

# Module 1 – Part I

## Welcome to the Deep forecasting course!

### What is Time Series Forecasting?

---



# ➔ What is Forecasting?

- Forecasting has fascinated people for thousands of years!

*Tell us what the future holds, so we may know that you are gods.  
Isaiah 41:23 700 BC*

- Forecasting is about predicting the future as accurately as possible, given all of the information available, including historical data and knowledge of any future events that might impact the forecasts.
- Forecasts could be **short-term**, **medium-term** or **long-term**.





# Quantitative vs Qualitative Forecasting

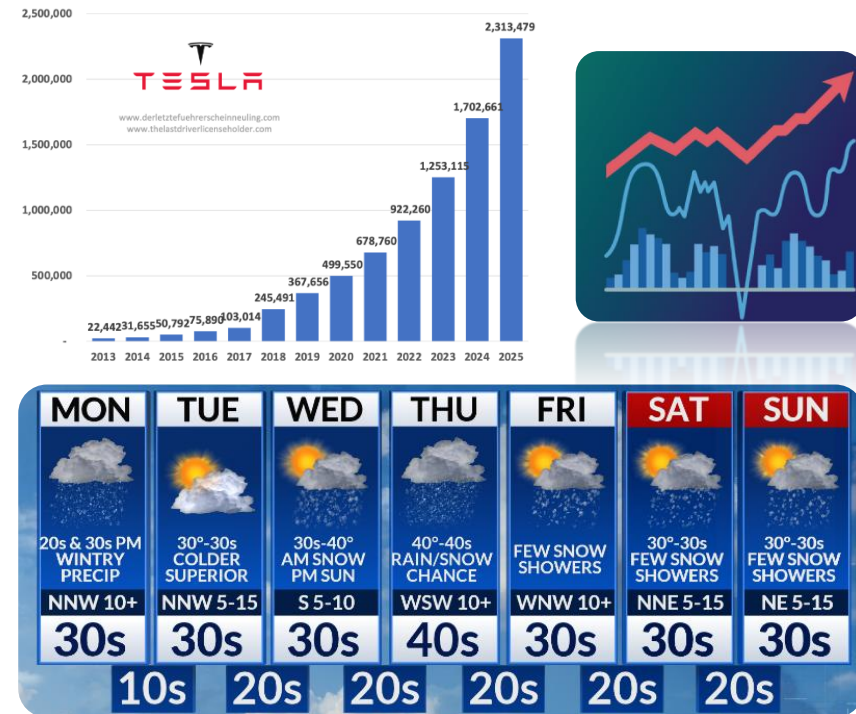


	Quantitative Forecasting	Qualitative Forecasting
Data sources	<b>Numeric</b> data and <b>statistical</b> analysis	<b>Expert opinions</b> , <b>subjective</b> judgment, or other non-numeric data
Accuracy	Can be <b>more accurate</b> when large amounts of reliable data are available	May be less accurate than quantitative methods, but can still be useful when data is difficult to measure numerically
Suitability	Suitable for forecasting phenomena that can be easily measured and tracked	Suitable for forecasting subjective or hard-to-measure phenomena
Flexibility	Rigid and <b>less flexible</b> ; relies on specific data and statistical techniques	<b>More flexible</b> and allows for the incorporation of expert judgment and other subjective factors
Examples	Stock prices, sales revenue, market demand for a stablished product, weather forecast, population growth	Fashion trends, market demand for a new product, employee performance

# ➔ More Forecasting examples

- Which of the following examples are easier to forecast?

1. Time of sunset this day next month
2. Apple stock price in 6 months
3. Apple stock price tomorrow
4. EURUSD exchange rate next week
5. Airline ticket demand next year
6. Airline ticket prices next year
7. US presidential election 2024
8. Monthly rainfall in Utah next winter

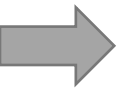




# ➔ What Impacts Forecastability?

- How do we say something is easier to forecast?
- Forecastability factors are:
  - Data Availability
  - How similar the future is to the past!
  - Good understanding of the underlying factors





# Explanatory vs Timeseries vs Mixed models

Model	Example
Explanatory (Cross sectional)	$P = f\left(\frac{P}{E}, \frac{P}{S}, size, \frac{B}{M}, GDP, CPI, \dots, u\right)$
Timeseries	$P_{t+1} = f(P_t, P_{t-1}, P_{t-2}, \dots, u)$
Mixed (dynamic regression, panel)	$P_{t+1} = f\left(\frac{P_t}{E_t}, \frac{P_t}{S_t}, size_t, \frac{B_t}{M_t}, GDP_t, CPI_t, \dots, u_t\right)$

- In this course we focus on **Timeseries models** because:
  1. We may **not know all the underlying factors!**
  2. Extremely difficult to **know or forecast the future value of many factors** when forecasting the variable of interest
  3. We are more interested in **predictability** rather than **explanatory** power



# Basic steps in a forecasting task

## Step 1: Problem definition

- Forecasting **type** and **horizon** (one-step, multi-step, multi-output forecasts), ...

## Step 2: Data Collection

- Time horizon, structural changes, **data type**, ...

## Step 3: Exploratory Analysis

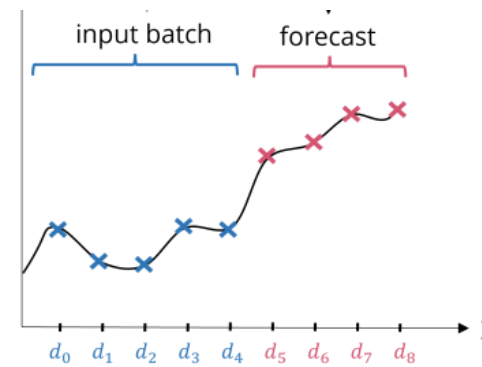
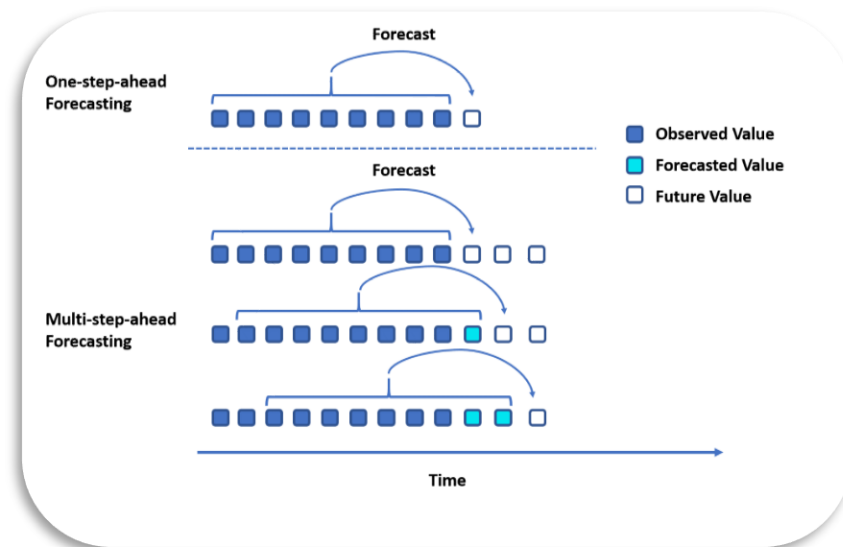
- Trend, seasonality, outliers, ...

## Step 4: Model Selection and Training

- **Traditional** vs **machine learning** vs **deep learning**

## Step 5: Model Evaluation and Comparison

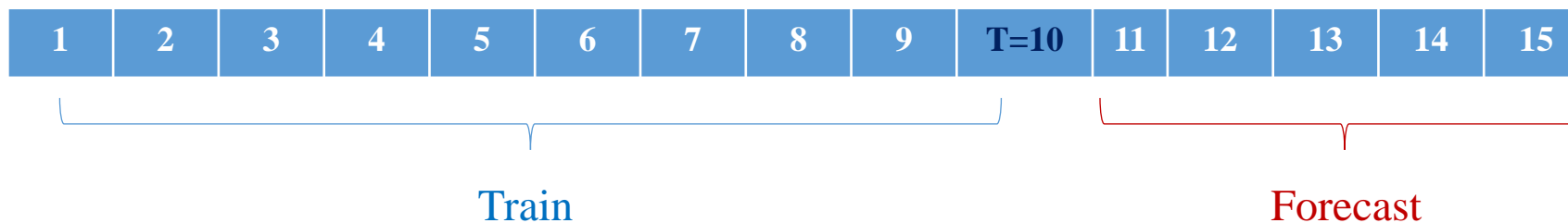
- MSE, RMSE, MAE,  $R^2$ , MAPE, sMAPE, ...



# Forecasting notation

$$\hat{y}_{t+h|t} = f(y_t)$$

- $y_t$  itself can be decomposed into different components (level, trend, seasonality)
- Fitted values at time  $t = 1 \dots T$ , are  $\hat{y}_{t|t-1}$  ( $h = 0$ )
- One-step ahead forecast at time  $T + 1$  ( $T$  last observation in train data) and  $h = 1$ .
- Multi-step ahead forecast:  $h = 2, 3, 4, \dots$ 
  - One-output at a time
  - Multi-output at once







# What is Time Series analysis and why it matters?

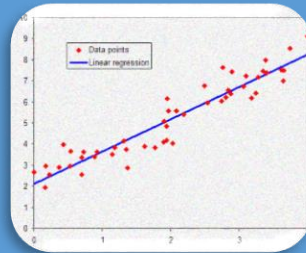
- Time series analysis is a powerful tool for **understanding** and **predicting** trends and patterns in data that are collected **over time**.
- Time series analysis is also useful for business **decision makers**, as it can help them to **forecast** future trends and make informed decisions based on data trends and patterns.
- **Why?**
  - Time series data is everywhere!
  - Better **career opportunities**,
    - Up to 90% of companies need better forecasting and
    - Less than 5% of data scientists are competent in time series analysis.
  - **Hedge against next recession!**



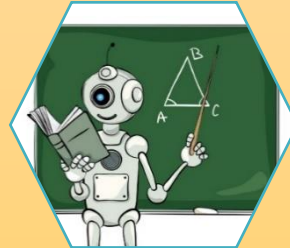


# What is our approach to time series analysis?

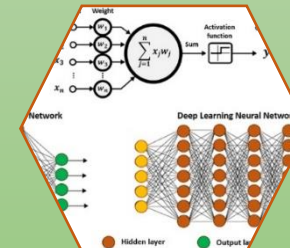
Econometrics



Machine Learning



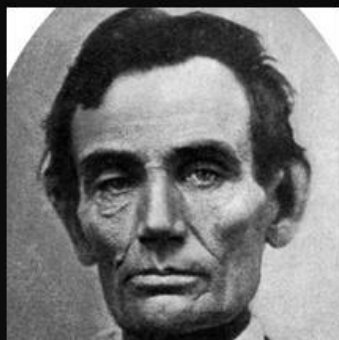
Deep Learning



## Deep Forecasting



# Forecasting is not always easy!



The most reliable way to predict  
the future is to create it.

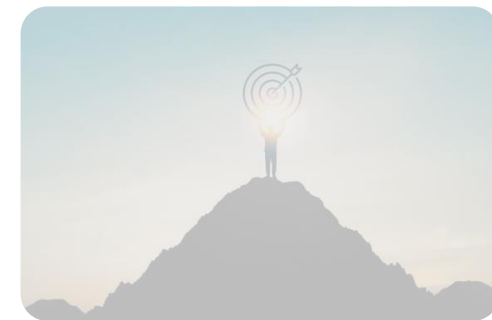
~ Abraham Lincoln

AZ QUOTES



# Road map!

- Module 1- Introduction to Deep Forecasting
- Module 2- Setting up Deep Forecasting Environment
- Module 3- ETS and Exponential Smoothing
- Module 4- ARIMA models
- Module 5- Machine Learning for Time series Forecasting
- Module 6- Deep Neural Networks
- Module 7- Deep Sequence Modeling (RNN, LSTM)
- Module 8- Transformers (Attention is all you need!)
- Module 9- Prophet and Neural Prophet

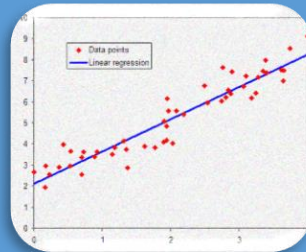


# Module 1 – Part II

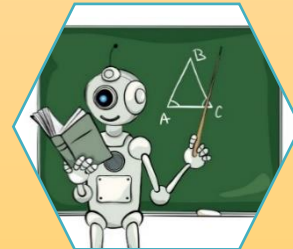
## What is Deep Forecasting?



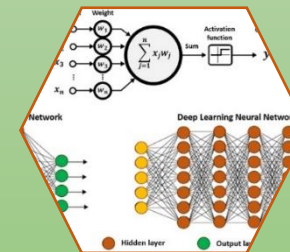
Econometrics



Machine Learning



Deep Learning



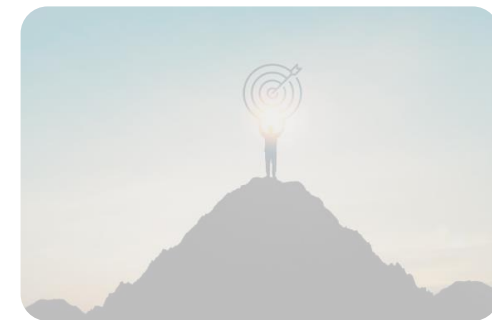
# Deep Forecasting





# Road map!

- **Module 1- Introduction to Deep Forecasting**
- Module 2- Setting up Deep Forecasting Environment
- Module 3- ETS and Exponential Smoothing
- Module 4- ARIMA models
- Module 5- Machine Learning for Time series Forecasting
- Module 6- Deep Neural Networks
- Module 7- Deep Sequence Modeling (RNN, LSTM)
- Module 8- Transformers (Attention is all you need!)
- Module 9- Prophet and Neural Prophet





# Big picture: Econometrics vs Machine Learning and Deep Learning



What are we trying to do as researchers? Solve real world problems, right?

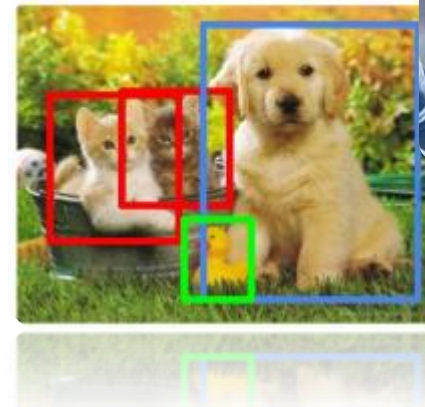


Is there a theory?

1- What is the **relationship** between

- Quantity demanded and price / income / technology / price of competitors / ... ?
- Wage and education/ age/ gender/ experience/ ...?

2- How about these problems? Object detection, Image Captioning, voice recognition, machine translation, and ...



# ➔ A simple example

- Quantifying wage components! (is there a theory?)
- What are the drivers:
  - Demographic variables: Education, age, experience, IQ, ...
  - Social and cultural variables: Ethnicity, race, gender, ...
  - Job characteristic variables: Industry, location, working hours, ...
- Let's build a model (**assuming** a linear functional form!)



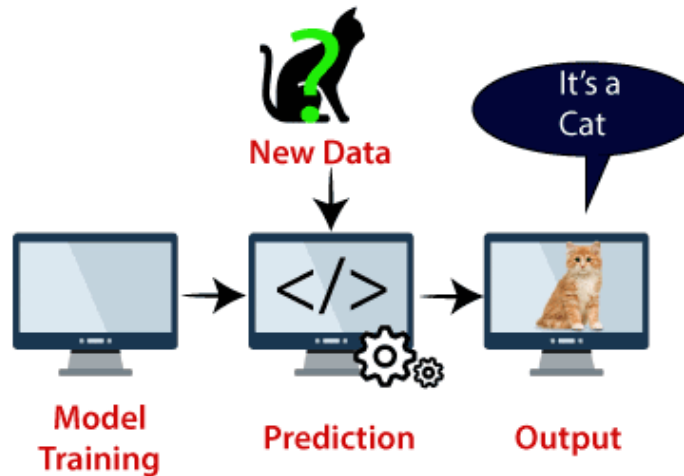
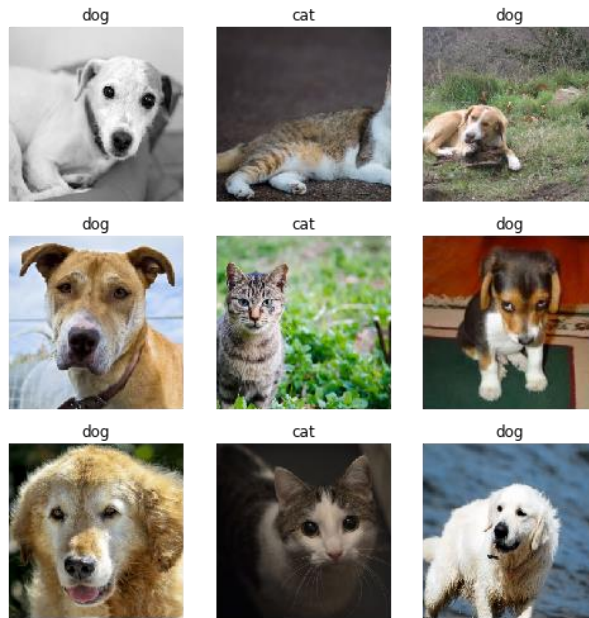
$$wage = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 exper + \beta_4 IQ + \dots + \beta_k hours + u$$

- Can you **interpret** this model? Do you care about the interpretability?
- Can you make **predictions** using your model?
- Can you make this functional form more flexible? What are the caveats?



# A different example

- Cat vs dog classification problem (image recognition)



- Do you really care about **interpretability** of the model here?
- What about accuracy of your **predictions**?



# Statistical learning vs machine learning

	Statistical Learning	Machine Learning / Deep Learning
Focus	Hypothesis testing & interpretability	Predictive accuracy and extracting complex patterns
Driver	Math, theory, hypothesis	Fitting data
<b>Data size</b>	Any reasonable set	Big data
<b>Data type</b>	Structured	Structured, unstructured, semi-structured
Dimensions / scalability	Mostly <b>low</b> dimensional data	<b>High</b> dimensional data
Strength	Understand <b>causal</b> relationship & behavior	Prediction (forecasting and nowcasting)
<b>Interpretability</b>	<b>High</b>	<b>Medium to Low</b>



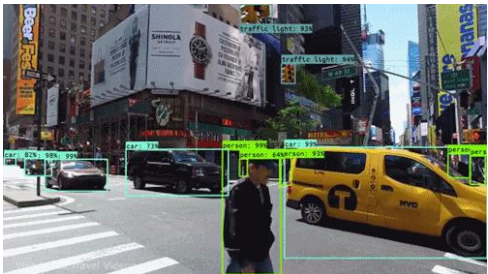


# Limitations of Econometrics/Structured ML

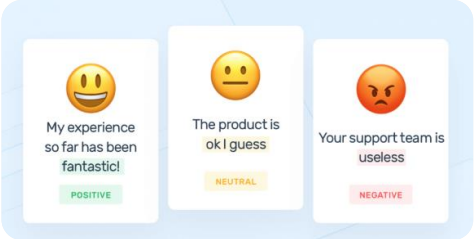
Econometrics/structured ML can only handle structured data (tabular data)!

## Structured Data

	A	B	C	D
1	Date	Account	Transaction Type	Amount
2	2017-01-12	123	Credit	6089.78
3	2017-01-12	123	Fee	9.99
4	2017-01-12	456	Debit	1997
5	2017-01-12	123	Debit	20996.12
6	2017-01-13	123	Debit	17
7	2017-01-13	123	Debit	914.36
8	2017-01-14	789	Credit	11314
9	2017-01-14	789	Fee	9.99
10	2017-01-14	456	Debit	15247.89
11	2017-01-14	123	Debit	671.28
12	2017-01-15	456	Credit	5072.1
13	2017-01-15	456	Fee	9.99
14	2017-01-16	456	Debit	5109.07
15	2017-01-19	123	Credit	482.01



Unstructured Data  
(everything else!!)



# ➔ A more complex example

## Stock price prediction \$\$\$

- What are the classical drivers:
  - Company's fundamentals (balance sheet, income statement, cash flow statement)
  - Competitors (comparing multiples)
  - Technical analysis!
  - Seasonality (holidays, months, days, ...)



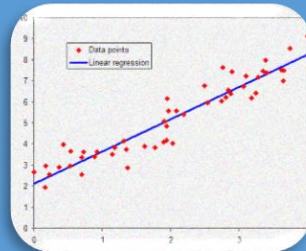
## What else?

- Market sentiment (news, tweets, blogger opinions, conference calls, ...)
- Satellite images from parking lots!

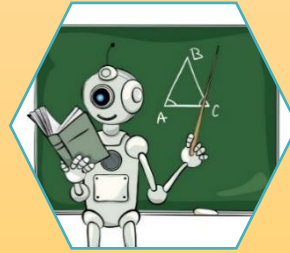


# What is our approach to time series analysis?

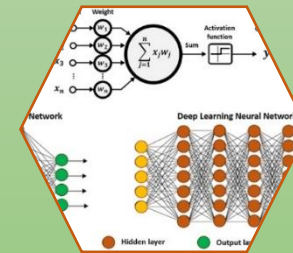
Econometrics



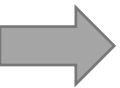
Machine Learning



Deep Learning



## Deep Forecasting



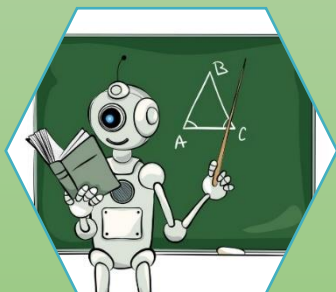
# Artificial intelligence vs Machine learning vs Deep learning

Artificial intelligence: Any technique which enables machines to mimic human behavior



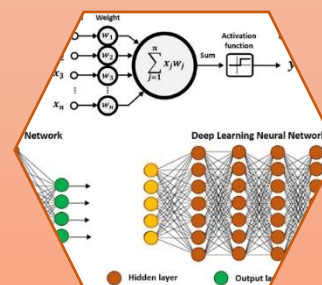
1950's

Machine Learning: Subset of AI that enables computers to learn from data. the model is trained with a set of algorithms



1980's

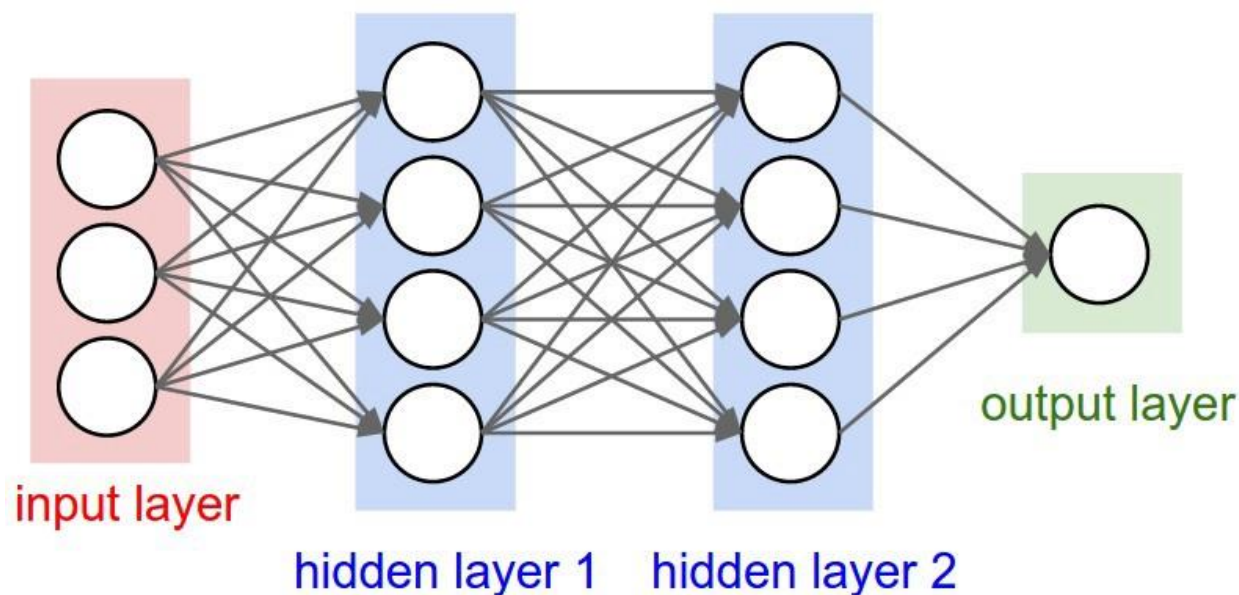
Deep Learning: Subset of ML that extract patterns from data using neural networks.



2010's

# ➔ What is Deep Learning?

- Deep learning is a type of machine learning that uses **multiple layers of neurons** to process data
- The goal of deep learning is to build a model that can **automatically learn complex patterns** from the data and make **accurate predictions** or decisions

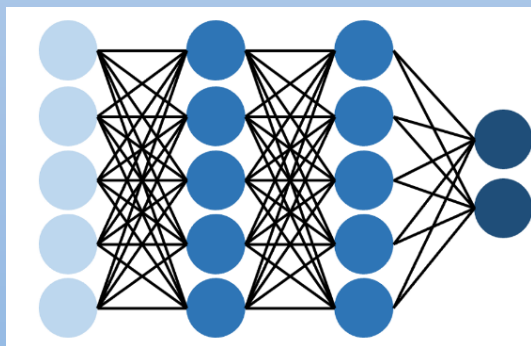




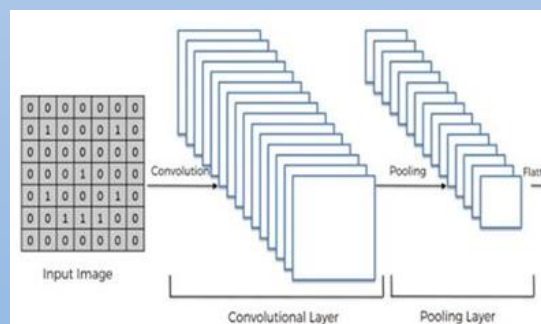


# Network examples

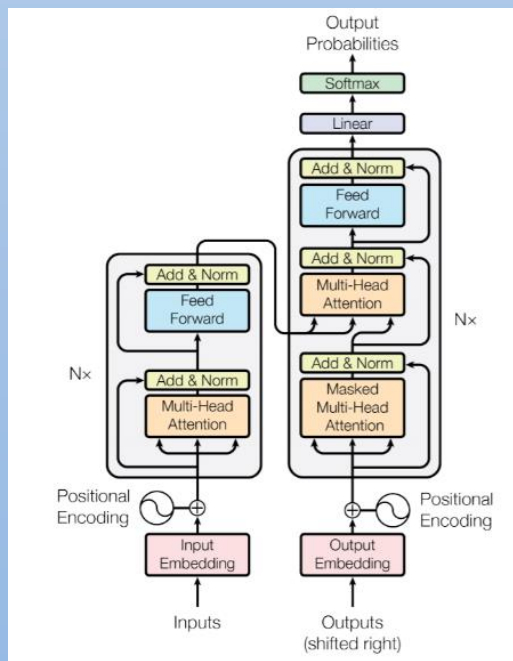
## Standard NN



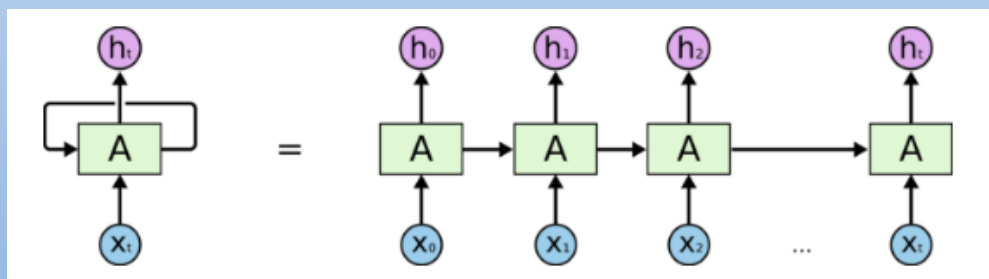
## Convolutional NN



## Transformers

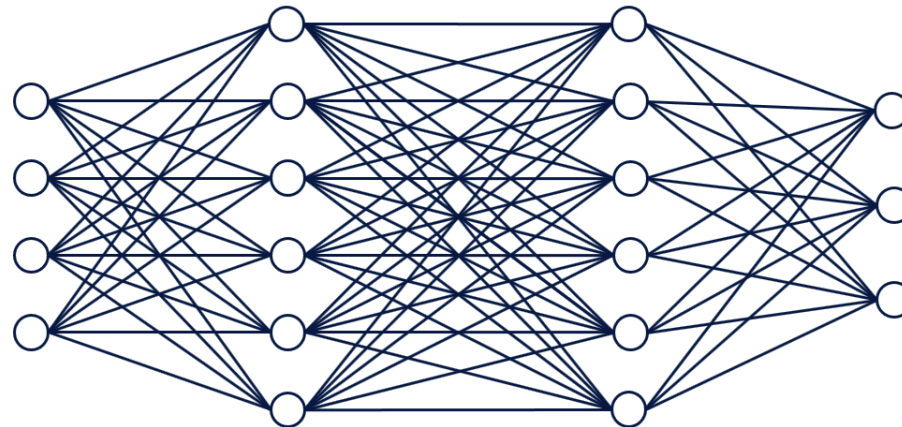


## Recurrent NN



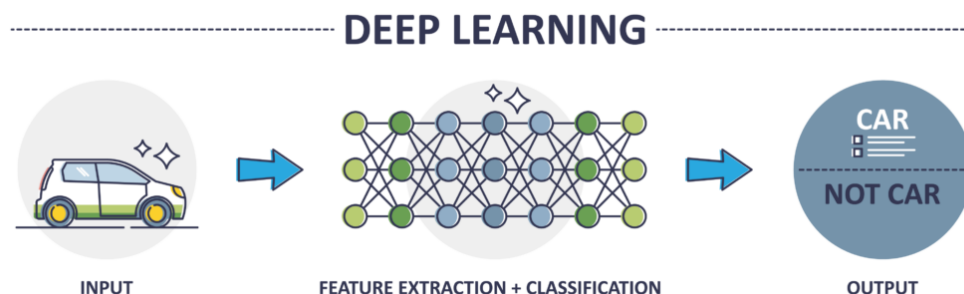
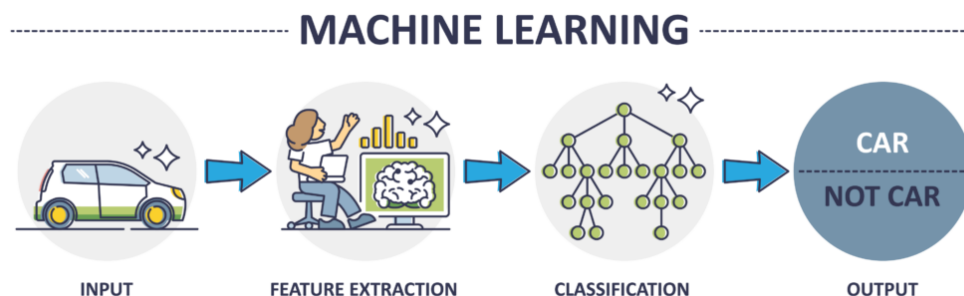
# ➔ How Deep Learning Works?

- The network uses a series of interconnected layers, with each layer transforming the input and **passing** it on to the next layer
- The **final layer of the network produces the output**, which is compared to the **true output** label to evaluate the performance of the network
- The network is then **adjusted** to minimize the difference between the predicted and true output labels, and the process is **repeated until the network has learned** to make accurate predictions on new, unseen data



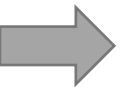
# ➔ Why Deep Learning?

- In deep learning, the model is **not explicitly programmed** with a set of rules or algorithms, but instead learns to recognize patterns and make predictions by adjusting the connections between the neurons in the network!
- Can learn the underlying features **directly from data** and in a hierarchical manner. **No feature extraction/engineering required!**



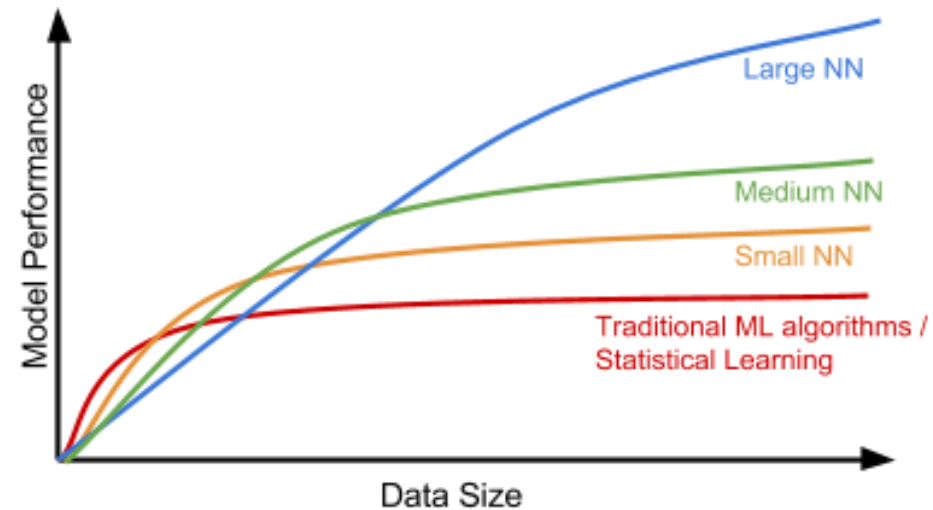
# → Why Deep Learning?

	Machine Learning	Deep Learning
Focus	General-purpose predictions	Extracting <b>high-level features</b> from data
Level of <b>Abstraction</b>	Explicitly programmed (rules/algos)	Learn by adjusting neurons
Examples	KNN, SVM, DT, RF, XGBoost, ...	CNN, RNN, LSTM, GAN, Transformers
<b>Requires Feature extraction</b>	Yes	No
Requires high processing power	Not necessarily	Yes
Interpretability	Medium to low	Very low to <b>none</b>
Execution Time on training	Less (compared to DL)	A lot
Execution Time on testing	More (compared to DL)	Little



# Why Now?

- In small data sets, perhaps traditional machine learning is more effective!
- The performance of deep learning is better when the **data is large**.
- Training large NN is **computationally expensive**.
- **Algorithms** are playing important role in speeding up the training process.





# ➔ Road map!

## ✓ Module 1- Introduction to Deep Forecasting

- Module 2- Setting up Deep Forecasting Environment
- Module 3- ETS and Exponential Smoothing
- Module 4- ARIMA models
- Module 5- Machine Learning for Time series Forecasting
- Module 6- Deep Neural Networks
- Module 7- Deep Sequence Modeling (RNN, LSTM)
- Module 8- Transformers (Attention is all you need!)
- Module 9- Prophet and Neural Prophet

