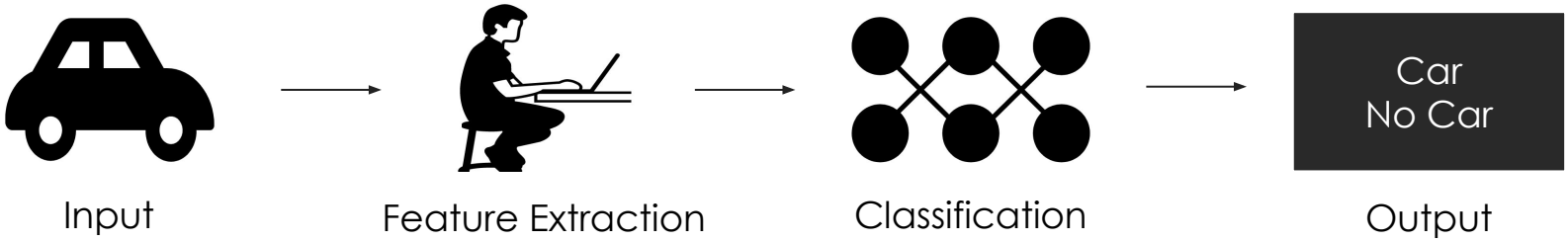
Overview:
Deep Learning - Part II

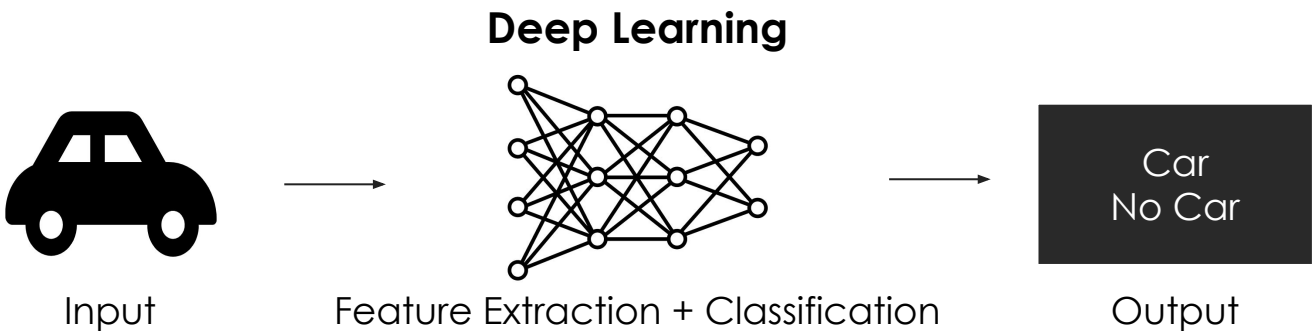
The Deep Learning Workflow

- We have understood the typical Machine Learning workflow as resembling the below schematic:



The Deep Learning Workflow

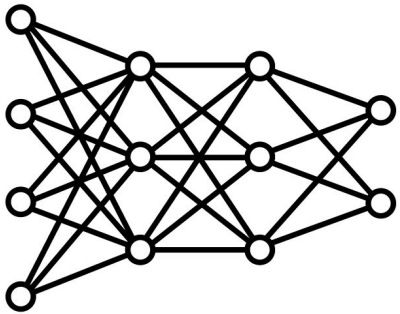
- The Deep Learning workflow, on the other hand, does this slightly differently.
- **It combines the Feature Extraction and Classification stages of the workflow into the same model architecture:**



The Deep Learning Workflow

- The way this works in Deep Learning, is that we would feed in the input data, **construct numerous layers of hierarchical, non-linear features out of the data** (which is how Deep Learning performs the Feature Extraction step), and lastly **provide these features of the image to the final layer of the model**, which is then responsible for the actual classification / predictive task. Typically, in the case of classification, this final layer would output the probabilities of every possible discrete class, and **the class with the highest probability would be considered the prediction of the model.**

- So the question then is, **why does this work?** Why do we need these layers of abstraction to represent the features of our dataset? Why should feature extraction be done in this manner?



Schematic of a Deep Neural Network with Multiple Layers

The Deep Learning Workflow

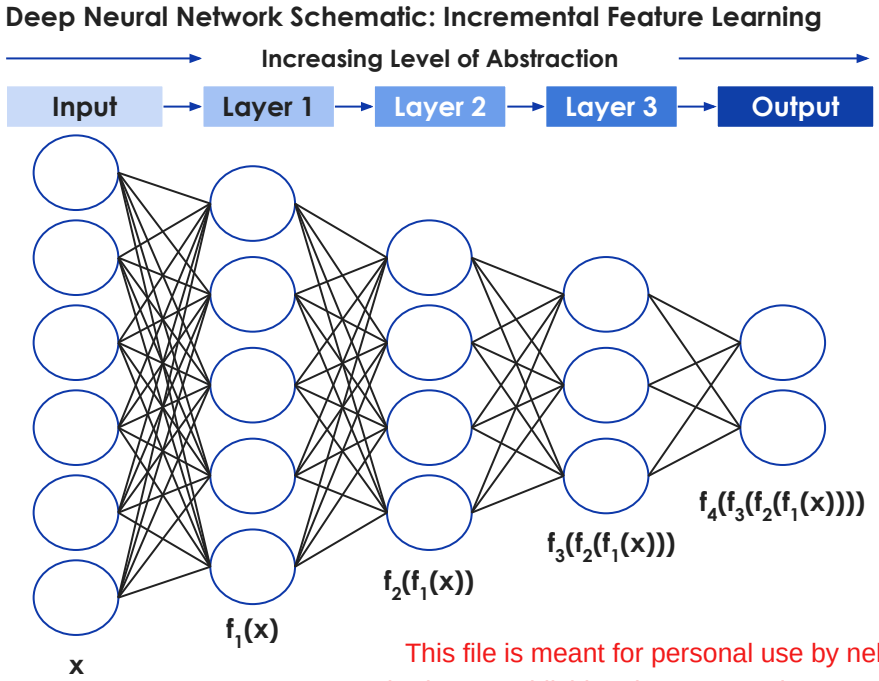
- The reason is that making output predictions directly from unstructured input data like images, text, audio or graphs is a complex, difficult task for Machine Learning to achieve in one go.

- Rather, to maximize our chances of success with unstructured data, it is necessary to **break down the complex task of finding this input-output mapping into an incremental sequence of simpler tasks.** This is what Deep Learning does, and why it's been so successful on unstructured data.

- Now let's understand **how** it accomplishes this.

Deep Learning: A Nested Hierarchy of Concepts

- Deep Learning represents every idea as a **nested hierarchy of concepts** - where every concept is defined in relation to simpler concepts, and more abstract representations are computed in terms of less abstract ones.

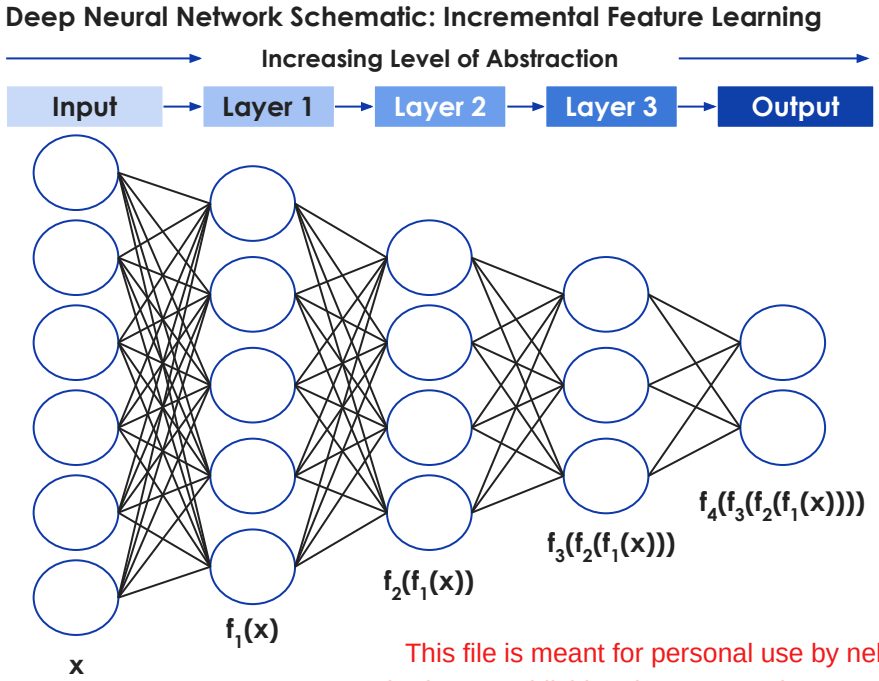


Nested because each layer of the neural network builds on the output of the previous layer - essentially a mathematical function of a function.

And **Hierarchy** because there is an order to the layers - the deeper we go, the more the level of abstraction and the more complex the features learned.

Deep Learning: A Nested Hierarchy of Concepts

- Deep Learning represents every idea as a **nested hierarchy of concepts** - Where every concept is defined in relation to simpler concepts, and more abstract representations are computed in terms of less abstract ones.



By building a model with successive layers, taking outputs from the previous layer as new inputs, **Deep Learning models are able to incrementally learn higher-level characteristics** directly from the data, and this is what eliminates the need for manual feature engineering, which can be domain specific and a constraint on model performance.

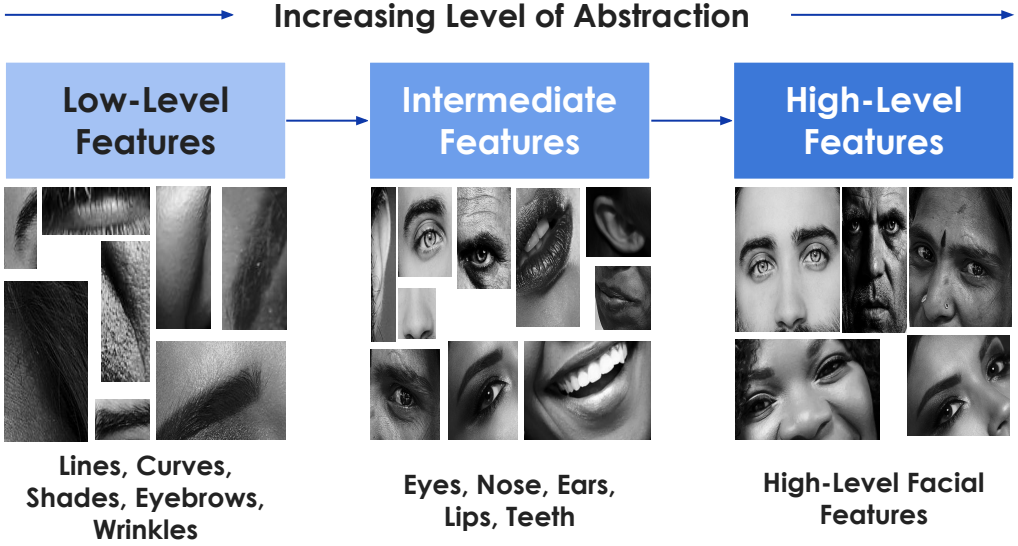
Deep Learning: A Nested Hierarchy of Concepts

- Let's understand this with a more concrete example.
- Let's assume we're building a **Deep Neural Network to perform Facial Recognition**, an important use case of Deep Learning that has now become widely adopted across the globe due to Apple's Face ID security authentication and similar software applications from other providers.

Let's assume these are faces of different people that the Neural Network must learn to recognize and tell apart.

Deep Learning: A Nested Hierarchy of Concepts

- What happens is that **the Deep Neural Network incrementally learns the features that characterize a person's face in a hierarchical manner**, and through this multi-layered approach, eventually learns what one person's face looks like and how it's different from that of someone else.



Deep Learning: A Nested Hierarchy of Concepts

- The layers between the input and output, called the **Hidden Layers**, are what allow Deep Neural Networks to learn features in this incremental manner. Each hidden layer takes the quality of the feature learning of the Neural Network to a different level, and so every layer has a role to play in creating the complex Deep Learning function that finally maps input to output.

- As we move forward, we will understand the mathematical and programmatic details of how exactly this nested hierarchy of features is built, and how we can create our own neural networks using publicly-available Deep Learning frameworks, to make predictions on unstructured data modalities.

Summary

So in order to summarize:

- **Deep Learning** is different from Machine Learning because **it combines the Feature Extraction and Classification stages of a prediction problem into the same model architecture**, where the hidden layers of the model are responsible for learning features, and the output layer is where the final prediction is computed.

- The reason this is required, is because making output predictions directly from unstructured input data like images, text, audio or graphs is complex and difficult to do. Rather, if we were to break the problem down into a series of smaller and simpler tasks, that would give the model a better chance of finding the right mapping from input to output.

Summary

So in order to summarize:

- Deep Learning is able to do this because it represents every idea as a **nested hierarchy of concepts**, and this allows it to incrementally learn more complex features through the hidden layers in its architecture.

- This characteristic is what makes Deep Learning so powerful and flexible, and allows it to mostly eliminate any specialized feature extraction techniques, which can be manual and domain-specific.

Great Learning

Happy Learning !