Chapter 1:

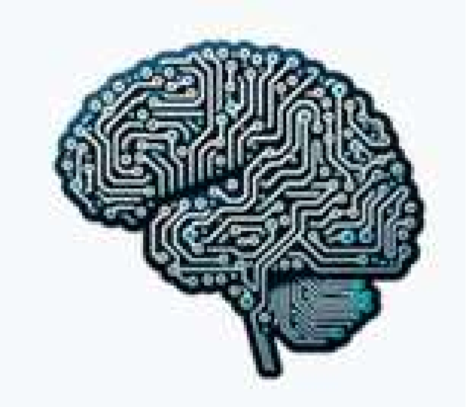
What is deep learning?

Outlines

1. Basic concepts
2. Why deep learning? Why now?

Basic concepts

1. Artificial intelligence:



- Artificial intelligence (AI) emerged in the 1950s when computer scientists began exploring whether machines could think.
- In 1956, John McCarthy proposed studying how machines could simulate human intelligence by learning, abstracting, and problem-solving.

Basic concepts

1. Artificial intelligence (cont):
   - Initially, AI focused on symbolic AI, where explicit rules and databases were used to mimic intelligence.
   - Concepts like machine learning and deep learning were later integrated, broadening AI beyond rule-based systems.
   - Overall, AI aims to automate intellectual tasks typically performed by humans.

Basic concepts

2. Machine learning:

   - Traditional programming involves humans writing rules to process input data and produce outputs,
   - In contrast, machine learning allows computers to derive rules from data by identifying patterns in examples. For instance, a system can learn to tag vacation photos by analyzing pre-tagged images.
   - Emerging in the 1990s, machine learning has become a dominant AI subfield due to faster hardware and larger datasets.

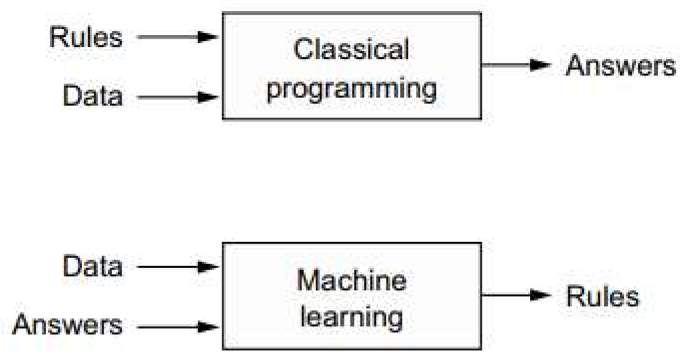
Basic concepts

2. Machine learning (cont):

   - Machine learning is related to statistics but differs significantly:
     - ✓ Machine learning handles large, complex datasets impractical for traditional statistical methods and focuses on practical, engineering-driven solutions.
     - ✓ Unlike theoretical disciplines, machine learning emphasizes empirical findings and relies heavily on advancements in software and hardware, especially in areas like deep learning..

Basic concepts

2. Machine learning (cont):



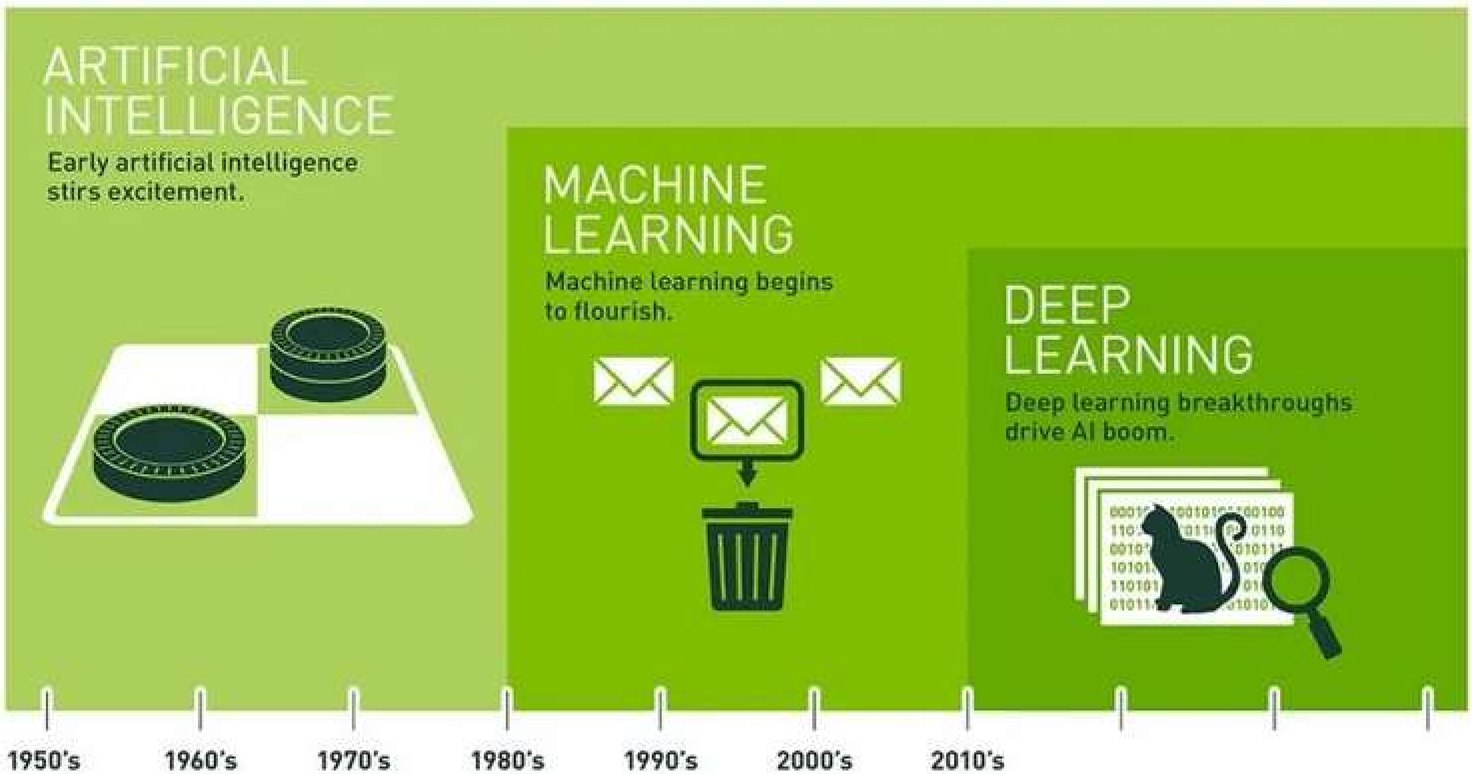
- The usual way:
  - ✓ A human programmer writes down rules (a computer program) to be followed to turn input data into appropriate answers.
- Machine learning:
  - ✓ The machine looks at the input data and the corresponding answers and figures out what the rules should be.

Basic concepts

2. Machine learning (cont):
   - Learning rules and representations from data:
     - ✓ Input data points.
       - For instance, if the task is speech recognition, these data points could be sound files of people speaking. If the task is image tagging, they could be pictures.
     - ✓ Examples of the expected output.
       - In an image task, expected outputs could be tags such as “dog,” “cat,” and so on.
     - ✓ Measure whether the algorithm is doing a good job.
       - The measurement is used as a feedback signal to adjust the way the algorithm works.
       - This adjustment step is what we call learning

Basic concepts

3. Deep learning:
   - Deep learning is a subfield of machine learning.



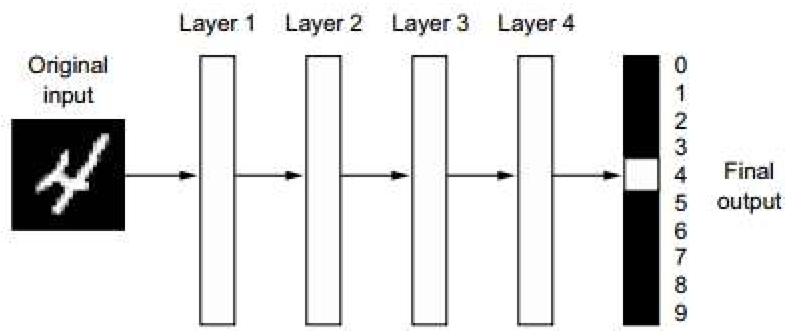
Source: NVIDIA

Basic concepts

3. Deep learning (cont):
   - Deep learning is a subfield of machine learning focused on learning multiple layers of increasingly meaningful representations from data.
   - The term "deep" refers to the number of layers, known as the model's depth, not a deeper understanding.
   - Unlike shallow learning, which uses one or two layers, deep learning involves tens or hundreds of layers, all learned automatically from training data.
   - It could also be called "layered" or "hierarchical representations learning.".

Basic concepts

3. Deep learning (cont):
   - Deep learning uses neural networks, structured in stacked layers, to learn representations from data.

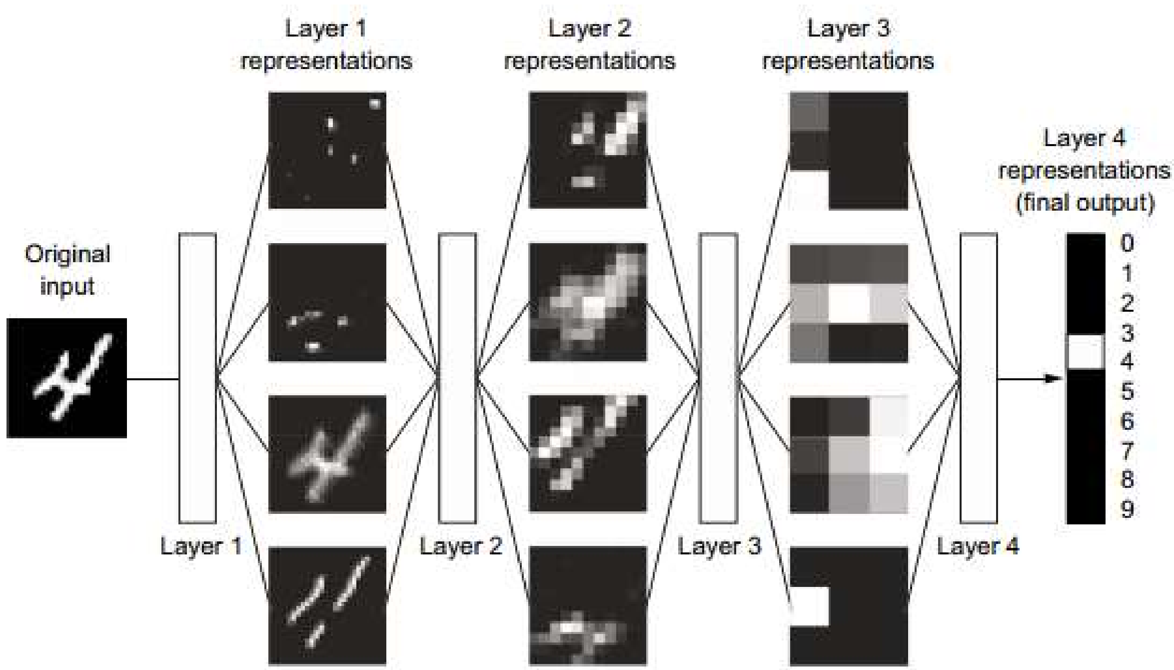


Illustrate a deep neural network for digit classification.

   - Although inspired by the brain's visual cortex, these models are not biologically accurate and do not mimic how the brain functions.

Basic concepts

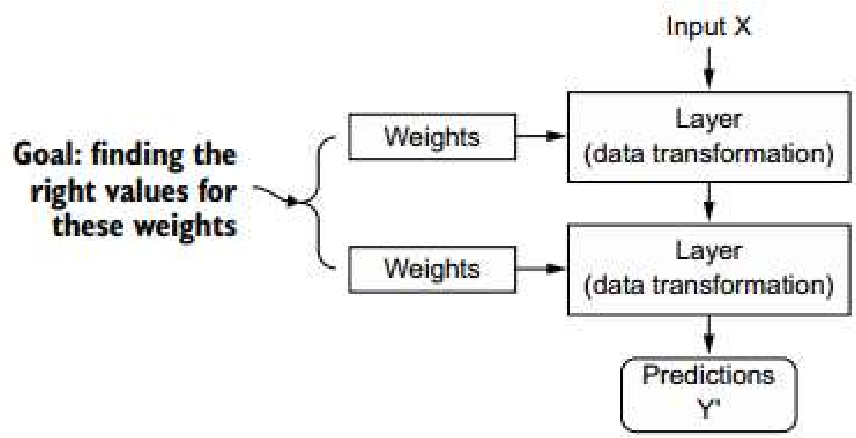
3. Deep learning (cont):
   - Deep learning is merely a mathematical framework for transforming data into increasingly informative representations through a multistage information-distillation process.



Illustrate the data representations learned by a digit classification model.

Basic concepts

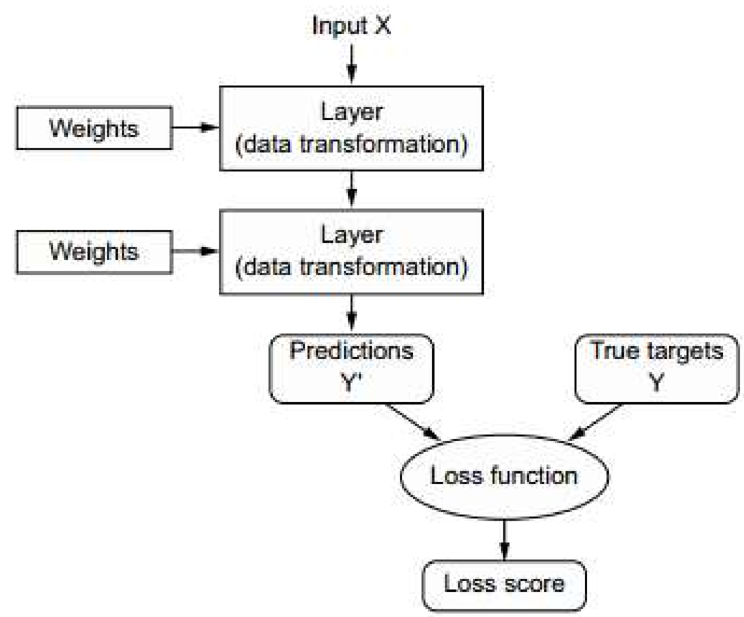
3. Deep learning (cont):
   - Deep neural networks map inputs (like images) to targets (like labels) through layers of simple data transformations.
   - Each layer's behavior is determined by its weights (parameters).



   - Learning means finding the optimal values for all these weights to correctly map inputs to targets.
   - Deep networks often contain millions of parameters; finding the right values is a complex and interconnected process.

Basic concepts

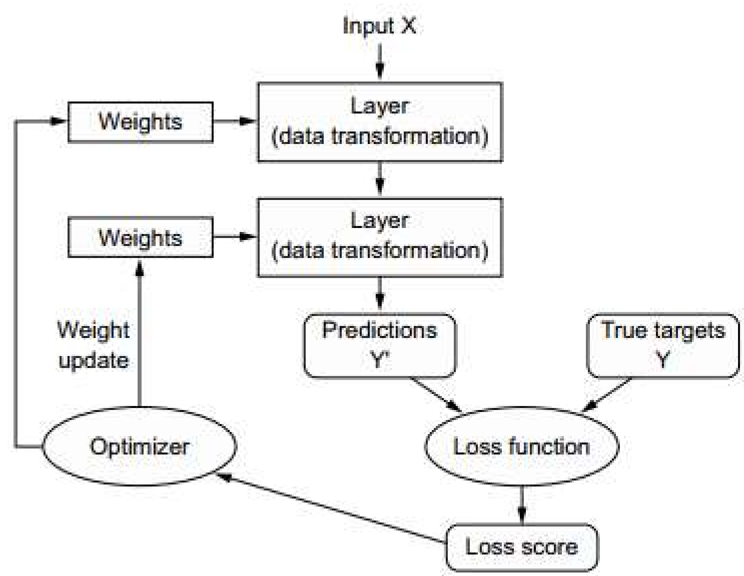
3. Deep learning (cont):
   - The quality of a neural network is measured by comparing the output of the neural network with what we expected.
   - The loss function takes the predictions of the network and the true target and computes a distance score, capturing how well the network.
   - The loss function of the network is sometimes called the objective function or cost function.



The loss function measures the quality of the network's output.

Basic concepts

3. Deep learning (cont):
   - The fundamental trick in deep learning is to use this score (loss value) as a feedback signal to adjust the value of the weights a little, in a direction that will lower the loss score for the current sample data.
   - This adjustment is the job of the optimizer.



The loss score is used as a feedback signal to adjust the weights.

   - The optimizer implements what's called the Backpropagation algorithm to optimize the result.
   - The Backpropagation algorithm is the central algorithm in deep learning.

Basic concepts

3. Deep learning (cont):
   - Achievements through deep learning
     - ✓ Near-human-level image classification.
     - ✓ Near-human-level speech transcription.
     - ✓ Near-human-level handwriting transcription.
     - ✓ Dramatically improved machine translation.
     - ✓ Dramatically improved text-to-speech conversion.
     - ✓ Digital assistants such as Google Assistant and Amazon Alexa.

Basic concepts

3. Deep learning (cont):
   - Achievements through deep learning
     - ✓ Near-human-level autonomous driving.
     - ✓ Improved ad targeting, as used by Google, Baidu, or Bing.
     - ✓ Improved search results on the web.
     - ✓ Ability to answer natural language questions.
     - ✓ Superhuman Go playing.

Why deep learning? Why now?

1. What makes deep learning different:

- Deep learning's rapid success stems from two key advantages:
  - ✓ Superior performance,
  - ✓ Simplified workflows.
- Deep learning doesn't relied heavily on manual "feature engineering" to transform input data for processing.
- Deep learning automates this process by learning multiple layers of representation in a single pass

=> eliminates the need for complex, multi-stage pipelines and replaces them with end-to-end models..

Why deep learning? Why now?

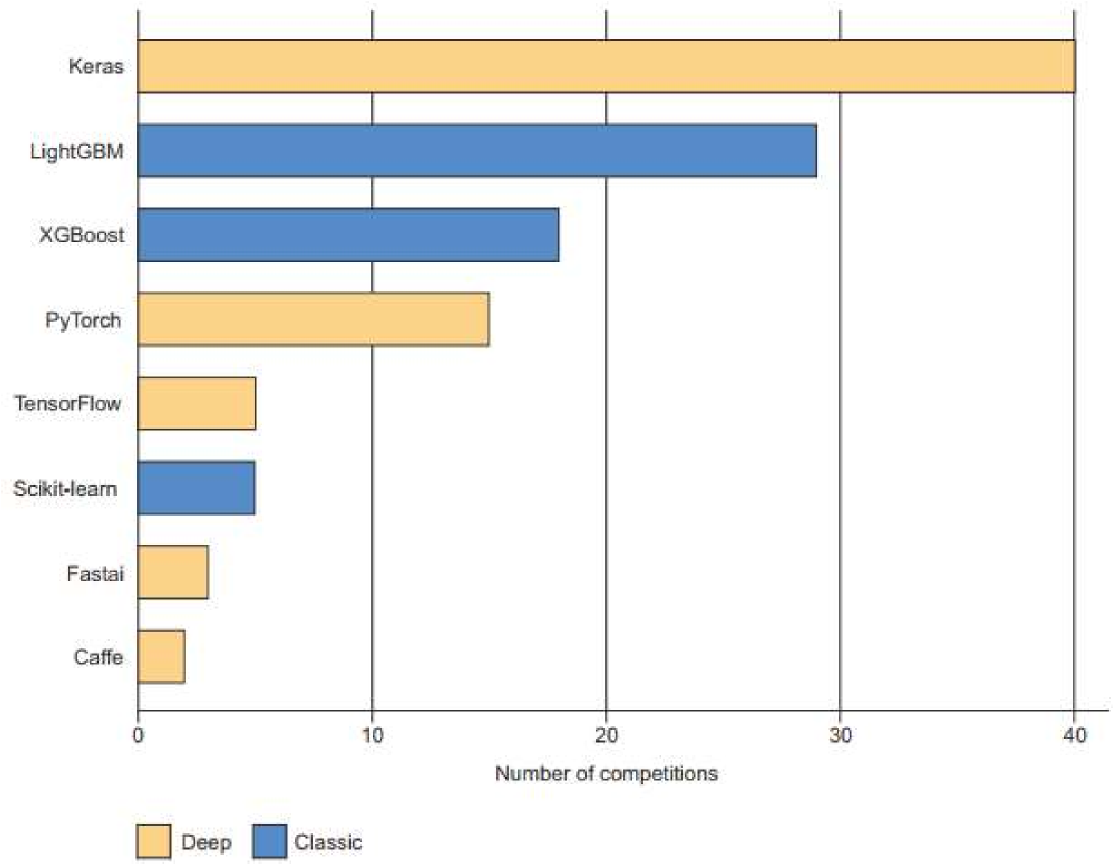
1. What makes deep learning different (cont):
   - Shallow learning methods fail to match deep learning's capability because they cannot jointly optimize successive layers of representations.
   - Deep learning uniquely updates all layers simultaneously, allowing for complex, abstract representations to emerge through incremental transformations.
   - This joint learning, guided by a unified feedback signal, adapts features across layers automatically, enabling deep learning to outperform earlier approaches in solving complex problems.

Why deep learning? Why now?

2. The modern machine-learning landscape:
   - Kaggle competitions provide insight into the effectiveness of machine learning algorithms and tools due to their competitive nature and diverse problems.
   - A 2019 survey of top-performing teams revealed that the most commonly used tools are deep learning methods (often via the Keras library) and gradient-boosted trees (typically implemented with LightGBM or XGBoost).
   - These tools consistently help teams achieve success in competitions..

Why deep learning? Why now?

2. The modern machine-learning landscape:



Machine learning tools used by top teams on Kaggle

Why deep learning? Why now?

- The two key ideas of deep learning for computer vision—convolutional neural networks and backpropagation—were already well understood by 1990.
- The Long Short-Term Memory (LSTM) algorithm, which is fundamental to deep learning for timeseries, was developed in 1997 and has barely changed since.
- So why did deep learning only take off after 2012?
- What changed in these two decades?.

Why deep learning? Why now?

- In general, three technical forces are driving advances in machine learning:
  - ✓ Hardware
  - ✓ Datasets and benchmarks
  - ✓ Algorithmic advances.

Why deep learning? Why now?

- Hardware:
  - Between 1990 and 2010, CPU speeds increased dramatically, enabling small deep learning models to run on laptops.
  - During the 2000s, companies like NVIDIA and AMD invested heavily in developing GPUs, originally designed for rendering complex 3D graphics in real-time video games.

Why deep learning? Why now?

- **Hardware**:
  - The GPUs, optimized for massively parallel tasks, proved useful for scientific computing.

Why deep learning? Why now?

- Hardware:
  - In 2007, NVIDIA introduced CUDA, a programming interface for GPUs, allowing for parallelizable tasks like physics modeling.
  - In 2011, researchers began using CUDA for deep neural networks, taking advantage of GPUs' ability to efficiently handle matrix multiplications.
  - Modern GPUs, like the NVIDIA Titan RTX deliver 16 teraFLOPS of computing power—500 times more than the fastest 1990, the Intel Touchstone Delta.
  - Beyond GPUs, the AI industry has developed specialized chips, such as Google's Tensor Processing Units (TPUs) in 2016.
  - TPUs are faster and more energy-efficient than GPUs.

Why deep learning? Why now?

❖ Hardware:

CPU
- Small models
- Small datasets
- Useful for design space exploration

GPU
- Medium-to-large models, datasets
- Image, video processing
- Application on CUDA or OpenCL

TPU
- Matrix computations
- Dense vector processing
- No custom TensorFlow operations

Why deep learning? Why now?

- Data:
  - Beside advancements in storage hardware that have been significant, the internet has revolutionized data collection and distribution, enabling large datasets critical for machine learning.
  - Platforms like Flickr, YouTube, and Wikipedia have become invaluable sources of data for computer vision and natural language processing.
  - The ImageNet dataset, containing 1.4 million annotated images across 1,000 categories, has been a pivotal resource for deep learning's growth.
  - The Kaggle's competitions since 2010 have motivated progress and established benchmarks that showcase deep learning's capabilities.

Why deep learning? Why now?

- Algorithms:
  - Until the late 2000s, deep neural networks couldn't be trained effectively due to issues with gradient propagation, limiting their depth and performance compared to shallow methods like SVMs and random forests.
  - Breakthroughs around 2009–2010, such as improved activation functions, weight-initialization schemes, and optimization algorithms (e.g., RMSProp and Adam), enabled training deeper models with 10+ layers.

Why deep learning? Why now?

- Algorithms:
  - Later innovations, including batch normalization, residual connections, and depthwise separable convolutions (2014–2016), allowed for arbitrarily deep models.
  - These advances led to modern deep learning's scalability, enabling large architectures (e.g., ResNet, Inception, BERT, GPT-3) that have revolutionized fields like computer vision and NLP.

Why deep learning? Why now?

3. Available resources:
   - Keras library.
   - Scikit-learn.
   - TensorFlow.
   - Kaggle.
   - Github.
   - …