



HACK 2.0 PITCH



Exploring Data to rock your Career

Date : 22nd March 2023 | Speaker : Ayon Roy

HACK 2.0 PITCH

- Executive Data Scientist @ NielsenIQ
- Z by HP Global Data Science Ambassador
- Mentored/Judged 100+ Hackathons
- Delivered 60+ Technical Talks
- Brought Kaggle Days Meetup Community in India for the 1st time



Ayon Roy

HACK 2.0 PITCH



Agenda

- A primer to Hackathon
- Introduction to Data domains
- Kick start your Data journey
- Things to focus on, while making a Data project
- How & Why organize your Data project
- Exciting projects to learn from
- A primer to Internships in Data domain

HACK 2.0 PITCH



HACK 2.0 PITCH



HACK 2.0 PITCH



HACK 2.0 PITCH



HACK 2.0 PITCH

What do you do in a Hackathon?

- You sit down and identify the problems as a team. Define the problem statement accurately, clearly and discuss its impact on society.
- You learn the different components of the problem, objectively. Ask questions like:
 - What can be done effectively?
 - How, why did it happen?
 - Where do we start?
- You generate different ideas and decide on one final idea to present to experts and SMEs for feedback

HACK 2.0 PITCH

What do you do in a Hackathon?

- Based on the feedback, You develop a crude prototype of the idea.
- You iterate, iterate, and then pivot if needed - then iterate, iterate and again iterate!
- You present your last iteration, rough and crude but in the best light possible
- Learn through the process, what worked and what didn't work, what you as a team agreed on and what you as a team didn't agree on

HACK 2.0 PITCH



What do you get out of participating in a hackathon?

- You learn to become a problem solver!
- You learn to work with a team of interdisciplinary people.
- You learn to give voice to the problems of society. you live in.

HACK 2.0 PITCH



What do you get out of participating in a hackathon?

- You learn to become part of the solution making progress and find out the real struggles behind it.
- You or your team may win the hackathon and get recognized as trailblazers in the challenges facing society now, but most importantly you will gain inside knowledge of how technology works, and how it can be leveraged to benefit humanity.

HACK 2.0 PITCH



Why the domain of Data matters in 2023 ?

To analyze, extract information from huge datasets which maybe beyond the ability of general tools to manage, process data.

- Volume : Scale of Data
- Variety : Different types of Data
- Velocity : Speedy Ingestion of new Data
- Veracity : Uncertainty in the Data

HACK 2.0 PITCH

So, what is AI, ML, DL ?

Artificial Intelligence

A technique for incorporating human intelligence to machine

Machine Learning

ML is a subset of AI that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. ML is about learn from past to predict the future.

Deep Learning

DL is a subset of ML where artificial neural networks, algorithms inspired by the human brain, learn from large amounts of data.

HACK 2.0 PITCH

A glowing yellow lightbulb icon with rays emanating from it, positioned centrally between the words 'HACK' and 'PITCH' in the title.

How to start your
Data journey

HACK 2.0 PITCH

A lightbulb icon with the number '2.0' inside it, positioned between the words 'HACK' and 'PITCH' in the title.

Start with
Maths

HACK 2.0 PITCH

- Week 1 : Linear Algebra [B] <https://www.khanacademy.org/math/linear-algebra>
- Week 2 : Calculus [B] <https://www.youtube.com/playlist?list=PLZHQObOWTQDMsr9K-rj53DwVRMYO3t5Yr> or <https://www.mathsisfun.com/calculus/> ; want theoretical notes , find it at <https://the-learning-machine.com/article/machine-learning/calculus> .
- Week 3 : Probability [B] <https://www.edx.org/course/introduction-probability-science-mitx-6-041x-2>
- Week 4 : Statistics [B] <http://alex.smola.org/teaching/cmu2013-10-701/stats.html>
- Algorithms (Only if you want to learn proper software development) [Highly optional]
This is an overview of what the students study as the subject Data Structures & Algorithm . So if you are fluent with this part , you can skip this !! <https://www.edx.org/course/algorithm-design-analysis-pennx-sd3x>

HACK 2.0 PITCH

A lightbulb icon with the number '2.0' inside it, positioned between the words 'HACK' and 'PITCH' in the title.

Start with Python

&

try to **implement** those
Mathematical Concepts

HACK 2.0 PITCH

A glowing yellow lightbulb icon with rays emanating from it, positioned centrally between the words 'HACK' and 'PITCH' in the title.

Start exploring Libraries
& then start related
courses

HACK 20 PITCH

- Introduction to python for data science [B] <https://www.datacamp.com/courses/intro-to-python-for-data-science>
- Want to dive deeper into Data Visualization & Pre-Processing ? Look into Data Visualization & Pre-Processing section in miscellaneous resources . [Highly optional]
- Want to explore the field of Deep Learning ? See the Deep Learning Section in miscellaneous resources . [Highly optional]
- Want to explore the field of Natural Language Processing [NLP } ? See the Natural language Processing Section in miscellaneous resources . [Highly optional]
- See how ML codes are written and made to work at - > <https://github.com/maykulkarni/Machine-Learning-Notebooks> or <https://github.com/GokuMohandas/practicalAI/blob/master/README.md> . [Highly optional]
- Find useful resources here at <https://github.com/ujjwalkarn/Machine-Learning-Tutorials/blob/master/README.md> . [Highly optional]

HACK 2.0 PITCH

A glowing yellow lightbulb icon with rays emanating from it, positioned centrally between the words 'HACK' and 'PITCH' in the title.

Don't rush behind
completing Courses
&
adding them to Resume

HACK 2.0 PITCH

A glowing yellow lightbulb icon with rays emanating from it, positioned between the words 'HACK' and 'PITCH' in the title.

Understand the concepts well
before starting Projects

HACK 2.0 PITCH

Start with
Projects

HACK 2.0 PITCH



Things to focus on
while making a
Data Project

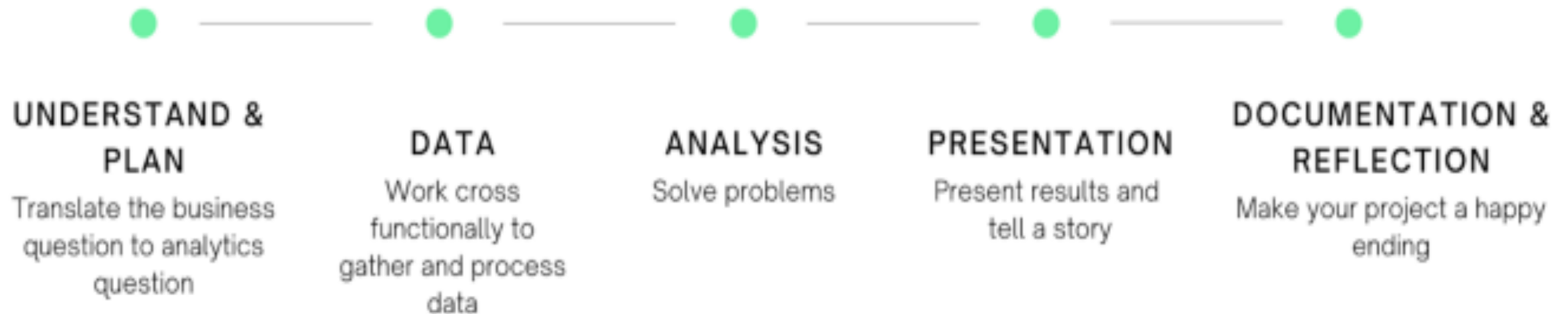
HACK 2.0 PITCH



HACK 2.0 PITCH

Analytics Project Life Cycle

The 5 Phases



HACK 2.0 PITCH

A glowing yellow lightbulb icon with rays emanating from it, positioned centrally between the words 'HACK' and 'PITCH' in the title.

**How to organize
your
Data Project?**

HACK 2.0 PITCH

Local Project Directory	Github Repository
<ul style="list-style-type: none">▪ Project plans/objectives▪ Project datasets▪ Project codes<ul style="list-style-type: none">○ Jupyter notebook○ R scripts○ Python scripts▪ Output files<ul style="list-style-type: none">○ Visualizations○ Tables○ Other useful outputs▪ Project report	<ul style="list-style-type: none">▪ README file▪ Project datasets▪ Project codes<ul style="list-style-type: none">○ Jupyter notebook○ R scripts○ Python scripts▪ Output files<ul style="list-style-type: none">○ Visualizations○ Tables○ Other useful outputs▪ Project report

<https://gist.github.com/ericmjl/27e50331f24db3e8f957d1fe7bbbe510>

<https://gist.github.com/ericmjl/27e50331f24db3e8f957d1fe7bbbe510>

HACK 2.0 PITCH

A lightbulb icon with the number '2.0' inside it, positioned between the words 'HACK' and 'PITCH' in the title.

But **why organize**
your
Data Project?

HACK 20 PITCH

- Organization increases productivity as avoid wasting time searching for project files such as datasets, codes, output files, and so on.
- A well-organized project helps you to keep and maintain a record of your ongoing and completed data science projects.
- Completed data science projects could be used for building future models.
- A well-organized project can easily be understood by other data science professionals when shared on platforms such as Github.

HACK 2.0 PITCH

But **which projects to**
start learning with ?

HACK 2.0 PITCH

- Beginners Section [B] : Brush your basic concepts and revise them to start doing projects

Titanic Dataset

Iris Dataset

Stock Price Prediction

Stores Sales Forecasting

Housing Price Prediction

Guide for Beginner Projects:

First of all see Below 2 videos to get an idea on how to make projects of Data Science and Machine Learning And then Move to Kaggle for Making your own project.Its is Good if you Make Minimum 2-3 Projects on your own.

- Titanic Survivor : <https://www.youtube.com/watch?v=fS70iptz-XU&t=>
- Credit Card Fraud Detection : <https://www.youtube.com/watch?v=gCWBFyFTxVU>

HACK 2.0 PITCH

Intermediate & Advanced Section

- Learn libraries like Opencv , Tensorflow , SkLearn
- 1) Natural Language Processing : MNIST Handwritten Digit Classification , Twitter Sentiment Analysis
 - 2) Email Spam Classifier
 - 3) Fraud Detection System
 - 4) Computer Vision : Face Recognition , Face Detection

HACK 20 PITCH

" I am a beginner in AI,ML,Data Science & trying to do projects; but not succeeding as I get stuck more often "

Here is my way ahead if you are facing the same.

" Start Simple Projects & Be Motivated "

We all usually want to do the best projects & showcase them in our [resume](#) & hence sometimes end picking up a complex project at the first go. But do understand that while it's very normal to pick complex projects as a beginner, because we can't analyze the scale of project at first go. Picking up a complex project at a first sight may demotivate you as they have a lot of details, requires lot of studies to progress, thus a beginner ends up leaving the project midway & be traumatized. So start your journey with Simpler & Smaller projects as they require comparatively less details & can be achieved over a short period of time, thus helping us to stay motivated & keep doing projects. And as the learning in AI, [machinelearning](#), [datascience](#) never stops, so as we get motivated with completion of small projects; we learn & practice more while increasing the complexity of our upcoming projects. Still waiting to start ? Start today !! All the best !!

HACK 20 PITCH

Ideas to take reference to build in Hackathon using data

- A system that detect fires and Smoke, It has two key services: fire tracking and alerting the emergencies

Video : <https://youtu.be/L5eUPXxJVdI>

Github : <https://github.com/Fellah-wassim/IgnitionGuard>

- Help visually impaired individuals cross the street using machine learning

Video : https://youtu.be/o_cyuqQoDiU

Github : <https://github.com/Szugalew/PedestrianTrafficLightDetectorRaspberryPi>

HACK 2.0 PITCH

Ideas to take reference to build in Hackathon using data

- A YouTube sentiment analysis app built using Node.js, Cohere API, and Google Cloud's YouTube Data API

Video : <https://youtu.be/ZrQemkKDYxU>

Github : <https://github.com/Ryan-Diep/ChatRoller>

- Using Image Analysis to predict what color palettes for clothing would look best on the user.
- Using neural networks for facial recognition to make paying easier.

HACK 20 PITCH



Participate in Hackathons
but

**Take the next step
with
Internships**

HACK 2.0 PITCH

A glowing yellow lightbulb icon with rays emanating from it, positioned between the words 'HACK' and 'PITCH' in the title.

What are the **requirements of a decent Internship Opportunity** in Data field ?

Let's divide the variety of Internships out there in the Data industry



Requirements for Entry Level Internships

- Ideal for Second-Third Year Students of a 4 year Undergraduate course
- Basic Level of Python, OOPs, File Systems
- Good knowledge of Scraping, Numpy, Pandas & Data Visualization libraries
- High level overview of Machine Learning algorithms
- Few basic Data Pre-Processing & Exploratory Data Analysis Projects



Requirements for Middle Level Internships

- Ideal for Third-Fourth Year Students of a 4 year Undergraduate course
- Sound knowledge of Data Science concepts & other ML algorithms with good grasp on statistics & concept of maths, SQL
- Good knowledge of Deep Learning concepts, DBMS, API development
- Self projects using ML algorithms



Requirements for Advanced Level Internships

- Ideal for Final Year Students of a 4 year Undergraduate course
- Specialised domain of expertise like Computer Vision, Natural Language Processing etc.
- Proficient with whiteboarding of ML algos along with explanation of the basics
- Basic knowledge of Docker, Cloud
- 4-5 very good projects using the complex Deep Learning concepts like LSTMs, Transformers etc.

Best way to get a Pre-Placement Offer before campus hiring starts

HACK 20 PITCH

Why should students do Internships in data domain ?

- To learn dealing with Messy, unstructured, incomplete data (This is real industry)
- To experiment & learn new things as an Intern so that you can save time as a Full Timer excluding the mistakes you did as an intern.
- To understand how end to end real world Data Science applications works.
- To network with people who look like Future of You.
- To work under pressure & learn how to deliver in tough deadlines too.

To make like minded friends & throw a party at Starbucks

HACK 2.0 PITCH

How to apply for Internships ?

- LinkedIn Jobs [Worked for Me]
- Angel.co [Worked for Me]
- Internshala [Worked for Me]
- Through Organization's Careers Page [Worked for Me]
- Commenting Interested on someone's LinkedIn Posts [Not Worked for Me]
- Career/ Internship Fairs [Never Tried]
- Via Winning Hackathons [Worked for Me]
- Asking for Referrals [Worked for Me]
- Community Events [Have seen it work]

HACK 2.0 PITCH

A few useful resources

- <https://ayonroy.ml/mentor>
- <https://github.com/iAyon/100DaysOfMLCode/blob/master/Coursework.md>

HACK 2.0 PITCH

A lightbulb icon with the number '2.0' inside it, positioned between the words 'HACK' and 'PITCH' in the title.

Let me answer your Questions
now.

It's your time to speak

HACK 2.0 PITCH

Danke Schoen

Questions ? Any Feedbacks ? Did you like the talk?
Tell me about it.

If you think I can help you,
connect with me via

Email : ayon-roy@outlook.com

LinkedIn : <https://www.linkedin.com/in/ayon-roy>

Website : <https://AYONROY.ML/>



Download the slides