ambiguous**
- **B. Subjective**
- **C. Mixed sentiment**
- **D. Neutral**
- **Answer: C**

14. Which Python library supports sentiment analysis via TextBlob?

- **A. Numpy**
- **B. Scikit-learn**
- **C. TextBlob**
- **D. PyTorch**
- **Answer: C**

15. An F1-score is especially important when:

- **A. Data is balanced**
- **B. Classes are imbalanced**
- **C. Only accuracy is needed**
- **D. Tokenization fails**
- **Answer: B**

16. What role does tokenization play in sentiment analysis?

- **A. Summarizes paragraphs**

- **B. Breaks text into manageable units (e.g. words)**
- **C. Builds graphs**
- **D. Normalizes values**
- **Answer: B**

17. Sentiment intensity is:

- **A. A score for text objectivity**
- **B. Degree of emotion (e.g. very positive)**
- **C. Probability of grammar correctness**
- **D. Measure of entity linking**
- **Answer: B**

18. Which of these can improve sentiment accuracy?

- **A. POS tagging**
- **B. Coreference resolution**
- **C. Negation handling**
- **D. Lemmatization only**
- **Answer: C**

19. The phrase “Not bad” is:

- **A. Negative sentiment**
- **B. Ambiguous**
- **C. Context-free**
- **D. Positive sentiment (due to negation)**
- **Answer: D**

20. Emojis in sentiment analysis are:

- **A. Ignored by all models**

- **B. Used to improve context in lexicon models**
- **C. Treated as punctuation**
- **D. Always negative**
- **Answer: B**

21. Zero-shot classification enables:

- **A. Learning without any data**
- **B. Labeling new categories not seen in training**
- **C. Entity linking**
- **D. Parsing trees**
- **Answer: B**

22. Transformer-based models in sentiment analysis are preferred because:

- **A. They ignore rare words**
- **B. They are less accurate**
- **C. They capture context better**
- **D. They are rule-based**
- **Answer: C**

23. Fine-tuning a pre-trained BERT for sentiment analysis requires:

- **A. Grammar tree**
- **B. Rule set**
- **C. Sentiment-labeled dataset**
- **D. Visualizer**
- **Answer: C**

24. Which of the following handles sarcasm poorly?

- **A. Rule-based sentiment models**

- **B. Deep contextual models**
- **C. BERT fine-tuned models**
- **D. LSTM with attention**
- **Answer: A**

25. Which metric is best when false positives and negatives are costly?

- **A. Accuracy**
- **B. Recall**
- **C. Precision**
- **D. F1-score**
- **Answer: D**

🧩 Section 10: Advanced Parsing & Semantics

1. **Semantic parsing transforms natural language into:**

- **A. Vectors**
- **B. Logical forms or machine-understandable queries**
- **C. Chunked phrases**
- **D. Morphological roots**
- **Answer: B**

2. **Which of the following helps with word sense disambiguation?**

- **A. POS tagging**
- **B. Syntax trees**
- **C. Contextual embeddings**
- **D. Lemmatization**

- **Answer: C**

3. **A sentence with multiple interpretations is said to be:**

- **A. Ambiguous**
- **B. Unsupervised**
- **C. Recursive**
- **D. Predictive**
- **Answer: A**

4. **The output of semantic role labeling includes:**

- **A. Intent labels**
- **B. Word vectors**
- **C. Arguments for predicates**
- **D. Entity names**
- **Answer: C**

5. **Which of these best captures meaning across sentence structure?**

- **A. POS tags**
- **B. Syntax trees**
- **C. Semantic dependency graphs**
- **D. Lemma chains**
- **Answer: C**

6. **WordNet is used for:**

- **A. POS tagging**
- **B. Syntax parsing**
- **C. Lexical semantics**
- **D. Coreference resolution**

- **Answer: C**

7. **Thematic roles in SRL include:**

- **A. Question-answering types**
- **B. Named entities**
- **C. Agent, theme, instrument**
- **D. Token lengths**
- **Answer: C**

8. **Inference in semantics often deals with:**

- **A. Morphological variants**
- **B. Logical entailment**
- **C. Vector normalization**
- **D. Sentiment labels**
- **Answer: B**

9. **Which of the following enhances semantic accuracy?**

- **A. Stemming**
- **B. Sentence segmentation**
- **C. Coreference resolution**
- **D. Frequency counts**
- **Answer: C**

10. **Dependency structures map:**

- **A. Topics to sentences**
- **B. Entity relations only**
- **C. Words and their grammatical functions**
- **D. Token counts**

- **Answer: C**

11. Which type of parsing is most useful for semantics?

- **A. Surface parsing**
- **B. Dependency parsing**
- **C. Syntactic chunking**
- **D. POS tagging**
- **Answer: B**

12. FrameNet is used to:

- **A. Parse syntax**
- **B. Label words with semantic frames**
- **C. Normalize embeddings**
- **D. Classify texts**
- **Answer: B**

13. Predicate-argument structures are key to:

- **A. Tokenization**
- **B. Sentiment scoring**
- **C. Semantic parsing**
- **D. Grammar correction**
- **Answer: C**

14. Compositional semantics refers to:

- **A. Sentence structure only**
- **B. Meaning derived from combining word meanings**
- **C. Random word combinations**
- **D. Syntax-only rules**

- **Answer: B**

15. Semantic entailment means:

- **A. A sentence implies another**
- **B. Synonyms are replaced**
- **C. Entity is named twice**
- **D. Meaning is ambiguous**
- **Answer: A**

16. The sentence “John gave Mary a book” assigns roles:

- **A. Verb-object-subject**
- **B. Agent-recipient-theme**
- **C. Subject-object-action**
- **D. Theme-agent-recipient**
- **Answer: B**

17. Pragmatic analysis focuses on:

- **A. Word structure**
- **B. Contextual interpretation of meaning**
- **C. Syntax tags**
- **D. Frequency analysis**
- **Answer: B**

18. Ambiguity in semantics can be resolved using:

- **A. WordNet**
- **B. Parse tree depth**
- **C. Token frequency**
- **D. N-gram models**

- **Answer: A**

19. The symbol '→' in logical form indicates:

- **A. Syntax agreement**
- **B. Mapping rules**
- **C. Implication**
- **D. Sentiment polarity**
- **Answer: C**

20. Semantic parsing is used in:

- **A. POS tagging**
- **B. Knowledge base querying**
- **C. N-gram expansion**
- **D. Morphological analysis**
- **Answer: B**

21. Which NLP task converts questions into SQL-like queries?

- **A. Chunking**
- **B. Named entity tagging**
- **C. Semantic parsing**
- **D. Sentiment classification**
- **Answer: C**

22. Which type of ambiguity involves conflicting sentence roles?

- **A. Lexical**
- **B. Structural**
- **C. Scope**
- **D. Thematic**

- **Answer: D**

23. A limitation of rule-based semantic systems is:

- **A. Accuracy**
- **B. Scalability and coverage**
- **C. Interpretability**
- **D. Determinism**
- **Answer: B**

24. Temporal semantics deals with:

- **A. Locations**
- **B. Word embeddings**
- **C. Time-based interpretations**
- **D. Grammar corrections**
- **Answer: C**

25. Which tool is used to visualize dependency semantics?

- **A. TensorBoard**
- **B. NLTK Tree**
- **C. displaCy by spaCy**
- **D. Matplotlib**
- **Answer: C**

🧠 Section 11: Deep Learning in NLP

1. **Deep learning models learn features from:**

- **A. Manual rules**

- **B. Structured trees**
- **C. Raw data automatically**
- **D. Lexicons**
- **Answer: C**

2. **Which neural network is best suited for sequential data like text?**

- **A. CNN**
- **B. RNN**
- **C. Decision Tree**
- **D. Random Forest**
- **Answer: B**

3. **The vanishing gradient problem affects:**

- **A. CNNs**
- **B. Short RNNs only**
- **C. Deep RNNs and LSTMs**
- **D. Shallow MLPs**
- **Answer: C**

4. **LSTM improves on RNN by:**

- **A. Increasing depth**
- **B. Introducing attention**
- **C. Using memory gates to manage long dependencies**
- **D. Using wider layers**
- **Answer: C**

5. **The output of a softmax layer is:**

- **A. Word vector**

- **B. Normalized probabilities**
- **C. Distance score**
- **D. Attention weight**
- **Answer: B**

6. **Which of these is not a typical activation function?**

- **A. ReLU**
- **B. Sigmoid**
- **C. Tanh**
- **D. Pool**
- **Answer: D**

7. **A CNN in NLP is used for:**

- **A. Entity linking**
- **B. Parsing**
- **C. Capturing local n-gram features**
- **D. Summarizing documents**
- **Answer: C**

8. **Attention in deep learning allows:**

- **A. Tokenization**
- **B. Stemming**
- **C. Focusing on important words in context**
- **D. Tree-based parsing**
- **Answer: C**

9. **Which layer reduces overfitting by randomly disabling neurons?**

- **A. Pooling**

- **B. Dropout**
- **C. Dense**
- **D. Softmax**
- **Answer: B**

10. What type of learning does BERT use initially?

- **A. Supervised**
- **B. Reinforcement**
- **C. Unsupervised (masked language modeling)**
- **D. Semi-supervised**
- **Answer: C**

11. The hidden state in RNN captures:

- **A. Output label**
- **B. Syntax rules**
- **C. Accumulated context**
- **D. Gradient descent**
- **Answer: C**

12. GRU stands for:

- **A. Gated Recurrent Unit**
- **B. Gradient Residual Unit**
- **C. Generalized ReLU Unit**
- **D. Grammar Reduction Unit**
- **Answer: A**

13. In LSTM, which gate decides how much to forget?

- **A. Input gate**

- **B. Output gate**
- **C. Forget gate**
- **D. Context gate**
- **Answer: C**

14. Which component in LSTM controls information retention?

- **A. Cell state**
- **B. Dropout**
- **C. ReLU**
- **D. Embedding layer**
- **Answer: A**

15. Which neural model is bidirectional?

- **A. Vanilla RNN**
- **B. CNN**
- **C. BiLSTM**
- **D. GRU**
- **Answer: C**

16. The number of trainable parameters in deep NLP models grows with:

- **A. Input length**
- **B. Number of layers and embedding size**
- **C. Document ID**
- **D. Vocabulary size only**
- **Answer: B**

17. Sequence-to-sequence models are used in:

- **A. Parsing trees**

- **B. Text classification**
- **C. Machine translation and summarization**
- **D. Tokenization**
- **Answer: C**

18. A model that converts audio to text is:

- **A. Text classifier**
- **B. Speech recognizer**
- **C. Token generator**
- **D. Syntax parser**
- **Answer: B**

19. Which deep learning model performs best on long-term dependencies?

- **A. CNN**
- **B. LSTM**
- **C. BiLSTM**
- **D. Transformer**
- **Answer: D**

20. Batch size in training affects:

- **A. Dropout rate**
- **B. Inference only**
- **C. Memory usage and convergence speed**
- **D. Vocabulary**
- **Answer: C**

21. Which loss function is used for multi-class classification?

- **A. MSE**

- **B. Binary cross-entropy**
- **C. Categorical cross-entropy**
- **D. Hinge loss**
- **Answer: C**

22. Which layer follows the embedding layer in typical NLP models?

- **A. Input layer**
- **B. Attention layer or RNN layer**
- **C. Dropout**
- **D. Output softmax**
- **Answer: B**

23. Which technique accelerates convergence in deep models?

- **A. Stemming**
- **B. Learning rate scheduling**
- **C. POS tagging**
- **D. Token pruning**
- **Answer: B**

24. Pretrained word embeddings help by:

- **A. Saving computation**
- **B. Boosting generalization with prior knowledge**
- **C. Decreasing vocabulary size**
- **D. Reducing parse tree depth**
- **Answer: B**

25. In NLP pipelines, deep learning is most beneficial for:

- **A. Grammar rules**

- **B. Rule-based NER**
- **C. Semantic understanding and generation tasks**
- **D. Token count**
- **Answer: C**

🗂️ Section 12: Text Classification & Topic Modeling

1. **Text classification is a type of:**

- **A. Unsupervised learning**
- **B. Reinforcement learning**
- **C. Supervised learning**
- **D. Deep clustering**
- **Answer: C**

2. **Naive Bayes works best for:**

- **A. Contextual embedding**
- **B. Word2Vec**
- **C. High-dimensional sparse data**
- **D. Syntax parsing**
- **Answer: C**

3. **Which classifier separates data using a hyperplane?**

- **A. Naive Bayes**
- **B. Decision Tree**
- **C. SVM**
- **D. Random Forest**

- **Answer: C**

4. **Which method is suitable for multi-class text classification?**

- **A. One-vs-all SVM**
- **B. BoW only**
- **C. Skip-gram**
- **D. Syntax trees**
- **Answer: A**

5. **In classification, overfitting means:**

- **A. Model underperforms**
- **B. Model performs well on training but poorly on new data**
- **C. Model generalizes well**
- **D. Model has high bias**
- **Answer: B**

6. **Which evaluation metric is not ideal for imbalanced classes?**

- **A. Accuracy**
- **B. F1-score**
- **C. Precision**
- **D. Recall**
- **Answer: A**

7. **Precision is defined as:**

- **A. TP / (TP + FN)**
- **B. TP / (TP + FP)**
- **C. TN / (TN + FP)**
- **D. FN / (TP + FN)**

- **Answer: B**

8. **LDA in topic modeling stands for:**

- **A. Long Dependency Analysis**
- **B. Latent Dirichlet Allocation**
- **C. Layered Deep Attention**
- **D. Local Distance Algorithm**
- **Answer: B**

9. **Topic modeling is:**

- **A. Supervised**
- **B. Unsupervised**
- **C. Reinforced**
- **D. Manual**
- **Answer: B**

10. **Which technique visualizes document-topic distribution?**

- **A. TF-IDF**
- **B. LDA**
- **C. NER**
- **D. BoW**
- **Answer: B**

11. **Gensim is used for:**

- **A. POS tagging**
- **B. Word embedding only**
- **C. Topic modeling and document similarity**
- **D. Syntax trees**

- **Answer: C**

12. Which metric is best for evaluating topic coherence?

- **A. Precision**
- **B. Perplexity**
- **C. BLEU**
- **D. ROUGE**
- **Answer: B**

13. K-means clustering can be used for:

- **A. Supervised learning**
- **B. Sentiment labeling**
- **C. Document clustering**
- **D. Syntax tagging**
- **Answer: C**

14. A topic is defined as:

- **A. Sentence of nouns**
- **B. Word frequency score**
- **C. Probability distribution over words**
- **D. Vector similarity**
- **Answer: C**

15. Which model provides probabilities of topics in a document?

- **A. TF-IDF**
- **B. CountVectorizer**
- **C. LDA**
- **D. SVM**

- **Answer: C**

16. Dimensionality reduction in topic modeling can be done using:

- **A. PCA**
- **B. LDA**
- **C. Random Forest**
- **D. POS tagging**
- **Answer: A**

17. Bag-of-Words in topic modeling provides:

- **A. Contextual info**
- **B. Token vectorization**
- **C. Term-document matrix**
- **D. Entity linking**
- **Answer: C**

18. In unsupervised text clustering, labels are:

- **A. Pre-defined**
- **B. Generated after model**
- **C. Derived using POS**
- **D. Fixed manually**
- **Answer: B**

19. What does the silhouette score evaluate?

- **A. Clustering quality**
- **B. Model overfitting**
- **C. Token length**
- **D. Parsing depth**

- **Answer: A**

20. Which algorithm is best suited for exploratory topic discovery?

- **A. SVM**
- **B. LDA**
- **C. FastText**
- **D. RoBERTa**
- **Answer: B**

21. Topic modeling outputs are typically:

- **A. Topic embeddings**
- **B. Word counts**
- **C. Probability distributions over topics and words**
- **D. Syntax trees**
- **Answer: C**

22. Which NLP task is often combined with topic modeling?

- **A. Summarization**
- **B. POS tagging**
- **C. Co-reference resolution**
- **D. NER**
- **Answer: A**

23. Which evaluation metric is common for text classification?

- **A. BLEU**
- **B. F1-score**
- **C. WER**
- **D. CER**

- **Answer: B**

24. Document similarity can be computed using:

- **A. Word frequency only**
- **B. Cosine similarity over TF-IDF vectors**
- **C. Syntax tree depth**
- **D. POS overlap**
- **Answer: B**

25. Which tool helps visualize topic clusters?

- **A. Matplotlib**
- **B. pyLDAvis**
- **C. Seaborn**
- **D. SciPy**
- **Answer: B**