

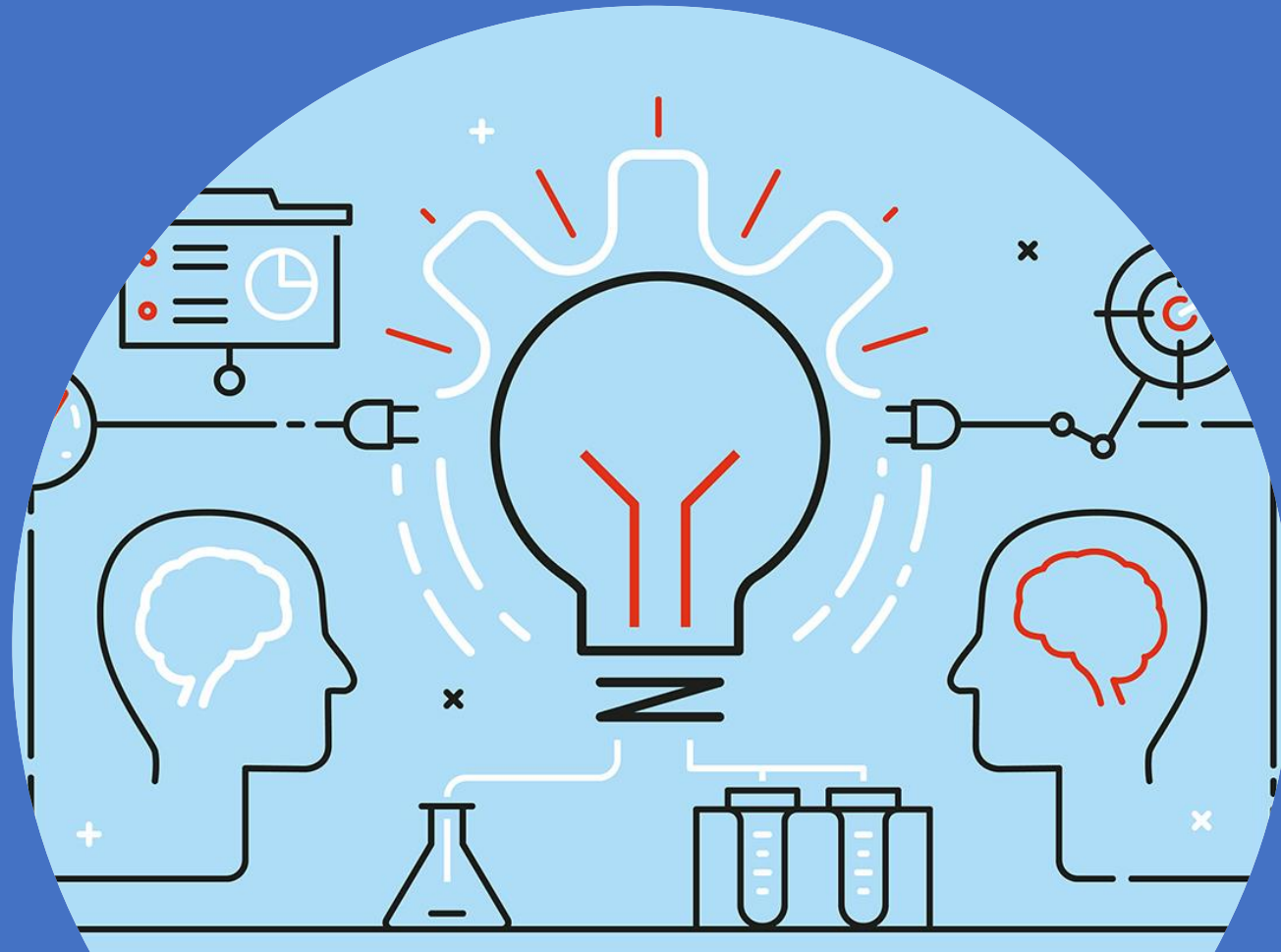
# ML-MACHINE LEARNING DL-DEEP LEARNING AI – ARTIFICIAL INTELLIGENCE DS – DATA SCIENCE

---

VENKATA REDDY KONASANI

# PART-1

## What is Machine Learning?



# Activity

Close your eyes and think about two terms – Machine Learning and Artificial Intelligence.

- What is the first thing that comes to your mind when you hear these terms
  - Machine Learning
  - Artificial Intelligence



# WHAT ARE YOU THINKING ABOUT – ROBOTS?



marking your homework ...



AI Robots Kill 29 Scientists in Japan ...  
snopes.com



Open Letter Slamming Europe's Proposal ...  
gizmodo.com



RoboBusiness Conferen ...  
zdnet.com



...t replace humans in decisi...



The Day Humans Taught Robots to Fight ...  
livescience.com



Humanoid robot takes over as teacher ...  
sciencenordic.com



Toyota Gets Back Into Humanoid ...  
spectrum.ieee.org



ABB, Kawasaki join hands for ...  
moneycontrol.com



About 800 million workers worldwide may ...  
economictimes.indiatimes.com



robots in Star Wars ...  
techxplore.com



# OR ANY AI BASED MOVIE?



Artificial Intelligence - Official ...  
a.com



ALITA BATTLE ANGEL Official Trailer ...  
youtube.com



Transcendence movie has Johnny Depp ...  
robhub.org



Heroes and Villains ...  
m



I, Robot (film) - Wikipedia  
en.wikipedia.org



Robot Heroes and Villains ...  
etonline.com



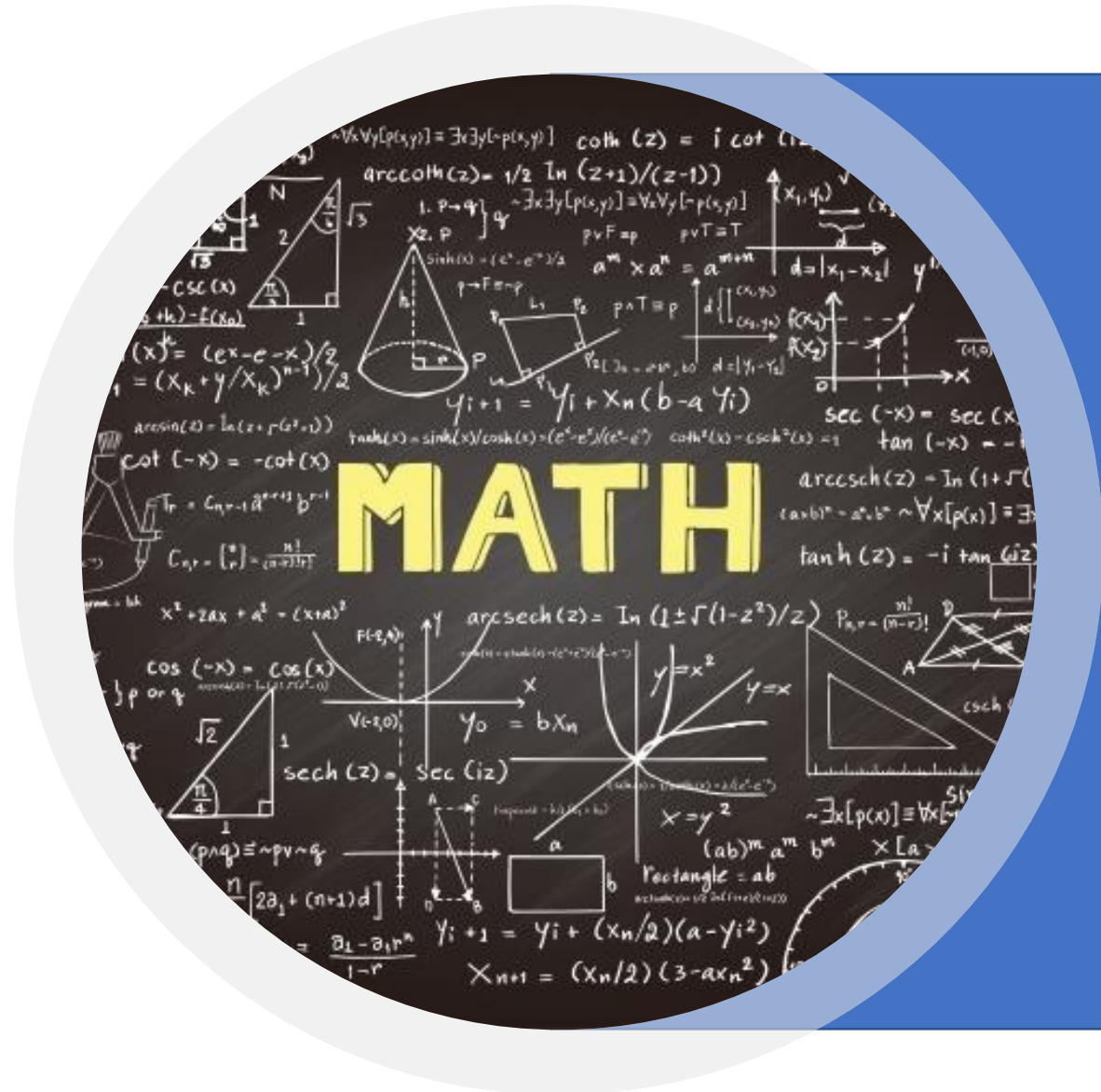
Bicentennial Man (film) - W...  
wikipedia.org



Steven Spielberg  
avclub

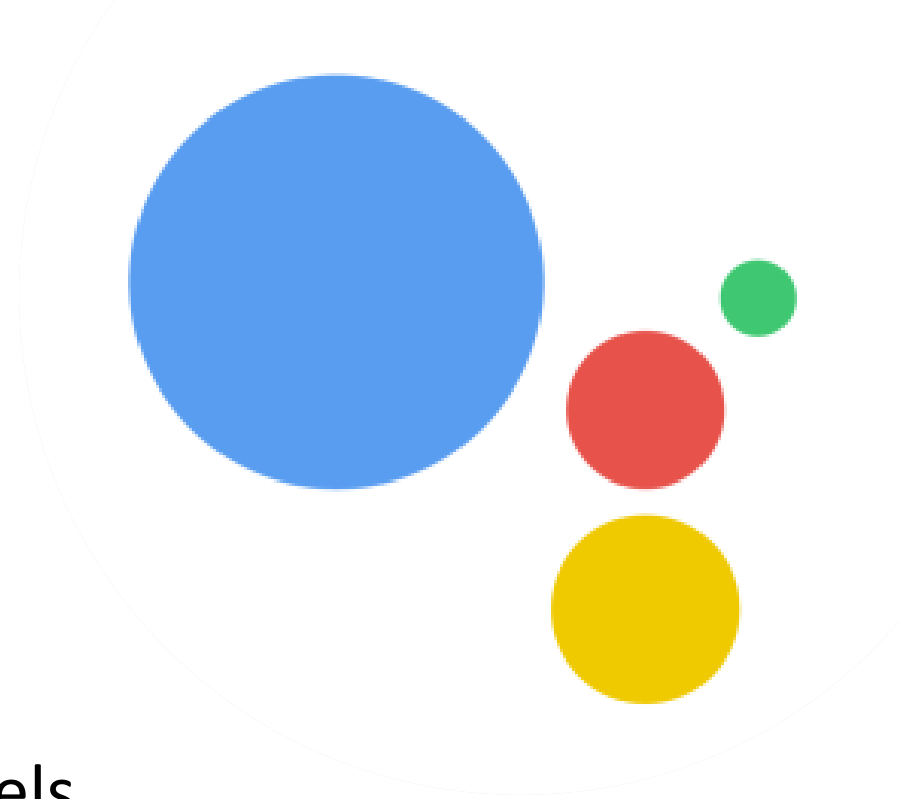
# HAVE YOU THOUGHT ABOUT

- Mathematics
- Statistics
- Datasets
- Data Analysis
- Optimization
- Algorithms
- Data Mining



# OK GOOGLE ...WHAT IS MACHINE LEARNING?

Machine learning (ML) is the scientific study of algorithms and **statistical models** that **computer systems** use to effectively **perform a specific task** without using explicit instructions, relying on models and inference instead.



# WHAT REALLY IS MACHINE LEARNING ?- WIKIPEDIA

**Machine learning** (ML) is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on models and inference instead. It is seen as a subset of artificial intelligence. Machine learning algorithms build a mathematical model of sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task.<sup>[1][2]</sup> Machine learning algorithms are used in the applications of email filtering, detection of network intruders, and computer vision, where it is infeasible to develop an algorithm of specific instructions for performing the task. Machine learning is closely related to computational statistics, which focuses on making predictions using computers. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a field of study within machine learning, and focuses on exploratory data analysis through unsupervised learning.<sup>[3][4]</sup> In its application across business problems, machine learning is also referred to as predictive analytics.



## IN SIMPLE TERMS ..

- Using historical data to make future predictions
- Building models on historical data to predictions
- Taking training data, building models on the training data using the models to make the future predictions
- Making the machine learn the patterns in the data



# DATA IS IN DIFFERENT FORMS



Numerical data

Image data (pixel intensities)

Video data (frames per second)

Sound data (waves)

Text data (tweets, comments, feedback)

# APPLICATIONS OF NUMERICAL DATA



## CREDIT RISK MODELS

Identifying risky customers before offering a loan



## MARKETING AND SALES ANALYTICS

Identify potential buyers from the population



## RETAIL SALES ANALYTICS

Promotional offers and customer attrition management



## FRAUD ANALYTICS

Detecting fraud transactions from millions of transactions

# APPLICATIONS OF MACHINE LEARNING – IMAGES AND VIDEO DATA

- Face recognition – Using image as input data
- Object recognition – Pixels is the input data
- Digit recognition – Using text as image
- Self Driving Cars – Using video data as input





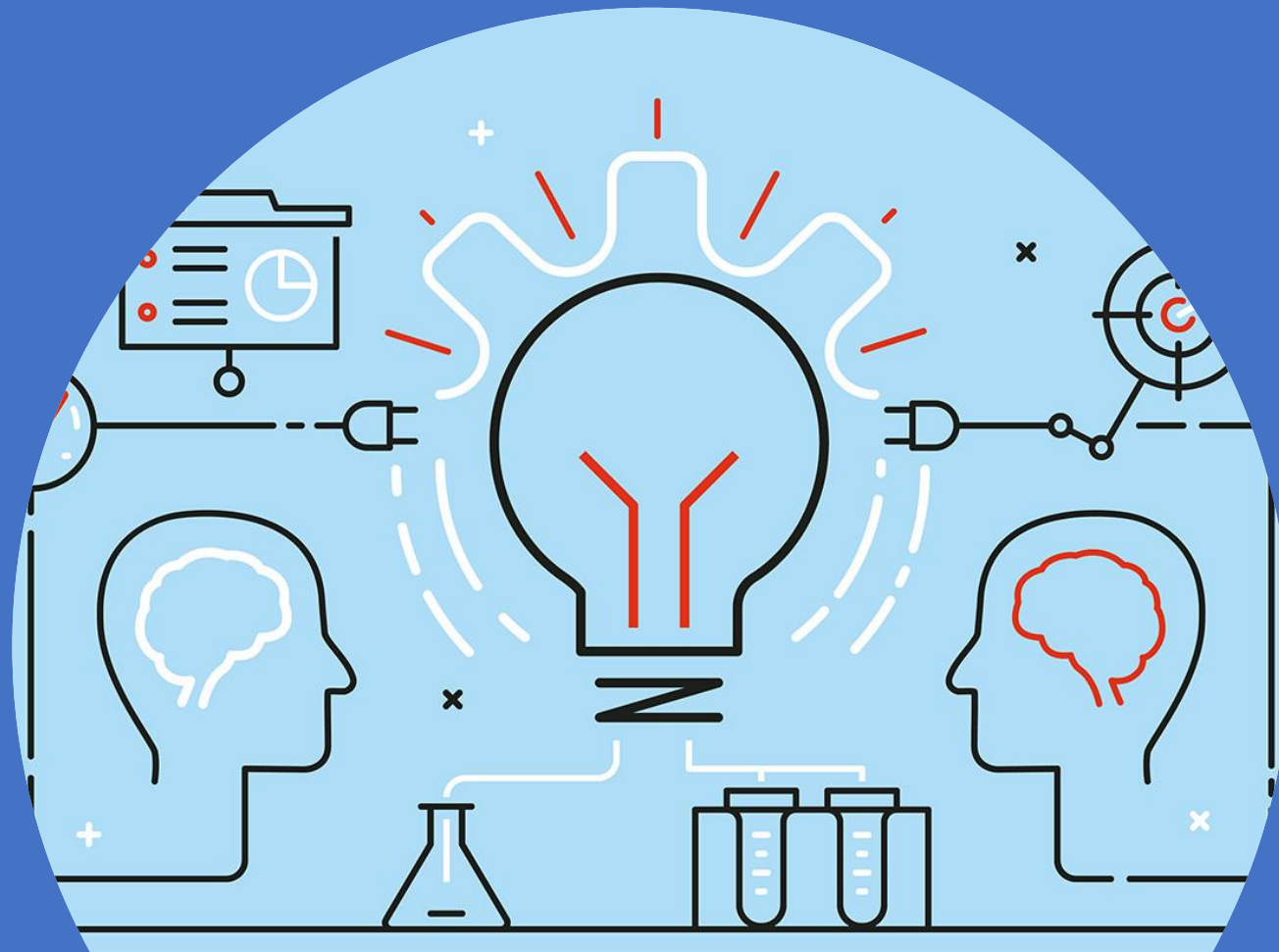
# APPLICATIONS ON TEXT DATA

- Sentiment Analysis
- Extraction of key topics in the data
- Document Classification
- Context understanding



# PART-2

## What is Deep Learning?



# ANN

- ANN- Artificial Neural Network
- ANN is one of the technique in Machine Learning
- ANN has input layer , hidden layer and output layer
- For a really complex and non liner datasets we need several hidden layers
- ANN with multiple hidden layers is known as deep neural network

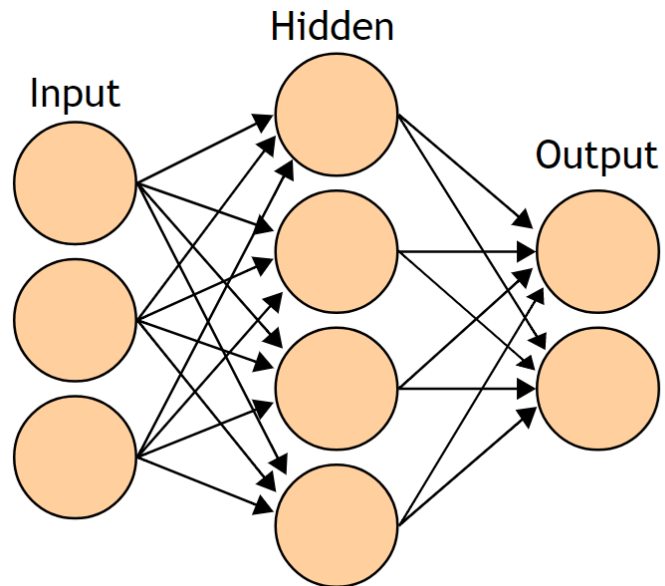
# DEEP LEARNING

- ANN with a single layer is known as shallow network
- ANN with multiple hidden layers is known as deep neural network
- Not just multiple hidden layers sometimes the type of hidden layer is also different.
- This concept of solving problems with multiple hidden layers is known as **deep learning**

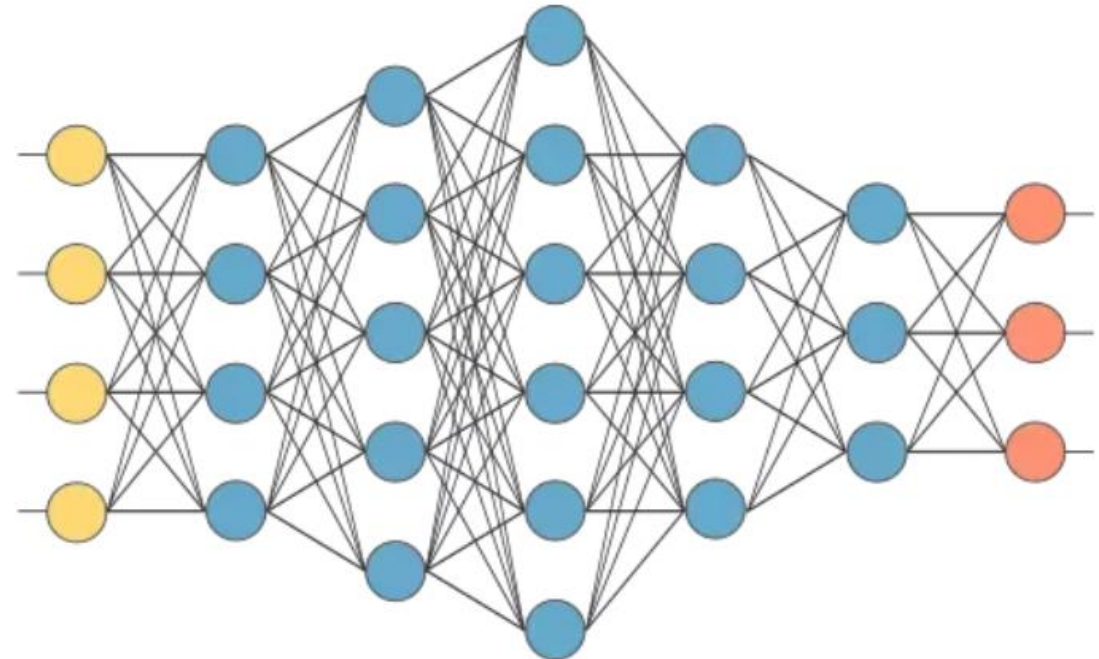


# DEEP VS SHALLOW NETWORKS

A neural network with single hidden layer is called a shallow network

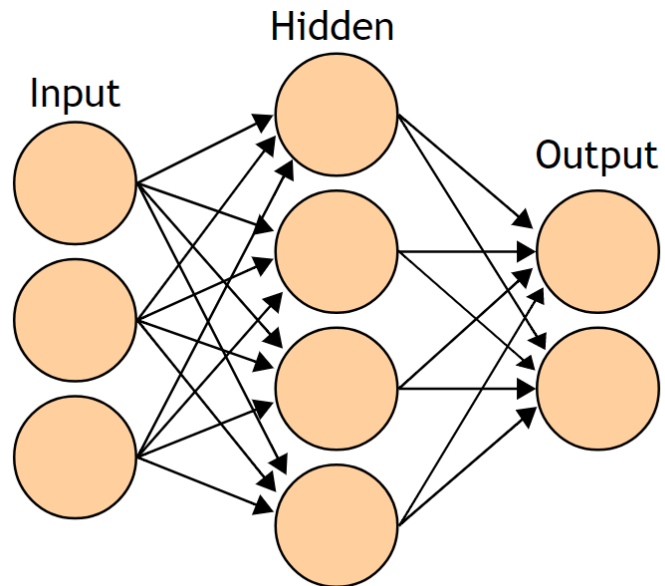


A neural network with more than one hidden layer is called deep neural network

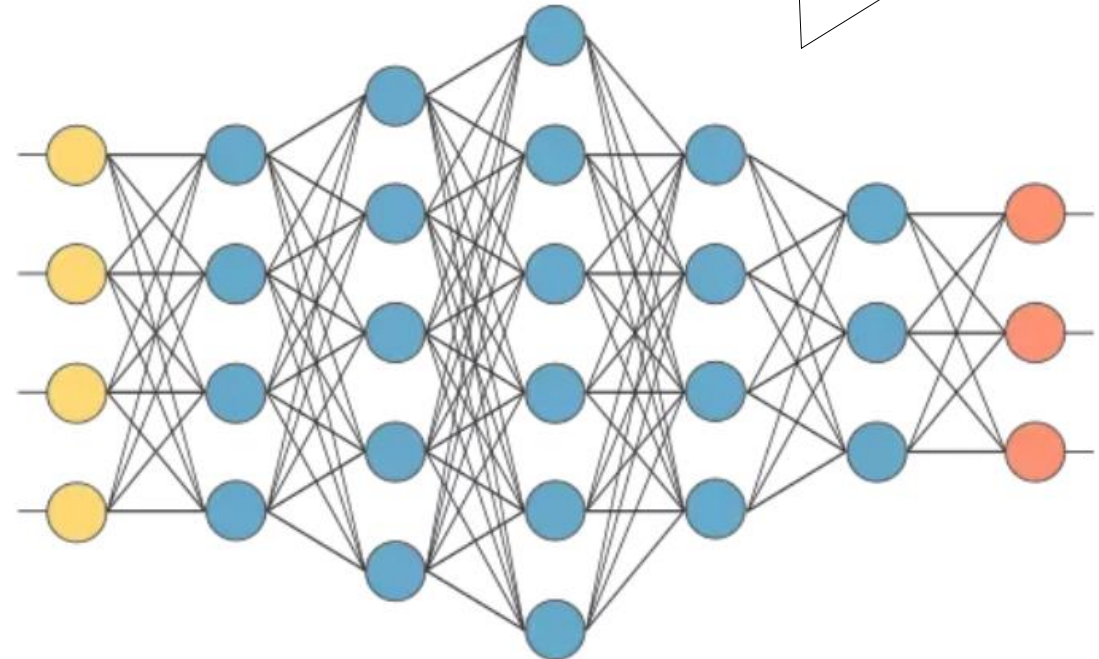


# DEEP VS SHALLOW NETWORKS

A single layer might not have the flexibility to capture all the non linear patterns in the data



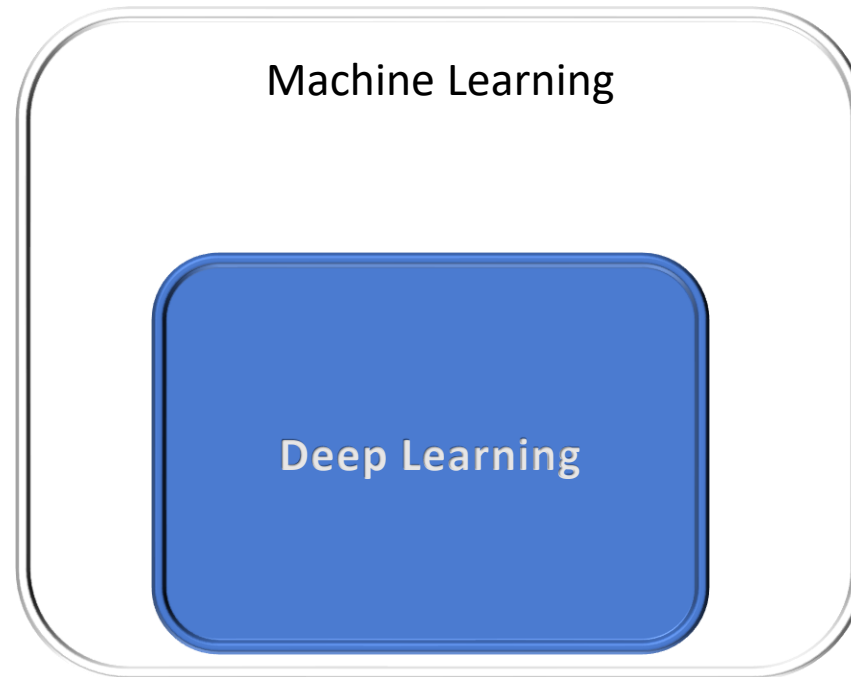
A deep network first learns the primitive features followed by high level features. This helps in building efficient models



# DEEP NEURAL NETWORKS

- Lot of experiments have shown that a deep network with less parameters performs better than a shallow network
- For example deep network with hidden nodes  $[10, 10, 10, 10]$  might perform better than shallow network with  $[80]$  hidden nodes
- Deep neural networks are amazingly powerful.
- With sufficient number of hidden layers and nodes, we can fit a model to any type of data
- They have the power to capture any amount of non linearity

# DEEP LEARNING IS A SUBSET OF MACHINE LEARNING





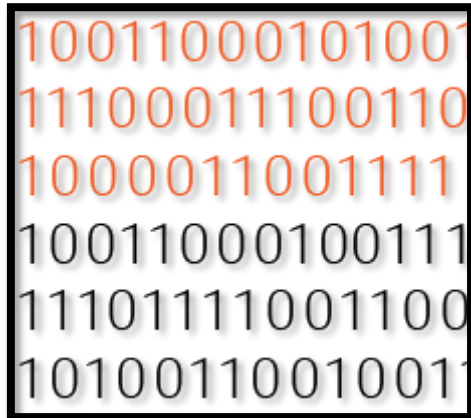
# PART-3

## What is Artificial Intelligence?

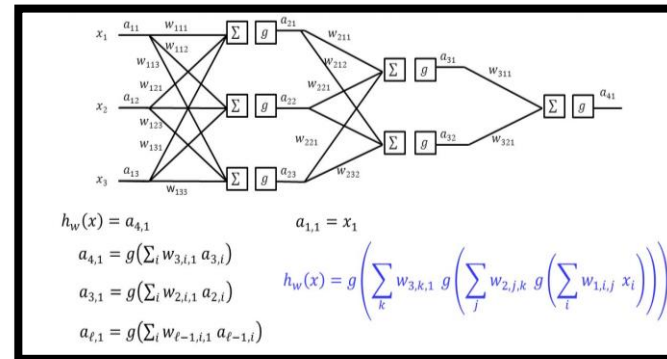


# MACHINE LEARNING MODELS

Training data




Build Model

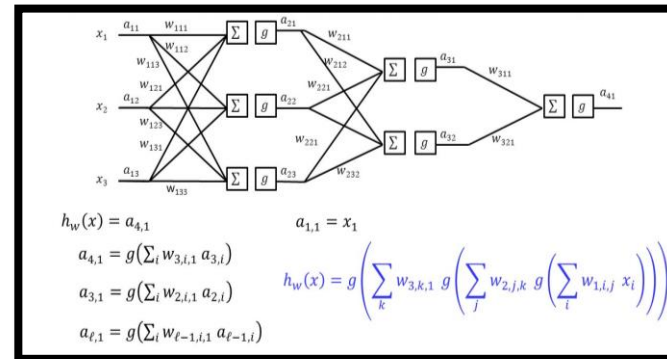


# MACHINE LEARNING MODELS

New data

11100011100110  
1000011001111

Apply Model



Get Prediction

Class1

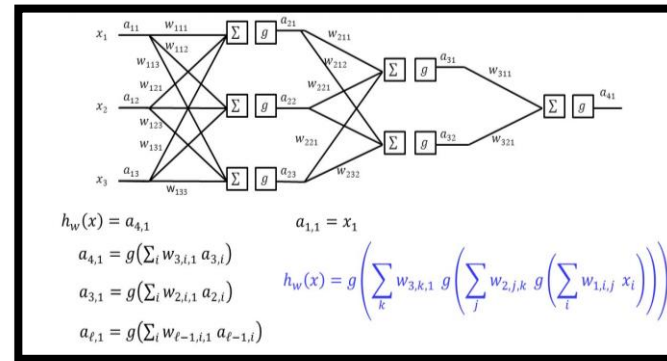
This prediction can be right or wrong

# MACHINE LEARNING MODELS

New data

11100011100110  
1000011001111

Apply Model



Get Prediction

Class2

This prediction can be right or wrong

One way models



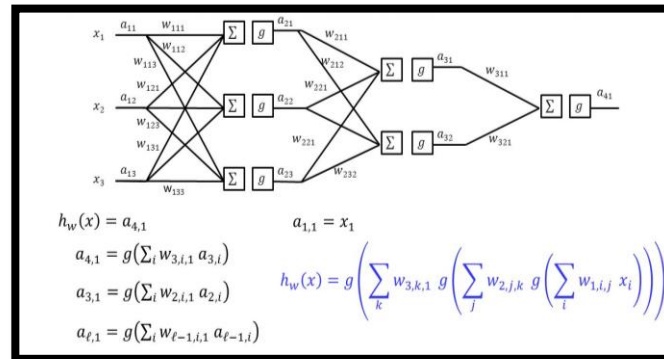
# AI = MACHINE LEARNING MODELS + FEEDBACK LOOP

Training data

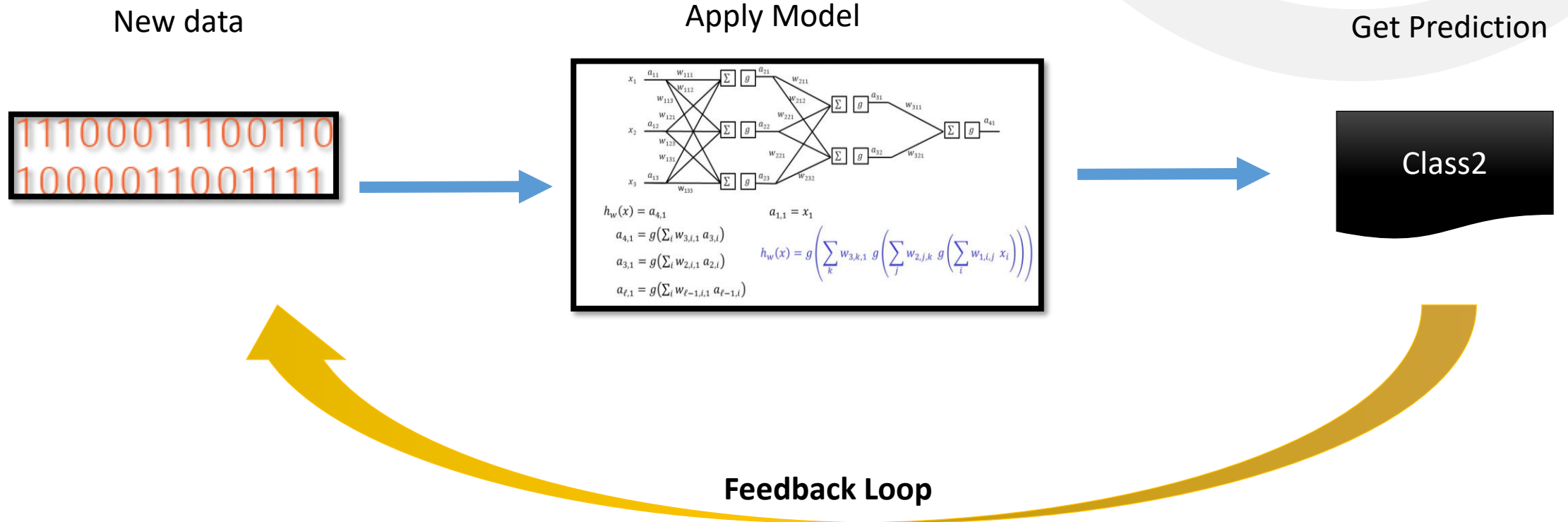


10011000101001  
11100011100110  
1000011001111  
10011000100111  
11101111001100  
10100110010011

Model

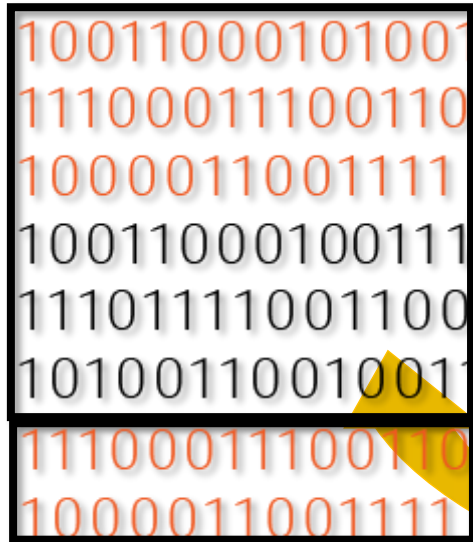


# AI = MACHINE LEARNING MODELS + FEEDBACK LOOP

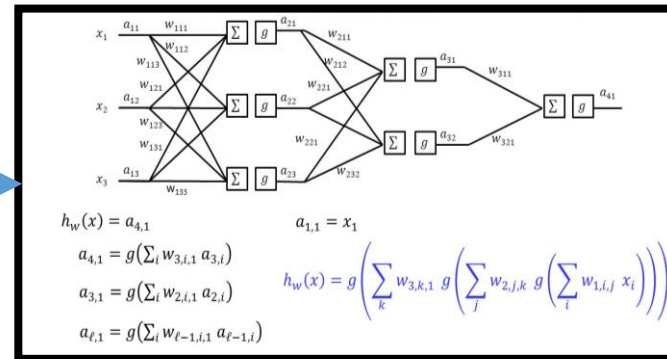


# AI = MACHINE LEARNING MODELS + FEEDBACK LOOP

Update Training  
data based on  
feedback



Update the Model  
based on data



Prediction

Class2

Feedback Loop

# HOW IS FEEDBACK COLLECTED

- Manual entry after going through test cases – Google maps
- Indirect feedback collection based on user actions for - User click vs not click on your YouTube ad
- Indirect feedback collection based on actions – In case of self driving car, hitting a wall is an action.



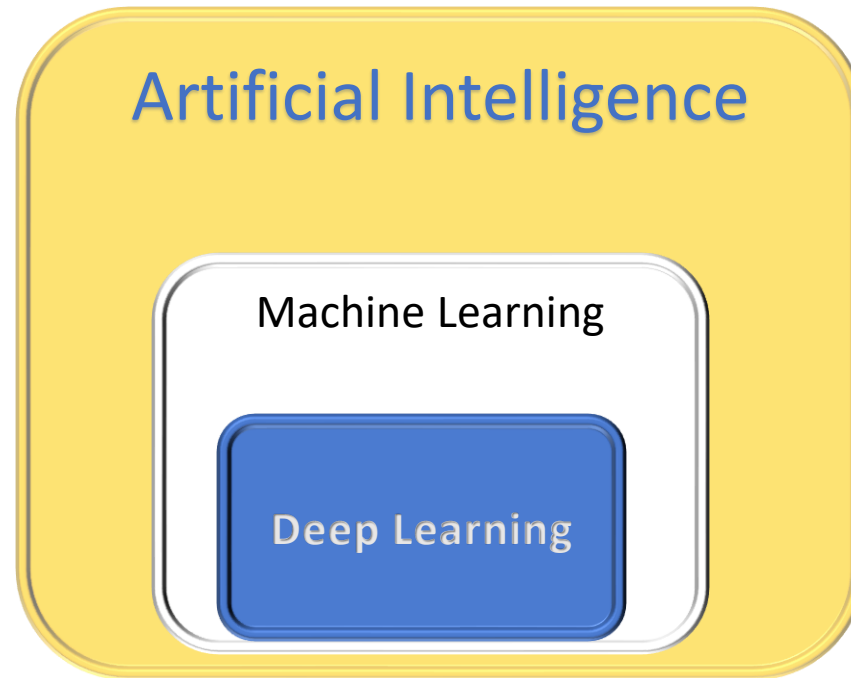
# APPLICATIONS OF AI

- Self driving cars
- SIRI / Ok-google
- Alexa /Google home
- Recommendation systems
- Image recognition
- Speech recognition
- Spam filtering





# MACHINE LEARNING IS A SUBSET OF ARTIFICIAL INTELLIGENCE



# PART-4

## What is Data Science?



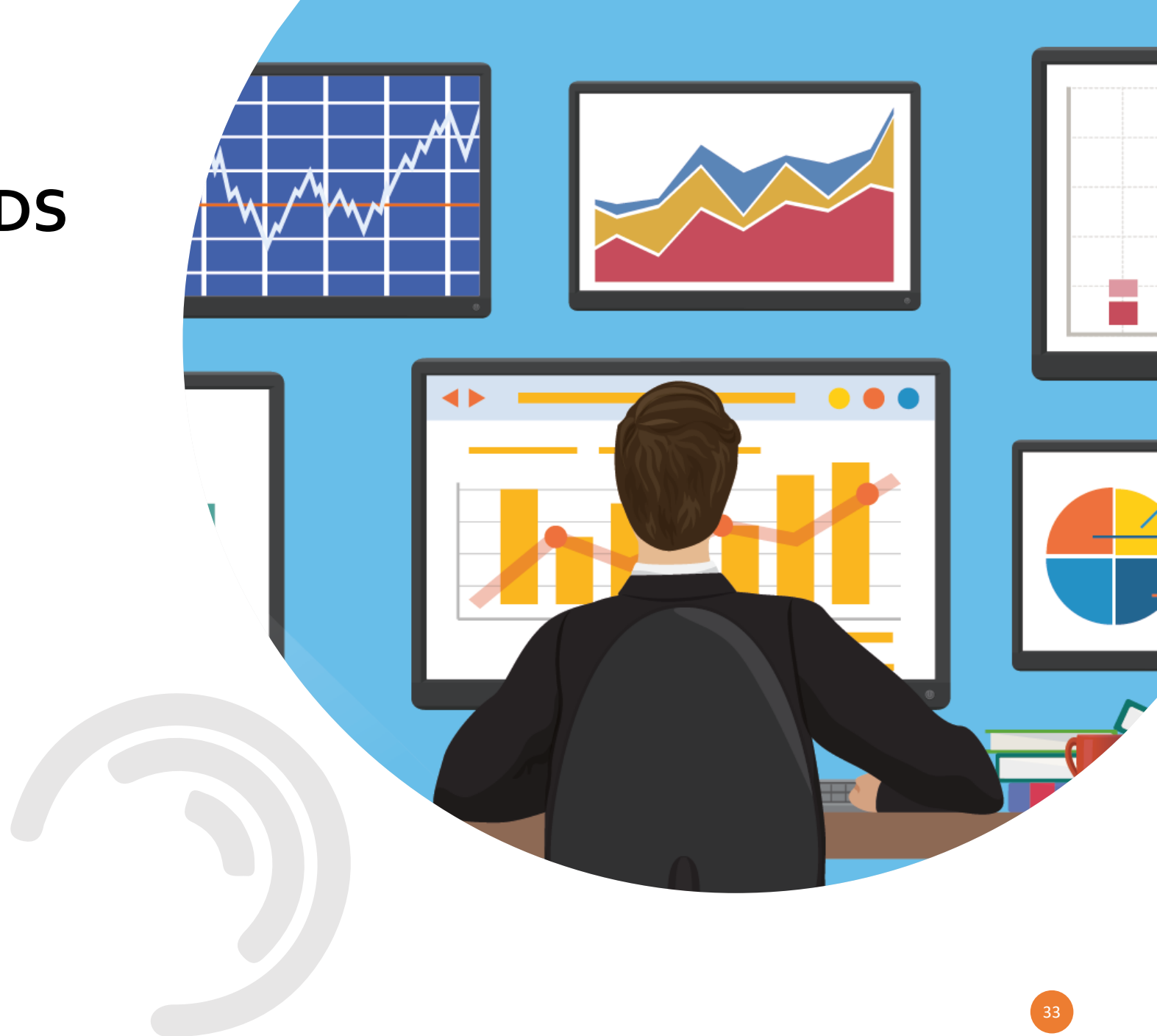
# WHAT IS DATA SCIENCE?

- Data Driven Decision making
- Making sense out of data
- Finding hidden patterns in the data
- Analysis using not just machine learning models but also using data visualizations, intelligent reports
- Most of the techniques and tools seen in data analysis in early days are now falling under data science

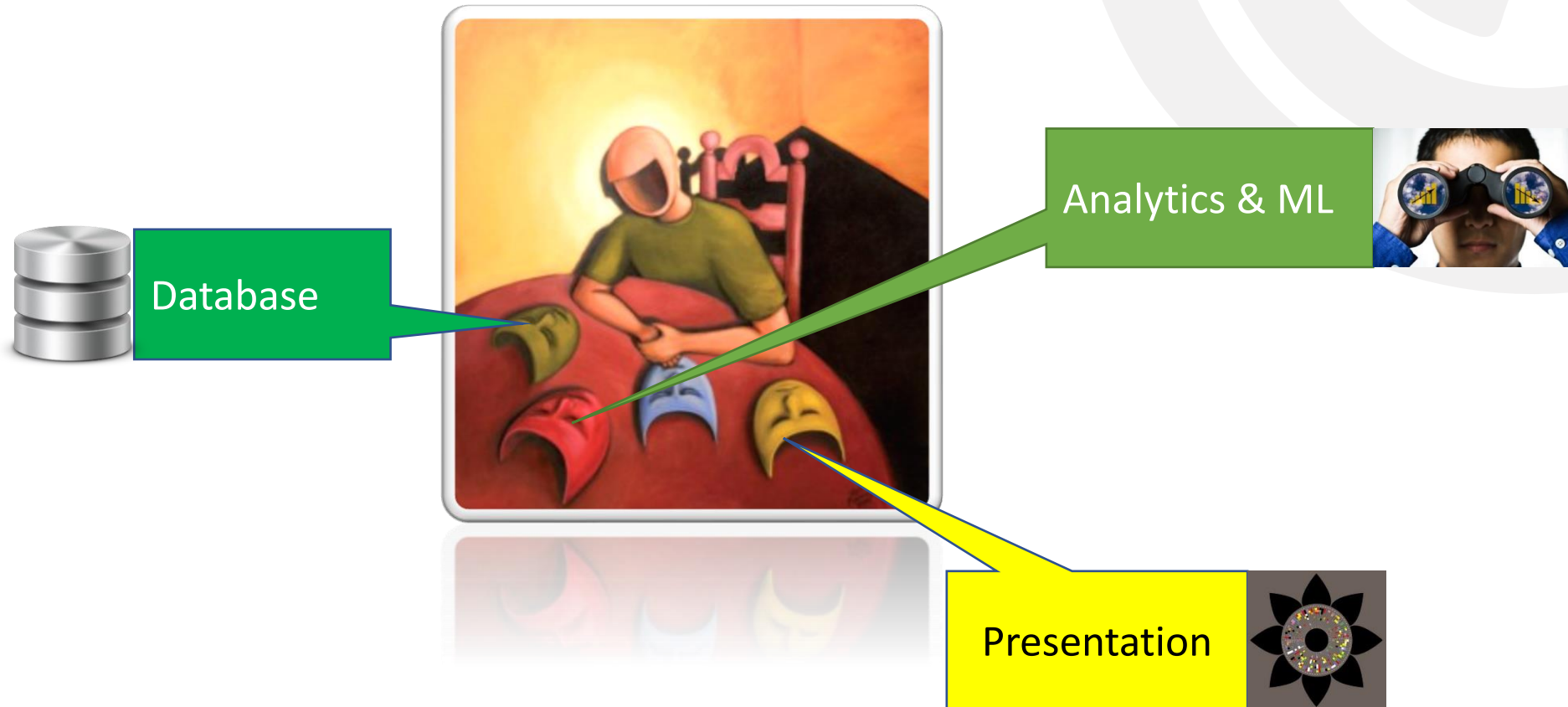


# DATA SCIENCE IS A FUSION OF MANY FIELDS

- Mathematics
- Statistics
- Coding
- Database management
- Data Analytics
- Predictive modelling
- Machine Learning
- Deep Learning



# DATA SCIENCE – THREE MAJOR TYPE OF SKILLS



# THE TECHNIQUES YOU NEED TO KNOW

## Database Knowledge

- Data base Management
- Data blending
- Querying
- Data manipulations
- ETL

## Analytics and Machine Learning

- Basic descriptive statistics
- Advanced analytics
- Predictive modeling
- Machine Learning
- Deep Learning

## Presentation Skill

- Data visualizations
- Report design
- Insights presentation



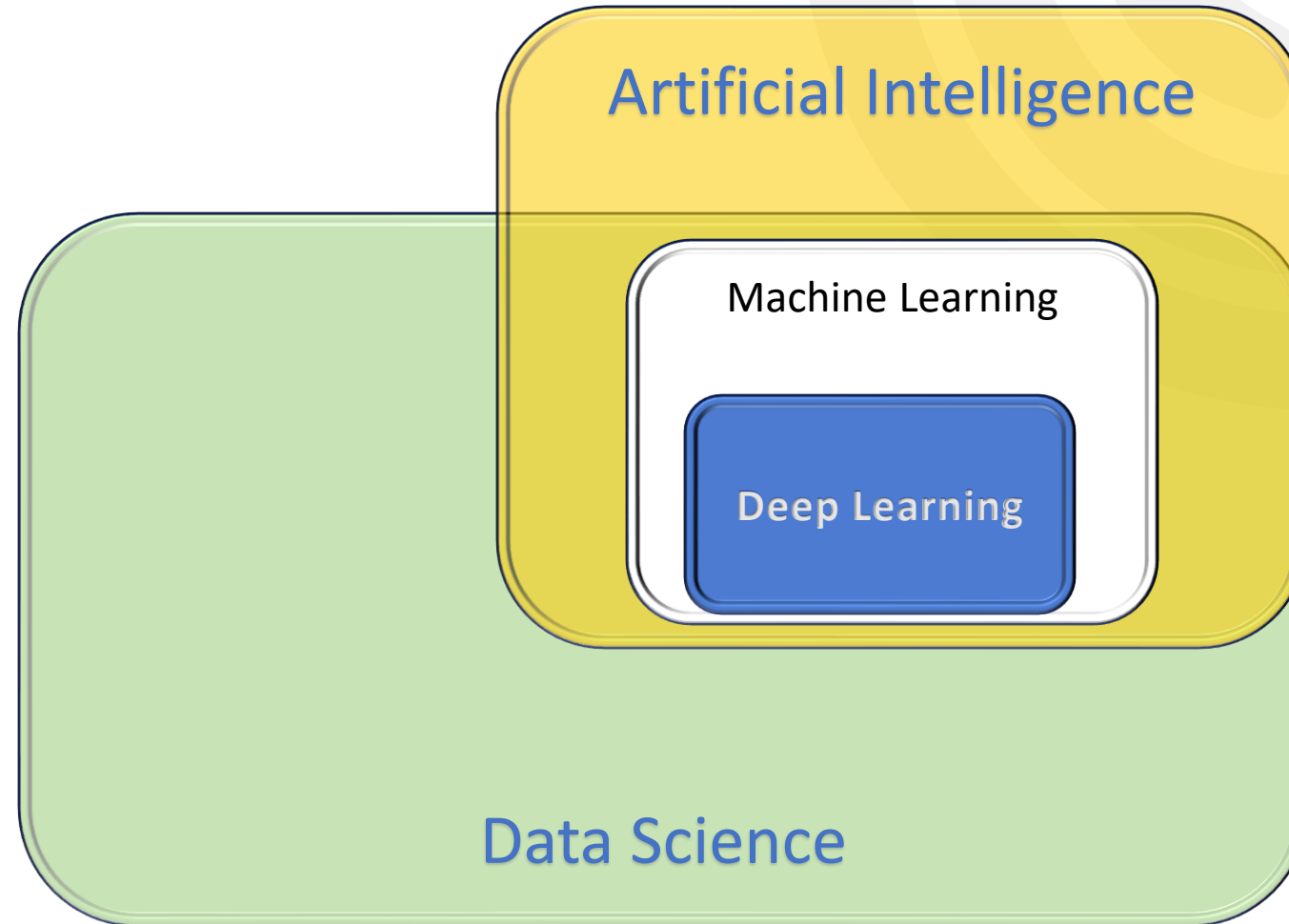
# DATA SCIENCE TOOLS AND SOFTWARE'S

Database tools  
SQL/MySQL  
Teradata  
DB2/Oracle/  
Informix/Exadata

Analytical tools  
SAS/R/SPSS/Python  
Weka/MATLAB

Presentation Tools  
Excel  
Tableau, Qlikview

# MACHINE LEARNING IS A PART OF DATA SCIENCE



\* These are individual interpretations

# PART-5

## The Learning Path



# FAQ BY DATA SCIENCE ASPIRANTS

- I want to be data scientist what training should I take?
- I already have knowledge on few tools, what are my next steps?
- What skill should I add to my profile to make it to next level?
- I am new to data science, where can I start ?



# CATEGORIES OF PROFILES

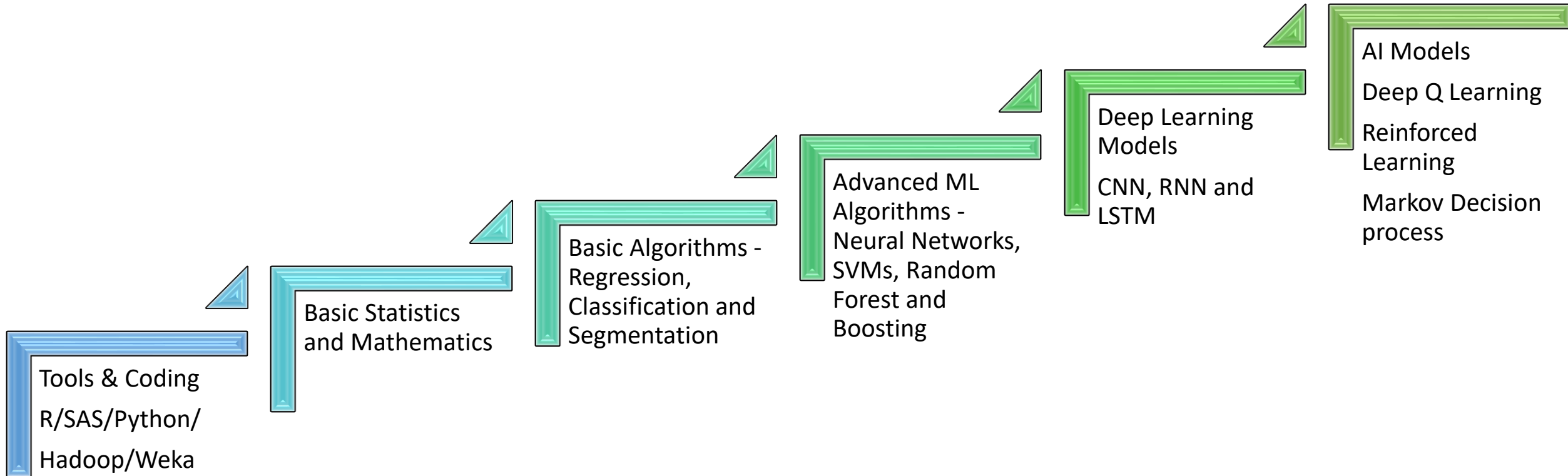
You need training based on your skill level.  
Based on skill set we can divide the whole data science aspirants into four categories

1. Beginner - Completely new to Data Science and ML
2. Intermediate - MIS and Reporting Analyst
3. Advanced – Data Analyst and Predictive Modeler
4. Complete Data Scientist – ML, Hadoop, R, Python, DL, AI





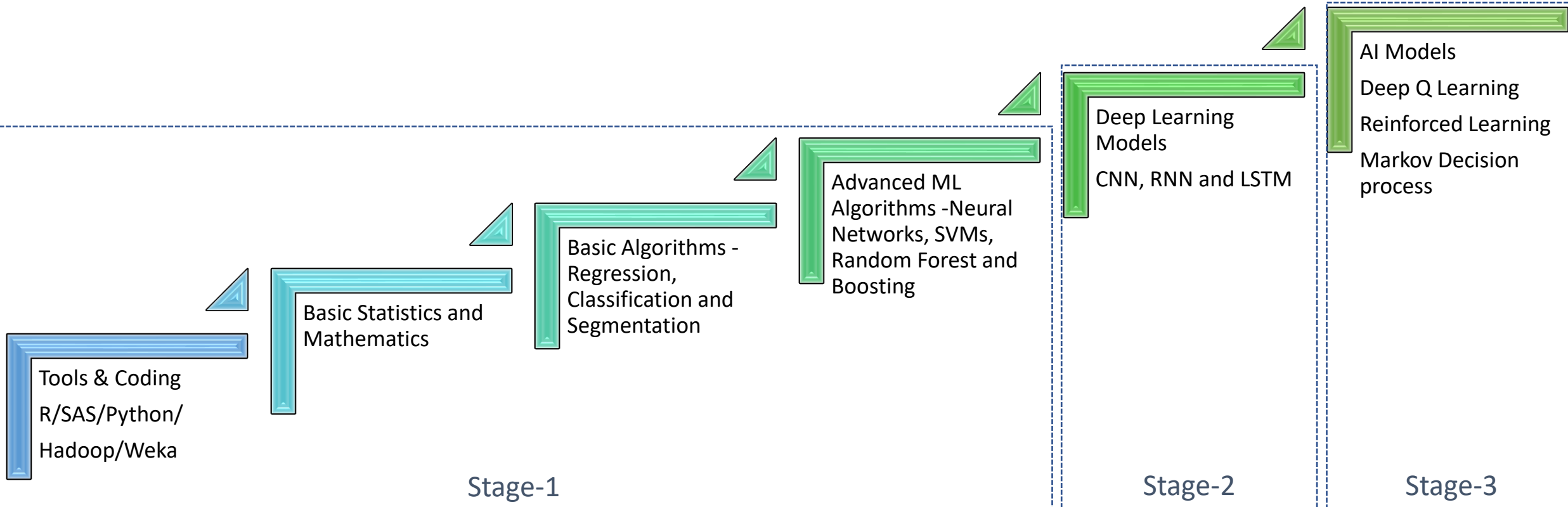
# THE LEARNING PATH





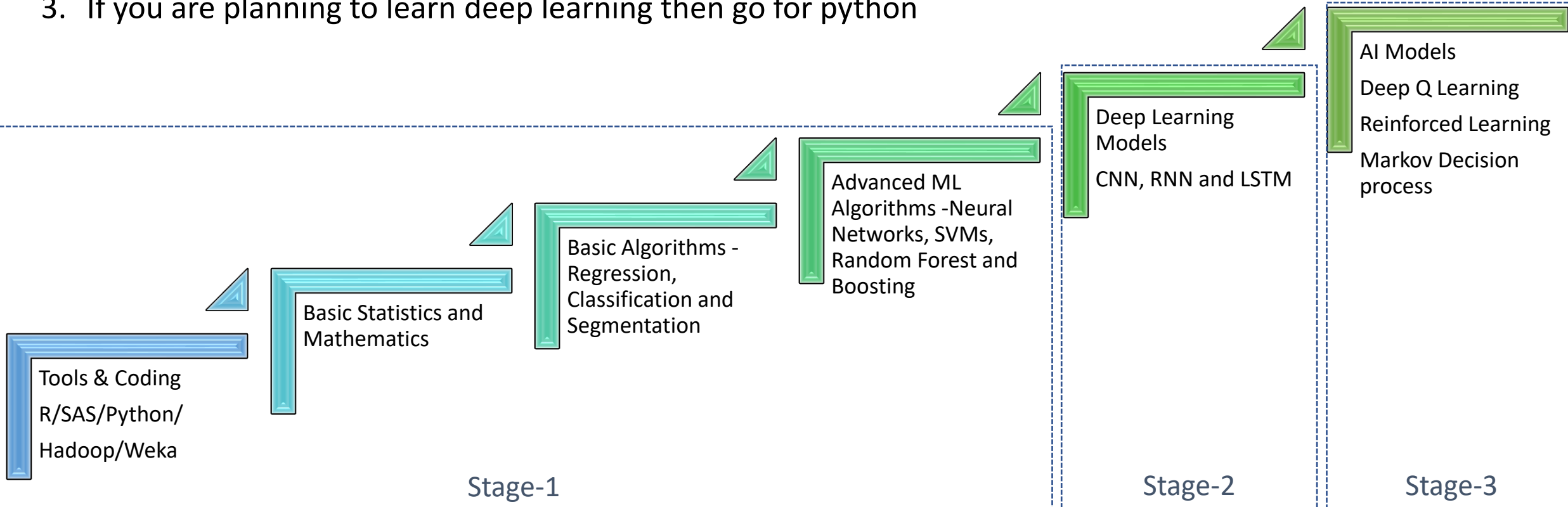
# THE LEARNING PATH – SUGGESTIONS

1. Do not try to learn all the steps in one sitting.
2. You need to learn, absorb and then practise before you reach the next step



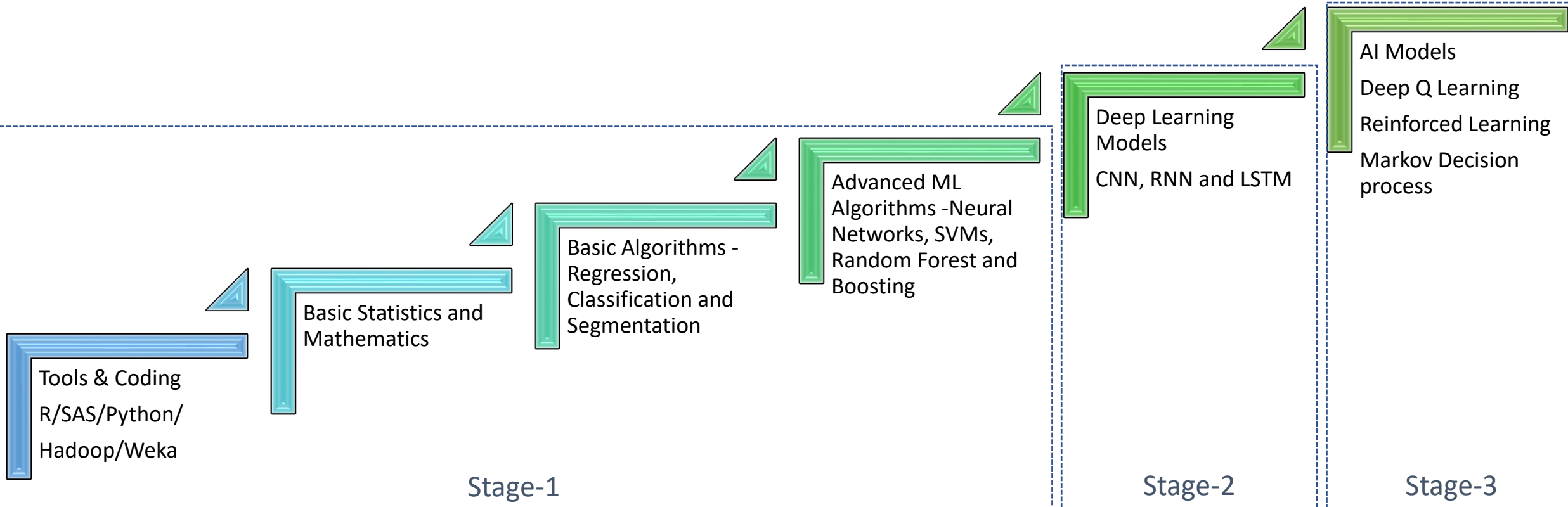
# THE LEARNING PATH – SUGGESTIONS

1. R or Python. Both are really good. Pick any one of them
2. It also depends on your business problem
3. If you are planning to learn deep learning then go for python



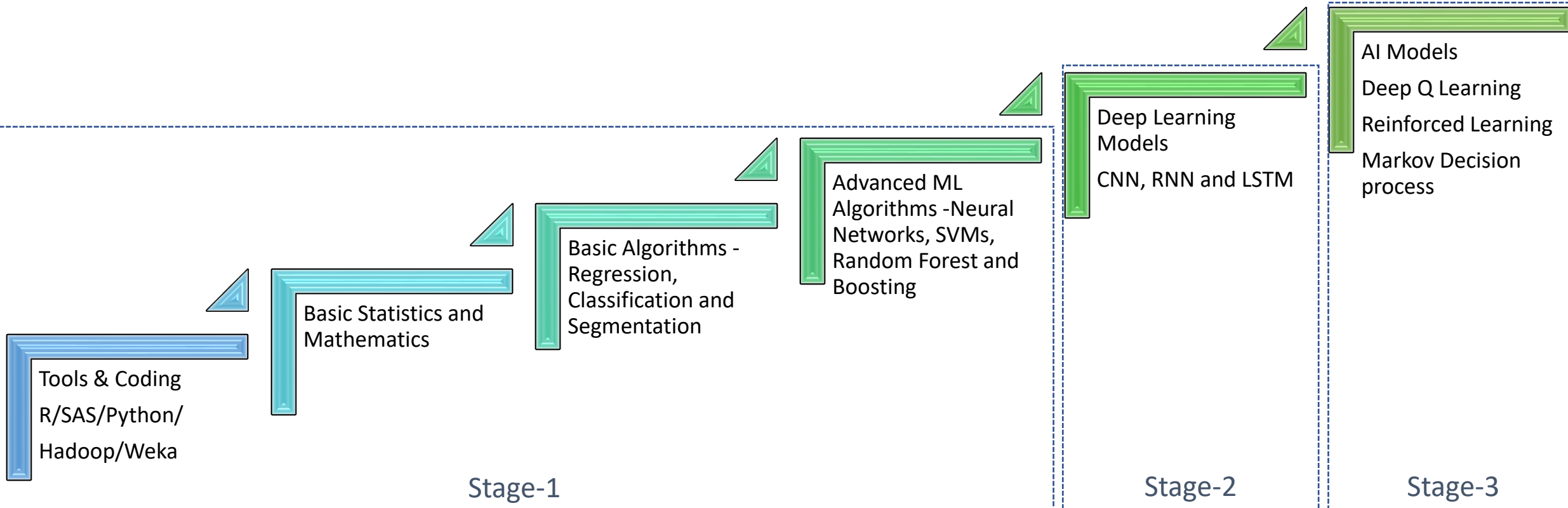
# THE LEARNING PATH – SUGGESTIONS

1. Do not start with stage-2 or stage-3 directly.
2. Strong fundamentals will make the learning easy in later stages.



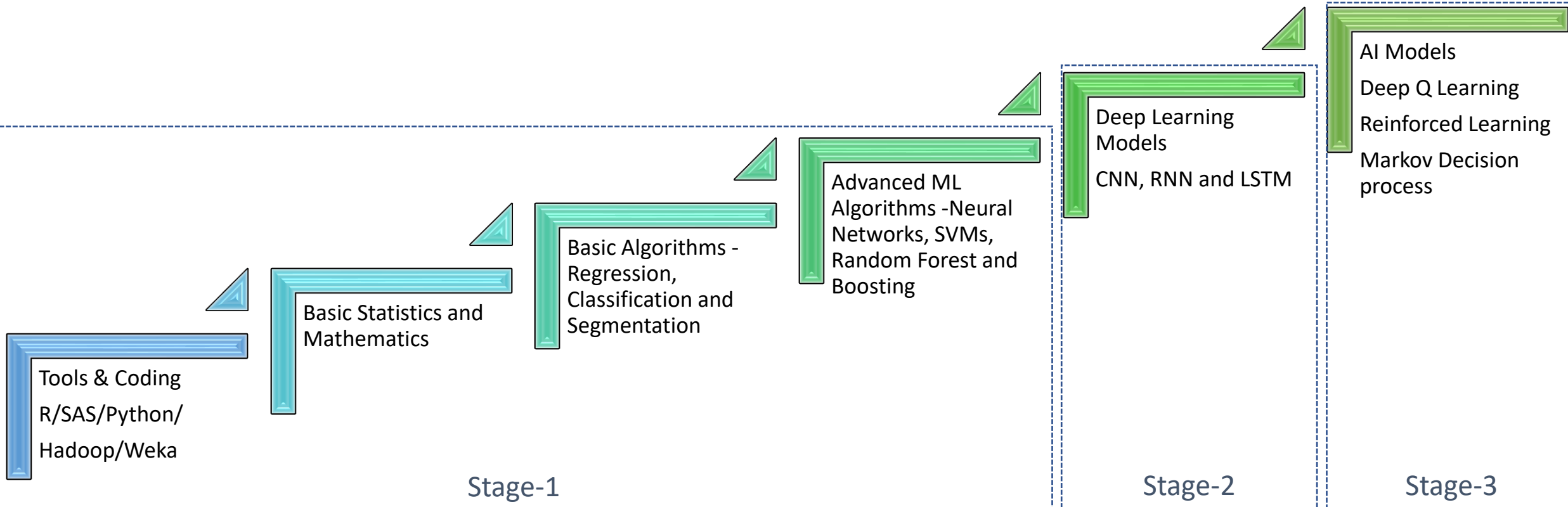
# THE LEARNING PATH – SUGGESTIONS

1. While learning these concepts, try to avoid academic style courses.
2. Look for the courses with lot of hands-on exercises and case studies



# THE LEARNING PATH – SUGGESTIONS

1. Do not focus on the tool, focus on the technique and algorithm
2. Learning python or R tool, will not make you a data scientist



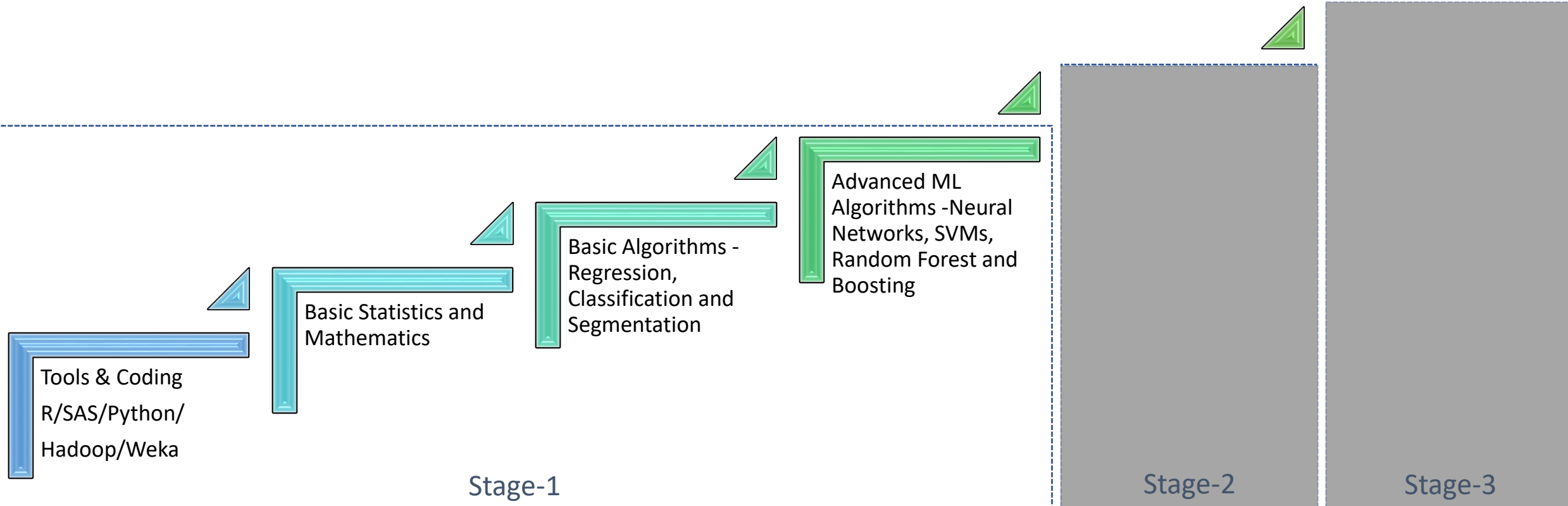
# PART-6

## Course Curriculum

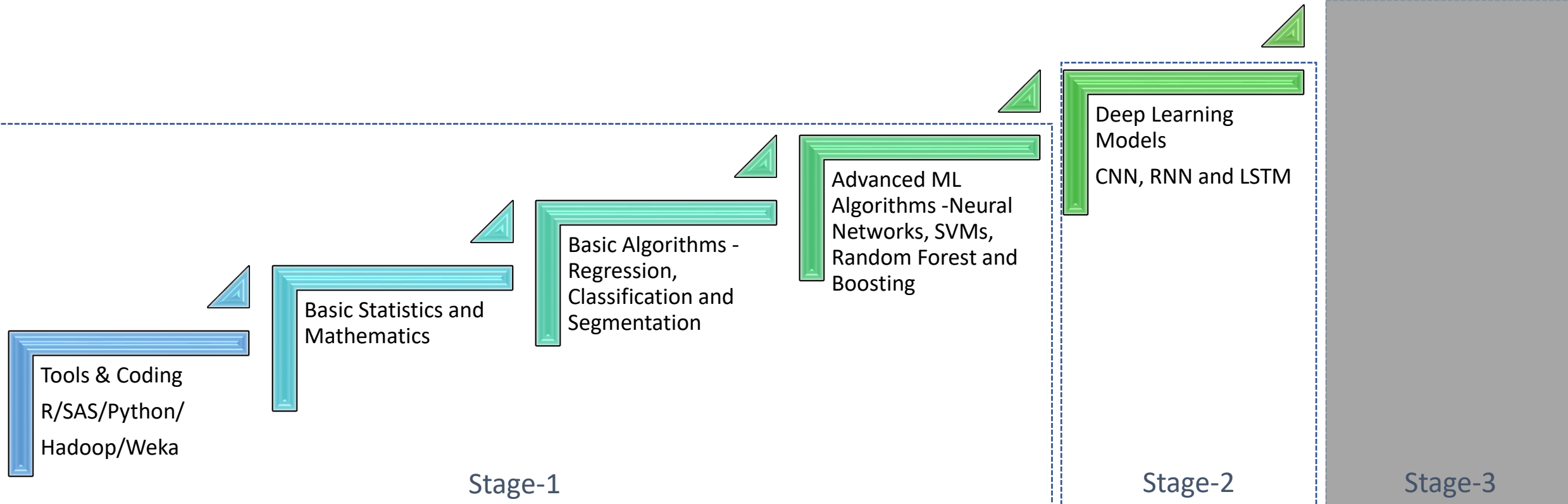




# FOCUS ON FIRST STAGE - ML



# FOCUS ON FIRST TWO STAGES - DL



# PART-7

## Data Science and Machine Learning Myths



# MYTH-1 : MATHEMATICS

Myth-1 : To be a good data scientist, you need to be exceptional at statistics, mathematics, calculus, algorithms etc.,

**Not necessarily.**

## MYTH-2 : PROGRAMMING

Myth-2 : To be a good data scientist, you need to have exceptional coding skills like Python, Java, C++ etc.,

**Not necessarily.**

## MYTH-3 : COMPLICATED MODELS

Myth-3 :Data science is all about building complex predictive and machine learning models to solving business problems

**Not necessarily.**



## MYTH-4 : MODEL BUILDING

Myth-4 :After collection of the data, most of the time is spent on model building process.

**Not necessarily.**

## MYTH-5 : LARGE DATASETS

Myth-5: While solving machine learning problems we need to handle really large datasets or most of the datasets are really large

**Not necessarily.**

## MYTH-6 : MACHINE LEARNING IN BUSINESS

Myth-6: Companies use really advanced deep learning and AI models for while building all their business strategies

**Not necessarily.**

## MYTH-7 : DIVERSE ALGORITHMS

Myth-7: A data scientist will be using all the models in their day to day life

**Not necessarily.**



# THANK YOU

---



VENKAT@STATINFER.COM



[WWW.STATINFER.COM](http://WWW.STATINFER.COM)