

Kaggle Vs Real World Projects

Ankit Rathi

(ankitrathi.com)

AGENDA

- **About Me**
- **Context Building**
- **Hackathons**
- **Real-World Projects**
- **Head to Head Comparison**
- **General Differences**
- **Q&A**

About Me

 **ankitrathi.com**



Ankit Rathie

Lead Architect - DS/AI | Author | Speaker | All views personal

Gautam Buddha Nagar, Uttar Pradesh, India · [500+ connections](#) ·

[Contact info](#)



[Add profile section](#)

[More...](#)

 SITA

 HBTI, Kanpur

Articles & activity

30,123 followers

[Manage followers](#)



How to launch your DS/AI career in 12 weeks?

 **Ankit Rathie**
Published on LinkedIn

From a time around when DS/AI field started picking up, every other day I get at least 8-10 messages from DS/AI starters & enthusiasts on 'How can I get into DS/AI field?'. Over th ...see more

70 Likes · 4 Comments



Like



Comment



Share

[See all articles](#)



Crossed 30K followers few thanks for all the encoura

Ankit shared this
62 Reactions · 4 Comments



Great Tirthajyoti, looks interesting, will try it shortly.

Ankit commented



It doesn't detail out the each and every algorithm, but covers the journey of ...

Ankit replied to a comment

[See all activity](#)

Experience



Lead Architect

SITA

Dec 2017 – Present · 1 yr 9 mos



Principal Consultant

Genpact Headstrong Capital Markets

Jun 2016 – Dec 2017 · 1 yr 7 mos

Noida Area, India



Tech Lead

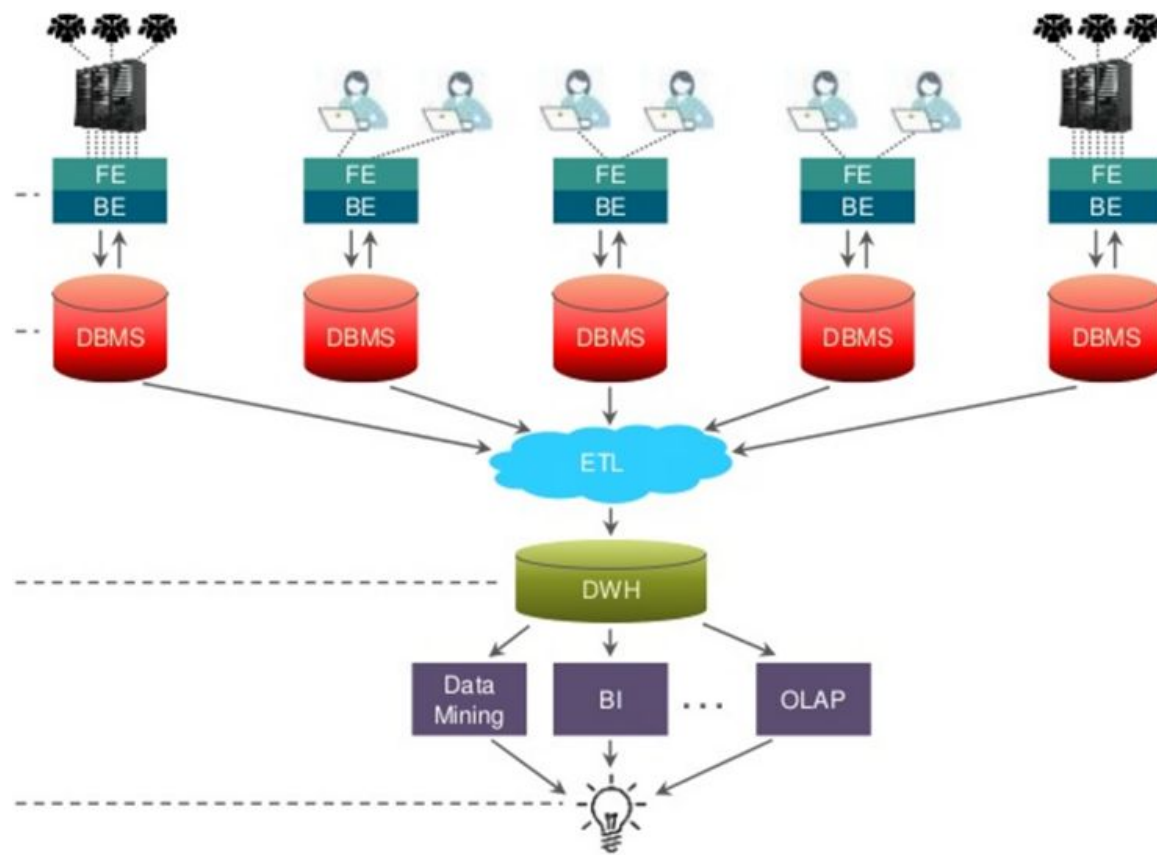
Royal Bank of Scotland

Nov 2010 – Jun 2016 · 5 yrs 8 mos

New Delhi Area, India

- **14+ years in Data & Analytics**
- **Currently working in SITA.aero**
- **Worked on DBs/ETL/DWH/BI**
- **Transitioned into DS/AI in 2014**
- **Infusing intelligence in products**
- **Blogger, Author & Speaker**

Let me start with a story...



Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE



kaggle.com

Home > Data Science > Machine Learning

Machine Learning

Created by:

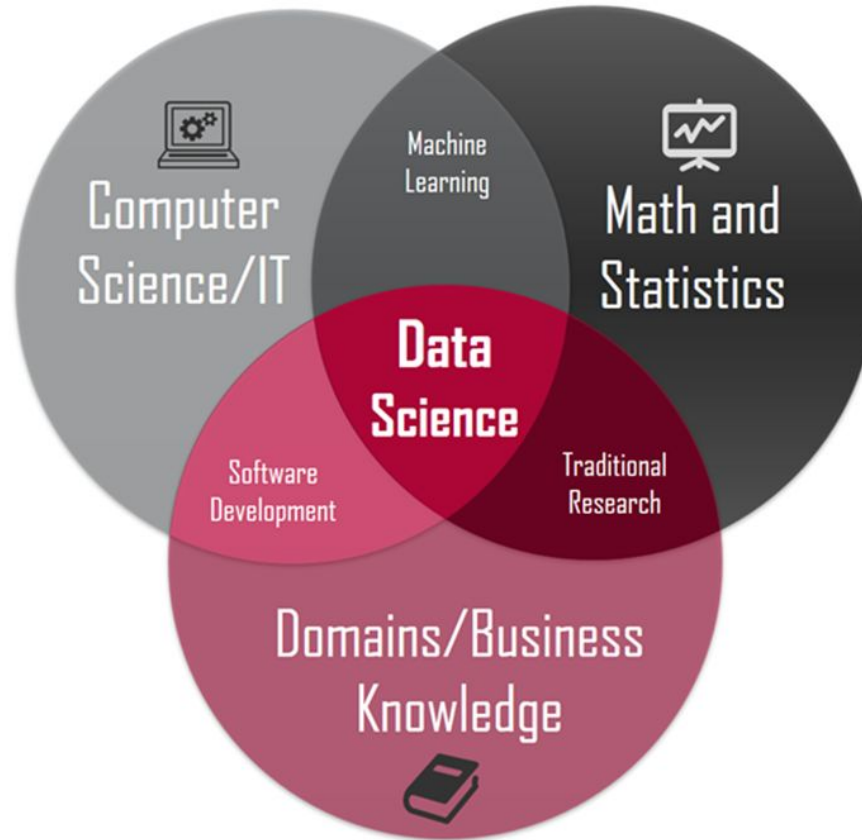


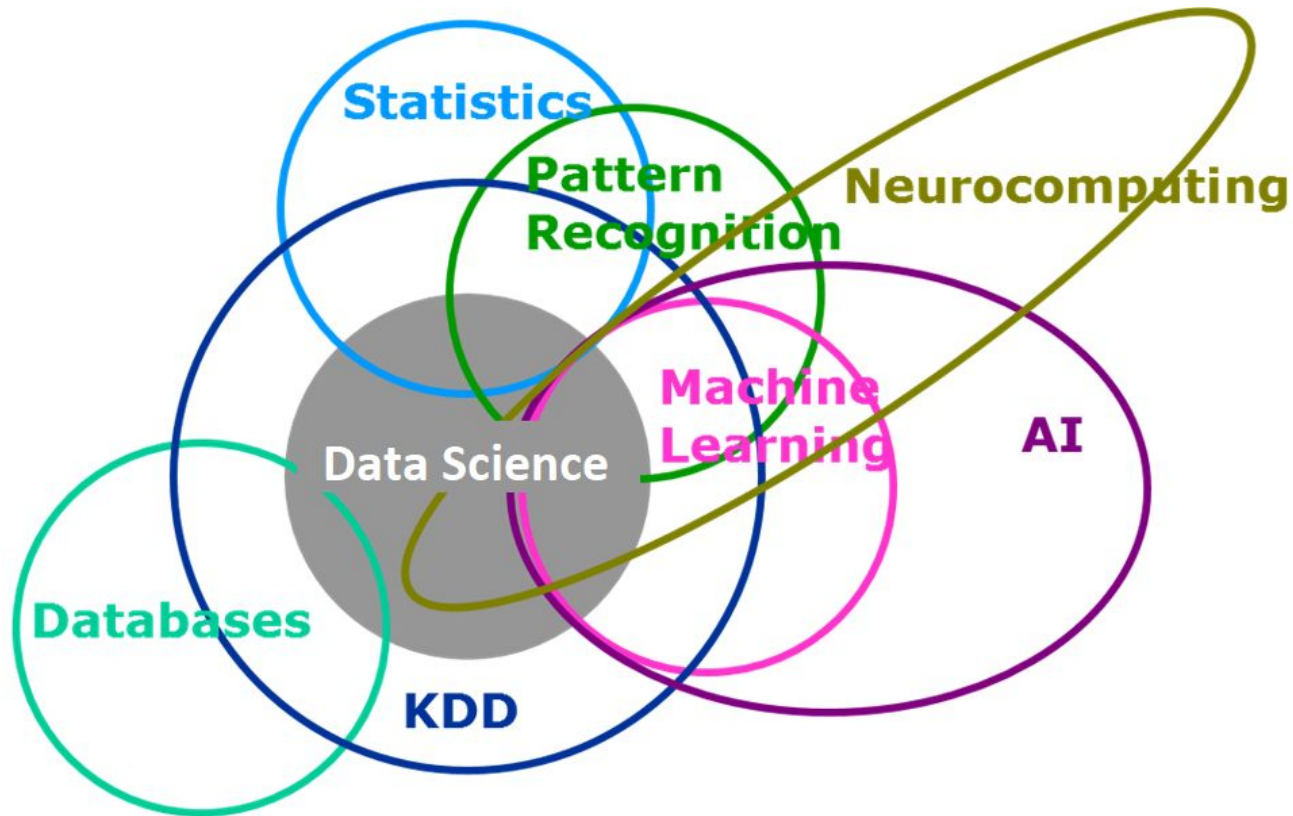
Andrew Ng
American computer scientist



Context Building







kaggle
DAYS

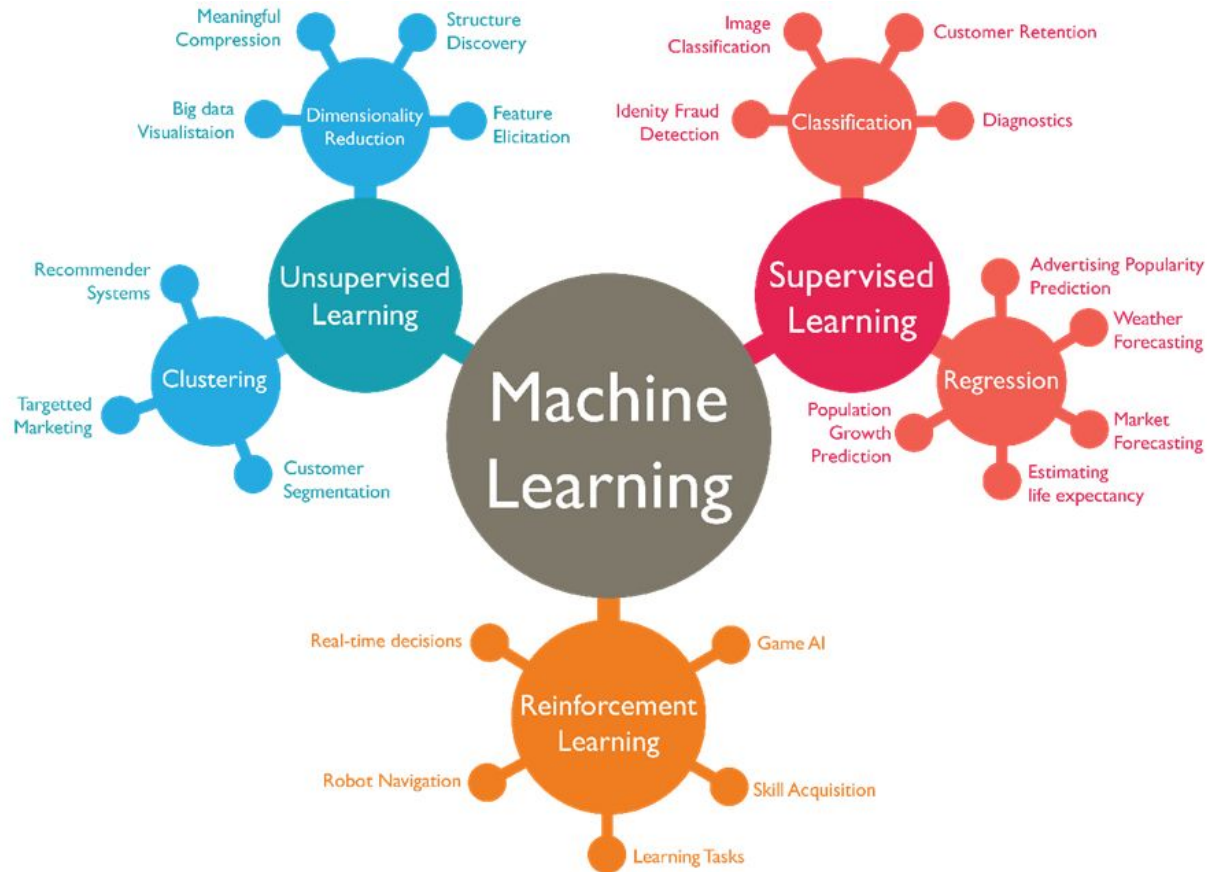
MEETUP

atlan

LOGICW

kaggle

bitgrit



ARTIFICIAL
INTELLIGENCE

Early artificial intelligence
stirs excitement.

MACHINE
LEARNING

Machine learning begins
to flourish.

DEEP
LEARNING

Deep learning breakthroughs
drive AI boom.



1950's

1960's

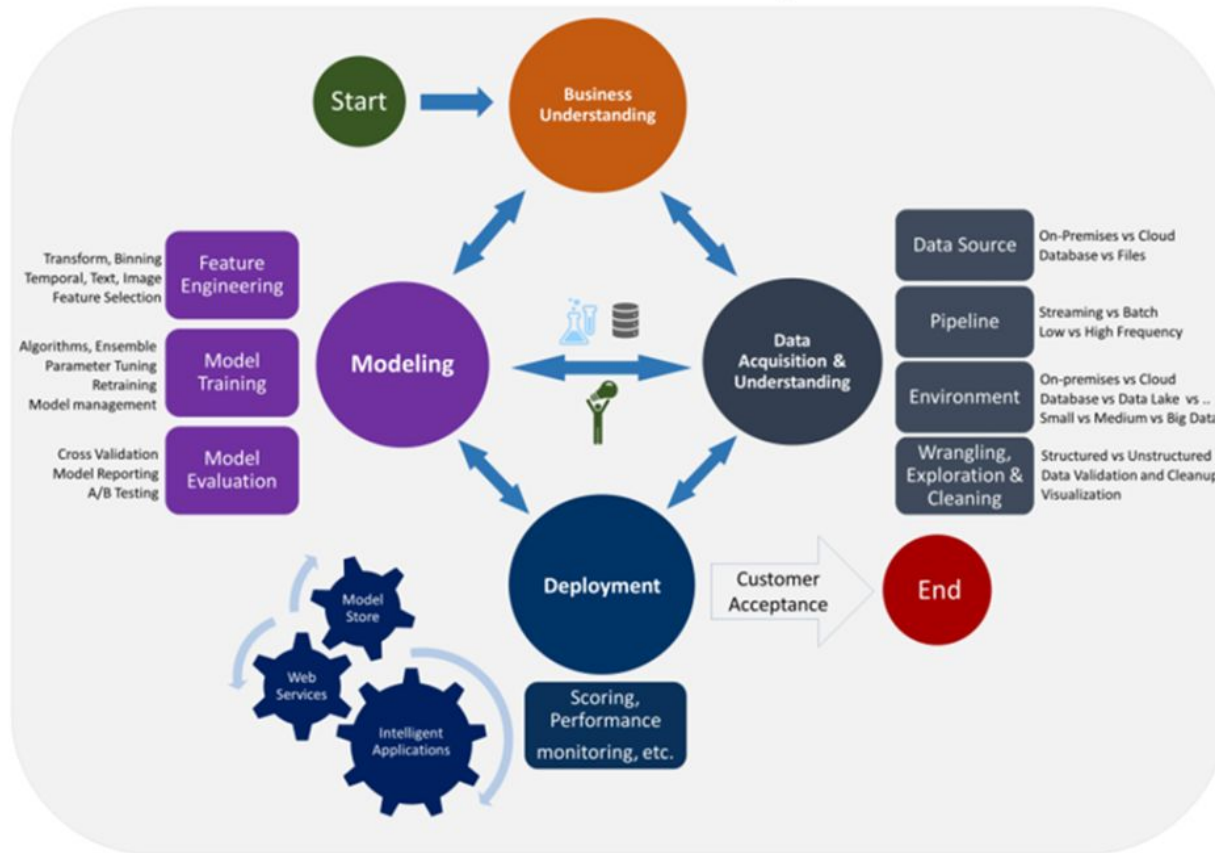
1970's

1980's

1990's

2000's

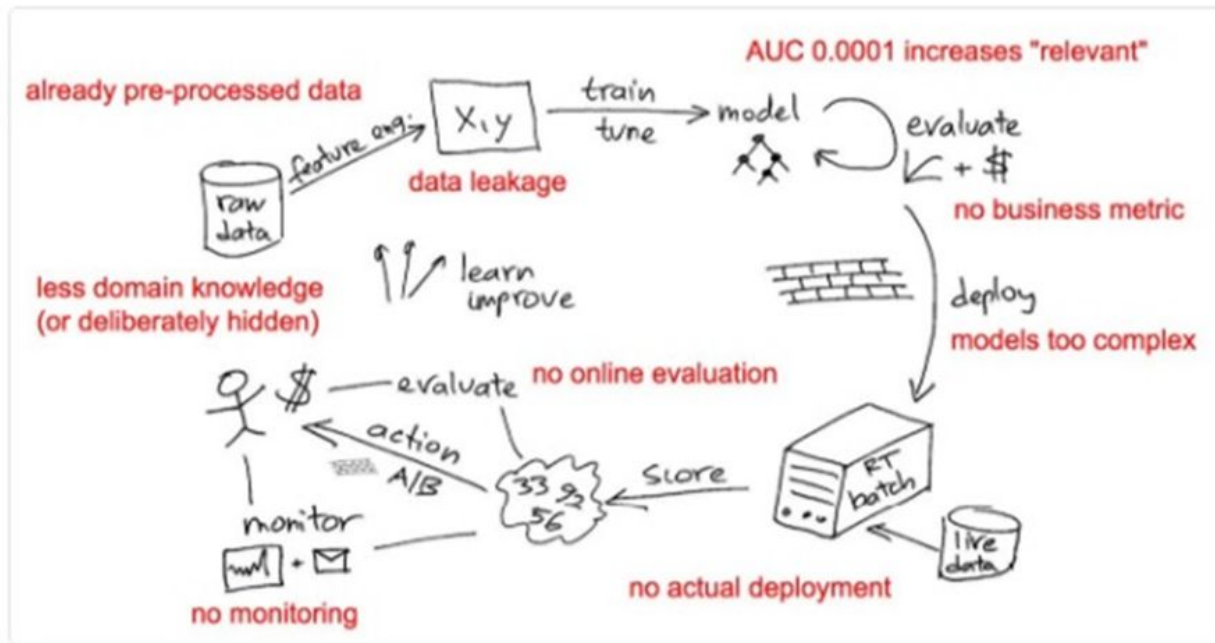
2010's



If you do [#kaggle](#) to learn [#machinelearning](#), you are missing on 80% of things you need for ML in real life/production

kaggle
DAYS

MEETUP



atlan

LOGICW

kaggle

bitgrit

Hackathons

- **Get the Overview**
- **Description, Evaluation, Timeline, Specific Req.**
- **Explore the data**
- **Explore other kernels**
- **Build a baseline**
- **Keep improvising till deadline**

Real World Projects

- **Identify & evaluate the Opportunity**
- **Develop Business Understanding**
- **Fetch, qualify & analyze available data**
- **Build a Prototype/POC**
- **Follow CRISP-DM methodology**
- **Deploy, Host & Monitor**

Head to Head Comparison



Problem Statement

- Kaggle: Problem Statement is well defined
- Project: Need to identify & formulate Problem Statement

Data Availability

- Kaggle: Data-sets are available
- Project: Need to identify & fetch relevant data

Evaluation Sets

- Kaggle: Train-Test-Real data are already segregated
- Project: Need to segregate Train-Test-Real data

Additional Data

- Kaggle: You may or may not use additional data
- Project: You can always identify & use relevant additional data

Evaluation Criteria

- Kaggle: Evaluation criteria is available
- Project: You need to define evaluation criteria

Deployment

- Kaggle: You need to submit the results in specific format
- Project: You need to deploy & host the model for business

Timelines

- Kaggle: You get a deadline to submit
- Project: You can carry on as long as project funds permit

General Differences

Model Performance

- Kaggle: You have the leaderboard to know where you are
- Project: You are the best as long as you are not challenged

Expectations

- Kaggle: Expectation is move higher the leaderboard
- Project: You can manage expectations of the stakeholders

Business Value

- Kaggle: You can use all the resources you can
- Project: You take decisions in terms of business value

Time to Market

- Kaggle: Competition timeline is important
- Project: Time to market is an important aspect

Collaboration

- Kaggle: You can collaborate with other competitors to form a team
- Project: You need T-shaped professionals to deliver

Model Complexity

- Kaggle: Models can be as complex as they can
- Project: Practical deployment aspects are considered while increasing the complexity

Concluding Thoughts

Q&A



kaggle
DAYS

MEETUP

amazonkindle
(ebook)

notionpress
(India)

amazon.com[®]
(Outside)

atlan

LOGICW

kaggle[™]

bitgrit