Topper's Solutions

....In Search of Another Topper

MCQ Edition

NATURAL LANGUAGE PROCESSING

(BE - COMPUTER)

8 SEM

As per Revised Syllabus w.e.f 2019-20

TOPPER'S SOLUTIONS

....In Search of Another Topper

There are many existing paper solution available in market, but Topper's Solution is the one which students will always prefer if they refer... ;) Topper's Solutions is not just paper solutions, it includes many other important questions which are important from examination point of view. Topper's Solutions are the solution written by the Toppers for the students to be the upcoming Topper of the Semester.

It has been said that "**Action Speaks Louder than Words**" So Topper's Solutions Team works on same principle. Diagrammatic representation of answer is considered to be easy & quicker to understand. So our major focus is on diagrams & representation how answers should be answered in examinations.

Why Topper's Solutions:

- Point wise answers which are easy to understand & remember.
- Diagrammatic representation for better understanding.
- Additional important questions from university exams point of view.
- Covers almost every important question.
- In search of another topper.

"Education is Free.... But its Technology used & Efforts utilized which we charge"

It takes lot of efforts for searching out each & every question and transforming it into Short & Simple Language. Entire Community is working out for betterment of students, do help us.

Thanks for Purchasing & Best Luck for Exams

BACKK BENCHERS

Handcrafted by BackkBenchers Community

“Have the courage to follow your **heart and intuition**.

They somehow know what **you truly want to become**.”

---- Steve Jobs.

Syllabus:

#	Module	Details Contents
1.	Introduction	History of NLP, Generic NLP system, levels of NLP , Knowledge in language processing , Ambiguity in Natural language , stages in NLP, challenges of NLP ,Applications of NLP
2.	Word Level Analysis	Morphology analysis –survey of English Morphology, Inflectional morphology & Derivational morphology, Lemmatization, Regular expression, finite automata, finite state transducers (FST) ,Morphological parsing with FST , Lexicon free FST Porter stemmer. N –Grams- N-gram language model, N-gram for spelling correction.
3.	Syntax analysis	Part-Of-Speech tagging(POS)- Tag set for English (Penn Treebank) , Rule based POS tagging, Stochastic POS tagging, Issues –Multiple tags & words, Unknown words. Introduction to CFG, Sequence labeling: Hidden Markov Model (HMM), Maximum Entropy, and Conditional Random Field (CRF).
4.	Semantic Analysis	Lexical Semantics, Attachment for fragment of English- sentences, noun phrases, Verb phrases, prepositional phrases, Relations among lexemes & their senses –Homonymy, Polysemy, Synonymy, Hyponymy, WordNet, Robust Word Sense Disambiguation (WSD) ,Dictionary based approach
5.	Pragmatics	Discourse –reference resolution, reference phenomenon , syntactic & semantic constraints on co reference
6.	Applications (preferably for Indian regional languages)	Machine translation, Information retrieval, Question answers system, categorization, summarization, sentiment analysis, Named Entity Recognition

Contact No: 7507531198

Email ID: Support@ToppersSolutions.com

Website: www.ToppersSolutions.com

Multiple Choice Questions (MCQ)

1. NLTK stands for _____.
   - **a. Natural Language Toolkit.**
   - b. Neutral Lingual Tool
   - c. Natural Language Tool
   - d. Neutral Language Toolkit

2. NLP is a subfield of _____.
   - **a. Artificial Intelligence**
   - b. Machine Learning
   - c. Deep Learning
   - d. None of Above

3. What is Sentiment Analysis?
   - a. Gathering data of emojis on social media posts.
   - b. None.
   - **c. recognizing the sentiment among several online posts and comments using NLP.**
   - d. recognizing the sentiment among several online posts and comments using NLTK.

4. Examples of NLP?
   - **a. Digital assistance, chatbots, Text summarization, text retrieval, sentiment analysis, translation etc.**
   - b. Clustering and differentiating patterns.
   - c. Deep Learning, Machine Learning, AI etc.
   - d. None of Above.

5. Likely, which languages can be used to work with NLP?
   - **a. Python & R language.**
   - b. JavaScript
   - c. Assembly
   - d. React Js.

6. When the first patents for "translating machines" were applied?
   - a. After 1945
   - **b. Mid 1930**
   - c. Mid 2000
   - d. Before 1930

7. Who discovered “Turing Test”?
   - **a. Alan Turing**
   - b. Venessa Turing

c. Leibniz
d. Descartes

8. NLP breaks down language into shorter, more basic pieces, called _____.
   a. Parameters
   b. Tokens.
   c. None.
   d. Arguments.

9. What are the components of NLP?
   a. Morphological and Lexical Analysis, Syntactic Analysis, Semantic Analysis, Discourse Integration, Pragmatic Analysis
   b. Only Morphological and Lexical Analysis.
   c. Only Semantic Analysis
   d. All of Above

10. What is Morphical and Lexical Analysis?
   a. It depicts analyzing, identifying and description of the structure of words. It includes dividing a text into paragraphs, words and the sentences.
   b. This component transfers linear sequences of words into structures.
   c. This only abstracts the dictionary meaning or the real meaning from the given context.
   d. All of Above.

11. Semantic Analysis means _____.
   a. It depicts analyzing, identifying and description of the structure of words. It includes dividing a text into paragraphs, words and the sentences.
   b. This component transfers linear sequences of words into structures. It shows how the words are associated with each other. And focuses only on the literal meaning of words, phrases, and sentences.
   c. deals with the overall communicative and social content.
   d. None of Above

12. What Pragmatic Analysis does?
   a. This component transfers linear sequences of words into structures.
   b. This only abstracts the dictionary meaning or the real meaning from the given context.
   c. This component transfers linear sequences of words into structures. It shows how the words are associated with each other.
   d. It deals with the overall communicative and social content and its effect on interpretation. It means abstracting or deriving the meaningful use of language in situations.

13. What is Syntax Analysis?
    a. This only abstracts the dictionary meaning or the real meaning from the given context.
    b. This component transfers linear sequences of words into structures. It shows how the words are associated with each other.
    c. It deals with the overall communicative and social content and its effect on interpretation. It means abstracting or deriving the meaningful use of language in situations.
    d. <u>It focuses about the proper ordering of words which can affect its meaning. This involves analysis of the words in a sentence by following the grammatical structure of the sentence. The words are transformed into the structure to show how the words are related to each other.</u>

14. Discourse Integration means _____.
    a. <u>It means a sense of the context. The meaning of any single sentence which depends upon those sentences. It also considers the meaning of the following sentence.</u>
    b. It depicts analyzing, identifying and description of the structure of words. It includes dividing a text into paragraphs, words and the sentences.
    c. This component transfers linear sequences of words into structures. It shows how the words are associated with each other. And focuses only on the literal meaning of words, phrases, and sentences.
    d. All of Above.

15. How to implement NLP?
    a. <u>Machine Learning & Statistical Inference.</u>
    b. Machine Learning & AI
    c. Deep Learning
    d. Python & R

16. What are the approaches of NLP?
    a. Morphological and Lexical Analysis,Syntactic Analysis, Semantic Analysis,Discourse Integration, Pragmatic Analysis
    b. <u>Symbolic, Statistical, Connectionist and Hybrid</u>
    c. Machine Learning, Deep Learning & AI
    d. None of These.

17. What Symbolic Approach performs?
    a. This component transfers linear sequences of words into structures. It shows how the words are associated with each other. And focuses only on the literal meaning of words, phrases, and sentences
    b. It harnesses various mathematical techniques and often uses large text corpora to develop approximately generalized models of linguistic phenomena based on actual examples.

c. <u>It performs extensive analysis of linguistic phenomena through explicit representation of facts about language and well-understood knowledge representation schemas and associated algorithms.</u>

d. It is based on the interconnection of networks having simple processing units with knowledge stored in weights to identify connections between units.

18. How does the Statistical Approach work?

a. <u>It uses statistical methods to resolve some of the difficulties in symbolic approach. It does this by harnessing various mathematical techniques and often using large text corpora to develop approximately generalized models of linguistic phenomena based on actual examples.</u>

b. It performs extensive analysis of linguistic phenomena through explicit representation of facts about language and well-understood knowledge representation schemas and associated algorithms.

c. It harnesses various mathematical techniques and often uses large text corpora to develop approximately generalized models of linguistic phenomena based on actual examples.

d. All of the above

19. Connectionist Approach is based on_____.

a. <u>The interconnection of networks having simple processing units with knowledge stored in weights to identify connections between units.</u>

b. It performs extensive analysis of linguistic phenomena through explicit representation of facts about language and well-understood knowledge representation schemas and associated algorithms.

c. It harnesses various mathematical techniques and often uses large text corpora to develop approximately generalized models of linguistic phenomena based on actual examples.

d. None of Above

20. Symbolic Approach is also called _____.

a. Convolutional Neural Networks.

b. <u>Rule based Approach.</u>

c. Corpus based.

d. Hybrid.

21. Statistical Approach is also called____.

a. <u>Corpus Based Approach.</u>

b. Rule Based Approach

c. CNN

d. K- nearest

22. Connectionist Approach is widely known as___.
    a. Statistical
    b. Symbolical
    c. <u>Neural Network</u>
    d. All of above

23. What kind of ambiguities are faced by NLP?
    a. Lexical and syntactical
    b. NLP does not face any ambiguity.
    c. semantical, discourse and Pragmatic.
    d. <u>Both a & c</u>

24. What is Lexical Ambiguity?
    a. <u>Ambiguity of a single word when it can be used as a verb, noun or an adjective.</u>
    b. Words having many meanings.
    c. Sentences and words are not aligned.
    d. All of the above.

25. What scope ambiguity involves?
    a. <u>Operators and quantifiers</u>
    b. Parameters and arguments
    c. Tokens
    d. None of Above.

26. When semantic ambiguity occurs?
    a. <u>when the meaning of the words themselves can be misinterpreted.</u>
    b. Words having many meanings.
    c. Both a & b
    d. None of the above.

27. What pragmatic ambiguity refers?
    a. <u>It refers to a situation where the context of a phrase gives it multiple interpretation</u>
    b. It refers to Statistical analysis
    c. It refers to only Misinterpreted words
    d. All of the above

28. What is corpus?
    a. A corpus is collection of Parameters and arguments
    b. **<u>A corpus is a large and structured set of machine-readable texts that have been produced in a natural communicative setting.</u>**
    c. It refers to a situation where the context of a phrase gives it multiple interpretation

d. All of the Above.

29. _____ depicts analyzing, identifying and description of the structure of words.
    a. Tokens
    b. Semantic Analysis
    c. Symbolic Analysis
    d. **Morphical And Lexical Analysis**

30. _____ includes dividing a text into paragraphs, words and the sentences.
    a. Morphological and Lexical Analysis
    b. Semantic Analysis
    c. Quantifiers
    d. None of the above.

31. _____ transfers linear sequences of words into structures.
    a. Semantic Analysis
    b. Tokens
    c. Lexical Analysis
    d. Discourse

32. _____ shows how the words are associated with each other.
    a. Semantic Analysis
    b. Tokens
    c. Lexical Analysis
    d. Discourse

33. _____ focuses only on the literal meaning of words, phrases, and sentences.
    a. Morphological and Lexical Analysis
    b. Semantic Analysis
    c. Quantifiers
    d. None of the above.

34. _____ deals with the overall communicative and social content and its effect on interpretation.
    a. Tokens
    b. Pragmatic Analysis
    c. Symbolic Analysis
    d. Morphical And Lexical Analysis

35. ____ means abstracting or deriving the meaningful use of language in situations.
    a. Semantic Analysis
    b. Tokens

c. Lexical Analysis

d. <u>Pragmatic Analysis</u>

36. It focuses about the proper ordering of words which can affect its meaning.

a. <u>Syntax Analysis</u>

b. Semantic Analysis

c. Lexical Analysis

d. Pragmatic Analysis

37. This involves analysis of the words in a sentence by following the grammatical structure of the sentence.

a. Tokens

b. Lexical Analysis

c. Discourse

d. <u>Syntax Analysis</u>

38. The words are transformed into the structure to show how the words are related to each other. This process is called as _____

a. <u>Syntax Analysis</u>

b. Semantic Analysis

c. Lexical Analysis

d. Pragmatic Analysis

39. ____means a sense of the context. The meaning of any single sentence which depends upon those sentences. It also considers the meaning of the following sentence.

a. <u>Discourse</u>

b. Semantic Analysis

c. Lexical Analysis

d. Pragmatic Analysis

40. Machine Learning & Statistical Inference are the popular methods for implementing___.

a. Lexical Analysis

b. Tokens and Quantifiers

c. <u>NLP</u>

d. None of the above.

41. It performs extensive analysis of linguistic phenomena through explicit representation of facts about language and well-understood knowledge representation schemas and associated algorithms. What is it?

a. Convolutional Neural Networks.

b. <u>Rule based Approach.</u>

c. Corpus based.
d. Hybrid.

42. It uses statistical methods to resolve some of the difficulties in symbolic approach. It does this by harnessing various mathematical techniques and often using large text corpora to develop approximately generalized models of linguistic phenomena based on actual examples.
    a. Convolutional Neural Networks.
    b. Rule based Approach.
    c. Corpus based.
    d. Statistical Approach

43. The interconnection of networks having simple processing units with knowledge stored in weights to identify connections between units.
    a. Connectionist Approach
    b. Neural Networks
    c. Hybrid approach
    d. Both a & b.

44. Rule Based Approach is also called _____
    a. Convolutional Neural Networks.
    b. Symbolic Approach.
    c. Corpus based.
    d. Hybrid.

45. Corpus Based Approach is also called _____
    a. Statistical Approach.
    b. Rule Based Approach
    c. CNN
    d. K- nearest

46. Neural Networks are also known as _____
    a. Statistical
    b. Symbolical
    c. Connectionist Approach
    d. All of above

47. Ambiguity of a single word when it can be used as a verb, noun or an adjective is called as _____
    a. Lexical Ambiguity
    b. Pragmatic Ambiguity
    c. Semantic Ambiguity
    d. None of These

48. Operators and Quantifiers are mostly responsible for _____
    a. Scope Ambiguity
    b. Pragmatic Ambiguity
    c. Semantic Ambiguity
    d. None of These

49. when the meaning of the words themselves can be misinterpreted then _____ ambiguity occurs.
    a. Scope Ambiguity
    b. Pragmatic Ambiguity
    c. Semantic Ambiguity
    d. None of These

50. What refers to a situation where the context of a phrase gives it multiple interpretations?
    a. Lexical Ambiguity
    b. Scope Ambiguity
    c. Semantic Ambiguity
    d. Pragmatic Ambiguity

51. What is Morphological Analysis?
    a. Morphological analysis is the process of providing grammatical information about the word on the basis of properties of the morpheme it contains.
    b. This component transfers linear sequences of words into structures.
    c. This only abstracts the dictionary meaning or the real meaning from the given context.
    d. All of Above.

52. What are the components of Morphological Analyzer acc., to Shrivastava et. al 2005?
    a. The recognition engine, identifying suffixes, and finding a stem within the input word algorithms
    b. Morpheme lexeme, Set of rules governing the spelling and composition of morphologically complex words & Decision algorithm
    c. The recognition engine, set of rules & Algorithm.
    d. All of the above.

53. Morphological analyzer is composed of the following Three parts according to Kumar, 2013?
    a. The recognition engine, identifying suffixes, and finding a stem within the input word algorithms
    b. Morpheme lexeme, Set of rules governing the spelling and composition of morphologically complex words & Decision algorithm
    c. The recognition engine, set of rules & Algorithm.
    d. All of the above.

54. Morphological analysis is also known as___.
    a. Sentiment Analysis
    b. Pragmatic Analysis
    c. CNN
    d. Lexical Analysis

55. What are the various methods of Morphological Analysis?
    a. Finite State Automata(FSA),Two level Morphology, Finite State Transducer(FST),
    b. Stemmer Algorithm, Corpus based Approach, DAWG(Directed Acyclic Word Graph)
    c. Paradigm based approach
    d. All of the above

56. What is FSA (Finite State Automata)?
    a. Finite state automaton is a model of behavior composed of state, transitions and actions.
    b. This consists of rules which map the two representations to eachother.Each rule is described through a finite-state transducer
    c. It takes raw corpus as input and produces a segmentation of the word forms observed in the text.
    d. None of the above

57. _____ is a model of behaviour composed of state, transition and actions.
    a. FST
    b. FSA
    c. DAWG
    d. Stemmer Algorithm

58. The lexical representation of a word-form is also called as_____.
    a. Morphophonemic
    b. Morphotacties
    c. Both a & b
    d. None of the above

59. _____has given a computational model of two-level morphology for word-form recognition and generation in his dissertation in 1983.
    a. Kumar
    b. Shrivastava
    c. Kimmo Koskenniemi
    d. Chomsky

60. The two-level morphology model consists of two representations and one rule. What are they?
    a. The surface representation of a word-form, The lexical representation of a word-form, rules which map the two representations.

b. The surface representation of a word-form, The statistical representation of a word-form, rules which map the two representations.
c. The surface representation of a word-form, The lexical representation of a word-form, rules which isolate the two representations.
d. None of the above

61. The surface representation of a word-form, The lexical representation of a word-form, rules which map the two representations refers to _____.
a. <u>Two-level morphology model</u>
b. Chomsky Model
c. Finite State Automata
d. Both a & c

62. _____is an advanced version of FSA(finite state automata)and is used to represent the lexicon computationally.
a. <u>FST</u>
b. FSA
c. DAWG
d. Stemmer Algorithm

63. Finite State Transducer is an advanced version of _____ and is used to represent the lexicon computationally.
a. FST
b. <u>FSA</u>
c. DAWG
d. Stemmer Algorithm

64. In FST what components are used to build morphological analyzers?
a. <u>lexicon, orthographic rules and spelling variations</u>
b. The surface representation of a word-form, The statistical representation of a word-form, rules which map the two representations.
c. Lexicon, The statistical representation & spelling variations
d. None of the above.

65. lexicon, orthographic rules and spelling variations are the components of _____.
a. Stemmer Algorithm
b. Two-level morphology
c. <u>FST</u>
d. FSA

66. An FST is simply a classical finite-state automaton whose transitions are ____ , rather than with single symbols.
    a. Labeled with tokens
    b. labeled with pairs
    c. Not labeled
    d. Both a & b

67. What is stemmer?
    a. This consists of rules which map the two representations to eachother.Each rule is described through a finite-state transducer
    b. It takes raw corpus as input and produces a segmentation of the word forms observed in the text.
    c. It is used for stripping of affixes. It uses a set of rules containing a list of stems and replacement rules.
    d. All of the above

68. The most widely used stemmer algorithm is _____.
    a. Potter Algorithm
    b. Porter Algorithm
    c. Decision Algorithm
    d. Both a & b

69. _____ used for stripping of affixes. It uses a set of rules containing a list of stems and replacement rules.
    a. Two-level morphology model
    b. Chomsky Model
    c. Finite State Automata
    d. Stemmer

70. What is corpus?
    a. A corpus is collection of Parameters and arguments
    b. Corpus is a large collection of written text belonging to a particular language
    c. It refers to a situation where the context of a phrase gives it multiple interpretation
    d. All of the Above.

71. DAWG stands for _____.
    a. Directed Acyclic World Graph
    b. Directed Acyclic Word Graph
    c. Directed Acrylic Word Graph
    d. Diverse Acyclic Word Graph

72. DAWG is ___.
    a. <u>It is a very efficient data structure for lexicon representation and fast string matching with a great variety of applications.</u>
    b. It is a very efficient data structure.
    c. It is used for representation.
    d. All of above

73. A ____ defines all the word forms of a given stem and also provides a feature structure with every word form. The _____ is efficient for inflectional rich languages.
    a. <u>Paradigm, paradigm-based approach</u>
    b. Paradigm based approach, paradigm
    c. Paradigm, rule-based Approach
    d. None of the above

74. The ANUSAARAKA research group has developed a language independent paradigm based morphological compiler program for ____. This or a variant of this scheme has been used widely in NLP.
    a. European Languages
    b. English Languages
    c. <u>Indian Languages</u>
    d. All of the above

75. The _____research group has developed a language independent paradigm based morphological compiler program for Indian languages. This or a variant of this scheme has been used widely in NLP.
    a. <u>ANUSAARAKA</u>
    b. Anusarka
    c. ANUSARKAR
    d. All of the above

76. Morphology is divided into two branches:
    a. Direct, Indirect
    b. Inflectional
    c. Derivational
    d. <u>Both b & c</u>

77. _____is a change in word form. This usually means the use of endings. For eg., He work*s*, he work*ed*, and he is working.
    a. Derivational Morphology
    b. <u>Inflectional Morphology</u>
    c. Both a & b
    d. None of the above

78. _____ creates new words. For example, *beauty* becomes *beautiful*. The affix *-ful* changes the word from a noun to an adjective.
    a. <u>Derivational Morphology</u>
    b. Inflectional Morphology
    c. Both a & b
    d. None of the above

79. _____ morphology uses many more affixes than ____morphology.
    a. Inflectional, derivational
    b. <u>Derivational, inflectional</u>
    c. Direct, Indirect
    d. None of the above

80. NLTK have following stemming classes:
    a. Porter Stemmer, Lancaster Stemmer
    b. Regexp Stemmer
    c. Snowball Stemmer
    d. <u>All of the above</u>

81. In _____, the words are replaced by the root words or the words with similar context.
    E.g.- Walking will be replaced by Walk(walk is the root word of walking)
    a. Stemming
    b. <u>Lemmatization</u>
    c. Both a & b
    d. None of the above

82. ____are created by removing the suffixes or prefixes used with a word. This process is called as____
    a. <u>Stems, Stemming</u>
    b. Lemma, Lemmatization
    c. Corpus
    d. Suffix stripping

83. NLTK requires Python versions above__.
    a. <u>2.7</u>
    b. 3.8
    c. 3.0
    d. 2.0

84. A _____, often called a pattern, specifies a set of strings required for a particular purpose. A simple way to specify a finite set of strings is to list its elements or members.
    a. <u>Regular Expression</u>

b. Non regular Expression
c. Finite Automata
d. None of the above

85. a|b* denotes
a. {ε, "a", "b", "aa", "ab", "ba", "bb", "aaa", ...}
b. {ε, "a", "b", "bb", "bbb", ...}
c. {"a", "ac", "ab", "abc", "abb", "abbc", ...}
d. None of the above

86. (a|b)* denotes
a. {ε, "a", "b", "aa", "ab", "ba", "bb", "aaa", ...}
b. {ε, "a", "b", "bb", "bbb", ...}
c. {"a", "ac", "ab", "abc", "abb", "abbc", ...}
d. None of the above

87. ab*(c|ε) denotes
a. {ε, "a", "b", "aa", "ab", "ba", "bb", "aaa", ...}
b. {ε, "a", "b", "bb", "bbb", ...}
c. {"a", "ac", "ab", "abc", "abb", "abbc", ...}
d. None of the above

88. (aa)*(bb)*b denotes
a. {b, aab, aabbb, aabbbbb, aaaab, aaaabbb, ...}
b. {ε, "a", "b", "aa", "ab", "ba", "bb", "aaa", ...}
c. {"a", "ac", "ab", "abc", "abb", "abbc", ...}
d. None of the above

89. (0 + 10*) denotes
a. {1, 01, 10, 010, 0010, ...}
b. { 0, 1, 10, 100, 1000, 10000, ... }
c. {ε, 0, 1, 01}
d. None of the above

90. ____is the simplest machine to recognize patterns. It has a set of states and rules for moving from one state to another but it depends upon the applied input symbol.
a. Finite Automata
b. DFA
c. NFA
d. None of the above

91. What is Q in the finite automata set?
    - **a. <u>Finite set of states.</u>**
    - b. set of Input Symbols.
    - c. Initial state.
    - d. set of Final States.

92. Σ denotes in Finite Automata set as:
    - a. Finite set of states.
    - **b. <u>set of Input Symbols.</u>**
    - c. Initial state.
    - d. set of Final States.

93. FA is characterized into two types & they are:
    - a. DA, NA
    - **b. <u>NFA, DFA</u>**
    - c. Both a & b
    - d. None of these

94. In ___null (or ε) move is not allowed.
    - a. NFA
    - **b. <u>DFA</u>**
    - c. Cyclic graphs
    - d. None of the above

95. In ____null (or ε) move is allowed i.e., it can move forward without reading symbols.
    - **a. <u>NFA</u>**
    - b. DFA
    - c. Cyclic graphs
    - d. None of the above

96. What FS transducers do?
    - **a. <u>It is a finite state automaton which produces output as well as reading input, it is useful for parsing.</u>**
    - b. It is the simplest machine to recognize patterns.
    - c. A simple way to specify a finite set of strings is to list its elements or members.
    - d. All of the above

97. Transducers work in 4 modes. What are they?
    - a. generation mode,recognition mode,
    - b. translation mode(left to right), translation mode(right to left)
    - **c. <u>Both a & b</u>**

d. None of the above

98. What is generation mode?
    a. It accepts when the word on the first tape consists of exactly as many as as the word on the second tape consists of bs.
    b. It reads as from the first tape and writes an b for every a that it reads onto the second tape.
    c. <u>It writes a string on one tape and a string on the other tape. Both strings have the same length.</u>
    d. It reads bs from the second tape and writes an a for every f that it reads onto the first tape.

99. What recognition mode does?
    a. <u>It accepts when the word on the first tape consists of exactly as many as the word on the second tape consists of.</u>
    b. It reads as from the first tape and writes an b for every a that it reads onto the second tape.
    c. It writes a string of as on one tape and a string bs on the other tape. Both strings have the same length.
    d. It reads bs from the second tape and writes an a for every f that it reads onto the first tape.

100. What are the advantages of Porter stemmer algorithm?
    a. Produces the best output as compared to other stemmers.Less error rate.
    b. Compared to Lovins it's a light stemmer.
    c. The Snowball stemmer framework designed by Porter is language independent approach to stemming
    d. <u>All of the above.</u>

101. The _____ algorithm is a process of removing suffixes from words in English.
    a. Lovins Stemmer
    b. <u>Porter Stemmer</u>
    c. paice/ Husk Stemmer
    d. None of the above

102. Disadvantages of Porter Stemming algorithm are:
    a. The stems produced are not always real words.
    b. It has at least five steps and sixty rules and hence is time consuming
    c. It writes a string on one tape and a string on the other tape. Both strings have the same length.
    d. <u>Both a & b</u>

103. The _____ is a very interesting method and it is language independent.
    a. Lovins Stemmer
    b. Porter Stemmer

c. paice/ Husk Stemmer

d. N-Gram Stemmer

104. N-Gram stemmer has following advantages:

a. Based on the concept of n-grams and string comparisons.

b. Language independent.

c. Both a & b

d. None of the above

105. Limitations of N-Gram Stemmer are:

a. Not time efficient.

b. Requires a significant amount of space for creating and indexing the n-grams.

c. Not a very practical method

d. All of the above

106. _____ is a vocabulary, a list of words, a dictionary.

a. Corpus

b. WordNet

c. Lexicon

d. None of the above

107. _____ is a lexical database for the English language.

a. Corpus

b. WordNet

c. Lexicon

d. None of the above

108. _____ basically, means a body, and in the context of Natural Language Processing (NLP), it means a body of text.

a. Corpus

b. WordNet

c. Lexicon

d. None of the above

109. What are the applications of FST?

a. Word Inflections. For example, pluralizing words (cat -> cats)

b. Morphological Parsing; i.e., extracting the “properties” of a word (e.g., computers -> computer + [Noun] + [Plural])

c. Simple Word Translation, e.g., translating US English to UK English

d. **All of the above**.

110. _____are general rules used when breaking a word into its stem and modifiers. An example would be: singular English words ending with -y, when pluralized, end with -ies.
    a. Morphological Rules
    b. Orthographic Rules
    c. Parsing
    d. None of the above

111. Types of stemming errors are:
    a. under stemming
    b. over stemming
    c. Mid stemming
    d. Both a and b

112. ____ is when two words with different stems are stemmed to the same root. This is also known as a _____.
    a. Under Stemming, False Positive
    b. Over Stemming, False Positive
    c. Under Stemming, False Negative
    d. Over Stemming, False Negative.

113. _____ is when two words that should be stemmed to the same root are not. This is also known as a____.
    a. Under Stemming, False Positive
    b. Over Stemming, False Positive
    c. Under Stemming, False Negative
    d. Over Stemming, False Negative.

114. It is possible to use backtracking in ____.
    a. NFA
    b. DFA
    c. DAG
    d. FA

115. It is not possible to use backtracking at all times in the case of _____.
    a. NFA
    b. DFA
    c. DAG
    d. FA

116. The full form of DFA is____.
    a. Derived Finite Automata

b. Deterministic Final Automata
c. <u>Deterministic Finite Automata</u>
d. Duplicate Final Automation

117. The full form of NFA is ____.
a. Non-derived Finite Automata
b. Non-deterministic Final Automata
c. <u>Non-deterministic Finite Automata</u>
d. Non-Duplicate Final Automation

118. The generally accepted approach to morphological parsing is through the use of a _____, which inputs words and outputs their stem and modifiers.
a. Non-deterministic Finite Automata(NFA)
b. Deterministic Finite Automata(DFA)
c. Directed Acyclic Graph(DAG)
d. <u>Finite StateTransducer(FST)</u>

119. FST as recognizer:
a. a machine that outputs pairs of strings of the language. Thus the output is a yes or no, and a pair of output strings.
b. A machine that reads a string and outputs another string.
c. <u>a transducer that takes a pair of strings as input and output accept if the string-pair is in the string-pair language, and a reject if it is not</u>
d. A machine that computes relation between sets

120. FST as generator:
a. <u>a machine that outputs pairs of strings of the language. Thus the output is a yes or no, and a pair of output strings.</u>
b. A machine that reads a string and outputs another string.
c. a transducer that takes a pair of strings as input and output accept if the string-pair is in the string-pair language, and a reject if it is not
d. A machine that computes relation between sets

121. FST as transducer:
a. a machine that outputs pairs of strings of the language. Thus the output is a yes or no, and a pair of output strings.
b. <u>A machine that reads a string and outputs another string.</u>
c. a transducer that takes a pair of strings as input and output accept if the string-pair is in the string-pair language, and a reject if it is not
d. A machine that computes relation between sets

122. FST as set relator:
    a. a machine that outputs pairs of strings of the language. Thus the output is a yes or no, and a pair of output strings.
    b. A machine that reads a string and outputs another string.
    c. a transducer that takes a pair of strings as input and output accept if the string-pair is in the string-pair language, and a reject if it is not
    d. <u>A machine that computes relation between sets.</u>

123. To construct a regular expression from a DFA, we replace each state in the ___one by one with a corresponding ___.
    a. NFA, regular expression
    b. <u>DFA, regular expression</u>
    c. Regular expression, DFA
    d. None of the above

124. If we can eliminate _____ from an FA, then our construction of an FA from a regular expression can be completed.
    a. <u>Epsilon transitions</u>
    b. States
    c. DFA
    d. NFA

125. In the fields of computational linguistics and probability, an_____ is a contiguous sequence of *n* items from a given sample of text or speech. The items can be phonemes, syllables, letters, words or base pairs according to the application.
    a. corpus
    b. Epsilon Transitions
    c. <u>N-grams</u>
    d. Lemma

126. Applications of n- gram:
    a. design kernels that allow machine learning algorithms such as support vector machines to learn from string data
    b. find likely candidates for the correct spelling of a misspelled word
    c. improve compression in compression algorithms where a small area of data requires *n*-grams of greater length
    d. <u>All of the above</u>

127. Syntactic *n*-grams are *n*-grams defined by paths in _____ dependency or constituent trees rather than the linear structure of the text.
    a. <u>Syntactic</u>

b. Semantic
c. Symbolic
d. None of the above

128. NLP is a subfield of _____, computer science, and artificial intelligence concerned with the interactions between computers and human language
a. Definitions
b. Texts
c. Contexts
d. Linguistics

129. NLP does not involves in-
a. Speech recognition
b. Language understanding
c. Language generation
d. Computer vision

130. Which is the method of NLP
a. Neural
b. Statistical
c. Symbolic
d. All of the above

131. Which is not NLP task?
a. Discourse
b. Morphological analysis
c. Object recognition
d. Relational semantics

132. Which is the goal of NLP?
a. Content generation
b. Machine translation
c. Paraphrasing
d. All of the above

133. Where NLP is not used?
a. Chat bots
b. Image/Video captioning
c. Language translator
d. Predictive analysis

134. What input we can process with NLP?
    a. Audio
    b. Video
    c. Text
    d. <u>All of the above</u>

135. Which is not a level of NLP process?
    a. Pragmatic
    b. Discourse
    c. Morphological
    d. <u>Textual</u>

136. Which difficulty occurs in NLP?
    a. Referential Ambiguity
    b. Lexical Ambiguity
    c. <u>Contextual Ambiguity</u>
    d. Syntax level Ambiguity

137. Which is not application of NLP?
    a. <u>OCR</u>
    b. Sentiment analysis
    c. Text classification
    d. Auto-correct

138. How many steps of NLP is there?
    a. 4
    b. <u>5</u>
    c. 2
    d. 6

139. _____ is the step in which an input sentence is converted into a hierarchical structure that corresponds to the units of meaning in the sentence.
    a. Graph Processing
    b. Syntactic Processing
    c. <u>Semantic Processing</u>
    d. All of the mentioned

140. Choose form the following areas where NLP can be useful.
    a. Information Retrieval
    b. Automatic Text Summarization
    c. Automatic Question-Answering Systems

d. All of the mentioned

141. What is the main challenge/s of NLP?
   a. Handling Ambiguity of Sentences
   b. Handling Tokenization
   c. Handling POS-Tagging
   d. All of the mentioned

142. What is Machine Translation
   a. Converts one human language to another
   b. Converts human language to machine language
   c. Converts any human language to English
   d. Converts Machine language to human language

143. In linguistic morphology _____ is the process for reducing inflected words to their root form.
   a. Rooting
   b. Stemming
   c. Text-Proofing
   d. Both Rooting & Stemming

144. Which is not a step in NLP?
   a. Lexical analysis
   b. Syntactic analysis
   c. Word analysis
   d. Pragmatic Analysis

145. Which of the following is demerits of Top-Down Parser?
   a. It is hard to implement.
   b. Slow speed
   c. inefficient
   d. Both B and C

146. Given a sound clip of a person or people speaking, determine the textual representation of the speech.
   a. Text-to-speech
   b. Speech-to-text
   c. Both A and B
   d. None of the Above

147. Which of the following is used to mapping sentence plan into sentence structure?
   a. Text planning

b. Sentence planning

c. <u>Text Realization</u>

d. None of the Above

148. Which is not a knowledge type of language processing?

a. Discourse

b. Pattern

c. <u>Symbol</u>

d. World

149. In morphology, we care about the _____ that make up the sentence

a. Characters

b. <u>Words</u>

c. Symbols

d. Lexicons

150. Which is not an example of morphology?

a. Prefix/suffix

b. Singularization/Pluralization

c. Lemmatization

d. <u>Word detection</u>

151. In parsing stage, we focus more on the _____ of the words within a sentence

a. Sequence

b. Group

c. <u>Relationship</u>

d. None of the above

152. Syntactical analysis is done at _____ level

a. <u>Sentence</u>

b. Word

c. Lexicon

d. Symbol

153. Morphological analysis is done at _____ level

a. Character

b. Lexicon

c. <u>Word</u>

d. Sentence

154. Morphemes that cannot stand alone and are typically attached to another to become a meaningful word is called,
    a. Free morphemes
    b. Bound morphemes
    c. Derived morphemes
    d. Lexical morphemes

155. Morphotactics is a model of
    a. Spelling modifications that may occur during affixation
    b. How and which morphemes can be affixed to a stem
    c. All affixes in the English language
    d. Ngrams of affixes and stems

156. Natural Language Processing, or NLP for short, is broadly defined as the automatic _____ of natural language, like speech and text, by software.
    a. Conversion
    b. Manipulation
    c. Correction
    d. None of the above

157. NLP is originated from the idea of _____ which came to existence during the second world war.
    a. Machine translation
    b. Turing machine
    c. Machine Intelligence
    d. None of the above

158. What is full form of NLG?
    a. Natural Language Generation
    b. Natural Language Genes
    c. Natural Language Growth
    d. Natural Language Generator

159. In linguistic morphology _____ is the process for reducing inflected words to their root form.
    a. Rooting
    b. Stemming
    c. Text-Proofing
    d. Both Rooting & Stemming

160. Which of the following is used study of construction of words from primitive meaningful units?
    a. Phonology
    b. Morphology
    c. Morpheme

d. Shonology

161. What is Morphological Segmentation?
    a. Does Discourse Analysis
    b. is an extension of propositional logic
    c. Separate words into individual morphemes and identify the class of the morphemes
    d. None of the Above

162. Word level anlysis helps to find _____ resolution according to the context
    a. Ambiguity
    b. Text
    c. Sentence
    d. None of the above

163. Word level analysis helps in _____ of given text
    a. Understanding
    b. Spellchecking
    c. Reducing size
    d. All of the above

164. _____ is the smallest part of a word
    a. Letter
    b. Morpheme
    c. Word sketches
    d. Verbs

165. The morphological level of linguistic processing deals with the study of word _____ and word _____
    a. Structure & Formation
    b. Joining & Cutting
    c. Equivalence & Opposite
    d. None of the above

166. Which is the Morpheme from following?
    a. Prefix
    b. Stem
    c. Root
    d. All of the above

167. Which is not used in Word level Analysis?
    a. Regular Expression

b. Decidability and Countability

c. Regular Grammar

d. Finite Automata

168. Which is not a approach of Morphology?

a. Morpheme based

b. Lexeme based

c. context based

d. Word based

169. In Morpheme based morphology, Word forms are analysed as _____ of morphemes?

a. Arrangements

b. Groups

c. Replacements

d. All of the above

170. Lexeme based morphology follows _____ and _____ approach.

a. Text & words

b. Letters & words

c. Item & Process

d. None of the above

171. Word based Morphology _____ that hold between the forms of inflectional paradigms

a. Generalizations

b. Separates

c. Groups

d. Categorised

172. Stemming is basically removing the _____ from a word and reduce it to its root word

a. Suffix

b. Root

a. Prefix

b. All of the above

173. Over-stemming is when two _____ with different stems are stemmed to the same root

a. Roots

b. Prefix

c. Suffix

d. Words

174. Lemmatization usually refers to doing things properly with the use of a _____ and _____ analysis of words
    a. Groups & Section
    b. Letter & Context
    c. Vobulary & Morphological
    d. None of the above

175. regular expression is _____ given to a function on what and how to match or replace a set of strings
    a. Word
    b. Instruction
    c. Limit
    d. Context

176. Which is not regex command from following?
    a. Strsplit ()
    b. Strjoin()
    c. Grep()
    d. Gsub()

177. _____ allows to mention and have control over how many times specific character(s) pattern should occur in the given text
    a. Selectors
    b. Commands
    c. Quantifiers
    d. None of the above

178. finite state automation is a model of behavior composed of state, _____ and _____
    a. Transitions & Actions
    b. Actions & Words
    c. Sequences & Sets
    d. None of the above

179. In Finite Automaton, transitions is process of _____ over from one state to another state.
    a. Adding
    b. Switching
    c. Merging
    d. Subtracting

180. To build Morphological Parser we need
    a. Lexicon
    b. Morphotactics

c. Orthographic rules

d. <u>All of the above</u>

181. Lexicon is repository of _____

a. <u>Words</u>

b. Letters

c. Sets

d. None of the above

182. A transducer maps between FST one set of _____ and another.

a. Words

b. Letters

c. <u>Symbols</u>

d. Sentences

183. Finite State Transducers is used to represent the _____ computationally

a. Context

b. Syntax

c. Morpheme

d. <u>Lexicon</u>

184. Porter stemmer is a process for removing the commoner _____ and _____ endings from words in English

a. Issue & complex

b. Words & eased

c. <u>Morphological & Inflexional</u>

d. None of the above

185. Porter Stemmer algorithm is part of a term _____ process

a. Standardization

b. Reduction

c. <u>Normalisation</u>

d. Cancellation

186. An n-gram is a contiguous _____ of n items from a given sample of text or speech

a. Set

b. Class

c. <u>Sequence</u>

d. None of the above

187. N-gram is used with

a. Turing machine
b. Finite Automata
c. <u>Markov Model</u>
d. All of the above

188. N-gram can improve _____ of auto completion system
a. Completion
b. <u>Prediction</u>
c. Understanding
d. None of the above

189. N-gram uses Maximum likelihood _____ to estimate probability
a. Completion
b. Addition
c. Reduction
d. <u>Estimation</u>

190. Which one is stemming algorithm?
a. Porter algorithm
b. Dawson algorithm
c. N-gram algorithm
d. <u>All of the above</u>

191. Orthographic rules are general rules used when breaking a word into its _____ and _____.
a. Letters, classifiers
b. Symbols, Quantifiers
c. <u>Stems, Modifiers</u>
d. None of the above

192. Derivational morphology changes both the meaning and the content of a listeme, while inflectional morphology doesn't change the meaning, but changes the function.
a. <u>Meaning, Content, Listeme, function</u>
b. Meaning, Listeme, content, function
c. Content, meaning, function, Listeme
d. Function, content, meaning, Listeme

193. Syntactic analysis or parsing may be defined as the process of _____ the _____ of symbols in natural language conforming to the rules of formal grammar
a. <u>Analyzing & Strings</u>
b. Defining & groups
c. Reducing & arrays

d. Reviewing & letters

194. Which is not role of parser?
   a. Report any syntax error
   b. Create parse tree
   c. Create symbol table
   d. **Correct any syntax error**

195. In Top-down parsing, the parser starts constructing the parse tree from the _____.
   a. Start letter
   b. Input symbol
   c. Start symbol
   d. End letter

196. In Bottom-up parsing, the parser starts constructing the parse tree from the _____.
   a. Input symbol
   b. Start symbol
   c. End letter
   d. Start letter

197. POS is the process of marking up a word in a text as corresponding to a particular part of speech, based on both its _____ and its _____
   a. Sets & meaning
   b. Definition & contexts
   c. Analysis & reporting
   d. Sets & definitions

198. Rule-based taggers use dictionary or _____ for getting possible tags for tagging each word.
   a. Set
   b. Array
   c. Lexicon
   d. Word

199. Following property is of - These taggers are knowledge-driven taggers.
   a. Rule based tagging
   b. Stochastic tagging
   c. All of the above
   d. None of the above

200. The model that includes _____ or probability (statistics) can be called stochastic
   a. **Frequency**

b. Action
c. Complexity
d. None of the above

201. Following property is of - This POS tagging is based on the probability of tag occurring.
a. Rule based tagging
b. **Stochastic tagging**
c. All of the above
d. None of the above

202. Transformation based tagging is _____ algorithm for automatic tagging
a. Information based
b. **Rule based**
c. Action based
d. Group based

203. Which is the step taken in Transformation based tagging?
a. Start with solution
b. Choosing most beneficial transformation
c. Applying to problem
d. **All of the above**

204. The Penn Treebank, or PTB for short, is a dataset maintained by the _____
a. **University of Pennsylvania**
b. Stanford University
c. Harward University
d. UC Berkley University

205. The job of a POS tagger is to resolve this _____ accurately based on the context of use.
a. Complexity
b. Simplicity
c. **Ambiguity**
d. Quality

206. In English, many common words have multiple meanings and therefore multiple POS.
a. Simple meanings
b. Complex meanings
c. **Multiple meanings**
d. None of the above

207. A context-free grammar (CFG) is a list of rules that define the set of all well-formed sentences in a language.
    a. List of actions
    b. **List of Rules**
    c. List of variables
    d. None of the above

208. In CFG, Each rule has a _____ side
    a. Right hand
    b. Left hand
    c. **Option A and B**
    d. None of the above

209. In CFG, Left hand side identifies _____ and right hand side defines _____
    a. **Syntactic Categories and component parts**
    b. Complex categories and Summation parts
    c. Simple Categories and Aggregate parts
    d. None of the above

210. subcategorization denotes the _____for lexical items (usually verbs) to require/allow the presence and types of the syntactic arguments with which they co-occur
    a. **Ability**
    b. Simplicity
    c. Complexity
    d. All of the above

211. Sequence labeling is a typical NLP task which assigns a _____ or _____ to each _____ in a given input sequence.
    a. Attribute, value, Symbol
    b. **Class, Label, token**
    c. Value, Class, Character
    d. None of the above

212. Types of Sequence labeling
    a. Token
    b. Span
    c. Raw
    d. **Option A and B**

213. HMMs are “a statistical Markov model in which the system being modeled is assumed to be a _____process with _____ states”.
    a. Convolution, completed
    b. **Markov, Unobservable**
    c. Analyzing, Categorized
    d. Complete, Observed

214. HMM are designed to model the joint distribution P(H , O) , where H is the hidden state and O is the observed state.
    a. **Hidden, Observed**
    b. Unobservable, Hidden
    c. Classified, Completed
    d. Open, Completed

215. HMM graphs consist of a Hidden Space and Observed Space, where the hidden space consists of the _____ and the observed space is the _____.
    a. Input, Categories
    b. Values, Variables
    c. **Labels, Input**
    d. Variables, Values

216. HMMs are limited to only _____ states and only take into account the last known _____.
    a. Complete, Value
    b. Unobserved, Variable
    c. Hidden, Attribute
    d. **Discrete, State**

217. Maximum Entropy Markov Models use a maximum entropy _____for _____ and local _____.
    a. **Framework, Features, Normalization**
    b. Rules, Variables, Classification
    c. Sets, Values, Distribution
    d. None of the above

218. In the context of POS tagging, the objective would be to build an HMM to model P(____ | ___) and compute the label probabilities given observations using _____ Rule.
    a. Value, Label, Markov
    b. **Word, Tag, Bayes**
    c. Attribute, Variable, Bayes
    d. Input, Label, Markov

219. In HMMs, spaces are connected via _____ matrices {T,A} to represent the probability of _____ from one state to another following their _____.
    a. **Transitions, Transitioning, Connections**
    b. Attribute, Changing, groups
    c. Label, moving, sets
    d. None of the above

220. Each connection in HMM represents a _____ over possible options; given our _____, this results in a large search space of the _____ of all words given the tag.
    a. Value, variables, associativity
    b. **Distribution, tags, probability**
    c. Variable, Labels, Transitivity
    d. Object, groups, associativity

221. In question answering and search tasks, we can use spans as entities to specify our search query
    a. Spans, Entities
    b. **Classes, Objects**
    c. Sequences, Variables
    d. All of the above

222. The label bias problem was introduced due to MEMMs applying local normalization.
    a. Value bias, Global
    b. Attribute bias, local
    c. Set bias, global
    d. **Label bias, normalization**

223. This often leads to the model getting stuck in _____ during _____.
    a. Global maxima, encoding
    b. Global minima, decoding
    c. Local maxima, Encoding
    d. **Local minima, decoding**

224. The _____ minima trap occurs because the overall model favors _____ with the least amount of transitions.
    a. Local, groups
    b. **Local, nodes**
    c. Global, nodes
    d. Global, groups

225. Semantic analysis is the process of understanding the _____ and _____ of words, signs and sentence structure
    a. Using, describing
    b. Meaning, interpretation
    c. Complexity, Usability
    d. None of the above

226. Elements of Semantic analysis
    a. Hyponymy
    b. Homonymy
    c. Polysemy
    d. All of the above

227. In homonymy, the meanings of the words are not related
    a. Joined
    b. **Related**
    c. Analysed
    d. Changed

228. word sense disambiguation (WSD) is the problem of determining which "sense" (meaning) of a word is activated by the use of the word in a particular _____.
    a. Classification
    b. Analysis
    c. Context
    d. Usability

229. Which is not method of WSD?
    a. Supervised learning
    b. Dictionary method
    c. Unsupervised learning
    d. Sem-supervised learning

230. Which algorithm is for WSD?
    a. Ceiling & Most Frequent Sense
    b. Simplified Lesk & Corpus Lesk
    c. Bootstrapping
    d. All of the above

231. WordNet is the _____ database
    a. Symbol
    b. Word
    c. Lexical
    d. Annotation

232. WordNet does not links words into semantic relations for?.
    a. Synonyms
    b. Hyponyms
    c. Meronyms
    d. <u>Homonyms</u>

233. WordNet can be used for
    a. Word Sense Disambiguation
    b. Information Retrieval
    c. Machine translation
    d. <u>All of the above</u>

234. ___ and ___ are the most common semantic roles.(No Answer)
    a. semantics
    b. antonyms
    c. <u>Agents and themes</u>
    d. Option A and option B

235. When two or more different forms have the same pronunciation, they are called ___.
    a. <u>Homophones</u>
    b. Hyponym
    c. Co-hyponyms
    d. Homonyms

236. compositional semantics deals with how those _____ meanings combine to form more _____ phrasal meanings.
    a. Contextual, simple
    b. <u>Lexical, complex</u>
    c. Symbolic, advanced
    d. None of the above

237. Which is a step of Lexical semantics?
    a. Classification of lexical items
    b. Decomposition of lexical items
    c. Analysing lexical items
    d. <u>All of the above</u>

238. semantic ambiguity happens when a sentence contains an _____ word or phrase.
    a. Mis-spelled
    b. Difficult
    c. <u>Ambiguous</u>
    d. Wrong

239. semantic analyzer would reject a sentence like _____
    a. Hot ice-cream
    b. Warm ice-cream
    c. Cold ice-cream
    d. Option A and B

240. Semantic Treebanks use a formal _____ of sentence's semantic structure.
    a. Definition
    b. Classification
    c. Segmentation
    d. Representation

241. Lexical analysis is based on _____ token but on the other side semantic analysis focuses on _____ chunks.
    a. Smaller, larger
    b. Similar, different
    c. Different, Similar
    d. Classified, Unsorted

242. In lexical semantics, we do study of _____
    a. Multiple words
    b. Group of sentances
    c. Individual words
    d. All of the above

243. In Hyponymy, It may be defined as the relationship between a generic ____ and ____ of that generic term.
    a. Words, sentences
    b. Term, instances
    c. Set, words
    d. None of the above

244. In Hyponymy, the generic term is called hypernym and its instances are called hyponyms.
    a. Hypernym, Hyponyms
    b. Hyponyms, Hypernyms
    c. Label, Tags
    d. None of the above

245. Semantic analysis creates a _____ of the meaning of a sentence.
    a. Representation
    b. Context
    c. Definition
    d. Classification

246. Which is the building block of semantic systems
    a. Entities, concepts
    b. Relations predicates
    c. Option A and B
    d. None of the above

247. What is reason for need of Meaning representations?
    a. Linking of linguistic elements to non-linguistic elements
    b. Representing variety at lexical level
    c. Can be used for reasoning
    d. All of the above

248. What is required for evaluation of WSD?
    a. Dictionary
    b. Test-corpus
    c. Sequences
    d. Option A and B

249. Which is a difficulty for WSD?
    a. Differences between dictionaries
    b. Different algorithm for different applications
    c. Inter-judge variance
    d. All of the above

250. polysemy has the same spelling but _____ meaning.
    a. Difficult and contextual
    b. Simple and Undestandable
    c. Different and related
    d. None of the above

251. Homonymy is defined as the words having _____ spelling or same form but having _____ and unrelated meaning.
    a. Different, different
    b. Same, different
    c. Same, same
    d. Different, same

252. Synonymy is the relation between two _____ items having different _____ but expressing the same or a close meaning.
    a. Same, meaning
    b. Different, context
    c. Lexical, forms
    d. None of the above

253. Antonymy is the relation between two lexical items having _____ between their semantic _____ relative to an axis.
   a. Symmetry, Components
   b. Relation, words
   c. Connection, groups
   d. Action, symbols

254. Which is a Application of WSD?
   a. Machine translations
   b. Information Retrieval
   c. Text mining
   d. All of the above

255. the sense of the word depends on the _____ words of that particular word.
   a. Relating
   b. Same
   c. Neighboring
   d. Dependent

256. Semantic analysis _____the text elements and assigns them to their logical and grammatical role.
   a. Reviews
   b. Identifies
   c. Checks
   d. Adds

257. Semantic analysis relates to concepts like _____ and _____, which is the particular combination of words that can be or frequently are surrounding a single word.
   a. Connotations, Collocation
   b. Addition, extraction
   c. Analysis, identification
   d. None of the above

258. Pragmatic Analysis is part of the process of extracting _____ from text.
   a. Context
   b. Words
   c. Letters
   d. Information

259. focuses on taking a _____ set of text and figuring out what the actual _____ was.
   a. Structured & Meaning
   b. Extracted & context
   c. Selected & Definition
   d. None of the above

260. deixis is the process of _____ via language
    - **a. Pointing**
    - b. Selecting
    - c. Extracting
    - d. Removing

261. The linguistic forms we use to accomplish this 'pointing' is called deictic _____.
    - a. Definition
    - b. Symbol
    - c. Context
    - **d. Expression**

262. Which is not a type of Deixis?
    - **a. Simple**
    - b. Person
    - c. Spatial
    - d. Temporal

263. “Implicature” denotes either the act of meaning or _____ one thing by saying something else, or the _____ of that act.
    - a. Adding, Set
    - **b. Implying, Object**
    - c. Extracting, Rule
    - d. Defining, Part

264. Implicatures can be determined by sentence _____ or by _____ context.
    - a. Definition, additional
    - **b. Meaning, Conversational**
    - c. Understanding, group
    - d. None of the above

265. Which is not a type of Implicature?
    - a. Conventional
    - b. Unconventional
    - **c. Conditional**
    - d. Conversational

266. A presupposition is an _____ about the world or background belief
    - a. Additional assumption
    - b. Defined assumption
    - c. Necessary assumption

d. Implicit assumption

267. A presupposition must be mutually known or assumed by the _____ and _____ for the utterance to be considered appropriate in context.
   a. Reader, Writer
   b. Actor, Object
   c. Speaker, Addressee
   d. None of the above

268. A presupposition trigger is a _____ item or linguistic construction which is responsible for the presupposition, and thus "triggers" it.
   a. Context, Well-defined
   b. Lexical, Linguistic
   c. Defined, Proper
   d. All of the above

269. speech act is something expressed by an individual that not only presents _____, but performs an _____ as well.
   a. Data, process
   b. Rules, Activity
   c. Information, Action
   d. None of the above

270. Which is not a level of Speech act?
   a. Locutionary
   b. Definite
   c. Indirect
   d. Performative

271. Which is primitive Speech act from following?
   a. Labelling
   b. Answering
   c. Requesting
   d. All of the above

272. Discourse deals with how the immediately preceding sentence can affect the _____ of the next sentence.
   a. Addition
   b. Extraction
   c. Interpretation
   d. Analysis

273. Which is the mode of discourse from following
    a. Narration, Description
    b. Exposition, Argument
    c. Option A and B
    d. None of the above

274. Feature of Discourse structure?
    a. Time
    b. Space
    c. Class
    d. All of the above

275. reference resolution may be defined as the task of _____what entities are referred to by which _____ expression.
    a. Explaining, Lexical
    b. **Determining, Linguistic**
    c. Analysing, Symbol
    d. None of the above

276. Which is type of reference resolution
    a. Definite and Indefinite
    b. Pronouns and Names
    c. Demonstratives
    d. All of the above

277. Which is not a task of reference resolution
    a. Coreference Resolution
    b. Constraint Resolution
    c. Constraint Coreference Resolution
    d. Pronomial Anaphora Resolution

278. Simple natural language phenomena (e.g., NP-NP, V-NP-NP patterns) can be described using _____
    a. CFG
    b. Turing machines
    c. DFA
    d. None of the above

279. Which is not a Semantic and Syntactic Constraint?
    a. Reflexive
    b. Pronoun

c. Animate

d. <u>Addition</u>

280. Which is not a distinction for Coreference?

a. Anaphora

b. Cataphora

c. Split Antecedents

d. <u>Combined Antecedents</u>

281. Which is not the application of Coreference resolution?

a. Text Classification

b. Textual Entailment

c. Automatic Summarization

d. <u>Text Recognition</u>

282. coreference resolution is a well-studied problem in _____.

a. <u>Discourse</u>

b. NLP

c. Text classification

d. None of the above

283. Algorithms intended to resolve coreferences commonly look first for the nearest _____ that is compatible with the referring _____.

a. Following Word, information

b. Group of sentences, Definitions

c. <u>Preceding individual, Expressions</u>

d. None of the above

284. 'anaphora' is an intra-linguistically _____ relation, whereas 'co-reference' necessarily requires access to 'extra-linguistic' _____.

a. Identified, quantity

b. Explained, expression

c. <u>Determinable, information</u>

d. Acceptable, symbols

285. co-reference always implies _____, whereas anaphora does not.

a. <u>Identity of reference</u>

b. Complexity of reference

c. Simplicity of reference

d. Quality of reference

286. Pragmatic Ambiguity arises when the _____ of words of a sentence is not specific; it concludes different meanings.
    a. Information
    b. Structure
    c. Context
    d. Meaning

287. Pragmatics is the study of the functions of the _____ and its use in context.
    a. Context
    b. Language
    c. Words
    d. Symbols

288. Reference may be defined as the _____ expression to denote an entity or individual.
    a. Symbolic
    b. Text
    c. Linguistic
    d. Lexical

289. The coherence of entire discourse can also be considered by _____ structure between _____ relations.
    a. Horizontal, Incoherence
    b. Vertical, all
    c. Hierarchical, Coherence
    d. Nonhierarchical, Complex

290. Pragmatic ambiguity refers to the situation where the context of a phrase gives it multiple _____.
    a. Assumptions
    b. Group of data
    c. Interpretations
    d. Classifications

291. Machine Translation (MT) is the task of automatically converting one natural _____ into another, preserving the meaning of the input text, and _____ fluent text in the output language.
    a. Definition, Assuming
    b. Context, Defining
    c. Language, Producing
    d. Symbols, Adding

292. Which is not MT approach?
    a. Rule-based
    b. Symbol
    c. Statistical
    d. Neural

293. MT performs _____ substitution of words in one language for words in another
    a. Mechanical
    b. Aesthetic
    c. Complex
    d. Contextual

294. Information retrieval (IR) is finding material (usually documents) of an _____ nature that satisfies an information need from within large _____.
    a. Categorized, words
    b. Defined, databases
    c. Unstructured, collections
    d. Correlated, datasets

295. Which is not a model of Information retrieval?
    a. Classical
    b. Non-classical
    c. Simple
    d. Alternative

296. Which is not a component of IR?
    a. Indexing system
    b. Collection of documents
    c. Defined set of Queries
    d. Reviewing system

297. A perfect IR system will retrieve only relevant documents.
    a. Relevant
    b. Accessible
    c. Large
    d. Small

298. Sentiment analysis is the _____ and _____ of emotions (positive, negative and neutral) within text data using text analysis techniques.
    a. Checking, reviewing
    b. Defining, Grouping

c. Interpretation, Classification

d. None of the above

299. Which is not type of Sentiment Analysis?

a. Emotion Detection

b. Aspect based

c. Word based

d. Bilingual

300. Which algorithm is used in Classification of SE?

a. SVM

b. Deep learning

c. Linear regression

d. All of the above

301. Which is a challenge for SE?

a. Subjectivity and Tone

b. Context and Polarity

c. Irony and Sarcasm

d. All of the above

302. Which is not a SE classification technique?

a. Automated

b. Hybrid

c. Semi-automated

d. Rule-based

303. Which is application of SE?

a. Customer feedback

b. Social media monitoring

c. Market Research

d. All of the above

304. Which is a step of finding Sentiment Polarity?

a. Sentiment identification

b. Feature selection

c. Sentiment classification

d. All of the above

305. Feature selection includes

a. Term presence and frequecy, Opinion words and phrases

b. Part of speech and negations

c. <u>Option A and B</u>

d. None of the above

306. Named entity recognition is a popular technique used in information extraction to _____ and _____ the named entities and _____ them under various predefined classes.

a. Define, group, differentiate

b. <u>Identify, Segment, Categorize</u>

c. View, Category, classify

d. None of the above

307. Entity extraction is really useful for analyzing _____ text

a. Structured

b. Categorised

c. Defined

d. <u>Unstructured</u>

308. Which is A method for Named-entity extractions

a. Hybrid

b. Lexicon

c. Machine learning based

d. <u>All of the above</u>

309. Rule-based systems for entity extraction employ a _____ of grammatical _____ hand-crafted by computational _____.

a. Set, Actions, scientists

b. <u>Series, Rules, Linguists</u>

c. Sequence, words, developers

d. None of the above

310. Which is not a application of Named-entity extractions

a. Customer feedback

b. Categorizing customer tickets for help

c. Analyze resumes

d. <u>Recognizing characters from images</u>

311. It is a process of generating a concise and meaningful _____ of _____ from multiple text resources such as books, news articles, blog posts, research papers, emails, and tweets.

a. Definition, words

b. Classification, symbols

c. <u>Summary, text</u>

d. Context, words

312. In extractive Summarization, we identify the important _____ or _____ from the original text and _____ only those from the text.
   a. Words, symbols, add
   b. Sentences, phrases, extracts
   c. Characters, context, remove
   d. Symbols, characters, join

313. In abstractive summarization, we _____ new sentences from the _____ text.
   a. Generate, original
   b. Extract, new
   c. Add, new
   d. None of the above

314. text summarization in NLP is treated as a _____ machine learning problem
   a. Supervised
   b. Unsupervised
   c. Semi-supervised
   d. None of the above

315. Text classification also known as text _____ or text _____ is the process of categorizing text into _____ groups.
   a. Reviewing, segmenting, unorganized
   b. Tagging, Categorization, organized
   c. Labelling, reviewing, organized
   d. Grouping, categorization, Unorganized

316. Text classification is used in
   a. Sentiment analysis
   b. Topic detection
   c. Language detection
   d. All of the above

317. Which is not a text classification technique?
   a. OCR
   b. Artifical Neural network
   c. Tf-idf
   d. Naive bayes classifier

318. Which is a type of text based classification?
    a. Request-based
    b. Content-based
    c. Option A and B
    d. Automatic

319. Sentiment analysis refers to the use of natural language processing to systematically _____, _____, _____, and _____ affective states and subjective information.
    a. Identify, extract, quantify, study
    b. Analyze, add, remove, segment
    c. View, extract, classify, study
    d. None of the above

320. Which is a example of Sentiment analysis?
    a. Coronet has the best lines of all-day cruisers.
    b. Bertram has a deep V hull and runs easily through seas.
    c. Pastel-colored 1980s-day cruisers from Florida are ugly.
    d. All of the above

321. Which model is not used in Information Retrieval?
    a. Extended Boolean model
    b. Vector space model
    c. Binary Independence model
    d. Bayesian model

322. What is pragmatic Analysis?
    a. Pragmatic Analysis deals with the overall communicative and social content and its effect on interpretation.
    b. It means abstracting or deriving the meaningful use of language in situations. In this analysis, the main focus is always on what was said in reinterpreted on what is meant.
    c. Pragmatic analysis helps users to discover this intended effect by applying a set of rules that characterize cooperative dialogues.
    d. All of the above

323. _____analysis deals with the overall communicative and social content and its effect on interpretation. It means abstracting or deriving the meaningful use of language in situations
    a. Morphological and Lexical Analysis
    b. Syntactic Analysis
    c. Semantic Analysis
    d. Pragmatic Analysis

324. We could also say that ____is the process of 'pointing' via language.
    a. Person deixis
    b. Spatial deixis
    c. Temporal deixis
    d. <u>Deixis</u>

325. Deictic expressions are among the first forms to be acquired and spoken by very young children. They can be used to point to a person via _____(*me*, *you*)
    a. <u>Person deixis</u>
    b. Spatial deixis
    c. Temporal deixis
    d. Deixis

326. Deictic expressions are among the first forms to be acquired and spoken by very young children. They can be used to point to a location via _____(*here*, *there*)
    a. Person deixis
    b. <u>Spatial deixis</u>
    c. Temporal deixis
    d. Deixis

327. Deictic expressions are among the first forms to be acquired and spoken by very young children. They can be used to point to time via_____ (*now*, *then*).
    a. Person deixis
    b. Spatial deixis
    c. <u>Temporal deixis</u>
    d. Deixis

328. An _____ is something the speaker suggests or implies with an utterance, even though it is not literally expressed.
    a. Deixis
    b. <u>Implicature</u>
    c. Presupposition
    d. Speech acts

329. _____can aid in communicating more efficiently than by explicitly saying everything we want to communicate.
    a. Deixis
    b. <u>Implicature</u>
    c. Presupposition
    d. Speech acts

330. A _____ is an implicit assumption about the world or background belief relating to an utterance whose truth is taken for granted in discourse.
    a. Deixis
    b. Implicature
    c. Presupposition(PSP)
    d. Speech acts

331. _____ represent a key concept in the field of pragmatics which can be broadly defined as language use in context taking into account the speaker's and the addressee's verbal and non-verbal contributions to the negotiation of meaning in interaction.
    a. Deixis
    b. Implicature
    c. Presupposition(PSP)
    d. Speech acts

332. What are the types of speech acts according to John Searle?
    a. Representatives
    b. Comisessives
    c. Directives
    d. All of the above

333. Speech Acts are of following types:
    a. Declarations
    b. Expressives
    c. Both a & b
    d. None of the above

334. Essentially, _____ is about the way **conversation** works in practice.
    a. Deixis
    b. Implicature
    c. Presupposition(PSP)
    d. Conversational Structure

335. _____ commit a speaker to the truth of an expressed proposition.
    a. Representatives
    b. Comisessives
    c. Directives
    d. All of the above

336. _____ commit a speaker to some future action.
    a. Representatives
    b. Comisessives
    c. Directives

d. All of the above

337. _____ are used by a speaker who attempts to get the addressee to carry out an action.
   a. Representatives
   b. Comisessives
   c. Directives
   d. All of the above

338. _____ affect an immediate change of affairs.
   a. Declarations
   b. Expressives
   c. Both a & b
   d. None of the above

339. _____ express some sort of psychological state.
   a. Declarations
   b. Expressive
   c. Both a & b
   d. None of the above

340. _____ may be defined as the linguistic expression to denote an entity or individual.
   a. Coherence
   b. Reference
   c. Discourse
   d. None of the above.

341. _____ may be defined as the task of determining what entities are referred to by which linguistic expression.
   a. Coherence resolution
   b. Reference Resolution
   c. Discourse
   d. None of the above.

342. The natural language expression that is used to perform reference is called a ____.
   a. **Referring expression**
   b. Corefer
   c. Antecedent
   d. Anaphora

343. It is the entity that is referred:
   a. Corefer
   b. Antecedent
   c. Anaphora
   d. Referent

344. When two expressions are used to refer to the same entity, they are called_____.
    a. <u>Corefers</u>
    b. Antecedent
    c. Anaphora
    d. Referent

345. The term has the license to use another term:
    a. Corefers
    b. <u>Antecedent</u>
    c. Anaphora
    d. Referent

346. ____may be defined as the reference to an entity that has been previously introduced into the sentence. And, the referring expression is called ____.
    a. Antecedent
    b. Anaphora
    c. Anaphoric
    d. <u>Both b & c</u>

347. The model that contains the representations of the entities that have been referred to in the discourse and the relationship they are engaged in.
    a. Coherence resolution
    b. Reference Resolution
    c. <u>Discourse Model</u>
    d. None of the above.

348. Referring Expressions have following types:
    a. Indefinite Noun Phrases
    b. Definite Noun Phrases
    c. Pronouns
    d. <u>All of the above</u>

349. Types of Referring expressions are:
    a. Demonstratives
    b. Names
    c. <u>Both a & b</u>
    d. None of the above

350. What are the reference tasks?
    a. Coreference Resolution
    b. Pronominal Anaphora resolution
    c. <u>Both a & b</u>
    d. Anaphora

351. What is Coreference resolution?
    a. It is defined as the task of finding the antecedent for a single pronoun.
    b. <u>It is the task of finding referring expressions in a text that refer to the same entity.</u>
    c. These demonstrate and behave differently than simple definite pronouns.
    d. All of the above

352. What is a Pronominal Anaphora expression?
    a. <u>It is defined as the task of finding the antecedent for a single pronoun.</u>
    b. It is the task of finding referring expressions in a text that refer to the same entity.
    c. These demonstrate and behave differently than simple definite pronouns.
    d. All of the above

353. What are the types of Coreference?
    a. Anaphora & Cataphora
    b. Split Antecedents
    c. coreferring noun phrases
    d. <u>All of the above</u>

354. In a coreference resolution task, the ____is the number of noun phrase pairs correctly labeled as coreferent (true positives) divided by the total number of pairs labeled as coreferent
    a. Recall
    b. <u>Precision</u>
    c. Anaphora
    d. None of the above

355. ____ in this context is defined as the number of true positives divided by the total number of pairs that actually corefer.
    a. <u>Recall</u>
    b. Precision
    c. Anaphora
    d. None of the above

356. What are the three machine learning approaches for coreference resolution?
    a. Clustering Approach
    b. Decision Tree
    c. Algorithm based on Bell tree
    d. <u>All of the above</u>

357. What is the concept of Clustering approach for coreference?
    a. <u>If the distance between two noun phrases is less than the clustering radius threshold r and their coreference equivalence classes are compatible, then the classes are merged</u>
    b. for coreference resolution requires a set of features describing pairs of noun phrases and recasting the coreference problem as a classification task.

c. In each step of the algorithm, one mention is added by either linking to each of existing entities, or starting a new entity. A new layer of nodes is created to represent all possible coreference outcomes by adding one mention. The number of tree leaves is the number of possible coreference outcomes and it equals the Bell number

d. All of the above

358. What is the Decision tree approach for Coreference?

a. If the distance between two noun phrases is less than the clustering radius threshold r and their coreference equivalence classes are compatible, then the classes are merged

b. Applying Decision tree for coreference resolution requires a set of features describing pairs of noun phrases and recasting the coreference problem as a classification task.

c. In each step of the algorithm, one mention is added by either linking to each of existing entities, or starting a new entity. A new layer of nodes is created to represent all possible coreference outcomes by adding one mention. The number of tree leaves is the number of possible coreference outcomes and it equals the Bell number

d. All of the above

359. What is the function of Algorithm based on a Bell tree?

a. If the distance between two noun phrases is less than the clustering radius threshold r and their coreference equivalence classes are compatible, then the classes are merged

b. Applying Decision tree for coreference resolution requires a set of features describing pairs of noun phrases and recasting the coreference problem as a classification task.

c. In each step of the algorithm, one mention is added by either linking to each of existing entities, or starting a new entity. A new layer of nodes is created to represent all possible coreference outcomes by adding one mention. The number of tree leaves is the number of possible coreference outcomes and it equals the Bell number

d. All of the above

360. An anaphora and all its antecedents form a coreference sequence called _____

a. Recall

b. Precision

c. Anaphora

d. Coreferential chain

361. A typical coreference resolution system takes an _____as input and produces the appropriate coreferential chains as output.

a. Coherence resolution

b. Reference Resolution

c. Discourse Model

d. Arbitrary document

362. Coreference is hard because _____
    a. Coreference is hard because of high ambiguity, and because you have to use many kinds of signals and knowledge to resolve the ambiguity.
    b. In particular, you need to know a lot about the world to be able to figure out what something refers to, knowledge which is difficult to encode into computers.
    c. Both a & b
    d. None of the above

363. Applications of Co-reference:
    a. Full text understanding: –understanding an extended discourse
    b. Machine translation (if languages have different features of gender, number, etc.)
    c. Text summarization, including things like web snippets & Tasks like information extraction and question answering, when some sentences have pronouns
    d. All of the above

364. If the distance between two noun phrases is less than the clustering radius threshold r and their coreference equivalence classes are compatible, then the classes are merged. This concept relates to____.
    a. Clustering Approach
    b. Decision Tree
    c. Algorithm based on Bell tree
    d. All of the above

365. In each step of this algorithm, one mention is added by either linking to each of existing entities, or starting a new entity. A new layer of nodes is created to represent all possible coreference outcomes by adding one mention. The number of tree leaves is the number of possible coreference outcomes and it equals the Bell number.
    a. Clustering Approach
    b. Decision Tree
    c. Algorithm based on Bell tree
    d. All of the above

366. Applying tree algorithms for coreference resolution requires a set of features describing pairs of noun phrases and recasting the coreference problem as a classification task.
    a. Clustering Approach
    b. Decision Tree Algorithm
    c. Algorithm based on Bell tree
    d. All of the above

367. Anaphora is a see also of coreference. As nouns the difference between anaphora and coreference is that _____ is (rhetoric) the repetition of a phrase at the beginning of phrases, sentences, or verses, used for emphasis while ____ is (grammar) the relationship between multiple terms that have a common referent.
    a. anaphora, coreference
    b. Coreference, anaphora
    c. Corpus
    d. Clustering

368. _____is the task of automatically converting one natural language into another, preserving the meaning of the input text, and producing fluent text in the output language.
    a. NLP
    b. Machine Translation
    c. Machine Learning
    d. All of the above

369. Machine translation algorithm are as follows:
    a. Hybrid machine translation approach
    b. Corpus-based
    c. Rule-based
    d. All of the above

370. What is Rule-based machine translation approach?
    a. a general term that denotes machine translation systems based on linguistic information about source and target languages basically retrieved from (bilingual) dictionaries and grammars covering the main semantic, morphological, and syntactic regularities of each language respectively.
    b. This representation is manipulated and transferred to a form suitable for the target language. Then at last output is generated in the target language
    c. as it name points, a bilingual parallel corpus to obtain knowledge for new incoming translation. This approach uses a large amount of raw data in the form of parallel corpora.
    d. All of the above.

371. What is corpus based machine translation?
    a. A general term that denotes machine translation systems based on linguistic information about source and target languages basically retrieved from (bilingual) dictionaries and grammars covering the main semantic, morphological, and syntactic regularities of each language respectively.
    b. This representation is manipulated and transferred to a form suitable for the target language. Then at last output is generated in the target language
    c. As its name points, a bilingual parallel corpus to obtain knowledge for new incoming translation. This approach uses a large amount of raw data in the form of parallel corpora.

d. All of the above.

372. What is hybrid based machine translation?
   a. A general term that denotes machine translation systems based on linguistic information about source and target languages basically retrieved from (bilingual) dictionaries and grammars covering the main semantic, morphological, and syntactic regularities of each language respectively.
   b. This representation is manipulated and transferred to a form suitable for the target language. Then at last output is generated in the target language
   c. <u>This approach to develop translation from source to target language, which is based on both rules and statistics.</u>
   d. All of the above.

373. Hybrid machine translation is a method of machine translation that is characterized by the use of _____ machine translation approaches within a single machine translation system.
   a. Single
   b. <u>Multiple</u>
   c. Both a & b
   d. None of the above

374. _____, also known as Knowledge-Based Machine Translation and Classical Approach of MT,
   a. Hybrid machine translation approach
   b. Corpus-based
   c. <u>Rule-based machine translation approach.</u>
   d. All of the above

375. RBMT methodology applies a set of linguistic rules in three different phases:
   a. Analysis
   b. Transfer
   c. Generation
   d. <u>All of the above</u>

376. A rule-based system requires:
   a. syntax generation and semantic generation.
   b. syntax analysis, semantic analysis,
   c. <u>Both a & b</u>
   d. None of the above

377. _____ may be defined as a software program that deals with the organization, storage, retrieval and evaluation of information from document repositories particularly textual information
    a. Analysis
    b. Transfer
    c. Generation
    d. Information retrieval

378. A _____that denotes machine translation systems based on linguistic information about source and target languages basically retrieved from (bilingual) dictionaries and grammars covering the main semantic, morphological, and syntactic regularities of each language respectively.
    a. Hybrid machine translation approach
    b. Corpus-based
    c. Rule-based machine translation approach.
    d. All of the above

379. As its name points, a bilingual parallel corpus to obtain knowledge for new incoming translation. _____approach uses a large amount of raw data in the form of parallel corpora.
    a. Hybrid machine translation approach
    b. Corpus-based
    c. Rule-based machine translation approach.
    d. All of the above

380. _____ approach to develop translation from source to target language, which is based on both rules and statistics.
    a. Hybrid machine translation approach
    b. Corpus-based
    c. Rule-based machine translation approach.
    d. All of the above

381. A_____ first analyses the source language input and creates an internal representation. This representation is manipulated and transferred to a form suitable for the target language.
    a. Hybrid machine translation approach
    b. Corpus-based machine translation approach
    c. Rule-based machine translation approach.
    d. Machine Translation system

382. In ____retrieval, the user must enter a query in natural language that describes the required information. Then the IR system will return the required documents related to the desired information.

a. Information Retrieval
b. <u>Ad-hoc retrieval</u>
c. Machine translation
d. None of the above

383. What are the types of Information retrieval model?
a. Classical IR model
b. Non-classical IR
c. Alternative IR model
d. <u>All of the above</u>

384. This model is based on mathematical knowledge that was easily recognized and understood as well.
a. <u>Classical IR model</u>
b. Non-classical IR
c. Alternative IR model
d. None of the above

385. What are the types of classical IR models?
a. Boolean,
b. Vector
c. Probabilistic
d. <u>All of the above</u>

386. Such kinds of IR models are based on principles other than similarity, probability, Boolean operations.
a. Classical IR model
b. <u>Non-classical IR</u>
c. Alternative IR model
d. None of the above

387. What are the examples of Non-classical IR models?
a. Information logic model
b. situation theory model
c. interaction models
d. <u>All of the above</u>

388. It is the enhancement of classical IR models making use of some specific techniques from some other fields.
a. Classical IR model
b. Non-classical IR
c. <u>Alternative IR model</u>

d. None of the above

389. Following are the examples of Alternative IR models:
   a. Cluster model,
   b. fuzzy model
   c. latent semantic indexing (LSI) models.
   d. <u>All of the above.</u>

390. _____ is a computer science discipline within the fields of information retrieval and natural language processing (NLP), which is concerned with building systems that automatically answer questions posed by humans in a natural language.
   a. Corpus-based machine translation approach
   b. Rule-based machine translation approach.
   c. Machine Translation system
   d. <u>Question answering (QA)</u>

391. What is Question answering in NLP?
   a. This representation is manipulated and transferred to a form suitable for the target language. Then at last output is generated in the target language
   b. as its name points, a bilingual parallel corpus to obtain knowledge for new incoming translation. This approach uses a large amount of raw data in the form of parallel corpora.
   c. <u>It is a computer science discipline within the fields of information retrieval and natural language processing (NLP), which is concerned with building systems that automatically answer questions posed by humans in a natural language</u>
   d. All of the above.

392. In information retrieval, an _____ question answering system aims at returning an answer in response to the user's question. The returned answer is in the form of short texts rather than a list of relevant documents.
   a. Close domain
   b. <u>Open domain</u>
   c. IR system
   d. None of the above

393. The open domain question answering system uses a combination of techniques from _____for finding answers.
   a. computational linguistics
   b. information retrieval
   c. knowledge representation
   d. <u>All of the above</u>

394. What Open system question answering system aims to?
    a. This representation is manipulated and transferred to a form suitable for the target language. Then at last output is generated in the target language
    b. as its name points, a bilingual parallel corpus to obtain knowledge for new incoming translation. This approach uses a large amount of raw data in the form of parallel corpora.
    c. <u>This system aims at returning an answer in response to the user's question. The returned answer is in the form of short texts rather than a list of relevant documents.</u>
    d. None of the above

395. An open source ____question answering system based on Ask Platypus and Wikidata was published in 2018.
    a. <u>math-aware</u>
    b. Close domain
    c. Open domain
    d. IR system

396. An open source math-aware question answering system takes an _____natural language question as input and returns a mathematical formula retrieved from Wikidata as a succinct answer.
    a. English or spanish
    b. English only
    c. <u>English or Hindi</u>
    d. English or German

397. _____ is the interpretation and classification of emotions (positive, negative and neutral) within text data using text analysis techniques.
    a. Lexical analysis
    b. Syntactical analysis
    c. Hybrid analysis
    d. <u>Sentiment Analysis</u>

398. _____ tools allow businesses to identify customer sentiment toward products, brands or services in online feedback.
    a. Lexical analysis
    b. Syntactical analysis
    c. Hybrid analysis
    d. <u>Sentiment Analysis</u>

399. Types of sentiment analysis are:
    a. Fine-grained Sentiment Analysis
    b. Emotion detection, Multilingual sentiment analysis
    c. Aspect-based Sentiment Analysis
    d. <u>All of the above</u>

400. What are the polarity categories?
    a. Positive
    b. Neutral
    c. Negative
    d. <u>All of the above</u>

401. This type of sentiment analysis aims at detecting emotions, like happiness, frustration, anger, sadness, and so on.
    a. Fine-grained Sentiment Analysis
    b. **<u>Emotion detection</u>**
    c. Multilingual sentiment analysis
    d. Aspect-based Sentiment Analysis

402. Many emotion detection systems use ____or complex machine learning algorithms.
    a. Tokens
    b. <u>Lexicons</u>
    c. Corpus
    d. Nothing

403. Usually, when analyzing sentiments of texts, let's say product reviews, you'll want to know which particular features people are mentioning in a positive, neutral, or negative way. That's where _____ can help,
    a. Fine-grained Sentiment Analysis
    b. Emotion detection
    c. Multilingual sentiment analysis
    d. <u>Aspect-based Sentiment Analysis</u>

404. Benefits of sentiment analysis includes:
    a. Sentiment analysis helps businesses process huge amounts of data in an efficient and cost-effective way.
    b. Sentiment analysis models can help you immediately identify these kinds of situations and gauge brand sentiment, so you can take action right away.

c. Tagging text by sentiment is highly subjective, influenced by personal experiences, thoughts, and beliefs. By using a centralized sentiment analysis system, companies can apply the same criteria to all of their data, helping them improve accuracy and gain better insights.

d. <u>All of the above</u>

405. Sentiment analysis uses various Natural Language Processing (NLP) methods and algorithms, The main types of algorithms used include:
    a. Rule-based
    b. Automatic
    c. Hybrid
    d. <u>All of the above</u>

406. What are the Sentiment analysis challenges?
    a. Subjectivity and Tone
    b. Context and Polarity
    c. Irony and Sarcasm
    d. <u>All of the above</u>

407. What are the Sentiment analysis challenges?
    a. Comparisons
    b. Emojis
    c. Defining Neutral
    d. <u>All of the above</u>

408. What are the applications of sentiment analysis?
    a. Social media monitoring
    b. Brand monitoring, Voice of customer (VoC)
    c. Customer service, Market research
    d. <u>All of the above</u>

409. _____also known as text tagging or text categorization is the process of categorizing text into organized groups.
    a. Rule based approach
    b. Lexicons
    c. Tokens
    d. <u>Text classification</u>

410. By using Natural Language Processing (NLP), _____can automatically analyze text and then assign a set of pre-defined tags or categories based on its content.
    a. Text classification
    b. <u>Text classifiers</u>
    c. Corpus
    d. None of the above

411. Some of the most common examples and use cases for automatic text classification include the following:
    a. Sentiment Analysis
    b. Topic Detection
    c. Language Detection
    d. <u>All of the above</u>

412. _____is the process of understanding if a given text is talking positively or negatively about a given subject (e.g. for brand monitoring purposes).
    a. Syntactical analysis
    b. Hybrid analysis
    c. <u>Sentiment Analysis</u>
    d. All of the above

413. _____ is the task of identifying the theme or topic of a piece of text (e.g. know if a product review is about Ease of Use, Customer Support, or Pricing when analyzing customer feedback).
    a. Sentiment Analysis
    b. <u>Topic Detection</u>
    c. Language Detection
    d. All of the above

414. The _____ is a procedure of detecting the language of a given text (e.g. know if an incoming support ticket is written in English or Spanish for automatically routing tickets to the appropriate team).
    a. Sentiment Analysis
    b. Topic Detection
    c. <u>Language Detection</u>
    d. All of the above

415. Text classification techniques include:
    a. Rule-based systems
    b. Machine Learning based systems
    c. Hybrid systems
    d. <u>All of the above</u>

416. _____classify text into organized groups by using a set of handcrafted linguistic rules.
    a. <u>Rule-based approach</u>
    b. Machine Learning based systems
    c. Hybrid systems
    d. All of the above

417. Some of the most popular machine learning algorithms for creating text classification models include:
    a. Naive Bayes
    b. Support Vector Machines
    c. Deep Learning
    d. <u>All of the above</u>

418. Generally, a classification technique could be divided into ____approaches.
    a. Rule-based approach
    b. Machine Learning approach
    c. Statistical approach
    d. <u>Both b & c</u>

419. It is a process of generating a concise and meaningful summary of text from multiple text resources such as books, news articles, blog posts, research papers, emails, and tweets.
    a. Sentiment Analysis
    b. Topic Detection
    c. Language Detection
    d. <u>Text Summarization</u>

420. Text summarization methods are as follows:
    a. <u>Extractive & Abstractive</u>
    b. Extractive Summarization
    c. Abstractive Summarization
    d. None of the above

421. _____ methods function by identifying the important sentences or excerpts from the text and reproducing them verbatim as part of the summary. No new text is generated; only existing text is used in the summarization process.
    a. Text Summarization
    b. <u>Extractive Summarization</u>
    c. Abstractive Summarization
    d. None of the above

422. What is extractive text summarization?
    a. <u>This method functions by identifying the important sentences or excerpts from the text and reproducing them verbatim as part of the summary. No new text is generated; only existing text is used in the summarization process.</u>
    b. methods employ more powerful natural language processing techniques to interpret text and generate new summary text, as opposed to selecting the most representative existing excerpts to perform the summarization.
    c. It is a process of generating a concise and meaningful summary of text from multiple text resources such as books, news articles, blog posts, research papers, emails, and tweets.

d. None of the above

423. _____ methods employ more powerful natural language processing techniques to interpret text and generate new summary text, as opposed to selecting the most representative existing excerpts to perform the summarization.
   a. Text Summarization
   b. Extractive Summarization
   c. <u>Abstractive Summarization</u>
   d. None of the above

424. What is abstractive text summarization?
   a. This method functions by identifying the important sentences or excerpts from the text and reproducing them verbatim as part of the summary. No new text is generated; only existing text is used in the summarization process**.**
   b. <u>methods employ more powerful natural language processing techniques to interpret text and generate new summary text, as opposed to selecting the most representative existing excerpts to perform the summarization.</u>
   c. It is a process of generating a concise and meaningful summary of text from multiple text resources such as books, news articles, blog posts, research papers, emails, and tweets.
   d. None of the above

425. _____ is the task of identifying and categorizing key information (entities) in text.
   a. Text Summarization
   b. Extractive Summarization
   c. Abstractive Summarization
   d. <u>Named entity recognition</u>

426. An _____ can be any word or series of words that consistently refers to the same thing. Every detected entity is classified into a predetermined category.
   a. Corpus
   b. <u>Entity</u>
   c. Tokens
   d. None of the above

427. At the heart of any NER model is a two-step process:
   a. Detect a named entity
   b. Categorize the entity
   c. Detecting the corpus
   d. <u>Both a & b</u>