GENERATIVE AI

**Can you explain the difference between generative and non-generative AI?**

* Generative AI creates new data similar to the training data (e.g., text, images), while non-generative AI focuses on tasks like classification, prediction, and decision-making.

**What machine learning techniques are commonly used for medical image analysis?**

* Techniques like convolutional neural networks (CNNs) are commonly used due to their effectiveness in processing and analyzing visual data.

The AI that detects breast cancer 4 years before it develops is an example of non-generative AI.

You have credit card history so based on this data non-generateive AI will decide where provide a loan or not so here not creating a new content you have a data based on that you are making certain decisions

GENERATIVE AI creates a new content / data example: chatgpt

RECURRENT NEURAL NETWORK: translation ex: google translate

LANGUAGE MODEL: is an AI model that can PREDICT THE NEXT WORD (OR SET OF WORDS.) for a given sequence of words.

When you want to train a LANGUAGE MODEL you don’t need lot of labeled data

Self-supervised learning is a type of **machine learning where the system learns to predict part of its input from other parts of the input without requiring labeled data.** It is a form of unsupervised learning where the model generates its own labels based on the inherent structure of the input data. This technique leverages the vast amount of unlabeled data available to create useful representations that can later be fine-tuned with smaller amounts of labeled data for specific tasks.

**Key Characteristics of Self-Supervised Learning:**

1. **No Labeled Data Required**: Self-supervised learning does not rely on manually labeled datasets. Instead, it generates pseudo-labels from the data itself.
2. **Pretext Tasks**: The model is trained on a pretext task, which is a proxy task designed to help the model learn useful representations. Examples include predicting the next word in a sentence or predicting missing parts of an image.
3. **Transfer Learning**: The representations learned from the pretext task can be transferred and fine-tuned for downstream tasks, such as classification or regression, often with limited labeled data.

**Examples of Self-Supervised Learning:**

**Natural Language Processing (NLP):**

* **Language Modeling**: **Training models like GPT and BERT** to predict the next word or the masked word in a sentence. For instance, BERT uses a masked language model (MLM) objective, where some words in a sentence are randomly masked, and the model learns to predict these masked words.

Example:

* + Input: "The cat sat on the [MASK]."
  + Task: Predict the masked word "mat."

**Benefits of Self-Supervised Learning:**

1. **Utilizes Unlabeled Data**: Makes use of vast amounts of unlabeled data, which is easier and cheaper to obtain than labeled data.
2. **Improves Representation Learning: Helps in learning rich and meaningful representations** that can be useful for various downstream tasks.
3. **Reduces Dependence on Labeled Data**: **Reduces the need for large labeled datasets,** which can be expensive and time-consuming to create.

**Real-World Applications:**

1. **Natural Language Processing**: Pretrained language models like BERT, GPT-3, and T5 have revolutionized NLP by learning from vast amounts of text data in a self-supervised manner and achieving state-of-the-art results in various NLP tasks.
2. **Computer Vision**: Self-supervised techniques are used to pretrain models on large image datasets, which can then be fine-tuned for specific tasks like object detection and image classification with limited labeled data.
3. **Speech Processing**: Models can be trained to predict future audio frames from past frames, aiding in tasks like speech recognition and speaker identification.

**Interview Questions and Answers:**

1. **What is self-supervised learning?**
   * **It is a type of machine learning where the system learns to predict part of its input from other parts of the input without requiring labeled data**
2. **How does self-supervised learning differ from supervised learning?**
   * Unlike supervised learning, which requires labeled data, self-supervised learning generates its own labels from the input data, leveraging unlabeled data.
3. **What are pretext tasks in self-supervised learning?**
   * Pretext tasks are proxy tasks designed to help the model learn useful representations from the data. Examples include predicting the next word in a sentence or completing a missing part of an image.
4. **Can you give an example of a pretext task in NLP?**
   * In NLP, a common pretext task is masked language modeling, where certain words in a sentence are masked, and the model learns to predict these masked words.
5. **What is contrastive learning in the context of self-supervised learning?**
   * Contrastive learning involves training the model to distinguish between different views or augmentations of the same data, aiming to bring representations of similar data closer and push representations of different data apart.
6. **How is self-supervised learning beneficial in real-world applications?**
   * It utilizes vast amounts of unlabeled data to learn rich representations, reduces dependence on labeled data, and improves performance on various downstream tasks through fine-tuning.

📚 **Supervised Learning:** Supervised learning involves training a model on labeled data to learn a mapping from inputs to outputs. For example, in image classification, a model learns to classify images based on predefined labels such as "cat" or "dog". **Similarly, in spam detection, the model identifies whether an email is spam or not based on labeled training data**. Real-World Application: In healthcare, supervised learning can be used to predict patient outcomes based on historical data, helping doctors make informed decisions. Interview Questions:

**Interviewer: "Can you explain how supervised learning works in image classification?"** Answer: "In image classification, supervised learning involves training a model with labeled images, where the model learns to classify new images into predefined categories like 'cat' or 'dog', based on similarities with the training data."

Interviewer: **"Give an example where supervised learning can be applied effectively."** Answer: "Supervised learning is crucial in medical diagnostics, where algorithms use labeled patient data to predict disease risks or treatment outcomes, aiding healthcare professionals in decision-making."

📊 Self-Supervised Learning: Self-supervised learning involves training a model on unlabeled data by generating pseudo-labels from the data itself. For instance, in language modeling, a model predicts the next word in a sentence without explicit labels. In image inpainting, a model fills in missing parts of an image based on surrounding context. Real-World Application: Self-supervised learning is used in autonomous driving for predicting future paths of vehicles based on historical data, enhancing safety measures. Interview Questions:

Interviewer: "Explain how self-supervised learning differs from supervised learning." Answer: "Self-supervised learning doesn't require explicit labels; instead, it generates pseudo-labels from unlabeled data. For example, in image inpainting, the model fills in missing parts of an image based on surrounding context, without needing predefined labels."

Interviewer: "Where can self-supervised learning be beneficial in real-world applications?" Answer: "In autonomous driving systems, self-supervised learning can predict future movements of vehicles by analyzing patterns in historical data, thereby improving navigation and safety on the roads."