What is Deep Learning?

Deep learning is subfield of AI and machine learning that is inspired by human brain.

Deep Learning algorithms attempt to draw similar conclusions as human do by continually analyzing data with a given logical structure called Neural network.

AI is a umbrella term that contains both ML and DL. The goal of AI is to build intelligent machines like humans. AI scientists want to build algorithms that mimic human thinking and human decision making. One big development in the field of AI is machine learning (ML). ML algorithms learns from data (inputs and outputs).

ML algorithms depend on statistical techniques to figure out/capture relationship between inputs and output. Whereas deep learning depends on logical structure called Neural network (inspired from human brain) to figure out relationship between input and output mostly in supervised learning. Neural Network is designed in a way that attempt to replicates human brain. CS scientist felt if intelligent machines are to developed, the we have to copy human brain, the most intelligent form on earth.

Different types of Neural Networks

Artificial Neural Network (ANN); simple and basic Neural Network

Convolutional Neural Network (CNN): performs well on image data

Recurrent Neural Network (RNN): performs well on speech and textual data

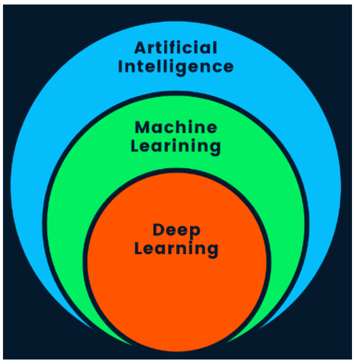
Generative adversial Network (GAN): generates text and images

Deep learning is a type of machine learning that teaches computers to perform tasks by learning from examples, much like humans do. Imagine teaching a computer to recognize cats: instead of telling it to look for whiskers, ears, and a tail, you show it thousands of pictures of cats. The computer finds the common patterns all by itself and learns how to identify a cat. This is the essence of deep learning. In simple terms, **deep learning** is like a virtual brain that helps computers learn from data so they can make decisions on their own.

In technical terms, deep learning uses something called "neural networks," which are inspired by the human brain. These networks consist of layers of interconnected nodes that process information. The more layers, the "deeper" the network, allowing it to learn more complex features and perform more sophisticated tasks.

"Is deep learning artificial intelligence?". The short answer is yes. Deep learning is a subset of machine learning, and machine learning is a subset of AI.

Artificial Intelligence is the idea that machines (or computers) can be built that has intelligence parallel (or greater) to that of a human, giving them capability to perform tasks that requires human intelligence to perform.



AI vs. ML vs. DL

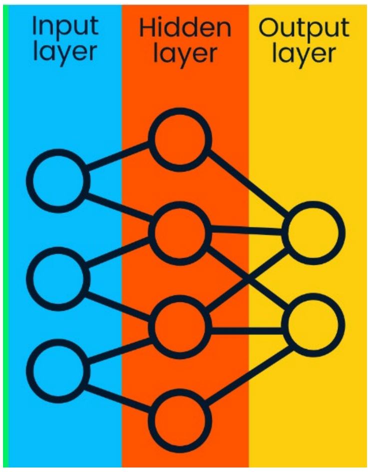
What is machine learning?

Machine learning is itself a subset of artificial intelligence (AI) that enables computers to learn from data and make decisions without explicit programming. It encompasses various techniques and algorithms that allow systems to recognize patterns, make predictions, and improve performance over time. Think of it as teaching computers to learn from experience, much like how humans do.

How deep learning differs from traditional machine learning

While machine learning has been a transformative technology in its own right, deep learning takes it a step further by automating many of the tasks that typically require human expertise.

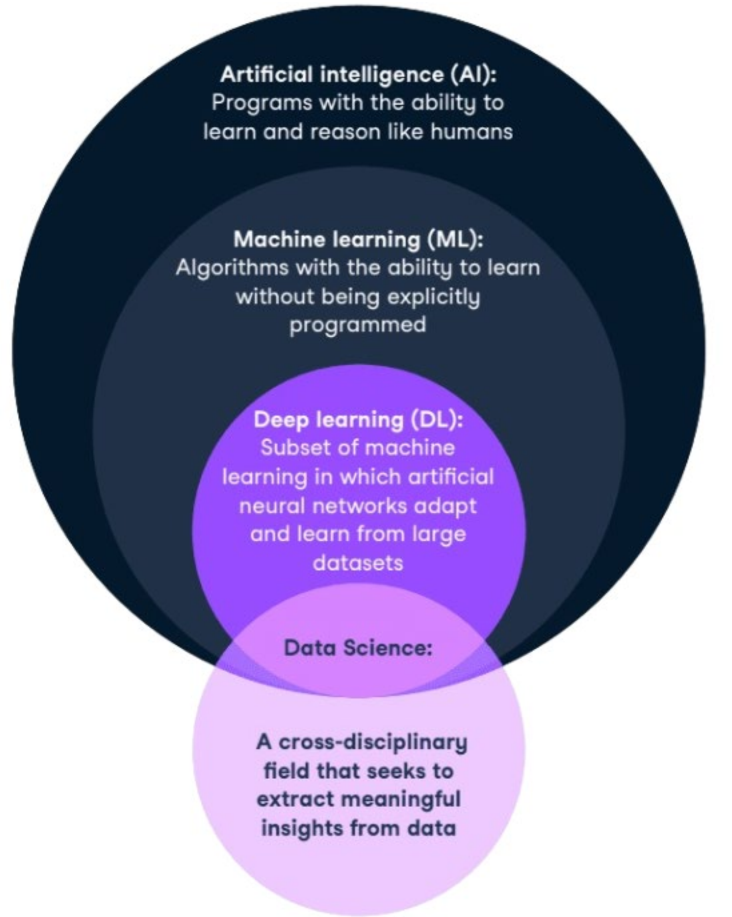
Deep learning is essentially a specialized subset of machine learning, distinguished by its use of neural networks with three or more layers. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—in order to "learn" from large amounts of data.



Simple Neural Network

Algorithm

A set of rules or instructions that a computer follows to perform a specific task. Algorithms are the building blocks of all AI systems.



AI in relation to data science and other key concepts

Data science is a multidisciplinary field of study that applies techniques and tools to draw meaningful information and actionable insights out of noisy data. Involving subjects like mathematics, statistics, computer science and artificial intelligence, data science is used across a variety of industries for smarter planning and decision making.

Data science is the realm (domain or field) of data scientists, who often rely on artificial intelligence, especially its subfields of machine learning and deep learning, to create models and make predictions using algorithms and other techniques.

Data science is an interdisciplinary field that combines techniques from statistics, computer science, domain expertise, and various data analysis methods to extract valuable insights and knowledge from data. The ultimate goal of data science is to turn raw data into actionable information, often with the help of tools and technologies like machine learning, data mining, and data analytics. These insights can be used for decision-making, problem-solving, and predictive modeling in a wide range of domains, from business and healthcare to social sciences and beyond

Scenario: Consider a manufacturing company that produces heavy machinery. The company has a fleet of trucks that are crucial for transporting raw materials and finished products. Downtime due to truck breakdowns can be costly in terms of delays and maintenance expenses.

Application of Data Science:

Data Collection: The company collects data from sensors installed on the trucks. These sensors monitor various parameters like engine temperature, oil pressure, vibration, and more. This data is collected regularly.

Data Analysis: Data scientists analyze the sensor data using statistical methods and domain expertise. They may use techniques like descriptive statistics to understand the normal operating conditions of the trucks and diagnostic analytics to identify patterns associated with breakdowns.

Predictive Modeling: Machine learning is employed to build predictive models. These models can forecast when a truck is likely to experience a breakdown based on historical data, patterns, and the real-time sensor readings. This is done using algorithms such as regression, decision trees, or neural networks.

Actionable Information: When the predictive model indicates that a truck is at risk of a breakdown in the near future, this information is relayed to maintenance teams in real-time. They can then schedule preventive maintenance, replacing parts or conducting repairs before the truck fails. This proactive approach minimizes downtime and reduces maintenance costs.

Decision-Making: The company uses the insights and actionable information to make decisions about when and how to maintain their truck fleet. This data-driven approach optimizes their operations and minimizes disruptions, ultimately improving productivity and cost-efficiency.

In this example, data science plays a crucial role in extracting valuable insights from sensor data and turning it into actionable information for predictive maintenance. The company uses data analytics and machine learning techniques to make informed decisions that save time and money while ensuring the continued functionality of their trucks.

Why deep learning is getting so famous: 2 Reasons

Applicability: DL algorithms are used in computer vision, natural language processing, speech recognition, machine translation, Bioinformatics, Drug Design, Medical image analysis, Climate Science, Material inspection, self-driving car, DL algorithms are designed in a manner that can be applied to wide domain of problems.

Performance: DL when applied to different problems provided state-of-the-art (that cannot be surpassed) results in different fields/domains. In most scenarios, DL beats human experts' performance. This is the reason deep learning is getting more attention from computer science or software community.

AlphaGo, the first computer program (AI agent) to defeat a professional human Go player (4 times out of 5 games) and a world champion, is arguably the strongest Go player in history. With origins in China over 3,000 years ago, Go is a profoundly complex board game, with more possible board configurations (10^170) than there are atoms in the known universe, making it a googol times more complex than chess. As a result, Go is renowned as the most challenging classical game for artificial intelligence due to its complexity.

Deep learning Definition:

Deep learning is part of a broader family of machine learning methods based on artificial neural networks with representation learning.

Feature Learning or Representation Learning: automatic features extraction from data

Feature engineering or feature extraction or feature discovery is the process of extracting features (characteristics, properties, attributes) from raw data.

Machine Learning requires feature engineering that is manually extract/create features from data.

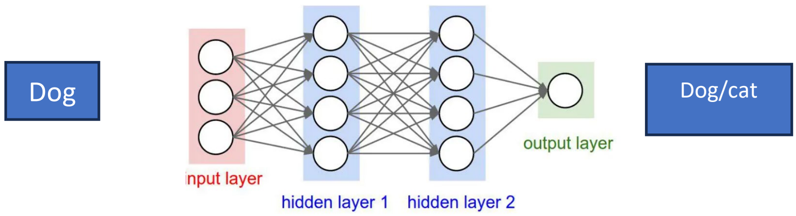
Image→ extract features(physical attributes: size, color)-→ dog/image classifier---output(dog or cat)

Deep Learning does not require feature engineering, DL automatically extracts features from data (image/text)based on its intelligence.

In machine learning, feature learning or representation learning[2] is a set of techniques that allows a system to automatically discover the representations needed for feature detection or classification from raw data. This replaces manual feature engineering and allows a machine to both learn the features and use them to perform a specific task.

Deep Learning Algorithms uses multiple layers to progressively extract higher-level features from the raw input. For example, in image processing, lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits or letters or faces.

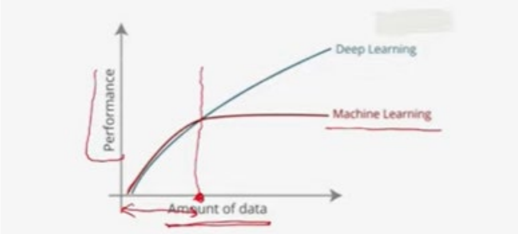
The data (image) is fed to the input layer, the starting hidden layer extracts basic or primitive features from data such as edges. The next hidden layer extracts complex features from data such as shapes. And the next layer extracts more complex features from data such as faces and eventually the result is given to output layer which decides whether image is dog or cat. The starting layers detects simple features while the deep layers detect complex features.



Deep Learning VS Machine Learning from practical perspective

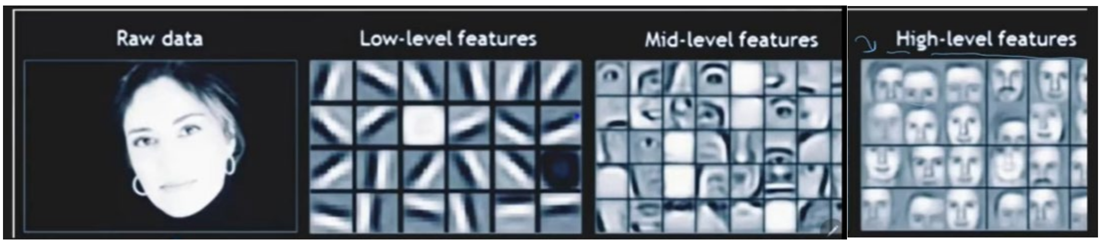
- Data dependency: deep learning requires more data in comparison to ML. if you are working on problem, then DL can only be applied if you have more data. The performance of DL is reliable when it sees or trained on more data. In the following graph, up to certain point, the performance of ML is better than DL, but after certain point, if u add more data the performance of ML will not improve any more and become stagnant as shown by flat line.

While following DL curve, the performance of DL linearly improve with the increasing/growing data. DL is data hungry. As soon as data grows, the performance of DL increase linearly.

- 

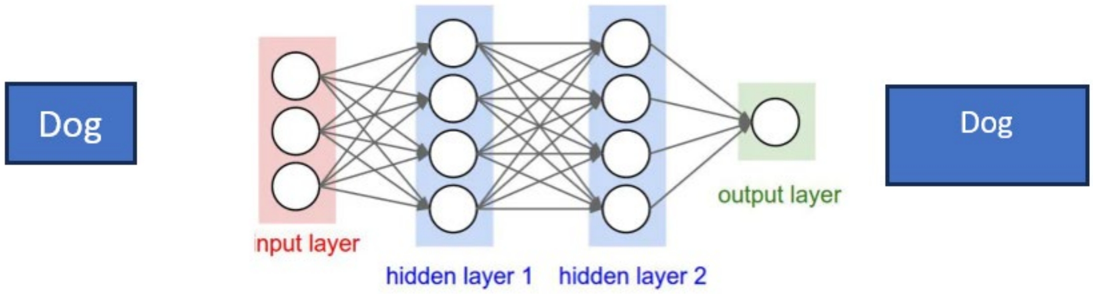
- Hardware dependency
- Machine learning models can be trained on simple machines using CPU. ML models need cheap hardware. While training DL models requires costly hardware i.e. powerful GPU with more memory. In DL models, a complex matrix multiplication is carried out, which can only be handled with GPU. If u run DL models on CPU, they will run slowly. In DL, hardware is costly.
-
- Training time: DL models are complex so training time for DL models is high while for Ml models, it is low. Sometime training DL models on high dataset needs weeks. For some researchers, training DL models took months for some specific purpose.
  The prediction time for DL models is fast while it varies for mL models for instance the prediction time for KNN algorithm is very slow.

- Feature Selection: In DL, the relevant features are automatically extracted from data, while in ML, features are manually extracted from the data. For example, You are working on a resume based project, where the placement/selection of candidate is predicted based on resume content. Resume content-→ predict-→yes/No
  If you are applying ML approach, you need to extract features from resume manually, and for this manual feature extraction, you need a domain expert in case a teacher,
  Features: 12th marks, 10th marks, No of achievements, No of courses, quality of college, based on these features, placement of candidate(y/N) will be predicted.
  If you apply DL approach, you just give the resume content to the algorithm, it will automatically extract features from it, and will predict whether the candidate will be placed or not.
- If You give an image to DL models, the starting hidden layers will extract low-level features such as edges, the next layers will extract mid-level features such as shapes and then last hidden layers extract high-level features such complete faces. This is how deep learning works



- Interpretability: Deep learning operates like a mysterious black box, concealing its internal thought processes and decision-making from us mere mortals. Only deep learning itself possesses this cryptic knowledge. Thus, when we find ourselves in situations where we need to provide not just the results but also the reasoning behind them, deep learning may not be the most suitable option.
- In DL models, the features are automatically discovered, and we donot what are those features(we do not understand), internally whats going on in DL is just like a black box. The

model will not tell me what it saw in the image. The model will not tell me why it call an image as dog. If someone asked me, why this image as dog, I donot have answer. Hene the model is not interpretable or explainable. The interpretability of DL models is very low.

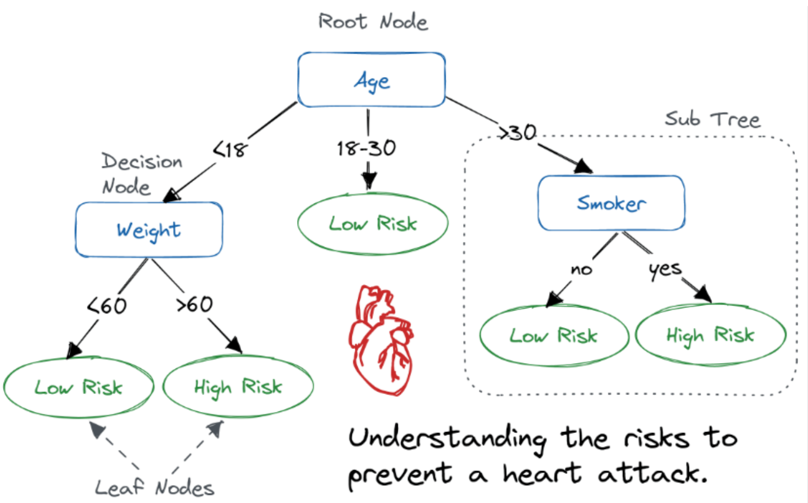


Dog/Cat Classifier

Suppose the DL model is trained on social networking website and it sees the comment and decide whether the user should be banned or not. Suppose the DL algo decides that the user should be banned. If the user ask about the reason, why am I banned? Since we applied DL so we do not know why or based on what reason, the algo has banned the user. So it is big flaw of DL. Explainability or interpretability is very important in several cases where DL fails. On the other hand, the interpretability of Ml models is very high.
Suppose u train logistic regression on the following data, based on weights, I can decide, which feature is more important and which feature is less important

CGPA w1	IQ w2	Placement
4	8	1
2.0	6	0

If the placement of certain candidate is 0, I can say that your weight w1 or CGPA is very low. Similary in decision tree, we make proper decisions and the flowchart tell us why certain point is given 1 or 0. So here the Explainability or interpretability is high. Explainability or interpretability is very important in cases where u have to tell others the reason why certain point/sample is given 1 or 0.



A decision tree is a white box type of ML algorithm. It shares internal decision-making logic, which is not available in the black box type of algorithms such as with a neural network. Its training time is faster compared to the neural network algorithm.

ML cannot be replaced by DL. If DL is so powerful why not use DL in every project.

Reason: ML is better than DL in few cases.

Where a needle is required, a sword is not suitable."

ML is needle and DL is sword, you have to decide whether u need ML or DL for a given problem.

Why Deep Learning is famous technology now

Work on DL stared in 1960 at the time of Allan turing but it became famous in 2012. What happened after 2010 that DL became a famous technology, there are multiple reasons;

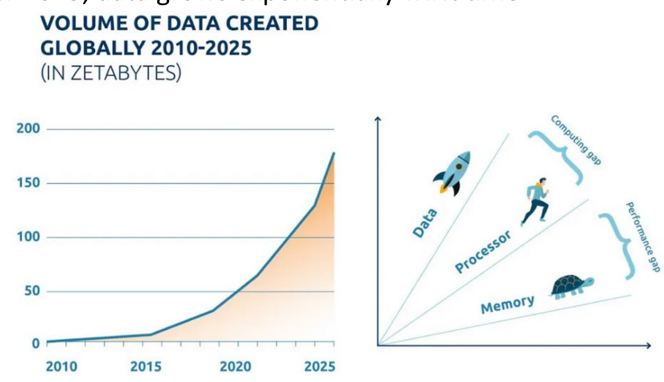
- Datasets
- Frameworks
- Architecture of models
- Hardware
- Community

Datasets: DL is data hungry. It needs more data for its prediction. After seeing more data, DL becomes sure about what patterns exist in the data. If u have 100 or 1000 rows of data, then u train ml to build models but not for DL. DL requires rows in lacs.

Around 2010, two revolution came, a) smartphone revolution b) internet pricing revolution and internet availability,means every individual had smartphone with high speed internet connection. Smartphones has social media platforms, facebook, twitter, tictoc, snapchat, and hence large amount of data is generated on daily basis. According to study,

Data generation begins in human history-------------XGB data-------------→ 2015

2015---------------------------------------XGB data -------------------------2016
2016---------------------------------------2XGB data -------------------------2017
2017--------------------------------------4XGB data -------------------------2018
After 2015, data grows exponentially w.r.t time



Volume of data created globally (2010–2025)

https://medium.com/grovf/embracing-the-exponential-growth-of-data-towards-the-breakthrough-of-memory-scaling-d29a94f2e45c

Deep learning so powerful today because of big dataset available. Dataset is big contributor in the boom of deep learning. Without data, there will be no deep learning.

Lots of data generated using smartphones. Work/research will not be done with just data generation. For instance, if u want to build image classification system, where u predict whether the given image is a cat or dog. First u need model training and at training time, u need to provide labelled data(image-label (cat or dog))., but the data generation using smart phones is unlabelled.. for instance, if u upload a picture a of cat, but u donot say it is a cat. In industry we need someone to create labels.

Companies like Microsoft, facebook, Google invested a lot of money to label the large amount of available unlabeled data. They observed that research in DL will not move forward if we restrict the data to their companies. So they open source the dataset, make the dataset public. Actual research in deep learning moved forward/carried out, when the datasets get available on github and Kaggle

Examples of public datasets:

Image data: Microsoft COCO , a powerful dataset for object detection. In each image, there are bounding boxes around the objects in the images, and also labelled the objects

Video data: Youtube 8M dataset, 6.1 million videos data

Text: SQUAD, 150,000 questions/answers dataset formed from Wikipedia

Audio: Google Audioset—2 lacs sound clips extracted from youtube that falls under 600 categories

Thousands of dataset available on Kaggle and github, it is not easy create these large datasets, research becomes easy when the datasets became available.

Hardware: Computing hardware is the second factor in pushing DL. To explain this reason, we refer to Moore's law which is famous law in computing, Moores is the co-founder of Intel. Law says that no of transistors on microchip doubles every two years, though the cost of hardware is halved.

Core i5 or core i7 chip---transistor—2 years(doubled)

Cost of making it is halved.

Means the performance of hardware will improve, while the its cost be cheaper. In 2005, phones will only 128 MB was available now smartphones with 128GB are available. Revolution in hardware is continuously happening and it pushed deep learning a lot .

DL requires lots of data, and it performs matrix operations on lots of data and we need processing power to handle operations on lots of numbers, if we use CPU, training time would be high. Research in 2010 realized that matrix operations can be converted into parallel tasks i.e. can be implemented via parallel processing. The got the idea that if the high quality image is read or rendered, we need parallel processing, for this purpose GPU is used. Similarly, why not to use GPU, for matrix operations.

This idea floated into DL community That GPU shoud be used to train our neural network models.

NVIDIA came and launched CUDA, a programming language to program GPU. So revolution came, rather than using conventional CPU, lets train models on GPU. Not only gamers, but DL researchers started buying graphic cards from NVIDIA stocks. GPU reduces training time from 10-20 times in comparison to CPU. Researcher realized that faster results can be achieved through parallel processing so research gained attention in this area and different types of hardware came into developed.

FPGA Field Programmable Gate Arrays (FPGAs), already existed, students with engineering background know about it. FPGA are programmable controllers, they are fast, reprogrammable (sit/setup custom solutions on it easily) and operates on low power but they are very expensive. The leading manufacturer of FPGA is XILINX

Field Programmable Gate Arrays (FPGAs) are integrated circuits often sold off-the-shelf. They're referred to as 'field programmable' because they provide customers the ability to reconfigure the hardware to meet specific use case requirements after the manufacturing process.

In today'ageFPGA is also widely used for training and running DL models. It is used in Micrsoft. Microsoft Bing' search engine AI part is run on FPGA.

ASIC: Application specific integrated circuits. They are custom made chips.

Different types of ASICs are Tensor Processing Unit (TPU), Edge TPU, Neural Processing Unit (NPU)

TPU: developed by google, a specialized hardware for training DL models. Google Colab has option of choosing TPU to run a piece of code;

Edge TPU: is a very small device came where u can perform calculations. Edge TPU developed for for edge devices for instance, u want to run DL model on Drone, Smart watch, smart glass. In such setup, u c an use Edge TPU to perform calculations.

NPU: came for smart phone applications which make use of ml and deep learning. NPU accelerate the Ml and DL operations on mobile devices.

In which situation, What type of hardware will work or Different hardwares for different requirements

DL Journey beginning-----------→ small projects---CPU

Training large Neural Networks----→ GPU or TPU RTX 2080

Smart phones-→ if u want to use deep learning in Apps in smartphones, u can use mobile CPU (snapdragon), Mobile GPU, Digital Signal Processor (DSP) or NPUs

Smart watch or smart glass-→ if u want to run DL models on smart watch or glass, use Edge TPU or NPU.

In last 10 years, focus shifted towards custom hardware for DL, so research in DL becomes fast.

Key Reasons for the Rise of Deep Learning:

1. **Data Growth:**
    - **Increased Data Volume:** Businesses now utilize software and mobile apps to manage operations, leading to an explosion of transactional data.
    - **Social Media:** Platforms like Twitter and Facebook generate immense data through user interactions (e.g., posts, likes, dislikes). This data supports applications like sentiment analysis.
    - **Relevance to Neural Networks:** Deep learning models require large datasets to perform effectively. The recent surge in available data aligns with these requirements, making deep learning feasible.

2. **Hardware Advancements:**
    - **Improved Computing Power:** Older computers with limited RAM (e.g., 500MB in 2003) made deep learning infeasible. Modern computers can execute deep learning tasks within seconds.
    - **Specialized Hardware:** GPUs (Graphical Processing Units) and TPUs (Tensor Processing Units) optimize parallel computation, crucial for deep learning tasks. Notable contributions include:
        - **NVIDIA GPUs:** Excellent for deep learning due to parallel processing capabilities.
        - **Google TPUs:** Specifically designed for TensorFlow operations, making training faster and efficient.
    - **Affordability:** These advancements have democratized access to the required computational power.
3. **Python and Open-Source Ecosystem:**
    - **Accessibility of Python:** Python is user-friendly and has a shallow learning curve compared to C++ or Java. It allows even non-programmers with a mathematics or statistics background to start deep learning quickly.
    - **Open-Source Frameworks:** Tools like PyTorch (by Facebook) and TensorFlow (by Google) provide free and accessible platforms for building and training neural networks. They simplify implementation and experimentation.
    - **Educational Democratization:** Python's simplicity and open-source tools have enabled a broader audience to explore deep learning.
4. **Cloud Computing and AI Boom:**
    - **Cloud Services:** Companies like AWS, Google Cloud, and Azure offer on-demand servers with GPUs/TPUs, eliminating the need for upfront investment in expensive hardware.
    - **Reduced Entry Barriers:** Renting computational resources makes experimentation and development accessible to individuals and small businesses.
    - **AI Integration in Business:** Major companies, such as Google, have transitioned from a "mobile-first" to an "AI-first" approach. This industry-wide focus on AI accelerates investment and innovation in deep learning.
5. **AI Boom and Industry Demand:**

- **Business Prioritization:** Artificial Intelligence has become essential for staying competitive, pushing organizations to adopt machine learning and deep learning solutions.
- **Global Shift:** AI is now a core focus for technological and business
- There is a prevalent AI boom in the businesses nowadays where all business executives want to benefit from artificial intelligence. This further accelerates the growth of deep learning.

major companies like Google have shifted their primary focus from prioritizing mobile technologies to emphasizing artificial intelligence (AI) as the core of their strategies. Here's a breakdown:

"Mobile-first" Approach:

- **Definition:** This refers to designing products, services, and strategies with mobile devices (like smartphones) as the primary platform. For example, in the past decade, companies focused on optimizing websites for mobile screens, developing mobile apps, and leveraging mobile-specific technologies.
- **Reason:** The rise of smartphones and their ubiquitous use drove companies to ensure their offerings were tailored to mobile users.

Transition to "AI-first" Approach:

- **Definition:** Now, the emphasis has shifted to AI technologies, where the primary goal is to integrate artificial intelligence into products and services. AI-first means designing with AI capabilities at the center, whether for personalization, automation, data-driven decision-making, or innovative features like voice assistants.
- **Examples of AI-first:**
    - Google Assistant in Android devices.
    - AI-driven search results and recommendations.
    - Automated tools like predictive typing and smart replies in Gmail.

Predicting Home Prices Using Linear Regression

- collected data on home prices based on the areas in Monroe Township, NJ.
- With this data, construct a machine learning model called linear regression
- LR will predicting homes prices with areas of 3,300 and 5,000 square feet.

Given these home prices find out prices of homes whose area is,

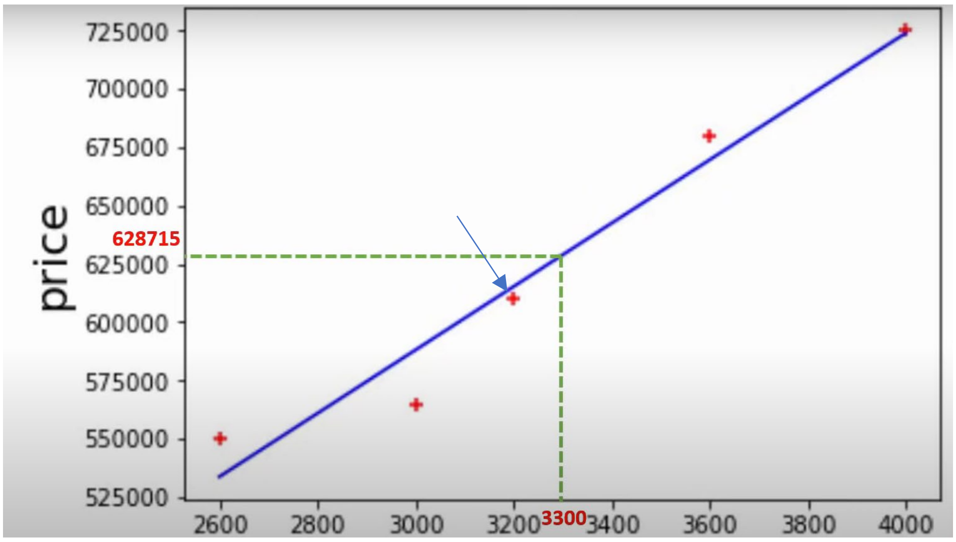
3300 square feet
5000 square feet

- Home prices in Monroe Twp, NJ (USA)

area	price
2600	550000
3000	565000
3200	610000
3600	680000
4000	725000

Data Visualization

- Visualize the available prices and areas with a scatter plot, where the red marker represents the data points.
- We'll then draw a blue line that best fits these data points.



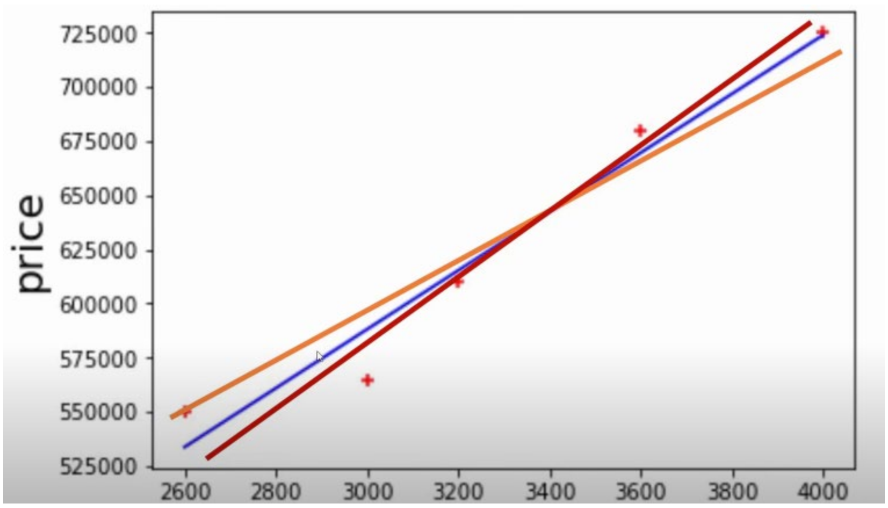
Purpose of the Blue Line

- Function: The blue line (represents a linear equation) estimate home prices for given areas.
- **Example**:

  - Predict price for 3,300 sq. ft. Home based on linear equation (Blue Line)
  - Predict price for for 5,000 sq. ft. Home based on linear equation (Blue Line)

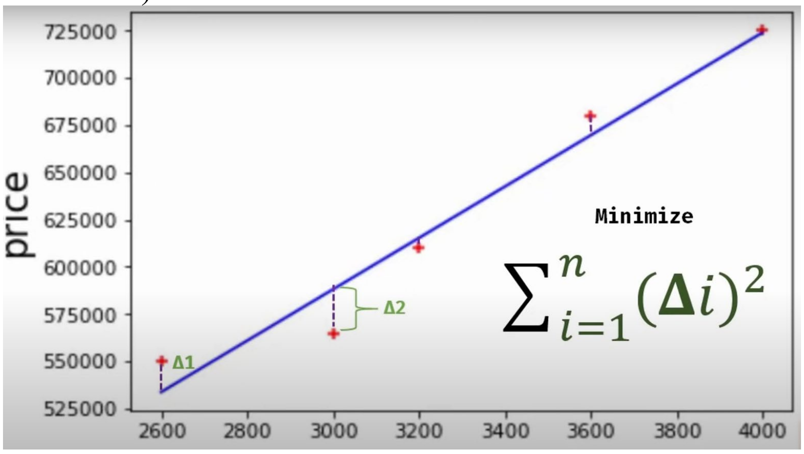
Choosing the Best Fit Line

- **Problem**: Many lines (e.g., red, orange) can be drawn in number of ways
- **Solution**:
  - Calculate errors (ΔDeltaΔ) between actual data points and points predicted by linear equation.
  - square individual errors, sum them up and then then try to minimize
  - We do this procedure for each line (red, orange, blue)

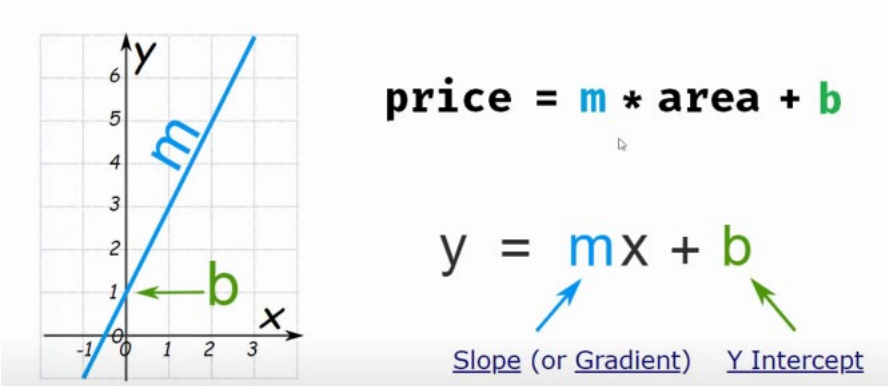


Minimizing Error

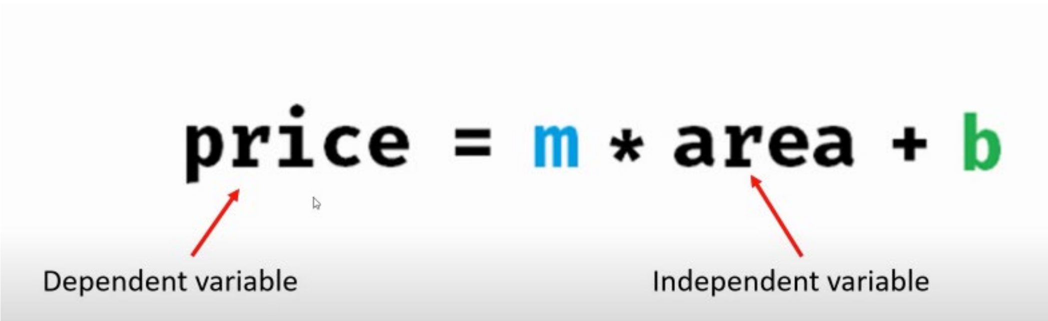
- **Approach**:
    - Minimize the total squared error.
    - Blue line is selected because found that it minimizes errors effectively (better than others.)



Area



Linear Equation looks like Y=mx +b

- 

- We calculate price based on area, y=price, x= area.

In the linear equation $Y = mx + b$, the two components m and b have specific meanings:

1. **Slope (m):**
   - The **slope** m represents the rate of change of Y with respect to X. It tells you how much Y will increase or decrease as X changes. If the slope is positive, Y increases as X increases, while if the slope is negative, Y decreases as X increases.
   - Mathematically, m is calculated as the "rise over run," which is the change in Y divided by the change in X between two points on the line. In other words, $m = \frac{\Delta Y}{\Delta X}$.
   - **Interpretation:** A slope of 2 means that for every 1 unit increase in x, y increases by 2 units.
2. **Y-intercept (b):**
   - The **Y-intercept** b is the point where the line crosses the Y-axis, which occurs when $x = 0$. In other words, it's the value of Y when X is zero.
   - The intercept represents the initial or starting value of Y before any changes in X occur. This is important in many practical applications where you might want to know the value of Y when X is zero.
   - **Interpretation:** If $b = 5$, this means that when $x = 0$, y will be 5.

- Linear regression helps in estimating continuous values like home prices.
  The best fit line minimizes prediction errors using mathematical optimization.

Linear Regression with Multiple Variables: Predicting Home Prices

- We will explore **linear regression with multiple variables**, also known as **multivariate regression**.
- Using this method, we aim to predict **home prices** in Monroe Township, New Jersey.

Key Factors in Home Pricing

The table provided includes various metrics such as:

- **Area (square footage)**
- **Number of bedrooms**
- **Age of the home**

area	bedrooms	age	price
2600	3	20	550000
3000	4	15	565000
3200		18	610000
3600	3	30	595000
4000	5	8	760000
4100	6	8	810000

These factors influence the ultimate price of a home.

- Previously, we examined **simple linear regression**, where the price depended only on the area.
- Now, we make the problem more complex by incorporating additional variables like **bedrooms** and **age**, reflecting real-world scenarios where home prices depend on multiple factors, not just square footage.
- After building our model, we will use it to predict the prices of two homes.

Data Analysis: A Crucial First Step

Before tackling any machine learning problem, it is essential to carefully analyze the dataset. Upon examining the data, here are the observations:

1. **Missing Data**: There is a data point missing in the table, which must be handled appropriately to ensure the model's accuracy.
2. **Linear Relationships**: A clear linear relationship exists between the independent variables (area, bedrooms, and age) and the target variable (price).
    - For instance, as a home ages, its price tends to decrease. Consider these examples:
        - A **3,200-square-foot home** with **18 years of age** is priced above **$600,000**.
        - A slightly larger **3,600-square-foot home**, due to its greater age, is priced lower than the smaller home.
    - Similarly, as the **area** and **number of bedrooms** increase, the price generally rises.

Based on the analysis, it is reasonable to apply **linear regression** to this dataset. Our linear equation will look like this where price is dependent on three features; area, bedrooms and age.

Dependent variable

Independent variables (**features**)

$$price = m_1 * area + m_2 * bedrooms + m_3 * age + b$$

Coefficients

$$y = m_1 x_1 + m_2 x_2 + m_3 x_3 + b$$

The above equation can be generalized into the following equation where u can have 'n' number of independent variables or features. In our case, we have 3 features

$$y = m_1 x_1 + m_2 x_2 + m_3 x_3 + b$$

Check python code for the remaining stuff