1. Which of the following is NOT a definition of language?
a) A system of communication
b) A biological entity
c) A form of thought
d) A collection of words

2. Language is considered a dynamic system because it:
a) Changes over time
b) Remains static
c) Is a fixed set of rules
d) Is only used in formal contexts

3. Which theory suggests that language is innate and human beings are born with the capacity for grammar?
a) Behaviourist theory
b) Interactionist theory
c) Nativist theory
d) Connectionist theory

4. Language is regarded as a biological phenomenon because it:
a) Exists only in human
societies
b) Has a genetic foundation in the human brain
c) Is learned socially
d) Functions through written symbols

5. Which of the following is an example of a spoken mode of language?
a) Text message
b) Sign language
c) **Audio recording**
d) Written text

6. The distinction between language as expression and content was proposed by:
a) **Noam Chomsky**
b) Ferdinand de Saussure
c) John Searle
d) Charles Peirce

7. Which of the following is a feature of artificial languages (like programming languages)?
a) **Ambiguity**
b) Dynamic syntax
c) Strict rules and structure
d) Emotional expression

8. Linguistics as a scientific study focuses on:
a) The artistic creation of language
b) The analysis of the structure, meaning, and usage of language
c) **The history of language**
d) The learning of foreign languages

9. Which of the following best describes a symbolic system in linguistics?
a) Natural language
b) **Gestures and facial expressions**
c) Computer programming languages
d) Both a and b

10. Which of the following is the primary function of language?
a) To enable communication
b) To express emotions
c) **To store memories**
d) To build social relationships

11. What does language analysis primarily focus on?
a) The emotional effects of language
b) The technicalities of language use in programming
c) **The structure and meaning of language**
d) The social contexts in which language is used

12. The paradigmatic relationship in linguistics refers to:
a) Words that are used together in the same sentence
b) The substitution of one word for another
c) The order of words in a sentence
d) **The sounds that words are made up of**

13. Which linguistic level focuses on the pronunciation of sounds?
a) Phonetics
b) **Semantics**
c) Syntax
d) Pragmatics

14. Which level of linguistic analysis deals with sentence structure and word order?
a) Phonology
b) **Morphology**
c) Syntax
d) Semantics

15. The process of deriving meaning from the structure and usage of language is called:
a) **Morphology**
b) Syntax
c) Semantics
d) Phonetics

16. Which is an example of an Artificial Intelligence (AI) subfield that focuses on understanding human language?
a) Computer Vision
b) Machine Learning
c) **Natural Language Understanding**
d) Robotics

17. Natural Language Generation (NLG) focuses on:
a) **Understanding text**
b) Creating human-like text
c) Learning language rules
d) Parsing sentences

18. The main difference between semantics and pragmatics is that:
a) **Semantics deals with word meanings, while pragmatics concerns context-based meaning**
b) Semantics involves word sounds, and pragmatics concerns grammar
c) Semantics is a subset of pragmatics
d) There is no difference

19. Which of the following is the primary function of Named-Entity Recognition (NER)?
a) Categorize words by their part of speech
b) **Identify and classify proper nouns and specific terms**
c) Separate text into sentences
d) Analyze sentence structure

20. Lemmatization involves:
a) Removing unnecessary characters from words
b) Reducing words to their root forms
c) **Assigning part-of-speech tags to words**
d) Identifying the subject-verb agreement in a sentence

21. Which machine learning model is primarily used for sequence data in NLP?
a) **Support Vector Machines**
b) Recurrent Neural Networks (RNN)
c) Decision Trees
d) Naive Bayes

22. Which type of neural network is particularly useful for processing grid-like data such as images and texts?
a) CNN (Convolutional Neural Network)
b) RNN (Recurrent Neural Network)
c) **LSTM (Long Short-Term Memory)**
d) MLP (Multilayer Perceptron)

23. Which of the following is NOT a challenge in training deep learning models for NLP?
a) Vanishing gradient problem
b) **Limited data availability**
c) Interpretability of models
d) High processing power required for training

24. The LSTM (Long Short-Term Memory) network resolves the vanishing gradient problem by:
a) Adding extra layers
b) **Using self-attention mechanisms**
c) Using gating mechanisms to retain information
d) Reducing the complexity of the network

25. In natural language processing (NLP), what does a lemmatizer do?
a) Assignsmeaning to a word
b) Identifies the syntactic structure of a sentence
c) Converts a word to its base or dictionary form
d) **Identifies the emotional tone of a word**

26. Which of the following neural networks is best suited for processing sequential data?
a) CNN
b) RNN
c) **DNN**
d) GAN

27. Which of the following is the main function of sequence-to-sequence (Seq2Seq) models in NLP?
a) **Categorizing text**
b) Translating one sequence to another (e.g., language translation)
c) Identifying entities in text
d) Generating summaries from text

28. Which neural network architecture is known for its self-attention mechanism in NLP tasks?
a) RNN
b) Transformer
c) **LSTM**
d) CNN

29. What is the primary advantage of using Transformer networks over RNNs?
a)Faster processing of sequential data
b) **Better performance on non-sequential tasks**
c) Ability to process sequences in parallel
d) Improved understanding of semantic meaning

30. BERT (Bidirectional Encoder Representations from Transformers) is mainly used for:
a) **Image recognition**
b) Text classification
c) Language translation
d) Speech synthesis

31. Which of the following technologies enables speech-to-text conversion?
a) NaturalLanguage Generation
b) Automatic Speech Recognition (ASR)
c) Text to Speech (TTS)
d) Sentiment Analysis

32. Which approach in NLP focuses on translating words from one language to another?
a) Sentiment analysis
b) Named-Entity Recognition
c) Machine translation
d) Text classification

33. Which of the following tasks is involved in information retrieval?
a) Identifying entities in text
b) Converting text to speech
c) Searching and retrieving relevant documents
d) Categorizing a set of documents

34. Which model is primarily used for text summarization?
a) Seq2Seq
b) Word2Vec
c) Transformer
d) LSTM

35. In automatic speech recognition (ASR), which of the following is a primary challenge?
a) Identifying sentiment
b) Handling different accents and noise
c) Understanding word meaning
d) Generating sentences from speech

36. Which of the following is a common method for recognizing words in handwritten documents?
a) Optical Character Recognition (OCR)
b) Machine Translation
c) Named-Entity Recognition
d) Text Summarization

37. Which of the following techniques is most used for text-based machine translation?
a) Neural machine translation (NMT)
b) Rule-based translation
c) Syntax-based translation
d) Statistical translation

38. Which NLP task would involve identifying the gender of a person from a name?
a)Named-Entity Recognition
b) Gender Prediction
c) Sentiment Analysis
d) Text Classification

39. Which of the following NLP tools is commonly used for tokenization and POS tagging?
a) TensorFlow
b) NLTK
c) OpenCV
d) PyTorch

40. Which of the following is an essential task for chatbots in NLP?
a) Speech synthesis
b) Sentiment analysis
c) Question answering and dialogue management
d) Named-Entity Recognition

41. Which of the following techniques is used for syntactic parsing in NLP?
a) Hidden Markov Models
b) Context-Free Grammar
c) Word2Vec
d) Convolutional Neural Networks

42. Semantic parsing involves:
a) Extracting meaning from spoken language
b) Converting structured data into unstructured data
c) Extracting relationships and meaning from a sentence
d) Identifying grammatical errors in a text

43. Which NLP technique is used to resolve ambiguity in a sentence by identifying the correct meaning of a word based on context?
a) Word Sense Disambiguation
b) Named-Entity Recognition
c) Tokenization
d) Dependency Parsing

44. Which approach focuses on extracting structured information from unstructured text data?
a) Information Extraction
b) Part-of-Speech Tagging
c) Text Summarization
d) Named-Entity Recognition

45. What is automatic summarization used for in NLP?
a) Translating languages
b) Reducing text length while preserving meaning
c) Identifying entities in text
d) Identifying the sentiment of the text

46. What is anaphora resolution in NLP?
a) Identifying similar words in a sentence
b) Resolving ambiguities in word meanings
c) Disambiguating the relationships between words in a sentence
d) Identifying which word a pronoun refers to

47. Ontology in NLP refers to:
a) A semantic dictionary of words
b) The structure of sentences
c) The study of speech sounds
d) A formal representation of knowledge

48. The semantic web is an extension of the current web, designed to:
a) Integrate text with multimedia content
b) Provide human-like interaction with machines
c) Allow data to be shared and reused across different applications
d) Use artificial intelligence for content generation

49. Which of the following algorithms is most commonly used for information extraction?
a) Hidden Markov Models
b) Support Vector Machines
c) Named-Entity Recognition
d) Word2Vec

50. The main challenge of deep parsing is:
a) Identifying the sequence of words
b) Managing a large vocabulary
c) The complexity of sentence structures
d) Assigning proper meanings to words

51. Which of the following statistical models is based on probabilities of state transitions?
a) Hidden Markov Models
b) Support Vector Machines
c) Naive Bayes
d) Decision Trees

52. Markov models are primarily used for:
a) Text classification
b) Language modeling and speech recognition
c) Named-Entity Recognition
d) Word embedding

53. In text classification, which of the following models is used for grouping similar text documents?
a) Clustering
b) Regression
c) Regression Trees
d) Naive Bayes

54. Which of the following is a common technique for categorizing text into different topics or genres?
a) Text Classification
b) Sequence Modeling
c) Named-Entity Recognition
d) Word Sense Disambiguation

55. Which of the following is the basis for the Centroid-based Classification method?
a) Dividing text based on its length
b) Grouping documents by central terms or themes
c) Training a model with labeled data
d) Identifying key phrases in text

56. Entropy models in statistical NLP are used to:
a) Measure the uncertainty in a dataset
b) Classify text based on predefined labels
c) Learn the meaning of sentences
d) Organize documents into clusters

57. Which of the following is an important concept in statistical parsing?
a) Assigning words to parts of speech
b) Reducing the number of words in a sentence
c) Estimating the probability of syntactic structures
d) Mapping one word to multiple possible meanings

58. What does text categorization aim to achieve?
a) Grouping text into predefined categories
b) Generating summaries of the text
c) Extracting named entities from text
d) Converting text into speech

59. The primary challenge in statistical text classification is:
a) Creating a large enough labeled dataset
b) Identifying the order of words in a sentence
c) Mapping words to their correct part of speech
d) Converting text into vectors

60. Which algorithm is often used in text classification to identify categories based on statistical patterns?
a) Support Vector Machine
b) Long Short-Term Memory
c) K-means clustering
d) Naive Bayes

61. Which of the following best distinguishes Machine Learning (ML) from Deep Learning (DL)?
a) ML focuses on shallow models, while DL uses deep neural networks
b) ML is based on simple algorithms, while DL requires large datasets
c) ML is better for complex tasks than DL
d) DL focuses on numerical data, while ML focuses on categorical data

62. Which of the following deep learning models is commonly used for sequential data like text?
a) RNN
b) CNN
c) GAN
d) MLP

63. Which of the following deep learning architectures is designed to handle long-term dependencies in sequential data?
a) LSTM (Long Short-Term Memory)
b) CNN (Convolutional Neural Network)
c) GAN (Generative Adversarial Network)
d) MLP (Multilayer Perceptron)

64. Which of the following models is used to process sequential data using a simple feedback mechanism?
a) RNN (Recurrent Neural Network)
b) CNN (Convolutional Neural Network)
c) LSTM (Long Short-Term Memory)
d) GAN (Generative Adversarial Network)

65. The most important challenge when training deep learning models for NLP is:
a)Overfitting to training data
b) Understanding the syntax of the language
c) Identifying the sentiment of the text
d) Reducing vocabulary size

66. Which type of neural network is most effective for understanding the context of words in long sentences?
a) Recurrent Neural Networks (RNN)
b) Convolutional Neural Networks (CNN)
c) Transformer Networks
d) Support Vector Machines

67. Which deep learning architecture is known for its self-attention mechanism, which is critical for NLP tasks?
a) RNN
b) CNN
c) Transformer
d) GAN

68. Which of the following models helps to overcome the vanishing gradient problem in RNNs?
a) LSTM
b) CNN
c) Decision Trees
d) **Naive Bayes**

69. Which machine learning technique is best suited for sentiment analysis?
a) Decision Trees
b) **Random Forest**
c) Support Vector Machines
d) Deep Learning Models

70. Which deep learning model architecture would you choose for large-scale language modeling tasks?
a) RNN
b) LSTM
c) **Transformer**
d) CNN

71. What is the first step in text pre-processing for NLP?
a) **Tokenization**
b) Lemmatization
c) Removing stopwords
d) Part-of-Speech Tagging

72. Which Python library is widely used for processing natural language text?
a)TensorFlow
b) OpenCV
c) NLTK
d) **Matplotlib**

73. The primary function of spaCy in NLP is to:
a) Build language models
b) Preprocess text and perform tokenization, POS tagging
c) Implement deep learning algorithms
d) Convert speech to text

74. Which of the following is a common challenge during pre-processing data for NLP?
a) Overfitting the model
b) Selecting relevant features for modeling
c) Balancing the dataset for equal representation of categories
d) Resolving sentence-level ambiguity

75. Why is pre-processing important in machine learning for NLP tasks?
a) To improvemodel interpretability
b) To ensure that the model has enough data
c) To reduce the complexity of the data
d) To ensure that data is in a format suitable for machine learning models

76. What does the "tokenization" process refer to in NLP?
a) Breaking down a sentence into its grammatical components
b) Converting text into numerical format
c) Breaking text into smaller units like words or phrases
d) Removing stopwords from the text

77. Which of the following techniques is used for identifying the base form of a word?
a) Lemmatization
b) Stemming
c) POS tagging
d) Tokenization

78. What is the purpose of Part-of-Speech (POS) tagging?
a) To identify the sentiment of a sentence
b) To break down text into sentences
c) To identify the grammatical role of each word in a sentence
d) To convert speech to text

79. Which of the following methods is used to convert continuous sequences of words into a fixed-size vector representation?
a) Word2Vec
b) One-Hot Encoding
c) GloVe
d) TF-IDF

80. Which of the following deep learning models is best for sequence-to-sequence tasks, like machine translation?
a) Transformer
b) LSTM
c) GAN
d) RNN

81. Which of the following is a key challenge in sequence modeling for NLP?
a) Handling out-of-vocabulary words
b) Balancing positive and negative samples
c) Feature extraction
d) Handling missing data

82. In Word2Vec, the "skip-gram" model tries to predict:
a) The surrounding words given a target word
b) A target word given the surrounding words
c) A word based on its position in a sentence
d) The grammatical role of a word in context

83. Which technique helps to represent a word as a dense vector of real numbers in NLP?
a) One-Hot Encoding
b) Word2Vec
c) N-grams
d) Bag of Words

84. The learning method in Word2Vec is based on:
a) Supervised learning
b) Reinforcement learning
c) Unsupervised learning
d) Semi-supervised learning

85. In a word embedding model, what is the purpose of a word embedding matrix?
a) To store vocabulary and their corresponding labels
b) To map words to their one-hot encoded vectors
c) To map words to continuous vector representations
d) To cluster words into predefined categories

86. Which of the following word representation models uses global co-occurrence statistics of words in a corpus?
a) Word2Vec
b) GloVe
c) One-Hot Encoding
d) FastText

87. Which word representation model is designed to handle out-of-vocabulary (OOV) words better? a) Word2Vec
b) GloVe
c) FastText
d) One-Hot Encoding

88. Which of the following is a common application of word embeddings?
a) Named-Entity Recognition
b) Sentiment Analysis
c) Tokenization
d) Word Sense Disambiguation

89. What is the key advantage of using embeddings over one-hot encoding?
a) Better accuracy in classification tasks
b) Reduced computational cost due to lower dimensionality
c) Better handling of synonyms and related words
d) Improved text summarization capabilities

90. Which of the following is used to measure the similarity between word embeddings?
a) Jaccard Similarity
b) Cosine Similarity
c) Euclidean Distance
d) Hamming Distance

91. In a sequence-to-sequence model, what is the main goal?
a) To predict the next word in a sentence
b) To translate one sequence of words into another sequence
c) To classify a document into categories
d) To generate word embeddings

92. Which of the following is a common architecture used for sequence-to-sequence tasks?
a) RNN
b) LSTM
c) GRU
d) All of the above

93. In a Sequence-to-Sequence model, the encoder processes:
a) The input sequence and encodes it into a fixed-size vector
b) The output sequence and decodes it
c) Both the input and output sequences simultaneously
d) None of the above

94. What is the main disadvantage of traditional sequence-to-sequence models?
a) They require an extremely large dataset
b) They are unable to handle long-term dependencies effectively
c) They cannot generate new sequences
d) They do not use embeddings for word representations

95. The Transformer model improves on sequence-to-sequence models by introducing:
a) Memory cells for storing long-term dependencies
b) A self-attention mechanism
c) Bidirectional encoding
d) Gated Recurrent Units (GRUs)

96. Which of the following is the main advantage of the Transformer model over traditional RNNs?
a) It processes sequences in parallel rather than sequentially
b) It uses less data for training
c) It can work without any labeled data
d) It is easier to train

97. Which mechanism in the Transformer model allows it to focus on different parts of the input sequence when processing each word?
a) Self-Attention
b) Backpropagation
c) Bidirectional Attention
d) Long Short-Term Memory

98. Which of the following is a key limitation of the Transformer model?
a) It cannot handle large datasets
b) It cannot handle long sequences efficiently
c) It requires a large amount of computational resources
d) It is not suitable for sequence-to-sequence tasks

99. Which of the following is the main advantage of BERT (Bidirectional Encoder Representations from Transformers)?
a) It only uses the decoder part of the Transformer
b) It is pre-trained on a large corpus and fine-tuned for downstream tasks
c) It generates a single vector for each input token
d) It performs better on smaller datasets

100. BERT's pre-training tasks involve which of the following?
a) Predicting the
next word in a sequence
b) Predicting missing words in a sequence
c) Generating word embeddings from scratch
d) Classifying sequences based on sentiment

101. Which of the following is a commonly used framework for deploying NLP models as web applications?
a) Flask
b) Django
c) Keras
d) PyTorch

102. In Flask, what is the primary function of routes in deploying an NLP model?
a) To train the model
b) To define endpoints for handling requests
c) To preprocess input data
d) To store model parameters

103. Which of the following is essential for deploying an NLP model in production environments?
a) Real-time inference capabilities
b) An offline database
c) A training pipeline

d) Labeling of large datasets

104. Which of the following is a common application of NLP in the healthcare industry?
a) Named-Entity Recognition for extracting medical terms
b) Predicting patient outcomes
c) Machine Translation for translating medical records
d) All of the above

105. Which of the following tools is primarily used for automatic speech recognition (ASR)?
a) Kaldi
b) SpaCy
c) NLTK
d) TensorFlow

106. What is the primary goal of Text-to-Speech (TTS) systems in NLP?
a) To convert written text into spoken words
b) To translate text into another language
c) To identify the emotional tone of a sentence
d) To identify named entities in speech

107. Which of the following is the most common approach for implementing machine translation in NLP?
a) Neural Machine Translation (NMT)
b) Rule-based Translation
c) Statistical Machine Translation
d) Both a and b

108. Which of the following is a primary goal of speech synthesis in NLP?
a) Identifying speakers
b) Converting text into natural-sounding speech
c) Analyzing the acoustic properties of speech
d) Translating speech into text

109. Which technology is commonly used for detecting the language of a given text?
a) Language Identification Models
b) Named-Entity Recognition
c) Part-of-Speech Tagging
d) Text Summarization

110. Which of the following tasks involves identifying the entities (e.g., names, dates) in a piece of text?
a**) Named-Entity Recognition (NER)**
b) Text Classification
c) Sentiment Analysis
d) Word Segmentation

111. Which of the following is a key component of automatic speech recognition (ASR)?
a) Signal processing
b) Speech segmentation
c) Acoustic modeling
d) All of the above

112. The technique used for representing phonetic sounds in written form is known as:
a**) Phonetic Transcription**
b) Speech Synthesis
c) Acoustic Modeling
d) Digital Signal Processing

113. Which technique is used in ASR to match speech signals with corresponding words?
a) Dynamic Time Warping
b) Hidden Markov Models
c) Convolutional Neural Networks
d) Support Vector Machines

114. Which of the following technologies is used to improve the clarity of speech in noisy environments?
a) Noise Filtering
b) Phonetic Transcription
c) Vowel Reduction
d) Speech Synthesis

115. Which process in speech synthesis converts linguistic input into prosody and timing?
a) Text Normalization
b) Phonetic Analysis
c) Prosodic Modeling
d) Speech Segmentation

116. Which is a common challenge in speech recognition systems?
a)Understanding multiple languages at once
b) Handling various accents and dialects
c) Synthesizing speech from text
d) All of the above

117. In speech synthesis, which approach focuses on constructing speech from pre-recorded human voices?
a) Concatenative synthesis
b) Articulatory synthesis
c) Parametric synthesis
d) Deep Learning-based synthesis

118. Which of the following is an important step in the Digital Signal Processing (DSP) for speech recognition?
a) Spectral analysis of speech signals
b) Lexical normalization of words
c) Sentence segmentation
d) Part-of-Speech tagging

119. The study of the physical production and perception of speech sounds is known as:
a) Acoustic Phonetics
b) Articulatory Phonetics
c) Computational Phonology
d) Digital Signal Processing

120. What is the purpose of speech recognition in NLP?
a) To generate human-like speech
b) To convert spoken language into text
c) To identify grammatical errors in spoken language
d) To identify emotions in spoken words

121. Which of the following is an important application of NLP in the e-commerce industry?
a) Chatbots for customer service
b) Machine Translation for product descriptions
c) Sentiment analysis for product reviews
d) All of the above

122. Which of the following technologies is used for Handwriting Recognition (HWR)?
a) Optical Character Recognition (OCR)
b) Deep Learning-based Recognition
c) Rule-based Recognition
d) Both a and b

123. Which Indian language script technology is used to process multiple languages in India?
a) Indic NLP
b) Devanagari Transliteration
c) Telugu Script Recognition
d) BERT for Indian Languages

124. Which of the following is a common use case of NLP in social media analysis?
a) Sentiment analysis of posts
b) Trend detection
c) Automated responses through chatbots
d) All of the above

125. Which of the following technologies enables automated correction of spelling and grammar in text?
a) Grammar Checkers
b) Spell Checkers
c) Part-of-Speech Tagging
d) Both a and b

126. What is the primary goal of machine translation?
a) To understand the sentiment of a text
b) To generate a translation of text from one language to another
c) To classify text into categories
d) To detect named entities in text

127. Which of the following is a critical challenge in Neural Machine Translation (NMT)?
a) Translating between languages with similar syntax
b) Handling idiomatic expressions in different languages
c) Overfitting to training data
d) Both a and b

128. Which of the following NLP tasks involves generating human-like conversations with users?
a) Machine Translation
b) Dialogue Systems
c) Named-Entity Recognition
d) Word Embedding

129. Which of the following approaches is used for improving machine translation accuracy?
a) Neural Machine Translation (NMT)
b) Rule-based Translation
c) Statistical Machine Translation
d) Both a and b

130. In which NLP application do we use Optical Character Recognition (OCR)?
a) Identifying the language of a document
b) Converting handwritten or printed text into machine-readable text
c) Translating text from one language to another
d) Analyzing sentiment in text

131. Which of the following is NOT typically a use case for NLP in healthcare?
a) Automatic patient record transcription
b) Diagnosing diseases based on medical images
c) Clinical decision support based on patient data
d) Text mining of electronic health records

132. In Natural Language Understanding (NLU), which task is aimed at determining the meaning of text in a specific context?
a) Sentiment Analysis
b) Text Classification
c) Named Entity Recognition
d) Semantic Role Labeling

133. Which of the following is an example of a named-entity in a text?
a) "The cat sat on the mat."
b) "Apple is located in Cupertino."
c) "Quickly running."
d) "There is a park in the city."

134. Which of the following approaches is used for extracting relevant information from unstructured text in NLP?
a) Information Retrieval
b) Information Extraction
c) Named-Entity Recognition
d) Text Classification

135. Which model is used for improving the generation of more coherent and accurate outputs in NLP tasks such as translation and summarization?
a) Recurrent Neural Networks (RNN)
b) Long Short-Term Memory (LSTM)
c**) Transformer**
d) Decision Trees

136. Which of the following NLP applications requires understanding the user's intent and context to generate a response?
a) Text Summarization
b) Machine Translation
c) Dialogue Systems
d) Named-Entity Recognition

137. In sentiment analysis, what is the typical goal of the analysis?
a) Classifying text as positive, negative, or neutral
b) Identifying key entities in a text
c) Translating text to another language
d) Summarizing the text content

138. Which of the following challenges is associated with speech recognition in noisy environments?
a) Phonetic confusion due to background noise
b) Difficulty in identifying the correct speaker
c) Processing different accents
d) All of the above

139. What is the primary benefit of neural machine translation (NMT) over statistical machine translation (SMT)?
a) NMT uses a rule-based approach for translations
b) NMT produces more fluent and natural translations
c) NMT does not require parallel corpora for training
d) NMT is faster in execution

140. Which of the following NLP tasks involves segmenting a large corpus of text into smaller, meaningful units?
a**) Tokenization**
b) Lemmatization
c) Word Sense Disambiguation
d) Named-Entity Recognition

141. Which of the following is a disadvantage of using traditional rule-based machine translation (RBMT)?
a) It requires a large amount of labeled data
b) It lacks flexibility and adaptability
c) It is computationally efficient
d) It works well for languages with complex grammar

142. Which of the following NLP tasks involves determining the most probable meaning of a word based on context?
a) Word Sense Disambiguation
b) Tokenization
c) Named-Entity Recognition
d) Sentiment Analysis

143. Which of the following tools is commonly used in NLP for performing tasks like POS tagging, dependency parsing, and named entity recognition?
a) TensorFlow
b) NLTK
c) Keras
d) Scikit-learn

144. Which of the following algorithms is commonly used in machine learning for text classification tasks?
a) K-Nearest Neighbors (KNN)
b) Support Vector Machines (SVM)
c) Random Forest
d) All of the above

145. In NLP, what is a "stop word"?
a) A word with no meaning
b) A word that is ignored during text preprocessing
c) A word that indicates the beginning of a sentence
d) A word with a strong emotional connotation

146. Which NLP technique is used to reduce words to their base or root form?
a) Stemming
b) Lemmatization
c) Tokenization
d) Chunking

147. Which of the following is the goal of automatic summarization in NLP?
a) To extract the most important parts of a text
b) To translate a text into a different language
c) To identify named entities in a text
d) To classify text into categories

148. Which of the following is a challenge faced by NLP in low-resource languages?
a) Lack of data for training models
b) Difficulty in identifying named entities
c) Complex sentence structures
d) All of the above

149. Which deep learning model is used for sequence modeling in NLP tasks?
a)Convolutional Neural Network (CNN)
b) Recurrent Neural Network (RNN)
c) Generative Adversarial Network (GAN)
d) Support Vector Machine (SVM)

150. Which of the following is a typical use case for neural networks in NLP?
a) Image classification
b) Sequence prediction in language generation
c) Facial recognition
d) Time-series forecasting