Natural Language Processing Engineer Interview Questions And Answers Guide.

Global Guideline.

https://globalguideline.com/

Natural Language Processing Engineer Job Interview Preparation Guide.

Question # 1

Tell me what is sequence learning?

Answer:-

Sequence learning is a method of teaching and learning in a logical manner.

Read More Answers.

Question # 2

Tell me what are the different methods for Sequential Supervised Learning?

Answer:-

The different methods to solve Sequential Supervised Learning problems are
* a) Sliding-window methods
* b) Recurrent sliding windows
* c) Hidden Markow models
* d) Maximum entropy Markow models
* e) Conditional random fields
* f) Graph transformer networks

Read More Answers.

Question # 3

Tell us what is bias-variance decomposition of classification error in ensemble method?

Answer:-

The expected error of a learning algorithm can be decomposed into bias and variance. A bias term measures how closely the average classifier produced by the learning algorithm matches the target function. The variance term measures how much the learning algorithm's prediction fluctuates for different training sets.

Read More Answers.

Question # 4

Tell me what are the two classification methods that SVM (Support Vector Machine) can handle?

Answer:-

* a) Combining binary classifiers
* b) Modifying binary to incorporate multiclass learning

Read More Answers.

Question # 5

Tell us what are the two methods used for the calibration in Supervised Learning?

Answer:-

The two methods used for predicting good probabilities in Supervised Learning are
* a) Platt Calibration
* b) Isotonic Regression

These methods are designed for binary classification, and it is not trivial.

Read More Answers.

Question # 6

Tell us in what areas Pattern Recognition is used?

Answer:-

Pattern Recognition can be used in
* a) Computer Vision
* b) Speech Recognition
* c) Data Mining
* d) Statistics
* e) Informal Retrieval
* f) Bio-Informatics

Read More Answers.

Question # 7

Explain me the function of 'Unsupervised Learning'?

Answer:-

* a) Find clusters of the data
* b) Find low-dimensional representations of the data
* c) Find interesting directions in data
* d) Interesting coordinates and correlations
* e) Find novel observations/ database cleaning

Read More Answers.

Question # 8

Tell me what are the three stages to build the hypotheses or model in machine learning?

Answer:-

* a) Model building
* b) Model testing
* c) Applying the model

Read More Answers.

Question # 9

Tell us what is inductive machine learning?

Answer:-

The inductive machine learning involves the process of learning by examples, where a system, from a set of observed instances tries to induce a general rule.

Read More Answers.

Question # 10

Explain me the difference between Data Mining and Machine learning?

Answer:-

Machine learning relates with the study, design and development of the algorithms that give computers the capability to learn without being explicitly programmed. While, data mining can be defined as the process in which the unstructured data tries to extract knowledge or unknown interesting patterns. During this process machine, learning algorithms are used.

Read More Answers.

Question # 11

Collaborative Filtering and Content Based Models are the two popular recommendation engines, what role does NLP play in building such algorithms.
A) Feature Extraction from text
B) Measuring Feature Similarity
C) Engineering Features for vector space learning model
D) All of these

Answer:-

D) All of these
NLP can be used anywhere where text data is involved - feature extraction, measuring feature similarity, create vector features of the text.

Read More Answers.

Question # 12

Tell me what are two techniques of Machine Learning?

Answer:-

The two techniques of Machine Learning are
* a) Genetic Programming
* b) Inductive Learning

Read More Answers.

Question # 13

Tell us what are the components of relational evaluation techniques?

Answer:-

The important components of relational evaluation techniques are
* a) Data Acquisition
* b) Ground Truth Acquisition

* c) Cross Validation Technique
* d) Query Type
* e) Scoring Metric
* f) Significance Test

Read More Answers.

Question # 14

Tell us the function of 'Supervised Learning'?

Answer:-

* a) Classifications
* b) Speech recognition
* c) Regression
* d) Predict time series
* e) Annotate strings

Read More Answers.

Question # 15

Please explain how can you avoid overfitting?

Answer:-

By using a lot of data overfitting can be avoided, overfitting happens relatively as you have a small dataset, and you try to learn from it. But if you have a small database and you are forced to come with a model based on that. In such situation, you can use a technique known as cross validation. In this method the dataset splits into two section, testing and training datasets, the testing dataset will only test the model while, in training dataset, the datapoints will come up with the model.
In this technique, a model is usually given a dataset of a known data on which training (training data set) is run and a dataset of unknown data against which the model is tested. The idea of cross validation is to define a dataset to "test" the model in the training phase.

Read More Answers.

Question # 16

Which of the following techniques can be used for the purpose of keyword normalization, the process of converting a keyword into its base form?
Lemmatization
Levenshtein
Stemming
Soundex
A) 1 and 2
B) 2 and 4
C) 1 and 3
D) 1, 2 and 3
E) 2, 3 and 4
F) 1, 2, 3 and 4

Answer:-

C) 1 and 3
Lemmatization and stemming are the techniques of keyword normalization, while Levenshtein and Soundex are techniques of string matching.

Read More Answers.

Question # 17

In Latent Dirichlet Allocation model for text classification purposes, what does alpha and beta hyperparameter represent-
A) Alpha: number of topics within documents, beta: number of terms within topics False
B) Alpha: density of terms generated within topics, beta: density of topics generated within terms False
C) Alpha: number of topics within documents, beta: number of terms within topics False
D) Alpha: density of topics generated within documents, beta: density of terms generated within topics True

Answer:-

D) Alpha: density of topics generated within documents, beta: density of terms generated within topics True

Read More Answers.

Question # 18

What is the right order for a text classification model components
Text cleaning
Text annotation
Gradient descent
Model tuning
Text to predictors
A) 12345
B) 13425
C) 12534
D) 13452

Answer:-

C) 12534
A right text classification model contains - cleaning of text to remove noise, annotation to create more features, converting text-based features into predictors, learning a model using gradient descent and finally tuning a model.

Read More Answers.

Question # 19

Social Media platforms are the most intuitive form of text data. You are given a corpus of complete social media data of tweets. How can you create a model that suggests the hashtags?
A) Perform Topic Models to obtain most significant words of the corpus
B) Train a Bag of Ngrams model to capture top n-grams - words and their combinations
C) Train a word2vector model to learn repeating contexts in the sentences
D) All of these

Answer:-

D) All of these
All of the techniques can be used to extract most significant terms of a corpus.

Read More Answers.

Question # 20

Do you know 'Overfitting' in Machine learning?

Answer:-

In machine learning, when a statistical model describes random error or noise instead of underlying relationship 'overfitting' occurs. When a model is excessively complex, overfitting is normally observed, because of having too many parameters with respect to the number of training data types. The model exhibits poor performance which has been overfit.

Read More Answers.

Question # 21

Explain me what is the difference between heuristic for rule learning and heuristics for decision trees?

Answer:-

The difference is that the heuristics for decision trees evaluate the average quality of a number of disjointed sets while rule learners only evaluate the quality of the set of instances that is covered with the candidate rule.

Read More Answers.

Question # 22

Tell me what is batch statistical learning?

Answer:-

Statistical learning techniques allow learning a function or predictor from a set of observed data that can make predictions about unseen or future data. These techniques provide guarantees on the performance of the learned predictor on the future unseen data based on a statistical assumption on the data generating process.

Read More Answers.

Question # 23

Tell me what are the different categories you can categorized the sequence learning process?

Answer:-

* a) Sequence prediction
* b) Sequence generation
* c) Sequence recognition
* d) Sequential decision

Read More Answers.

Question # 24

N-grams are defined as the combination of N keywords together. How many bi-grams can be generated from given sentence:
"Analytics Vidhya is a great source to learn data science"
A) 7
B) 8
C) 9
D) 10
E) 11

Answer:-

C) 9
Bigrams: Analytics Vidhya, Vidhya is, is a, a great, great source, source to, To learn, learn data, data science

Read More Answers.

Question # 25

You have created a document term matrix of the data, treating every tweet as one document. Which of the following is correct, in regards to document term matrix?
Removal of stopwords from the data will affect the dimensionality of data
Normalization of words in the data will reduce the dimensionality of data
Converting all the words in lowercase will not affect the dimensionality of the data
A) Only 1
B) Only 2
C) Only 3
D) 1 and 2
E) 2 and 3
F) 1, 2 and 3

Answer:-

D) 1 and 2
Choices A and B are correct because stopword removal will decrease the number of features in the matrix, normalization of words will also reduce redundant features, and, converting all words to lowercase will also decrease the dimensionality.

Read More Answers.

Question # 26

True or False: Word2Vec model is a machine learning model used to create vector notations of text objects. Word2vec contains multiple deep neural networks
A) TRUE
B) FALSE

Answer:-

B) FALSE
Word2vec also contains preprocessing model which is not a deep neural network

Read More Answers.

Question # 27

What are the possible features of a text corpus
Count of word in a document
Boolean feature - presence of word in a document
Vector notation of word
Part of Speech Tag
Basic Dependency Grammar
Entire document as a feature
A) 1
B) 12
C) 123
D) 1234
E) 12345
F) 123456

Answer:-

E) 12345
Except for entire document as the feature, rest all can be used as features of text classification learning model.

Read More Answers.

Question # 28

Retrieval based models and Generative models are the two popular techniques used for building chatbots. Which of the following is an example of retrieval model and generative model respectively.
A) Dictionary based learning and Word 2 vector model
B) Rule-based learning and Sequence to Sequence model
C) Word 2 vector and Sentence to Vector model
D) Recurrent neural network and convolutional neural network

Answer:-

B) Rule-based learning and Sequence to Sequence model
choice 2 best explains examples of retrieval based models and generative models

Read More Answers.

Question # 29

Basic Natural Language Processing Engineer Job Interview Questions:

Answer:-

* Do you know about latent semantic indexing? Where can you apply it?
* Is it possible to find all the occurrences of quoted text in an article? If yes, explain how?
* What is a POS tagger? Explain the simplest approach to build a POS tagger?
* Which is a better algorithm for POS tagging - SVM or hidden Markov models?
* What is the difference between shallow parsing and dependency parsing?
* What package are you aware of in python which is used in NLP and ML?
* Explain one application in which stop words should be removed.
* How will you train a model to identify whether the word "Raymond" in a sentence represents a person's name or a company?

Read More Answers.

Question # 30

Common Natural Language Processing Engineer Job Interview Questions:

Answer:-

* As a beginner in Natural Language processing, from where should I start?
* What is the relation between sentiment analysis, natural language processing and machine learning?
* What is the current state of the art in natural language processing?
* What is the state of the art in natural language understanding?
* Which publications would you recommend reading for someone interested in natural language processing?
* What are the basics of natural language processing?
* Could you please explain the choice constraints of the pros/cons while choosing Word2Vec, GloVe or any other thought vectors you have used?
* How do you explain NLP to a layman?
* How do I explain NLP, text mining, and their difference in layman's terms?
* What is the relationship between N-gram and Bag-of-words in natural language processing?

* Is deep learning suitable for NLP problems like parsing or machine translation?
* What is a simple explanation of a language model?
* What is the definition of word embedding (word representation)?
* How is Computational Linguistics different from Natural Language Processing?
* Natural Language Processing: What is a useful method to generate vocabulary for large corpus of data?
* How do I learn Natural Language Processing?
* Natural Language Processing: What are good algorithms related to sentiment analysis?
* What makes natural language processing difficult?
* What are the ten most popular algorithms in natural language processing?
* What is the most interesting new work in deep learning for NLP in 2017?
* How is word2vec different from the RNN encoder decoder?
* How does word2vec work?
* What's the difference between word vectors, word representations and vector embeddings?
* What are some interesting Word2Vec results?
* How do I measure the semantic similarity between two documents?
* What is the state of the art in word sense disambiguation?
* What is the main difference between word2vec and fastText?
* In layman terms, how would you explain the Skip-Gram word embedding model in natural language processing (NLP)?
* In layman's terms, how would you explain the continuous bag of words (CBOW) word embedding technique in natural language processing (NLP)?
* What is natural language processing pipeline?
* What are the available APIs for NLP (Natural Language Processing)?
* How does perplexity function in natural language processing?
* How is deep learning used in sentiment analysis?

Read More Answers.

Question # 31

General Natural Language Processing Engineer Job Interview Questions:

Answer:-

* Differentiate regular grammar and regular expression.
* How will you estimate the entropy of the English language?
* Describe dependency parsing?
* What do you mean by Information rate?
* Explain Discrete Memoryless Channel (DMC).
* How does correlation work in text mining?
* How to calculate TF*IDF for a single new document to be classified?
* How to build ontologies?
* What is an N-gram in the context of text mining?
* What do you know about linguistic resources such as WordNet?
* Explain the tools you have used for training NLP models?

Read More Answers.

Question # 32

Fresh Natural Language Processing Engineer Job Interview Questions:

Answer:-

* Artificial Intelligence: What is an intuitive explanation for recurrent neural networks?
* How are RNNs storing 'memory'?
* What are encoder-decoder models in recurrent neural networks?
* Why do Recurrent Neural Networks (RNN) combine the input and hidden state together and not seperately?
* What is an intuitive explanation of LSTMs and GRUs?
* Are GRU (Gated Recurrent Unit) a special case of LSTM?
* How many time-steps can LSTM RNNs remember inputs for?
* How does attention model work using LSTM?
* How do RNNs differ from Markov Chains?
* For modelling sequences, what are the pros and cons of using Gated Recurrent Units in place of LSTMs?
* What is exactly the attention mechanism introduced to RNN (recurrent neural network)? It would be nice if you could make it easy to understand!
* Is there any intuitive or simple explanation for how attention works in the deep learning model of an LSTM, GRU, or neural network?
* Why is it a problem to have exploding gradients in a neural net (especially in an RNN)?
* For a sequence-to-sequence model in RNN, does the input have to contain only sequences or can it accept contextual information as well?
* Can "generative adversarial networks" be used in sequential data in recurrent neural networks? How effective would they be?
* What is the difference between states and outputs in LSTM?
* What is the advantage of combining Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN)?
* Which is better for text classification: CNN or RNN?
* How are recurrent neural networks different from convolutional neural networks?

Read More Answers.

Question # 33

Professional Natural Language Processing Engineer Interview Questions:

Answer:-

* What is part of speech (POS) tagging? What is the simplest approach to building a POS tagger that you can imagine?
* How would you build a POS tagger from scratch given a corpus of annotated sentences? How would you deal with unknown words?
* How would you train a model that identifies whether the word "Apple" in a sentence belongs to the fruit or the company?
* How would you find all the occurrences of quoted text in a news article?
* How would you build a system that auto corrects text that has been generated by a speech recognition system?
* What is latent semantic indexing and where can it be applied?
* How would you build a system to translate English text to Greek and vice-versa?
* How would you build a system that automatically groups news articles by subject?

* What are stop words? Describe an application in which stop words should be removed.
* How would you design a model to predict whether a movie review was positive or negative?
* What is entropy? How would you estimate the entropy of the English language?
* What is a regular grammar? Does this differ in power to a regular expression and if so, in what way?
* What is the TF-IDF score of a word and in what context is this useful?
* How does the PageRank algorithm work?
* What is dependency parsing?
* What are the difficulties in building and using an annotated corpus of text such as the Brown Corpus and what can be done to mitigate them?
* What tools for training NLP models (nltk, Apache OpenNLP, GATE, MALLET etc...) have you used?
* Do you have any experience in building ontologies?
* Are you familiar with WordNet or other related linguistic resources?
* Do you speak any foreign languages?

Read More Answers.

Question # 34

While working with context extraction from a text data, you encountered two different sentences: The tank is full of soldiers. The tank is full of nitrogen. Which of the following measures can be used to remove the problem of word sense disambiguation in the sentences?
A) Compare the dictionary definition of an ambiguous word with the terms contained in its neighborhood
B) Co-reference resolution in which one resolute the meaning of ambiguous word with the proper noun present in the previous sentence
C) Use dependency parsing of sentence to understand the meanings

Answer:-

A) Compare the dictionary definition of an ambiguous word with the terms contained in its neighborhood
Option 1 is called Lesk algorithm, used for word sense disambiguation, rest others cannot be used.

Read More Answers.

Question # 35

Google Search's feature - "Did you mean", is a mixture of different techniques. Which of the following techniques are likely to be ingredients?
Collaborative Filtering model to detect similar user behaviors (queries)
Model that checks for Levenshtein distance among the dictionary terms
Translation of sentences into multiple languages
A) 1
B) 2
C) 1, 2
D) 1, 2, 3

Answer:-

C) 1, 2
Collaborative filtering can be used to check what are the patterns used by people, Levenshtein is used to measure the distance among dictionary terms.

Read More Answers.

Question # 36

Polysemy is defined as the coexistence of multiple meanings for a word or phrase in a text object. Which of the following models is likely the best choice to correct this problem?
A) Random Forest Classifier
B) Convolutional Neural Networks
C) Gradient Boosting
D) All of these

Answer:-

B) Convolutional Neural Networks
CNNs are popular choice for text classification problems because they take into consideration left and right contexts of the words as features which can solve the problem of polysemy

Read More Answers.

Question # 37

Solve the equation according to the sentence "I am planning to visit New Delhi to attend Analytics Vidhya Delhi Hackathon".
A = (# of words with Noun as the part of speech tag)
B = (# of words with Verb as the part of speech tag)
C = (# of words with frequency count greater than one)
What are the correct values of A, B, and C?
A) 5, 5, 2
B) 5, 5, 0
C) 7, 5, 1
D) 7, 4, 2
E) 6, 4, 3

Answer:-

D) 7, 4, 2
Nouns: I, New, Delhi, Analytics, Vidhya, Delhi, Hackathon (7)
Verbs: am, planning, visit, attend (4)
Words with frequency counts > 1: to, Delhi (2)
Hence option D is correct.

Read More Answers.

Question # 38

How many trigrams phrases can be generated from the following sentence, after performing following text cleaning steps:
Stopword Removal
Replacing punctuations by a single space
"#Analytics-vidhya is a great source to learn @data_science."
A) 3
B) 4
C) 5
D) 6
E) 7

Answer:-

C) 5
After performing stopword removal and punctuation replacement the text becomes: "Analytics vidhya great source learn data science"
Trigrams - Analytics vidhya great, vidhya great source, great source learn, source learn data, learn data science

Read More Answers.

Question # 39

In a corpus of N documents, one document is randomly picked. The document contains a total of T terms and the term "data" appears K times.
What is the correct value for the product of TF (term frequency) and IDF (inverse-document-frequency), if the term "data" appears in approximately one-third of the total documents?
A) KT * Log(3)
B) K * Log(3) / T
C) T * Log(3) / K
D) Log(3) / KT

Answer:-

B) K * Log(3) / T
formula for TF is K/T
formula for IDF is log(total docs / no of docs containing "data")
= log(1 / (â…“))
= log (3)
Hence correct choice is Klog(3)/T

Read More Answers.

Question # 40

What is the major difference between CRF (Conditional Random Field) and HMM (Hidden Markov Model)?
A) CRF is Generative whereas HMM is Discriminative model
B) CRF is Discriminative whereas HMM is Generative model
C) Both CRF and HMM are Generative model
D) Both CRF and HMM are Discriminative model

Answer:-

B) CRF is Discriminative whereas HMM is Generative model

Read More Answers.

Question # 41

Tell me what is 'Training set' and 'Test set'?

Answer:-

In various areas of information science like machine learning, a set of data is used to discover the potentially predictive relationship known as 'Training Set'. Training set is an examples given to the learner, while Test set is used to test the accuracy of the hypotheses generated by the learner, and it is the set of example held back from the learner. Training set are distinct from Test set.

Read More Answers.

Question # 42

Tell us why instance based learning algorithm sometimes referred as Lazy learning algorithm?

Answer:-

Instance based learning algorithm is also referred as Lazy learning algorithm as they delay the induction or generalization process until classification is performed.

Read More Answers.

Question # 43

Tell me what are the areas in robotics and information processing where sequential prediction problem arises?

Answer:-

The areas in robotics and information processing where sequential prediction problem arises are
* a) Imitation Learning
* b) Structured prediction
* c) Model based reinforcement learning

Read More Answers.

Question # 44

Which of the following regular expression can be used to identify date(s) present in the text object:
"The next meetup on data science will be held on 2017-09-21, previously it happened on 31/03, 2016"
A) d{4}-d{2}-d{2}

B) (19|20)d{2}-(0[1-9]|1[0-2])-[0-2][1-9] C) (19|20)d{2}-(0[1-9]|1[0-2])-([0-2][1-9]|3[0-1])
D) None of the above

Answer:-

D) None of the above
None if these expressions would be able to identify the dates in this text object.

Read More Answers.

Question # 45

Which of the following statement is(are) true for Word2Vec model?
A) The architecture of word2vec consists of only two layers - continuous bag of words and skip-gram model
B) Continuous bag of word is a shallow neural network model
C) Skip-gram is a deep neural network model
D) Both CBOW and Skip-gram are deep neural network models
E) All of the above

Answer:-

C) Skip-gram is a deep neural network model
Word2vec contains the Continuous bag of words and skip-gram models, which are deep neural nets.

Read More Answers.

Question # 46

While working with text data obtained from news sentences, which are structured in nature, which of the grammar-based text parsing techniques can be used for noun phrase detection, verb phrase detection, subject detection and object detection.
A) Part of speech tagging
B) Dependency Parsing and Constituency Parsing
C) Skip Gram and N-Gram extraction
D) Continuous Bag of Words

Answer:-

B) Dependency Parsing and Constituency Parsing
Dependency and constituent parsing extract these relations from the text

Read More Answers.

Question # 47

Explain me what is the general principle of an ensemble method and what is bagging and boosting in ensemble method?

Answer:-

The general principle of an ensemble method is to combine the predictions of several models built with a given learning algorithm in order to improve robustness over a single model. Bagging is a method in ensemble for improving unstable estimation or classification schemes. While boosting method are used sequentially to reduce the bias of the combined model. Boosting and Bagging both can reduce errors by reducing the variance term.

Read More Answers.

Question # 48

What percentage of the total statements are correct with regards to Topic Modeling?
It is a supervised learning technique
LDA (Linear Discriminant Analysis) can be used to perform topic modeling
Selection of number of topics in a model does not depend on the size of data
Number of topic terms are directly proportional to size of the data
A) 0
B) 25
C) 50
D) 75
E) 100

Answer:-

A) 0
LDA is unsupervised learning model, LDA is latent Dirichlet allocation, not Linear discriminant analysis. Selection of the number of topics is directly proportional to the size of the data, while number of topic terms is not directly proportional to the size of the data. Hence none of the statements are correct.

Read More Answers.

Question # 49

Explain what is algorithm independent machine learning?

Answer:-

Machine learning in where mathematical foundations is independent of any particular classifier or learning algorithm is referred as algorithm independent machine learning?

Read More Answers.

Question # 50

Do you know the two components of Bayesian logic program?

Answer:-

Bayesian logic program consists of two components. The first component is a logical one ; it consists of a set of Bayesian Clauses, which captures the qualitative structure of the domain. The second component is a quantitative one, it encodes the quantitative information about the domain.

Read More Answers.

Question # 51

Tell me what is an Incremental Learning algorithm in ensemble?

Answer:-

Incremental learning method is the ability of an algorithm to learn from new data that may be available after classifier has already been generated from already available dataset.

Read More Answers.

Question # 52

Do you know what is PAC Learning?

Answer:-

PAC (Probably Approximately Correct) learning is a learning framework that has been introduced to analyze learning algorithms and their statistical efficiency.

Read More Answers.

Question # 53

Which of the following features can be used for accuracy improvement of a classification model?
A) Frequency count of terms
B) Vector Notation of sentence
C) Part of Speech Tag
D) Dependency Grammar
E) All of these

Answer:-

E) All of these
All of the techniques can be used for the purpose of engineering features in a model.

Read More Answers.

Question # 54

While creating a machine learning model on text data, you created a document term matrix of the input data of 100K documents. Which of the following remedies can be used to reduce the dimensions of data -
Latent Dirichlet Allocation
Latent Semantic Indexing
Keyword Normalization
A) only 1
B) 2, 3
C) 1, 3
D) 1, 2, 3

Answer:-

D) 1, 2, 3
All of the techniques can be used to reduce the dimensions of the data.

Read More Answers.

Question # 55

Tell me what is the difference between artificial learning and machine learning?

Answer:-

Designing and developing algorithms according to the behaviours based on empirical data are known as Machine Learning. While artificial intelligence in addition to machine learning, it also covers other aspects like knowledge representation, natural language processing, planning, robotics etc.

Read More Answers.

Question # 56

Please explain what is PCA, KPCA and ICA used for?

Answer:-

PCA (Principal Components Analysis), KPCA (Kernel based Principal Component Analysis) and ICA (Independent Component Analysis) are important feature extraction techniques used for dimensionality reduction.

Read More Answers.

Question # 57

Do you know what is Model Selection in Machine Learning?

Answer:-

The process of selecting models among different mathematical models, which are used to describe the same data set is known as Model Selection. Model selection is applied to the fields of statistics, machine learning and data mining.

Read More Answers.

Question # 58

Which of the following models can be used for the purpose of document similarity?

A) Training a word 2 vector model on the corpus that learns context present in the document
B) Training a bag of words model that learns occurrence of words in the document
C) Creating a document-term matrix and using cosine similarity for each document
D) All of the above

Answer:-

D) All of the above
word2vec model can be used for measuring document similarity based on context. Bag Of Words and document term matrix can be used for measuring similarity based on terms.

Read More Answers.

Question # 59

Which of the following technique is not a part of flexible text matching?
A) Soundex
B) Metaphone
C) Edit Distance
D) Keyword Hashing

Answer:-

D) Keyword Hashing
Except Keyword Hashing all other are the techniques used in flexible string matching

Read More Answers.

Question # 60

Tell us what are the different Algorithm techniques in Machine Learning?

Answer:-

The different types of techniques in Machine Learning are
* a) Supervised Learning
* b) Unsupervised Learning
* c) Semi-supervised Learning
* d) Reinforcement Learning
* e) Transduction
* f) Learning to Learn

Read More Answers.

Question # 61

Tell me what are the two paradigms of ensemble methods?

Answer:-

The two paradigms of ensemble methods are
* a) Sequential ensemble methods
* b) Parallel ensemble methods

Read More Answers.

Question # 62

Tell me what are the five popular algorithms of Machine Learning?

Answer:-

* a) Decision Trees
* b) Neural Networks (back propagation)
* c) Probabilistic networks
* d) Nearest Neighbor
* e) Support vector machines

Read More Answers.

Question # 63

Tell me what is not Machine Learning?

Answer:-

* a) Artificial Intelligence
* b) Rule based inference

Read More Answers.

Question # 64

Tell me what is Inductive Logic Programming in Machine Learning?

Answer:-

Inductive Logic Programming (ILP) is a subfield of machine learning which uses logical programming representing background knowledge and examples.

Read More Answers.

Question # 65

Tell me what are support vector machines?

Answer:-

Support vector machines are supervised learning algorithms used for classification and regression analysis.

Read More Answers.

Question # 66

Tell me what is classifier in machine learning?

Answer:-

A classifier in a Machine Learning is a system that inputs a vector of discrete or continuous feature values and outputs a single discrete value, the class.

Read More Answers.

Question # 67

Do you know why overfitting happens?

Answer:-

The possibility of overfitting exists as the criteria used for training the model is not the same as the criteria used to judge the efficacy of a model.

Read More Answers.

Question # 68

Tell me what are Bayesian Networks (BN)?

Answer:-

Bayesian Network is used to represent the graphical model for probability relationship among a set of variables .

Read More Answers.

Question # 69

Tell me when to use ensemble learning?

Answer:-

Ensemble learning is used when you build component classifiers that are more accurate and independent from each other.

Read More Answers.

Question # 70

Tell me what are the advantages of Naive Bayes?

Answer:-

In Naive Bayes classifier will converge quicker than discriminative models like logistic regression, so you need less training data. The main advantage is that it can't learn interactions between features.

Read More Answers.

Question # 71

Do you know what is the standard approach to supervised learning?

Answer:-

The standard approach to supervised learning is to split the set of example into the training set and the test.

Read More Answers.

Question # 72

Tell us why ensemble learning is used?

Answer:-

Ensemble learning is used to improve the classification, prediction, function approximation etc of a model.

Read More Answers.

Question # 73

Tell me various approaches for machine learning?

Answer:-

The different approaches in Machine Learning are
* a) Concept Vs Classification Learning
* b) Symbolic Vs Statistical Learning
* c) Inductive Vs Analytical Learning

Read More Answers.

Question # 74

Tell us what is dimension reduction in Machine Learning?

Answer:-

In Machine Learning and statistics, dimension reduction is the process of reducing the number of random variables under considerations and can be divided into feature selection and feature extraction

Read More Answers.

Question # 75

Explain what is ensemble learning?

Answer:-

To solve a particular computational program, multiple models such as classifiers or experts are strategically generated and combined. This process is known as ensemble learning.

Read More Answers.

Question # 76

Tell me which method is frequently used to prevent overfitting?

Answer:-

When there is sufficient data 'Isotonic Regression' is used to prevent an overfitting issue.

Read More Answers.

Question # 77

Explain me what is Genetic Programming?

Answer:-

Genetic programming is one of the two techniques used in machine learning. The model is based on the testing and selecting the best choice among a set of results.

Read More Answers.

Question # 78

Please explain what is Perceptron in Machine Learning?

Answer:-

In Machine Learning, Perceptron is an algorithm for supervised classification of the input into one of several possible non-binary outputs.

Read More Answers.

Education and Science Most Popular Interview Topics.

1 : Mathematics Frequently Asked Interview Questions and Answers Guide.

2 : Biology Frequently Asked Interview Questions and Answers Guide.

3 : Metallurgy Frequently Asked Interview Questions and Answers Guide.

4 : Geography Frequently Asked Interview Questions and Answers Guide.

5 : Chemistry Frequently Asked Interview Questions and Answers Guide.

6 : Geology Frequently Asked Interview Questions and Answers Guide.

7 : Science Frequently Asked Interview Questions and Answers Guide.

8 : Ecology Frequently Asked Interview Questions and Answers Guide.

9 : Mass Communication Frequently Asked Interview Questions and Answers Guide.

10 : Biochemistry Frequently Asked Interview Questions and Answers Guide.

About Global Guideline.

Global Guideline is a platform to develop your own skills with thousands of job interview questions and web tutorials for fresher's and experienced candidates. These interview questions and web tutorials will help you strengthen your technical skills, prepare for the interviews and quickly revise the concepts. Global Guideline invite you to unlock your potentials with thousands of **Interview Questions with Answers** and much more. Learn the most common technologies at Global Guideline. We will help you to explore the resources of the World Wide Web and develop your own skills from the basics to the advanced. Here you will learn anything quite easily and you will really enjoy while learning. Global Guideline will help you to become a professional and Expert, well prepared for the future.

* This PDF was generated from https://GlobalGuideline.com at **November 29th, 2023**

* If any answer or question is incorrect or inappropriate or you have correct answer or you found any problem in this document then don't hesitate feel free and e-mail us we will fix it.

You can follow us on FaceBook for latest Jobs, Updates and other interviews material.
www.facebook.com/InterviewQuestionsAnswers

Follow us on Twitter for latest Jobs and interview preparation guides
https://twitter.com/InterviewGuide

Best Of Luck.

Global Guideline Team
https://GlobalGuideline.com
Info@globalguideline.com