The D ONE - Make Sense - Glossary

A

Autoencoder
An ML autoencoder compresses data into a smaller size (encoding) and then reconstructs it back to the original size (decoding). It learns to do this by minimizing the difference between the original and the reconstructed data. The encoder part reduces data dimensions, and the decoder tries to rebuild the original input from this compressed version. The goal is to capture the most important features of the data.

Artificial General Intelligence (AGI)
A (hypothetical) highly autonomous system that outperforms humans at most economically valuable work.

B

Backpropagation
A key method in training neural networks. Backpropagation adjusts the weights of a neural network by calculating the gradient of the loss function with respect to each weight, and using it to minimize error during training.

Bias
Bias refers to systematic errors in the model's predictions, for example caused by imbalanced or unrepresentative training data. Bias results in inaccurate or unfair predictions.

Black Box
Describes AI models where the internal workings and decision-making processes are not easily interpretable, even though they can provide accurate results.

C

CNN
A Convolutional Neural Network (CNN) processes images by detecting contrasts and patterns. Similar to an → autoencoder, it uses layers to learn important features, but with a focus on interpreting images rather than compressing them. First, convolutional layers use small grids (filters) that slide over the image to identify contrasts like edges or textures. Next, pooling layers focus on preserving the most important contrasts within each small region, reducing data size while keeping key features. Finally, fully connected layers use these features to classify the image or make predictions. CNNs effectively learn to recognize patterns in images automatically.

D

Diffusion
Diffusion in AI is like dipping a tea bag into hot water—initially clear, but it becomes darker as the tea diffuses. The AI model reverses this, gradually turning the tea back to clear water by removing the noise. The model is trained with pairs of noisy and original data, learning to subtract noise step by step to restore the original image. Like → CNNs and → autoencoders, diffusion models learn patterns in data but focus on restoring or generating clear images from noise.

To direct the process toward a specific result (e.g. generating a realistic human image), the model is given a description or prompt. This input is → embedded. The model combines the current state of the image with this embedding, adjusting the denoising process to ensure the final image aligns with the prompt. This way, the random noise gradually transforms into an image that matches the desired outcome.

DLRM
In a Deep Learning Recommendation Model (DLRM), matrix factorization and deep learning are combined to predict e.g. user-item interactions. Imagine a matrix where rows represent users, columns represent items, and entries show interactions like ratings or clicks. Factorization splits the matrix into two: one for user → embeddings and one for item embeddings. The dot product of these embeddings shows how well they match. For other interactions (e.g. time spent watching a film), a neural network creates embeddings of similar shape.

A second neural network takes the combined embeddings and learns their relationships. The original interactions (e.g. the ratings), are used as labels in supervised learning, guiding the model to minimize prediction errors.

E

Embedding
Embedding converts words or data into vectors that represent their meaning. These vectors capture similarities and relationships, such as the connections between ‚man' and ‚woman' or ‚king' and ‚queen.' Embedding is a key concept in AI. It makes higher-level data like images, sound, video, and text available for computation.

Explainability
The ability to understand and interpret the decisions made by AI models. Explainability is crucial for transparency and trust in AI.

F

Fine Tuning
Fine-tuning in AI is the process of taking a pre-trained model and adapting it for a specific task. You start with a model trained on a large, general dataset that has learned broad patterns. To make it work well on a new task, like identifying types of flowers, you continue training it on a smaller, task-specific dataset. This fine-tuning adjusts the model's knowledge to better suit the new task while avoiding overfitting. Fine-tuning is efficient because it builds on the pre-trained model, needing less data and training time for good performance on the new task.

Foundational Model
A foundational model is a large AI model that serves as a base for many tasks. Trained on diverse data, it learns general patterns and features. This broad knowledge allows it to be easily adapted or → fine-tuned for specific tasks, like text generation or image recognition, making it a powerful and efficient starting point for specialized models.

G

GAN
A Generative Adversarial Network (GAN) generates realistic data using two networks: a generator and a discriminator. The generator creates fake data, and the discriminator tries to detect it. They compete, improving each other until the generator produces data that's hard to distinguish from real.

GNN
A Graph Neural Network (GNN) processes data structured as graphs, like social networks or molecular structures. It learns by passing and aggregating node features (→ embeddings). The GNN updates these embeddings by combining them with those of neighboring nodes through operations like weighted sums or averaging. The combined result is passed through an activation function, like ReLU, which introduces non-linearity by keeping positive values and setting negatives to zero. This process refines the embeddings over multiple layers.

For predictions or classifications, the GNN uses the final embeddings. For node classification, it assigns labels to individual nodes. For graph classification, it aggregates node embeddings to classify the entire graph. For link prediction, it evaluates pairs of node embeddings to predict connections. The GNN outputs these results based on the learned patterns and relationships in the graph.

Gradient Descent
An optimization algorithm used to minimize the loss function by iteratively adjusting the model's parameters in the direction that reduces error the most (resulting in an optimal line, very similar to downhill skiing).

H

Hallucination
In AI, hallucination is when a model generates false or made-up information that seems plausible. This happens in text generation or translation when the model fills in gaps with inaccurate details. It occurs because the model predicts based on learned patterns which don't apply.

Hyperparameters
Settings that must be defined before training a machine learning model, such as the learning rate, the number of layers in a neural network, or the number of units in each layer. Hyperparameters are tuned to optimize model performance.

I

Inference
The process by which an AI model makes predictions or classifications based on new, unseen data, after it has been trained.

L

Latent Space
A representation of data in which similar inputs are located close to each other in a lower-dimensional space. This is often used in models like → autoencoders, where compressed representations of data are learned.

M

Matrix Factorisation
See → DLRM

N

Neural Network
A neural network in machine learning is a mathematical model inspired by biological brains. It takes an input vector, processes it through multiple layers, and produces an output vector. Each layer transforms the input based on previous layers. The coordinates in these layers represent neurons, and the connections between them represent synapses. Neural networks are essential in deep learning, including models like transformers.

Natural Language Processing (NLP)
A branch of AI that focuses on the interaction between computers and human language. NLP encompasses tasks such as text generation, translation, sentiment analysis, and question-answering.

O

Overfitting
A scenario where a model performs well on training data but poorly on new, unseen data because it has learned noise or specific patterns in the training data that don't generalize.

P

Post-Transfomer
Post-Transformer models build on the original Transformer architecture by improving efficiency and scalability. They still use attention mechanisms to focus on different parts of the input data, understanding relationships and context. However, they often implement sparse attention, focusing only on the most relevant parts of the input, which reduces computation and speeds up processing. These models may also have modular designs, allowing different parts to specialize in various tasks, making them more adaptable. With improved layer structures and training methods, Post-Transformer models capture complex patterns more effectively while being less resource-intensive.

Prompt
In AI, a prompt is the input given to a model to guide its response. It can be a question, a phrase, or a description that directs the model on what to generate or answer. The model uses the prompt to understand the context and produce a relevant output.

R

Reinforcement Learning (RL)
A machine learning paradigm where an agent learns to make decisions by interacting with an environment and receiving feedback in the form of rewards or punishments. It's often used for tasks where sequential decision-making is critical, such as game playing or robotics.

Retrieval Augmented Generation (RAG)
RAG (Retrieval-Augmented Generation) allows LLM-based chatbots to answer questions using your own data. It works by first retrieving relevant information from a collection of documents. This information is then used to augment the question, helping the chatbot generate an answer based only on the retrieved data. RAG ensures responses are grounded in the provided information, reducing errors and integrating proprietary knowledge without needing to retrain the model.

RNN
A Recurrent Neural Network (RNN) processes sequential data like time series or text by using loops that pass information from one step to the next. The loop connects each step's output back to the previous step, updating a hidden state that carries context through the sequence. This allows the RNN to remember important details from earlier steps and use them to influence future predictions. While RNNs are good at handling sequences, they can struggle with remembering information from much earlier in the sequence.

T

Tokenization
The process of breaking down text into smaller units, such as words, subwords, or characters, to feed into an AI model.

Transfer Learning
The practice of applying a model trained on one task to a different but related task. This leverages the knowledge gained from the first task to improve performance on the second, reducing the need for large amounts of data and computation.

Transformer
A transformer breaks a sentence into parts (→ tokenization) and → embeds each token, enabling the recognition of similar words and the assignment of attention scores. These scores highlight important words by converting them into probabilities. By focusing on key words, the model improves its understanding of context, helping it predict the next word, translate text, or summarize information.

V

Vector
In a plane, a point can be described with its coordinates. These coordinates form a 2-dimensional vector (a list of 2 values). Two points close to each other have similar vectors. More specifically, the angle between these vectors is small. In AI, vectors often have 700 or even 1500 dimensions, which is difficult to imaginge. But they still work in the same way.

Z

Zero-Shot Learning
The ability of a model to recognize and classify data it has never seen during training by relying on pre-learned features or knowledge transfer.