

CAREER IN DATA SCIENCE

Yogesh Haribhau Kulkarni



Outline

① DATA SCEINCE

② CAREER IN DATA SCIENCE

③ REFERENCES

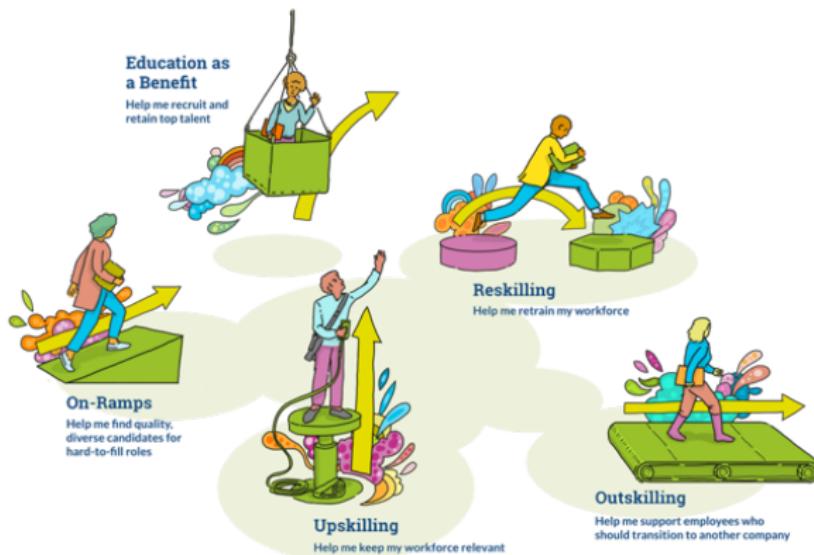
④ REFERENCES

Career in AI



Current State

- ▶ 44% of US workforce < \$18K/yr (< poverty line), works 80-100 hrs/week
- ▶ Automation CAGR 7% (as per BCG), to reach \$114B by 2025

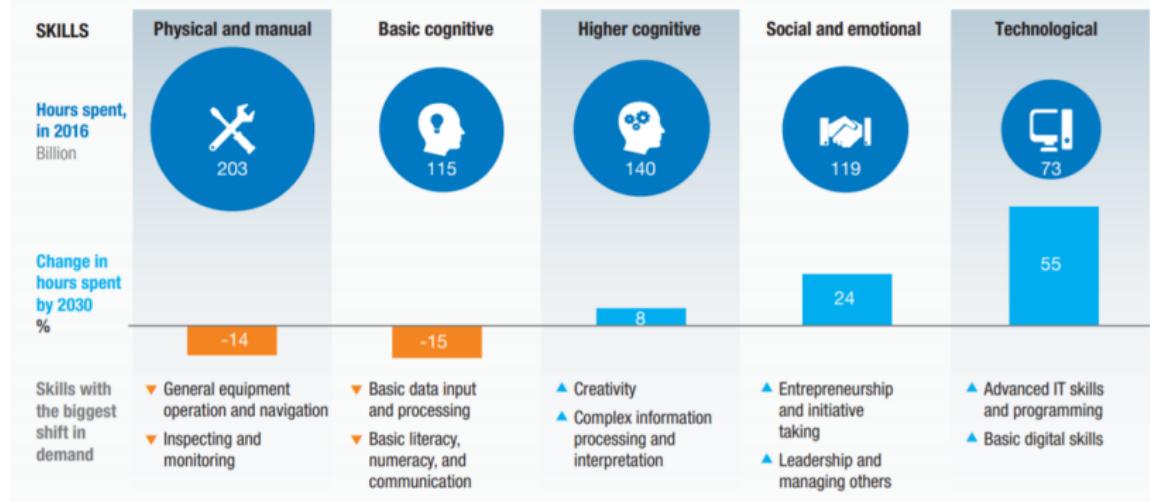


(Ref: As Pressure To Upskill Grows, 5 Models Emerge – Forbes.com)

McKinsey Global Institute Report – Discussion Paper 2018

Automation and AI will change the skills needed in the workforce

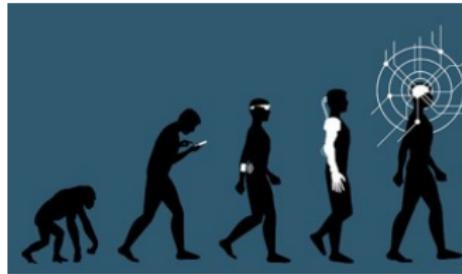
Total is for United States and 14 Western European countries



Examples

Less mechanical, automatable

- ▶ Bill and account collectors
- ▶ Data entry operators
- ▶ Computer network support
- ▶ Secretaries and administrative assistants
- ▶ Insurance sales agents
- ▶ Office clerks



More Cognitive, Creative, Human

- ▶ Software developers
- ▶ Customer service representatives
- ▶ General and operations managers
- ▶ Human resources managers
- ▶ Personal finance advisors
- ▶ Psychologists
- ▶ Artists
- ▶ Sportsman
- ▶ Researchers
- ▶ Nurses, care

(Source: The Simplistic Debate Over Artificial Intelligence – Preston
Estep)

Changes

Technological

- ▶ AI deluge
- ▶ Digitization → Data + APIs
- ▶ Remote *

Social

- ▶ Over interaction + isolation
- ▶ Obsolete roles, emergence of new
- ▶ Lifelong re-skilling



(Source: Rise of the Chatbots: How AI Changed Customer Service –
Salesforce.com)

Financial

- ▶ Widening gap
- ▶ Flatter world
- ▶ Gigs over jobs

Examples

Technological

- ▶ AI deluge
- ▶ Digitization → Data + APIs
- ▶ Remote *

Social

- ▶ Over interaction + isolation
- ▶ Obsolete roles, emergence of new
- ▶ Lifelong re-skilling



(Source: AI in healthcare – foreseemed.com)

Data Science is critical in bringing intelligent automation



What are Data Sciences?

ie What is Artificial Intelligence? Machine Learning? Deep Learning?



Data Science

- ▶ Science of Data (obviously)
- ▶ Use of Data for Applications
- ▶ Some parts of AI uses Data to find patterns and insights which are helpful in multiple applications
- ▶ Machine and Deep Learning that part of AI that leverages data.

So, more on AI-ML here ...



"Houston, we have a problem!!"



The Problem

- ▶ Every company is claiming to be working in AI-ML
 - ▶ Is it so?
 - ▶ What exactly is AI (ML)?
 - ▶ What is not AI?
- ▶ Or is it just a plain BIG hype?

What's the core idea?

- ▶ behind problem solving?
- ▶ behind writing software algorithms?
- ▶ solving research problems?

Desire

- ▶ To find a “function”
- ▶ To find a relation
- ▶ To find a transformation
- ▶ To build a model
- ▶ From given inputs to desired outputs.

That's it.



Functions

- ▶ Some functions are straight forward
- ▶ *"In summer, ice-cream sale goes up"*
- ▶ Cause and effect
- ▶ Relation (function, Mathematical model) is found out
- ▶ Here, simple rule based programming suffices

Functions

- ▶ But some functions are complex
- ▶ *"More you put efforts, your business flourishes."*
- ▶ Cause and effect again, but the relation is far to complex
- ▶ Too many variables
- ▶ Here, simple rule based programming not humanly possible.
- ▶ Lots of research needed to come up with equations.

Functions

- ▶ $E = mc^2$
- ▶ What's this? a function?
- ▶ Input variable(s)?
- ▶ Output variable(s)?
- ▶ Parameters?
- ▶ How's the relation? linear?

Functions

- ▶ But most real-life functions are not deterministic
- ▶ Some are probabilistic, some non-linear.
- ▶ *“Detecting if the tumor is benign or malignant”*
- ▶ *“At any state in the game of chess, what's the next move?”*

Chess: next move?

- ▶ Needs extreme expertise
- ▶ Needs “intelligence”
- ▶ How do you get that?
 - ▶ Built by lots of training.
 - ▶ By studying lots of past games.
- ▶ This is how Humans build intelligence

Intelligence

- ▶ Can machine (software/program) also do the same?
- ▶ Can it play chess?
- ▶ Can it build intelligence?
- ▶ By looking at past experiences (data),
- ▶ Training Data: games played, moves used, etc.

Yes, it can!! Thats Artificial Intelligence.

What is AI?



What is Artificial Intelligence (AI)?

My definition:

“If machines (or computer programs) start doing some/all of these “intelligent” tasks, then that’s Artificial Intelligence”



Intelligence: the differentiation

- ▶ Ability to think various domains
- ▶ Ability produce something new
- ▶ Ability to detect the unseen
- ▶ Ability to enhance knowledge (rules, patterns)

All these, AI has started doing. The AI era has arrived!!

Everyday usage

Artificial intelligence seems to have become ubiquitous.

- ▶ Replying to our emails on Gmail
- ▶ Learning how to drive our cars,
- ▶ Sorting our holiday photos.
- ▶ etc.

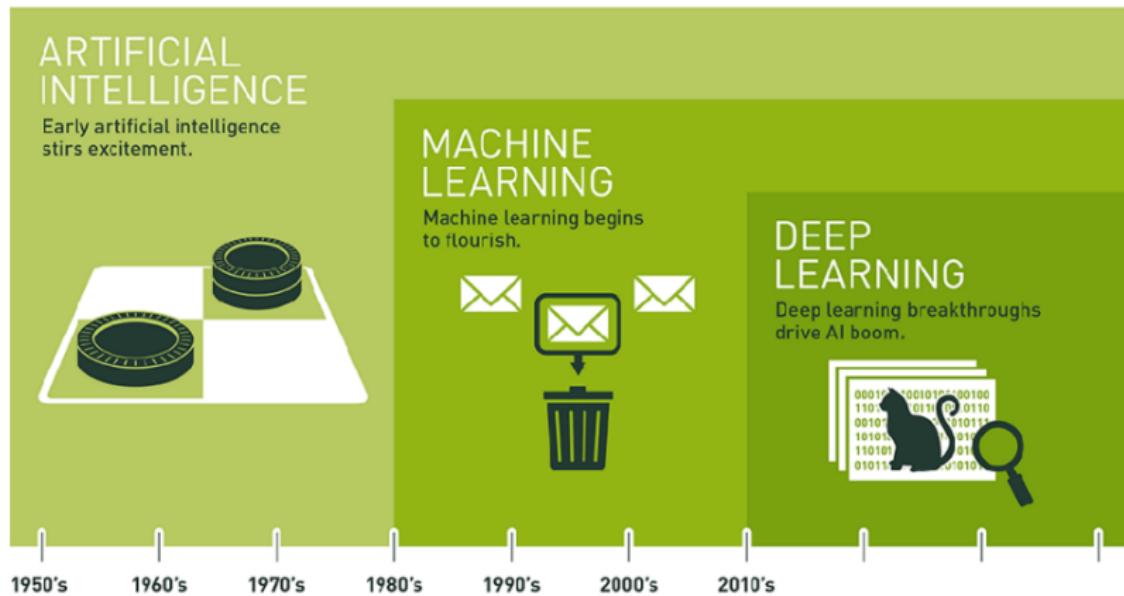
Too good to be true, isn't it, sort of Magical !!

But then ...

- ▶ When its too good, you start suspecting
- ▶ Is it for real!!
- ▶ How can such thing happen?
- ▶ How far will it go?

The next thing you know, people are worrying about exactly how and when AI is going to doom humanity.

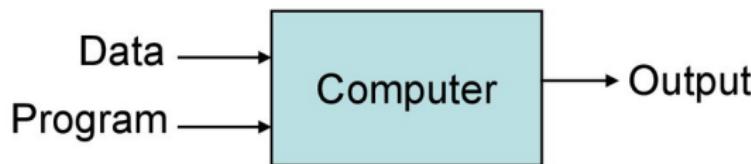
Relationship between AI, ML, DL



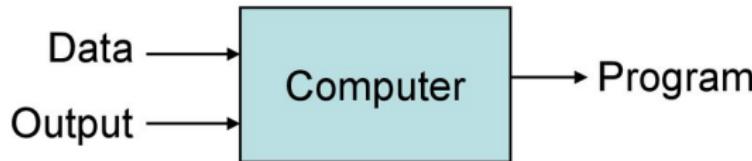
(Ref: <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>)

Traditional vs. Machine Learning?

Traditional Programming



Machine Learning



Why Machine Learning?

- ▶ Problems with High Dimensionality
- ▶ Hard/Expensive to program manually
- ▶ Techniques to model 'ANY' function given 'ENOUGH' data.
- ▶ Job \$\$\$



Why now?

- ▶ Flood of data (Internet, IoT)
- ▶ Increasing computational power
- ▶ Easy/free availability of algorithms
- ▶ Increasing support from industries

Preparing Career in Data Sciences



So, What is a Data Scientist?

- ▶ Software + Mathematics + ML/DL specific techniques
- ▶ Discovers patterns and trends in datasets to get insights.
- ▶ Creates forecasting algorithms like Classification, Clustering, etc

Business Intelligence vs Data Science

Business Intelligence	Data Science
Extract insights using past data	Predict future using past data
Uses structured data	Uses both structured and unstructured data
Use of basic statistics with emphasis on visualization (dashboards, reports)	Leverages more sophisticated statistical and predictive analysis and machine learning (ML)

Roles in Data Science

- ▶ Researcher: Advanced Mathematics, Programming, domain, Adv ML/DL, etc. You invent new things and write NeurIPS paper.
- ▶ Implementation:
 - ▶ Data Analyst: Bridge the gap between the data scientists and the business analysts, organizing and analyzing data to answer the questions the organization poses.
 - ▶ Data Engineer: Focus on developing, deploying, managing, and optimizing the organization's data infrastructure and data pipelines.
 - ▶ Data Scientist: Use/mine data, clean, build applications, reports, etc
- ▶ User level: Product Manager, Sales, Manager, etc

But . . . , But . . . , How to prepare?



What got you here, won't get you there!!
- Marshal Goldsmith



So, Well, you can't prepare!!
not everything, but certainly, specifically ...



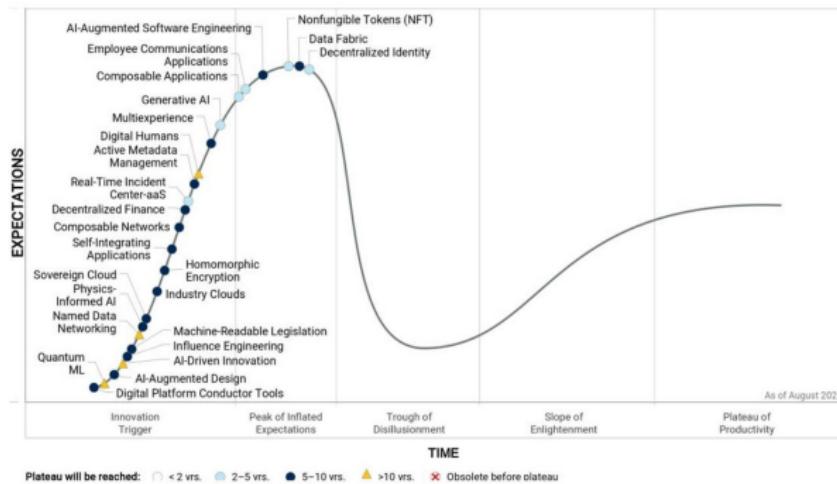
Gartner Hype Cycle Emerging Technologies 2020

Hype Cycle for Emerging Technologies, 2020



Gartner Hype Cycle Emerging Technologies 2021

Hype Cycle for Emerging Technologies, 2021



Source: Gartner (August 2021)

747576

Gartner Hype Cycle Emerging Technologies 2022

Hype Cycle for Emerging Tech, 2022



gartner.com

Source: Gartner

Career in Data Science

YHK

Gartner

40 / 54

Career in Data Science ...



Pros and Cons

Advantages:

- ▶ Domain Expertise
- ▶ Maturity, Communication, Soft Skills
- ▶ Problem Solving

Dis-advantages:

- ▶ Lost touch with Mathematics
- ▶ Un-Learning and Re-Learning inertia
- ▶ Starting from scratch? Seniority?

Why do you want to Switch?

- ▶ $\$_1\$_2\$_3 \dots \$_n$?
- ▶ Will remain in fashion forever?
- ▶ Hate my current job? no growth?

What's in it for me?

Current + ML combo?

- ▶ First : DON'T QUIT!!!
- ▶ Don't lose advantage due to domain expertise
- ▶ ML just another problem solving technique, IF DATA IS AVAILABLE
- ▶ Can you leverage domain expertise and apply ML there, a good/smooth transition?



Learning Path, Roadmap



Resources

- ▶ First : try Free Online resources, see how much you grasp
- ▶ No expensive (read, fees in lakhs) certification courses, to start with
- ▶ Test waters, gain some understanding of yourself then decide.

Steps

Prep:

- ▶ Mathematics: Statistics, Calculus, Linear Algebra
- ▶ Programming: Python, Data Structure & Algorithms, Tools
- ▶ ML/DL: algorithms & frameworks

Practice: Kaggle, Hackathons, projects on Github, blogs, Meetups-talks, etc.

Analytics Vidhya Learning Path 2017

- ▶ An year long schedule
 - ▶ Mostly free resources
 - ▶ Followed it myself
 - ▶ Separate paths for:
 - ▶ Beginner: Not much experience in programming but just college maths
 - ▶ Transitioner: Decent experience programming, but no ML and just college maths
 - ▶ Intermediate: Knows ML, comfortable with programming and maths.



<https://www.analyticsvidhya.com/blog/2017/01/the-most-comprehensive-data-science-learning-plan-for-2017/>

More Generally . . . For Career . . .



Ikigai



(Source: How To Find Your Ikigai And Transform Your Outlook On Life And Business - Chris Myers)

Generic Gyan

Specific Knowledge

- ▶ Unique, rare combination
- ▶ Un-trainable, un-scalable
- ▶ Acquired through apprenticeship

Leverage

- ▶ Permission-ed: Capital, Labor
- ▶ Un-permission-ed: Content, code
- ▶ Marginal cost of duplication

Wealth is a positive sum game

Foreword by
TIM FERRISS

THE ALMANACK OF NAVAL RAVIKANT

A guide to wealth and happiness
ERIC JORGENSEN

Own story

- ▶ Current jobs were not available 10 yrs back
- ▶ Pick difficult problems/domain, Job, RnD
- ▶ If not now, can change later, Mid career change, harder, but possible.

YHK

References

- ▶ Learning Plan 2017 for beginners in data science - Analytics Vidhya
- ▶ Mid-Career Transitions into ML-AI, with Yogesh Kulkarni - Choose To Think <https://www.youtube.com/watch?v=IQzWosVzkM4>
- ▶ How to become a Data Scientist? - Yogesh Kulkarni
<https://medium.com/technology-hits/how-to-become-a-data-scientist-f673a30cafcd>
- ▶ Finding Your Persona - Yogesh Kulkarni
<https://medium.com/technology-hits/finding-your-persona-ad433fca8c91>
- ▶ What is Data Science? - SimpliLearn
- ▶ Roadmap: How to Learn Machine Learning in 6 Months - Zach Miller, Senior Data Scientist at Metis
- ▶ Tetiana Ivanova: How to become a Data Scientist in 6 months — PyData London 2016
- ▶ Renee Teate — Becoming a Data Scientist Advice From My Podcast Guests
- ▶ How to switch career to data science from non computer science background - Codebasics
- ▶ Step by step roadmap for machine learning engineer - Codebasics
- ▶ And . . . Mid-Career Transitions into ML-AI, with Yogesh Kulkarni - Choose To Think

References

- ▶ Learning Plan 2017 for beginners in data science - Analytics Vidhya
- ▶ Mid-Career Transitions into ML-AI, with Yogesh Kulkarni - Choose To Thinq <https://www.youtube.com/watch?v=IQzWosVzkM4>
- ▶ How to become a Data Scientist? - Yogesh Kulkarni
<https://medium.com/technology-hits/how-to-become-a-data-scientist-f673a30cafcd>
- ▶ Finding Your Persona - Yogesh Kulkarni
<https://medium.com/technology-hits/finding-your-persona-ad433fca8c91>
- ▶ What is Data Science? - SimpliLearn
- ▶ Roadmap: How to Learn Machine Learning in 6 Months - Zach Miller, Senior Data Scientist at Metis
- ▶ Tetiana Ivanova: How to become a Data Scientist in 6 months — PyData London 2016
- ▶ Renee Teate — Becoming a Data Scientist Advice From My Podcast Guests
- ▶ How to switch career to data science from non computer science background - Codebasics
- ▶ Step by step roadmap for machine learning engineer - Codebasics
- ▶ And . . . Mid-Career Transitions into ML-AI, with Yogesh Kulkarni - Choose To Thinq

YHK

Thanks ...

- ▶ Search "**Yogesh Haribhau Kulkarni**" on Google and follow me on LinkedIn and Medium
- ▶ Office Hours: Saturdays, 2 to 5pm (IST); Free-Open to all; email for appointment.
- ▶ Email: yogeshkulkarni at yahoo dot com



(Generated by Hugging Face QR-code-AI-art-generator,
with prompt as "Follow me")

