# 100 QUESTIONS FOR DATA SCIENTIST INTERVIEW

## 1. Statistical Analysis

1. **What is hypothesis testing?**
   - **Answer**: Hypothesis testing is a method used to decide if there is enough evidence to reject a null hypothesis. It involves setting up two hypotheses (null and alternative), calculating a test statistic, and comparing it against a threshold (p-value) to determine the result.
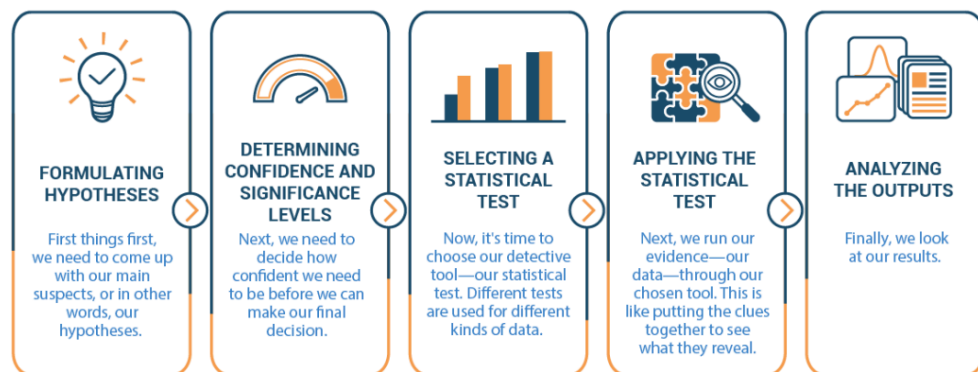


FORMULATING HYPOTHESES — First things first, we need to come up with our main suspects, or in other words, our hypotheses.

DETERMINING CONFIDENCE AND SIGNIFICANCE LEVELS — Next, we need to decide how confident we need to be before we can make our final decision.

SELECTING A STATISTICAL TEST — Now, it's time to choose our detective tool—our statistical test. Different tests are used for different kinds of data.

APPLYING THE STATISTICAL TEST — Next, we run our evidence—our data—through our chosen tool. This is like putting the clues together to see what they reveal.

ANALYZING THE OUTPUTS — Finally, we look at our results.

Image Source: Quanthub

2. **Explain the p-value in hypothesis testing.**
   - **Answer**: The p-value measures the probability of obtaining results at least as extreme as the observed ones, assuming the null hypothesis is true. A low p-value ($< 0.05$) indicates strong evidence against the null hypothesis.

3. **What is a z-test?**
   - **Answer**: A z-test is used to determine if there is a significant difference between sample and population means. It is often used when the sample size is large ($n > 30$) and the population variance is known.

4. **Describe the t-test.**
   - **Answer**: A t-test is used to determine if there is a significant difference between the means of two groups, especially when the sample size is small and the population variance is unknown.

5. **What is a Gaussian distribution?**
   - **Answer**: A Gaussian distribution, also known as the normal distribution, is a continuous probability distribution characterized by its bell-shaped curve, mean, and standard deviation. It is a symmetrical, bell-shaped distribution in

which the mean, median, and mode are all equal. It describes many natural phenomena.

**Probability density function**



The red curve is the *standard normal distribution*.
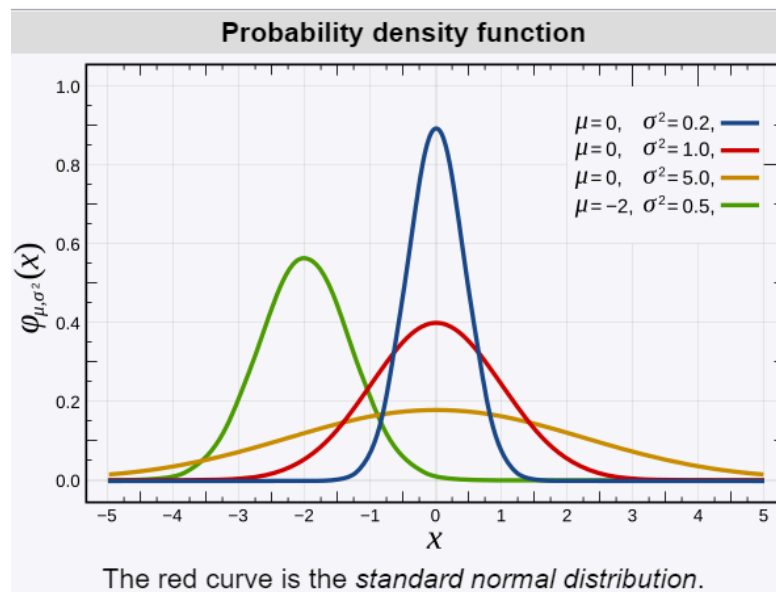
6. **Explain the Central Limit Theorem (CLT).**
   - **Answer**: The CLT states that the distribution of the sample mean approximates a normal distribution as the sample size becomes large, regardless of the population's distribution, provided the samples are independent and identically distributed.

7. **What is a binomial distribution?**
   - **Answer**: A binomial distribution describes the number of successes in a fixed number of independent Bernoulli trials, with each trial having two possible outcomes (success or failure) and a constant probability of success.

$$ P_x = \binom{n}{x} p^x q^{n-x} $$

| $P$ | = binomial probability |
|---|---|
| $x$ | = number of times for a specific outcome within n trials |
| $\binom{n}{x}$ | = number of combinations |
| $p$ | = probability of success on a single trial |
| $q$ | = probability of failure on a single trial |
| $n$ | = number of trials |

8. **What is the difference between Type I and Type II errors?**

- ○ **Answer**: A Type I error occurs when the null hypothesis is wrongly rejected (false positive), while a Type II error happens when the null hypothesis is wrongly accepted (false negative).

| Null hypothesis is ... | True | False |
|---|---|---|
| Rejected | Type I error<br>False positive<br>Probability = α | Correct decision<br>True positive<br>Probability = 1 – β |
| Not rejected | Correct decision<br>True negative<br>Probability = 1 – α | Type II error<br>False negative<br>Probability = β |

9. **Explain the concept of confidence intervals.**
   - ○ **Answer**: Confidence intervals provide a range of values within which the true population parameter is expected to lie, with a certain level of confidence (e.g., 95%).

$$CI = \bar{x} \pm z \frac{s}{\sqrt{n}}$$

$CI$ = confidence interval

$\bar{x}$ = sample mean

$z$ = confidence level value

$s$ = sample standard deviation

$n$ = sample size

10. **What is the Chi-square test?**
    - ○ **Answer**: A chi-square test is used to help determine if observed results are in line with expected results and to rule out that observations are due to chance. A chi-square test is appropriate for this when the data being analyzed is from a random sample, and when the variable in question is a categorical variable. The test assumes (null hypothesis) that the variables are independent. The the model does not fit, then that proves that the variables are dependent.

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$\chi^2$ = chi squared

$O_i$ = observed value

$E_i$ = expected value

## 2. Probability Questions

11. **What is Bayes' Theorem?**
    - **Answer**: Bayes' Theorem calculates the probability of an event based on prior knowledge of conditions related to the event. It is formulated as

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

$A, B$ = events

$P(A|B)$ = probability of A given B is true

$P(B|A)$ = probability of B given A is true

$P(A), P(B)$ = the independent probabilities of A and B

.

12. **Explain the concept of conditional probability.**
    - **Answer**: Conditional probability is the probability of an event occurring given that another event has already occurred, represented as P(A|B).

13. **What are the key assumptions of Naive Bayes?**
    - **Answer**: Naive Bayes assumes that features are conditionally independent given the class label and that each feature contributes equally to the outcome. The Naïve Bayes classifier is a supervised machine learning algorithm that is used for classification tasks such as text classification.

$$p(C_k \mid \mathbf{x}) = \frac{p(C_k)\, p(\mathbf{x} \mid C_k)}{p(\mathbf{x})}$$

$$\text{posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}}$$

14. **How does the Naive Bayes classifier handle continuous data?**
    - **Answer**: For continuous data, Naive Bayes typically assumes a Gaussian distribution and uses the probability density function to estimate the likelihood of the data given a class.

15. **What is the law of total probability?**
    ○ **Answer**: The law of total probability states that the total probability of an outcome can be found by considering all possible ways in which the outcome can occur. It is expressed as:

$$P(A) = \sum_{n} P(A \cap B_n)$$

$P$ = probability

$A$ = any event

$B_n$ = event

16. **Explain the difference between discrete and continuous random variables.**
    ○ **Answer**: Discrete random variables take on a finite or countable number of values (e.g., number of heads in coin tosses), while continuous random variables take on an infinite number of values within a range (e.g., heights of people).

17. **What is the difference between joint probability and marginal probability?**
    ○ **Answer**: Joint probability is the probability of two events occurring together, while marginal probability is the probability of a single event occurring, irrespective of other events.

18. **What is a Markov Chain?**
    ○ **Answer**: A Markov Chain is a stochastic process where the future state depends only on the current state and not on the sequence of events that preceded it, known as the memoryless property.

19. **What is the difference between prior and posterior probability?**
    ○ **Answer**: Prior probability is the initial probability of an event before new data is considered, while posterior probability is the updated probability after considering new data.

20. **Explain the concept of entropy in probability.**
    ○ **Answer**: Entropy is a measure of uncertainty or randomness in a probability distribution. Higher entropy indicates more unpredictability.

# 3. Supervised Learning

21. **What is supervised learning?**
    ○ **Answer**: Supervised learning is a type of machine learning where the model is trained on labeled data, the model learns to map input features to output labels.
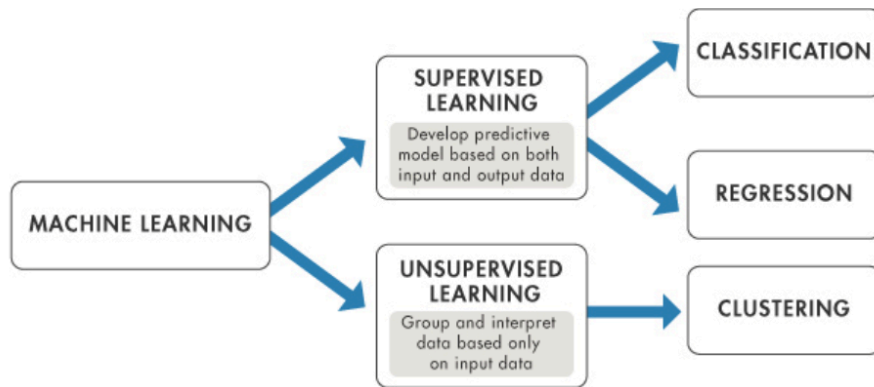
Image Source: Mathworks

22. **What are the main types of supervised learning models?**
    ○ **Answer**: The main types are regression (predicting continuous values) and classification (predicting discrete labels).

23. **Explain the concept of overfitting.**
    ○ **Answer**: Overfitting occurs when a model learns the training data too well, capturing noise and outliers, leading to poor generalization to new, unseen data.
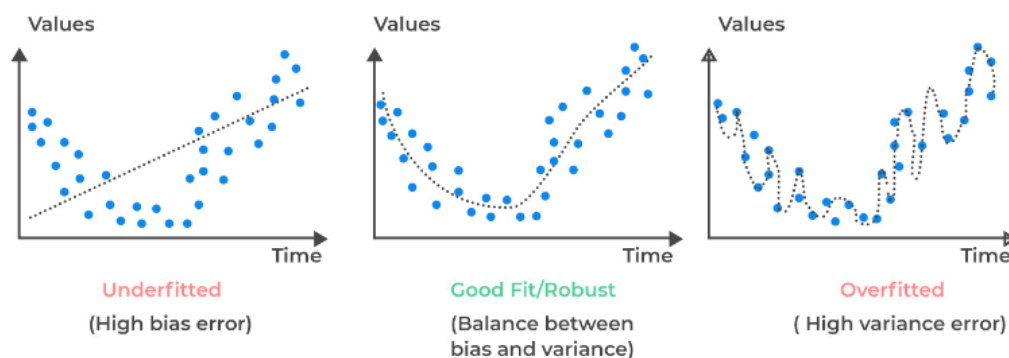


24. **What is feature engineering?**
    ○ **Answer**: Feature engineering involves creating or transforming features to improve the performance of a machine learning model, often using domain knowledge.

25. **What is a cost function?**
    ○ **Answer**: A cost function measures the error between the predicted values and the actual values, guiding the optimization process to minimize this error.

    **Example Cost Functions:**
    Regression: MSE, MAE, Huber Loss.
    Classification: Cross-Entropy Loss, Hinge Loss, Focal Loss.

26. **What are hyperparameters in machine learning?**
    ○ **Answer**: Hyperparameters are settings used to control the training process of a model (e.g., learning rate, number of trees in a random forest) and are typically set before training.

27. **Explain the concept of regularization.**
    ○ **Answer**: Regularization adds a penalty to the cost function to prevent overfitting by discouraging overly complex models, with techniques like L1 (Lasso) and L2 (Ridge) regularization.

28. **What is cross-validation?**
    ○ **Answer**: Cross-validation is a technique to assess the performance of a model by dividing the data into training and validation sets multiple times, ensuring the model generalizes well to unseen data.
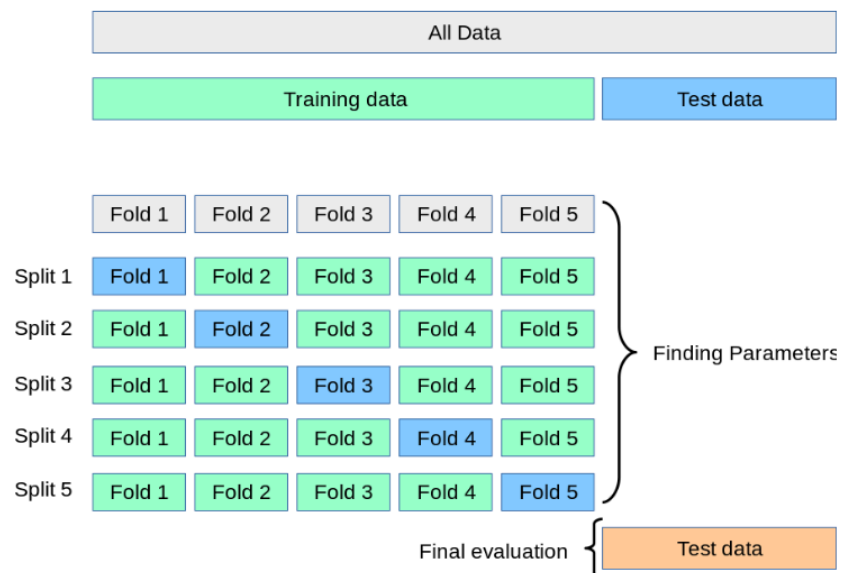


29. **How do you evaluate the performance of a classification model?**
    ○ **Answer**: Performance can be evaluated using metrics such as accuracy, precision, recall, F1 score, confusion matrix, and ROC-AUC curve.

30. **What is the R-squared value in regression?**
    ○ **Answer**: The R-squared value measures the proportion of variance in the dependent variable that is predictable from the independent variables, indicating the goodness of fit of the model. For example, a model with an R-squared value of 0.91 means that approximately 91% of the variance in the dependent variable is explained by the independent variables.

31. **What is the difference between precision and recall?**
    ○ **Answer**: Precision is the ratio of true positive predictions to the total predicted positives, while recall is the ratio of true positive predictions to the total actual

positives.

32. **Explain the concept of the F1 score.**
   - **Answer**: The F1 score is the harmonic mean of precision and recall, providing a balance between the two metrics and giving a single performance measure.

33. **What is a confusion matrix?**
   - **Answer**: A confusion matrix is a table that summarizes the performance of a classification model by showing the true positives, true negatives, false positives, and false negatives.



|  |  | Predicted Class | | |
|---|---|---|---|---|
|  |  | Positive | Negative |  |
| Actual Class | Positive | True Positive (TP) | False Negative (FN) <br> Type II Error | Sensitivity <br> $\dfrac{TP}{(TP+FN)}$ |
|  | Negative | False Positive (FP) <br> Type I Error | True Negative (TN) | Specificity <br> $\dfrac{TN}{(TN+FP)}$ |
|  |  | Precision <br> $\dfrac{TP}{(TP+FP)}$ | Negative Predictive Value <br> $\dfrac{TN}{(TN+FN)}$ | Accuracy <br> $\dfrac{TP+TN}{(TP+TN+FP+FN)}$ |

Confusion Matrix

34. **Describe the ROC-AUC curve.**
   - **Answer**: The ROC curve plots the true positive rate against the false positive rate at various threshold settings, and the AUC (Area Under the Curve) measures the overall performance of the model. In general, AUC values are interpreted as follows: 0.5-0.6 (failed), 0.6-0.7 (worthless), 0.7-0.8 (poor), 0.8-0.9 (good), > 0.9 (excellent).
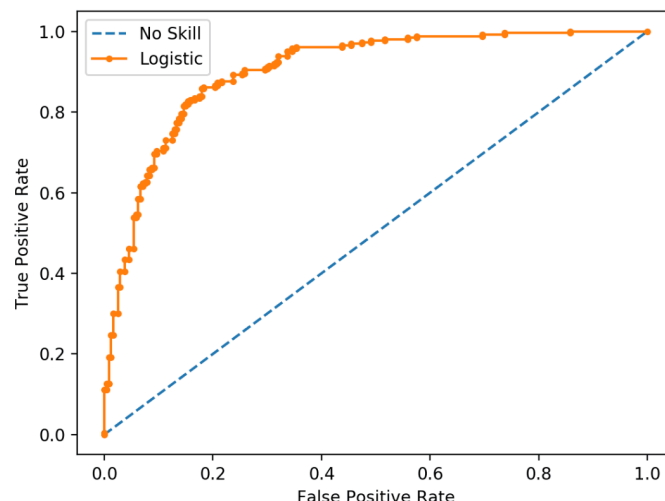
### 35. **What is the difference between linear and logistic regression?**
  - ○ **Answer**: Linear regression predicts continuous values using a linear relationship between input features and the target variable, while logistic regression predicts binary outcomes using a logistic function to model the probability of the target variable.

### 36. **What is gradient descent?**
  - ○ **Answer**: Gradient descent is an optimization algorithm used to minimize the cost function by iteratively updating the model parameters in the direction of the steepest descent of the cost function.

### 37. **Explain the concept of learning rate.**
  - ○ **Answer**: The learning rate is a hyperparameter that controls the step size during the optimization process, determining how quickly or slowly the model learns.

### 38. **What is the difference between batch and stochastic gradient descent?**
  - ○ **Answer**: Batch gradient descent updates the model parameters using the entire training dataset, while stochastic gradient descent updates the parameters using one sample at a time, providing faster but noisier updates.

### 39. **Describe the concept of early stopping.**
  - ○ **Answer**: Early stopping is a regularization technique where the training process is halted when the model's performance on the validation set starts to deteriorate, preventing overfitting.

### 40. **What is a decision tree?**
  - ○ **Answer**: A decision tree is a supervised learning model that splits the data into subsets based on feature values, forming a tree structure where each node represents a feature and each branch represents a decision.

# Elements of a decision tree
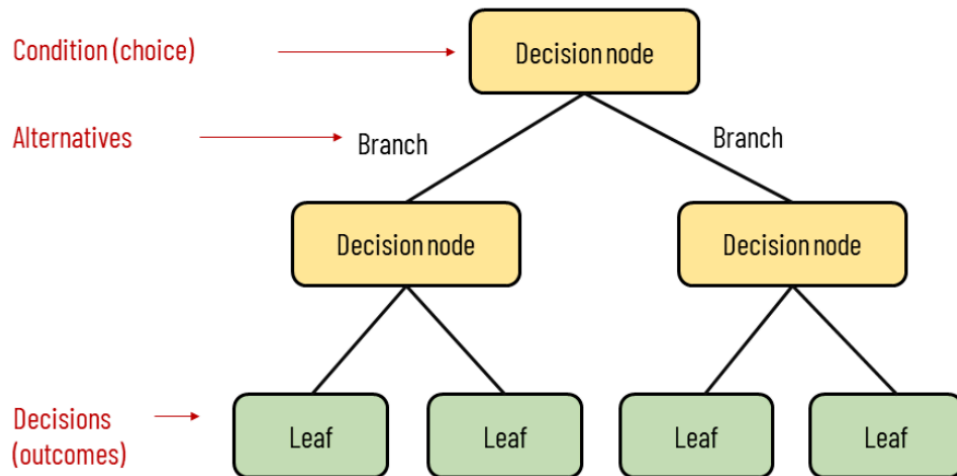


## 4. Unsupervised Learning

41. **What is unsupervised learning?**
    ○ **Answer**: Unsupervised learning is a type of machine learning where the model is trained on unlabeled data, discovering hidden patterns or structures in the data.

42. **What are the different types of clustering algorithms?**
    ○ **Answer**: Common clustering algorithms include K-means, hierarchical clustering, DBSCAN, and Gaussian Mixture Models.

43. **What is the K-means clustering algorithm?**
    ○ **Answer**: K-means is a clustering algorithm that partitions the data into K clusters by iteratively assigning data points to the nearest cluster center and updating the cluster centers based on the mean of the assigned points.
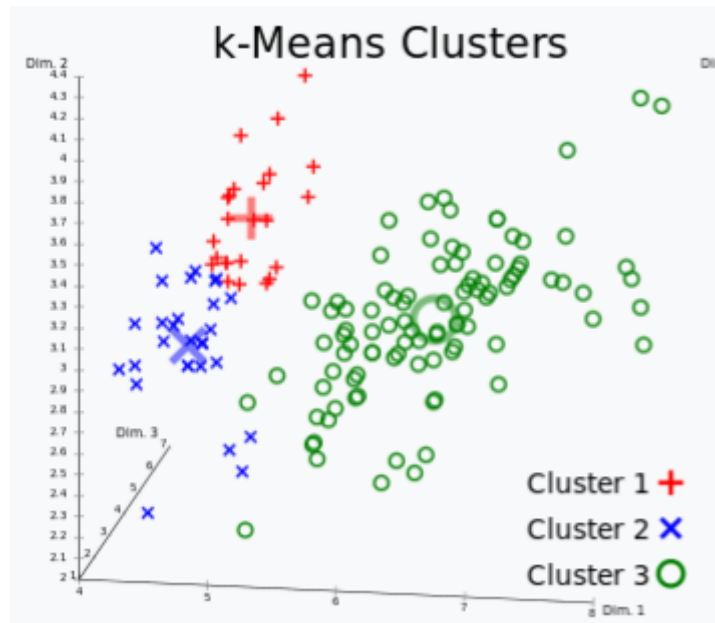
k-Means Clusters

Image Source: Wiki

44. **How do you determine the value of K in K-means clustering?**
    ○ **Answer**: The value of K can be determined using the Elbow method, where the sum of squared distances from each point to its assigned cluster center is plotted for different values of K, and the point where the curve bends (elbow) indicates the optimal K.
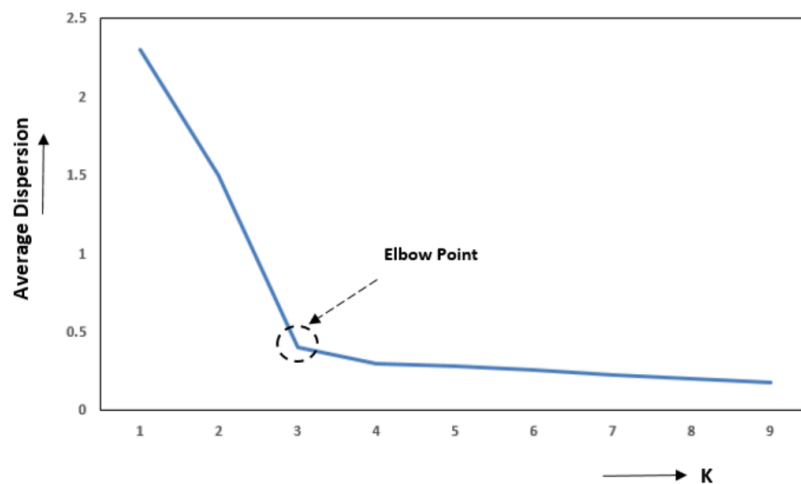


Image Source: O'reilly

45. **Explain hierarchical clustering.**
    ○ **Answer**: Hierarchical clustering builds a hierarchy of clusters by either iteratively merging smaller clusters into larger ones (agglomerative) or splitting larger clusters into smaller ones (divisive).

Agglomerative          Divisive

46. **What is the DBSCAN algorithm?**
   ○ **Answer**: DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a clustering algorithm that groups together points that are closely packed, marking points in low-density regions as outliers.
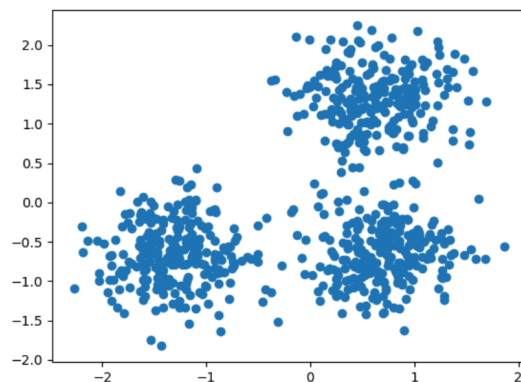


Image Source: Wiki

47. **What is a Gaussian Mixture Model (GMM)?**
   ○ **Answer**: A GMM is a probabilistic model that assumes the data is generated from a mixture of several Gaussian distributions, each representing a cluster.

48. **Explain the concept of the Silhouette Score.**
   ○ **Answer**: The Silhouette Score measures how similar an object is to its own cluster compared to other clusters, ranging from -1 to 1, with higher values indicating better clustering performance.
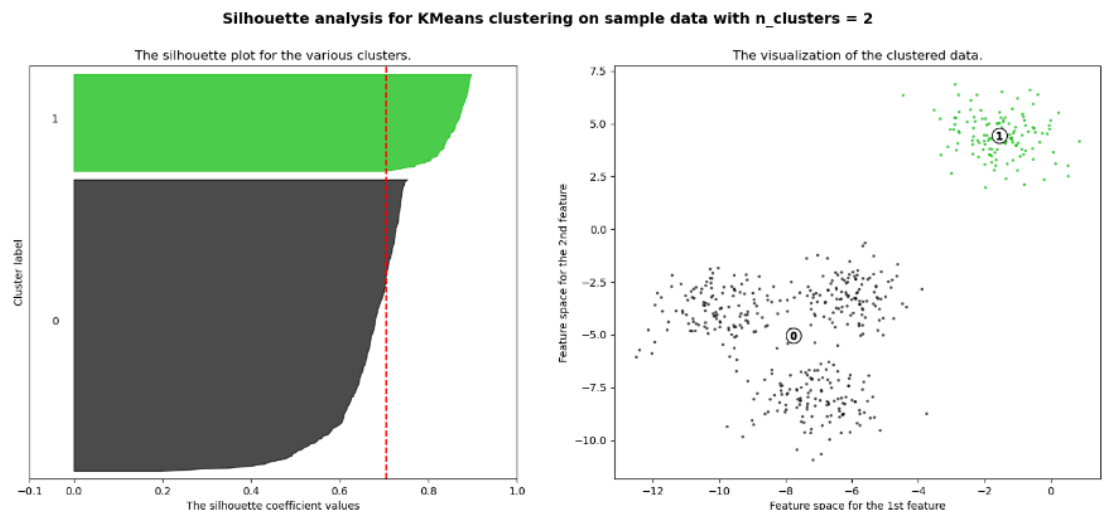
Image Source: Sci-kit Learn

49. **What is the Adjusted Rand Index?**
   ○ **Answer**: The Adjusted Rand Index measures the similarity between two data clusterings by considering all pairs of samples and counting pairs that are assigned in the same or different clusters in the predicted and true clusterings, adjusted for chance.

50. **Describe the Davies-Bouldin Index.**
   ○ **Answer**: The Davies-Bouldin Index evaluates clustering quality by measuring the average similarity ratio of each cluster with its most similar cluster, with lower values indicating better clustering.

51. **What is Principal Component Analysis (PCA)?**
   ○ **Answer**: PCA is a dimensionality reduction technique that transforms the data into a new coordinate system, where the first principal component has the largest variance, and each succeeding component has the highest variance possible under the constraint that it is orthogonal to the preceding components.

52. **Explain the concept of eigenvalues and eigenvectors in PCA.**
   ○ **Answer**: In PCA, eigenvalues represent the variance captured by each principal component, while eigenvectors define the direction of the principal components in the feature space.

53. **What is t-SNE?**
   ○ **Answer**: t-SNE (t-Distributed Stochastic Neighbor Embedding) is a dimensionality reduction technique that visualizes high-dimensional data by mapping it to a lower-dimensional space, preserving the local structure of the data.

54. **What is the difference between PCA and LDA?**

- ○ **Answer**: PCA is an unsupervised technique that reduces dimensionality by maximizing variance, while LDA (Linear Discriminant Analysis) is a supervised technique that maximizes the separation between classes.

55. **Explain the concept of clustering validation.**
    - ○ **Answer**: Clustering validation assesses the quality of clustering results using internal metrics (e.g., Silhouette Score), external metrics (e.g., Adjusted Rand Index), and relative validation (e.g., comparing different clustering methods).

56. **What is a dendrogram?**
    - ○ **Answer**: A dendrogram is a tree-like diagram that visualizes the arrangement of clusters produced by hierarchical clustering, showing how clusters are merged or split at each level of the hierarchy.
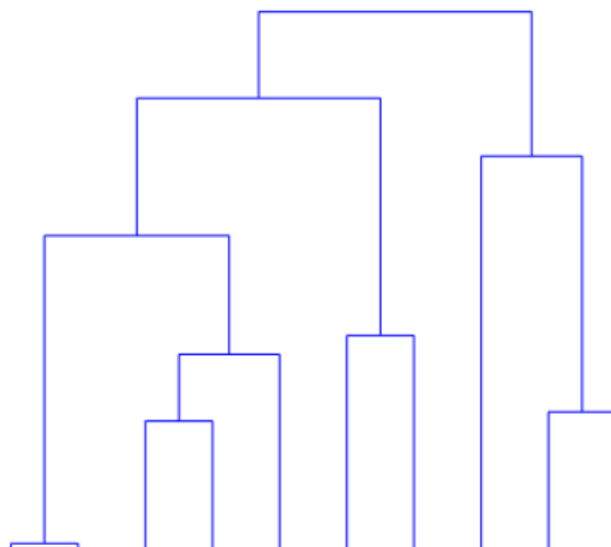


57. **How do you handle categorical data in clustering?**
    - ○ **Answer**: Categorical data can be handled using techniques such as one-hot encoding, label encoding, or using algorithms like k-prototypes that can handle mixed data types.

58. **What is the difference between hard and soft clustering?**
    - ○ **Answer**: In hard clustering, each data point is assigned to a single cluster, while in soft clustering, each data point can belong to multiple clusters with varying degrees of membership.

59. **Explain the concept of anomaly detection in unsupervised learning.**
    - ○ **Answer**: Anomaly detection identifies rare or unusual patterns in data that do not conform to expected behavior, often using clustering or density-based methods to detect outliers.

60. **What is the purpose of dimensionality reduction?**

- ○ **Answer**: Dimensionality reduction aims to reduce the number of features in a dataset while preserving important information, improving model performance, reducing computational complexity, and mitigating the curse of dimensionality.

# 5. Natural Language Processing (NLP)

61. **What are the common preprocessing steps in NLP?**
    - ○ **Answer**: Common preprocessing steps include tokenization, stemming, lemmatization, removing stop words, and converting text to lowercase.

62. **What is tokenization?**
    - ○ **Answer**: Tokenization is the process of splitting text into individual units, such as words or phrases, which are called tokens.

63. **Explain the concept of word embeddings.**
    - ○ **Answer**: Word embeddings are vector representations of words that capture their semantic meaning by placing similar words close to each other in a high-dimensional space.

64. **What is BERT?**
    - ○ **Answer**: BERT (Bidirectional Encoder Representations from Transformers) is a pre-trained NLP model that uses transformers to understand the context of a word based on all surrounding words in a sentence.

65. **Describe the difference between stemming and lemmatization.**
    - ○ **Answer**: Stemming reduces words to their base or root form by removing suffixes, while lemmatization reduces words to their base or dictionary form, considering the word's meaning and context.
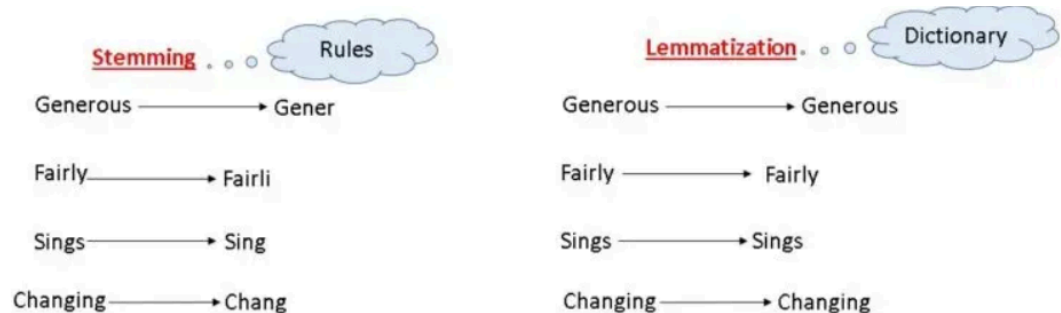


66. **What is a bag-of-words model?**
    - ○ **Answer**: A bag-of-words model represents text data as a collection of word frequencies, ignoring the order and context of the words, often used in text classification and information retrieval.

67. **Explain the term TF-IDF.**

○ **Answer**: TF-IDF (Term Frequency-Inverse Document Frequency) is a weighting scheme that reflects the importance of a word in a document relative to a collection of documents, calculated as the product of term frequency and inverse document frequency. An example of implementation TF-IDF is given below:

```python
from sklearn.feature_extraction.text import TfidfVectorizer

# Sample documents
documents = [
    "I love programming in Python",
    "Python is great for machine learning",
    "Machine learning is a growing field",
    "I enjoy learning new things"
]

# Create TF-IDF Vectorizer
tfidf_vectorizer = TfidfVectorizer()

# Fit and transform the documents
tfidf_matrix = tfidf_vectorizer.fit_transform(documents)

# Display TF-IDF matrix
print("TF-IDF Matrix:")
print(tfidf_matrix.toarray())

# Display feature names
print("Feature Names:")
print(tfidf_vectorizer.get_feature_names_out())
```

The output of the code will be:

```
TF-IDF Matrix:
[[0.         0.46979135 0.         0.         0.46979135 0.
  0.46979135 0.46979135 0.         0.         0.         0.46979135]
 [0.         0.         0.47363118 0.         0.47363118 0.47363118
  0.         0.         0.47363118 0.         0.         0.        ]
 [0.         0.         0.         0.6065106  0.         0.6065106
  0.         0.         0.         0.         0.         0.        ]
 [0.50576228 0.         0.         0.         0.         0.
  0.         0.         0.         0.50576228 0.50576228 0.        ]]
Feature Names:
['enjoy' 'for' 'great' 'growing' 'in' 'is' 'learning' 'love' 'machine'
 'new' 'programming' 'python']
```

68. **What is the purpose of text normalization?**
○ **Answer**: Text normalization standardizes text data by converting it to a consistent format, including lowercasing, removing punctuation, and handling contractions, improving the accuracy of NLP models.

69. **What is named entity recognition (NER)?**
    ○ **Answer**: NER is an NLP task that identifies and classifies named entities (e.g., people, organizations, locations) in text into predefined categories.

70. **Explain the concept of sentiment analysis.**
    ○ **Answer**: Sentiment analysis is the process of determining the emotional tone or opinion expressed in text, classifying it as positive, negative, or neutral.

71. **What is the difference between rule-based and machine learning-based NLP methods?**
    ○ **Answer**: Rule-based methods use handcrafted linguistic rules to process text, while machine learning-based methods use statistical models to learn patterns from labeled data.

72. **Describe the architecture of a transformer model.**
    ○ **Answer**: Transformer models use self-attention mechanisms to weigh the importance of different words in a sentence, consisting of an encoder to process input text and a decoder to generate output text, with parallel processing enabling efficient training.
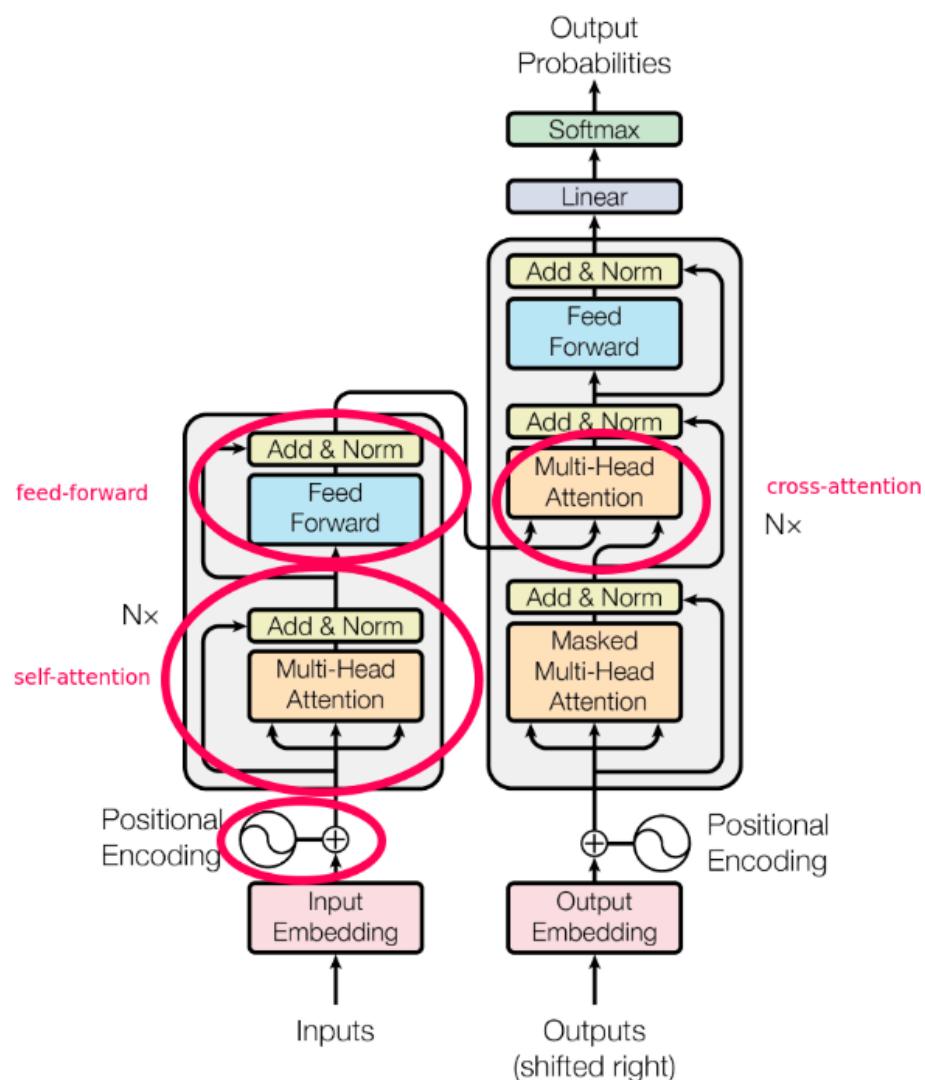
73. **What is transfer learning in NLP?**
    ○ **Answer**: Transfer learning in NLP involves using pre-trained models on large corpora and fine-tuning them on specific tasks, leveraging learned language representations to improve performance and reduce training time.

74. **Explain the concept of language modeling.**
    ○ **Answer**: Language modeling involves predicting the probability distribution of words or sequences of words in a language, often used for tasks like text generation and machine translation.

75. **What is the difference between generative and discriminative models in NLP?**
    ○ **Answer**: Generative models learn the joint probability distribution of input and output and can generate new data, while discriminative models learn the conditional probability distribution of the output given the input and are used for classification tasks.

76. **What are attention mechanisms in NLP?**
    ○ **Answer**: Attention mechanisms enable models to focus on relevant parts of the input sequence when generating output, improving the handling of long-range dependencies and enhancing performance on tasks like machine translation.

77. **Describe the sequence-to-sequence (Seq2Seq) model.**
    ○ **Answer**: Seq2Seq models are used for tasks where the input and output are sequences, such as machine translation and text summarization, consisting of an encoder to process the input sequence and a decoder to generate the output sequence.

78. **What is the purpose of word2vec?**
    ○ **Answer**: Word2vec is a word embedding model that learns vector representations of words from large text corpora, capturing semantic relationships by placing similar words close together in a continuous vector space.

79. **Explain the concept of an LSTM network.**
    ○ **Answer**: Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) designed to capture long-range dependencies and avoid the vanishing gradient problem, using memory cells and gating mechanisms to regulate information flow.

80. **What is the role of a tokenizer in NLP?**
    ○ **Answer**: A tokenizer converts text into tokens, which are the basic units of meaning (e.g., words, subwords) used by NLP models for processing and analysis.

# 6. Image Processing

81. **What is image processing?**
    ○ **Answer**: Image processing involves manipulating and analyzing images to extract meaningful information, enhance quality, and perform tasks like segmentation, detection, and recognition.

82. **Explain the concept of convolution in image processing.**
    ○ **Answer**: Convolution is a mathematical operation used to apply filters to images, detecting features like edges, textures, and patterns by sliding the filter across the image and computing the dot product.

83. **What is a convolutional neural network (CNN)?**
    ○ **Answer**: A CNN is a deep learning model designed for image processing tasks, using convolutional layers to automatically learn hierarchical features from raw pixel data, followed by pooling and fully connected layers for classification.

84. **Describe the role of pooling layers in CNNs.**
    ○ **Answer**: Pooling layers reduce the spatial dimensions of feature maps, retaining important information while reducing computational complexity and providing translation invariance.
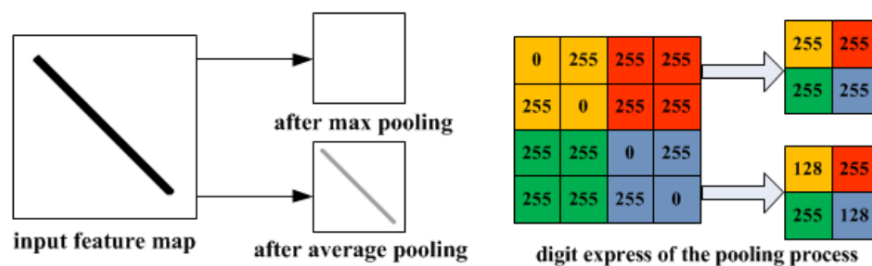


Image source: https://towardsai.net/p/l/introduction-to-pooling-layers-in-cnn
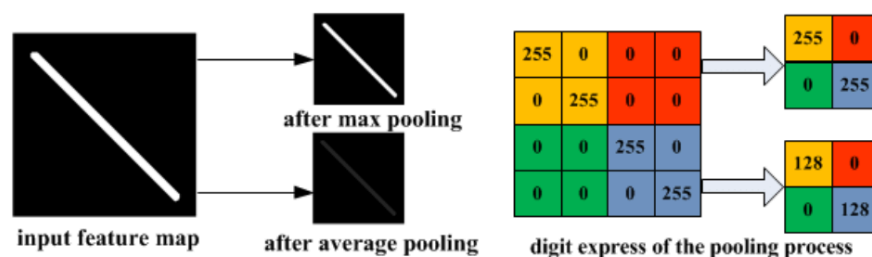
85. **What is the purpose of data augmentation in image processing?**
    ○ **Answer**: Data augmentation artificially increases the diversity of the training dataset by applying transformations like rotation, scaling, flipping, and

cropping, improving the generalization ability of image models.

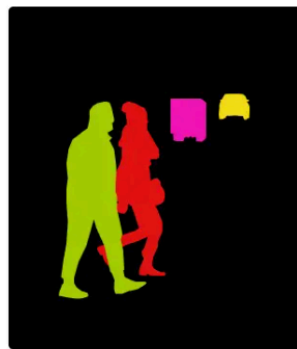86. **Explain the concept of transfer learning in image processing.**
    ○ **Answer**: Transfer learning involves using pre-trained CNN models on large image datasets and fine-tuning them on specific tasks, leveraging learned features to improve performance and reduce training time.
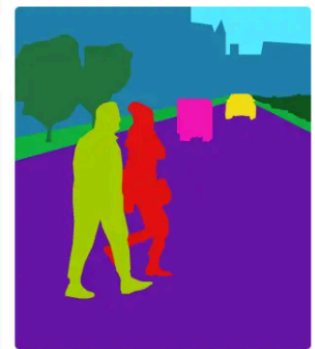
87. **What is image segmentation?**
    ○ **Answer**: Image segmentation is the process of dividing an image into meaningful regions or segments, often used for object detection, medical imaging, and scene understanding.



SEMANTIC IMAGE SEGMENTATION    INSTANCE SEGMENTATION    PANOPTIC SEGMENTATION

88. **Describe the difference between semantic and instance segmentation.**
    ○ **Answer**: Semantic segmentation classifies each pixel in an image into a predefined category, while instance segmentation distinguishes between different instances of the same category, providing both class and instance information.

89. **What is the purpose of object detection?**
    ○ **Answer**: Object detection identifies and localizes objects within an image, providing bounding boxes and class labels, enabling applications like autonomous driving, surveillance, and image search.
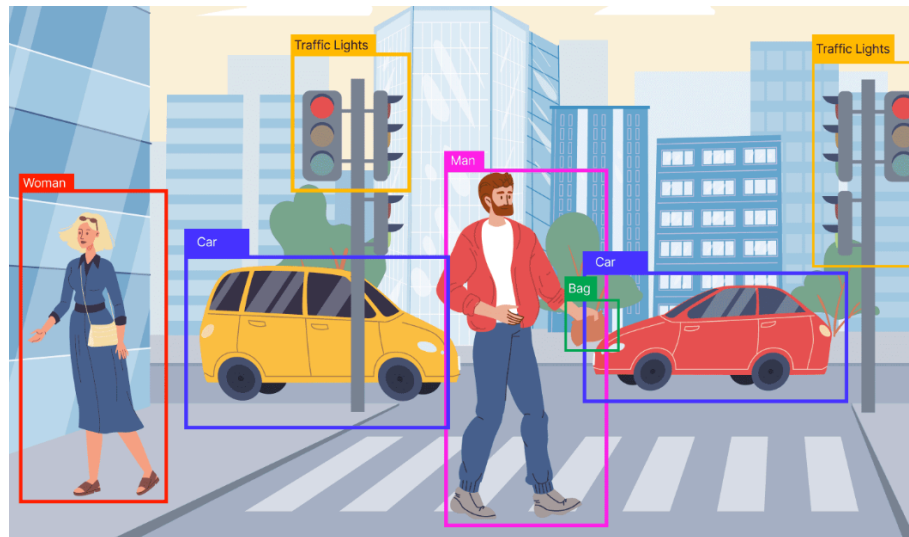
Image Source:
https://deeplobe.ai/exploring-object-detection-applications-and-benefits/

90. **Explain the concept of a sliding window in object detection.**
   - **Answer**: The sliding window technique involves moving a fixed-size window across the image, extracting features and applying a classifier at each position to detect objects, often used in traditional object detection methods.

91. **What is the YOLO algorithm?**
   - **Answer**: YOLO (You Only Look Once) is a real-time object detection algorithm that divides the image into a grid, predicting bounding boxes and class probabilities simultaneously for each grid cell, achieving high speed and accuracy.
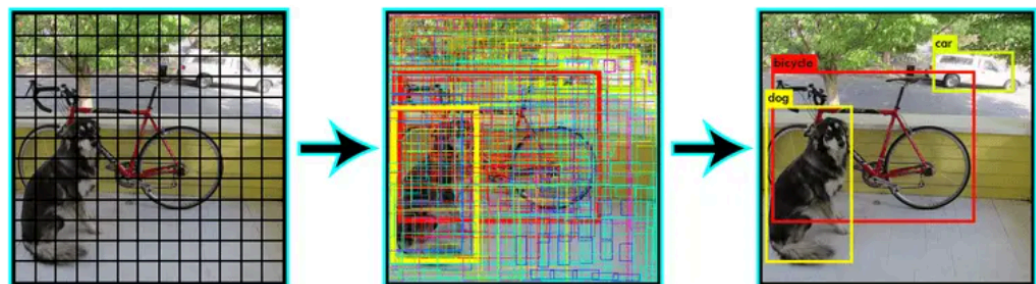


92. **Describe the concept of a generative adversarial network (GAN).**
   - **Answer**: A GAN is a deep learning model consisting of a generator and a discriminator, where the generator creates realistic data samples and the discriminator distinguishes between real and fake samples, training both components adversarially.

93. **What is style transfer in image processing?**
   - **Answer**: Style transfer is a technique that combines the content of one image with the style of another, generating a new image that preserves the content

while adopting the artistic style.

94. **Explain the purpose of edge detection.**
    ○ **Answer**: Edge detection identifies significant boundaries and transitions in an image, highlighting the edges of objects and structures, often used as a preprocessing step in image analysis tasks.

95. **What is the role of a filter in image processing?**
    ○ **Answer**: Filters are used to enhance or suppress specific features in an image, such as noise reduction, edge enhancement, and blurring, by applying convolution operations with predefined kernels.

96. **Describe the concept of histogram equalization.**
    ○ **Answer**: Histogram equalization improves the contrast of an image by redistributing the intensity values, creating a uniform histogram and enhancing details in areas with poor contrast.

97. **What is optical character recognition (OCR)?**
    ○ **Answer**: OCR is a technology that converts scanned images of text or printed documents into machine-readable text, enabling applications like digitization, text search, and automated data entry.

98. **Explain the concept of a heatmap in image processing.**
    ○ **Answer**: A heatmap visualizes the intensity or importance of different regions in an image, often used to highlight areas of interest, attention, or activation in models like CNNs.
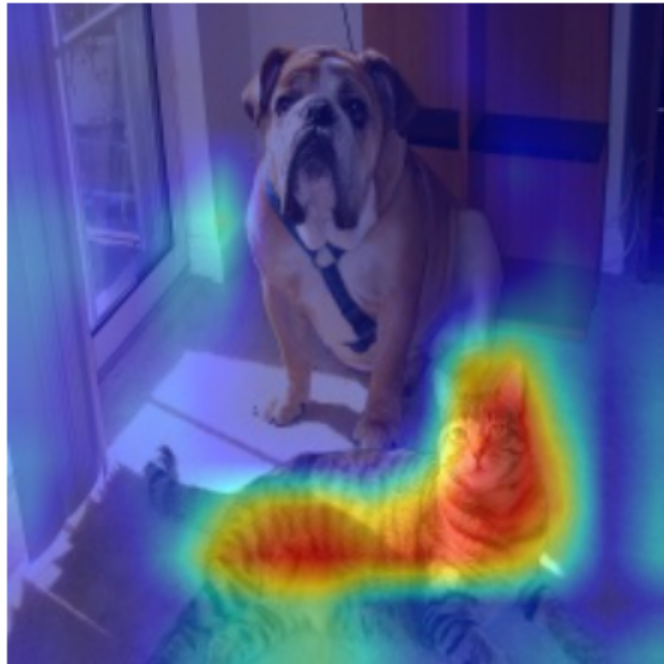


In this image, from jacobgil/pytorch-grad-cam, a cat is highlighted in red for the class "Cat," indicating that the network is looking at the right place when

making the classification decision.

99. **What is the purpose of image restoration?**
    ○ **Answer**: Image restoration aims to recover the original image from degraded or noisy versions, using techniques like deblurring, denoising, and inpainting to improve image quality.

100. **Describe the concept of super-resolution in image processing.**
    ○ **Answer**: Super-resolution is the process of enhancing the resolution of an image, generating high-resolution images from low-resolution inputs using techniques like interpolation, deep learning, and GANs.