

The Product Management for AI & Data Science course

⚠ A product manager is a multidisciplinary individual who identifies

- what product to build
- when to build it
- why it needs to be built

⚠ A great product manager has a broad knowledge of technical, user experience, and business domains, but they're not necessarily the expert in any of these areas

⚠ Challenges of AI projects / products

- uncertainty of time & performance
- challenge of communicating AI

⚠ project managers have

- a clear defined project
- timeline
- budget

⚠ project manager → tactical
product manager → strategic

⚠ Product manager's responsibilities

- ↳ gathering requirement from user feedback
- ↳ identifying problems & opportunities
- ↳ prioritizing initiatives
- ↳ building a roadmap
- ↳ creating a vision

⚠ Data analytics looks for answers and data science explores ideas

⚠ product managers still work very closely with data analysis to help them analyze results of experiments, learn from data for product research, or make some predictions about how their product will grow

⚠ Data Science

- ↳ How to build innovative AI & Data Products

⚠ Data Analytics

- ↳ determine what to build & how your product is performing

⚠ Algorithms are a well defined procedure that tells a computer how to make a particular action happen

⚠ Artificial Intelligence (or AI)

↳ technologies that perform cognitive tasks that normally only humans could perform

⚠ Some AI subjects

- computer vision
- natural language processing (NLP)
- speech recognition
- robotics

⚠ Artificial Narrow Intelligence

↳ AI that can match human performance in one specific domain

⚠ Artificial General Intelligence

↳ AI that can match human's overall intelligence across multiple domains

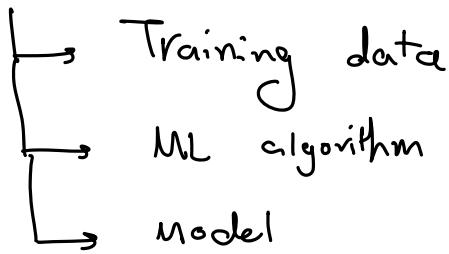
⚠ Artificial Super Intelligence

↳ an AI that can exceeds human intelligence

⚠ Machine Learning (ML)

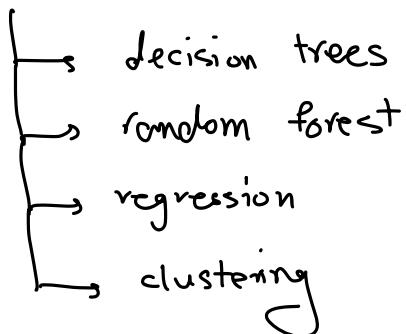
↳ a subset of AI that uses algorithms to teach machines how to learn through experience in large datasets

A Parts of Machine Learning



⚠ Machine learning will find patterns in the examples we present it with, and it will make predictions based on those patterns

⚠ Machine learning is powered by a lot of statistical thinking



⚠ Emphasize the following points with stakeholders

- ↳ garbage data in equals garbage data out
- ↳ the machine is only as good as the data it is trained on
- ↳ humans are still needed to both label the example data and / or validate the models' outputs.

⚠ Deep learning is a subset of machine learning and a much newer method of building AI than ML

- ↳ doesn't mean it's superior to ML

⚠ Three parts of Deep Learning Neural network

- an input layer
 - ↳ the raw data input, where your data goes in
- hidden layers
 - ↳ where at each layer a new analysis is applied to the data from the previous layer
- an output layer
 - ↳ where the final output is processed

⚠ Neuron

- ↳ where the data is being processed

⚠ channels

- ↳ the connection between layers

⚠ with more layers comes a greater need for data

⚠ convolutional neural network (CNN)

- ↳ commonly used in image recognition

⚠ Recurrent neural network (RNN)

- ↳ commonly used in text to speech applications

⚠ When to use machine learning vs. deep learning

- ↳ how much data you have
- ↳ hardware
- ↳ training
- ↳ input
- ↳ interpretability



⚠ Feature Selection

- ↳ helps reduce amount of time and complexity that comes with training and testing a model

⚠ supervised learning

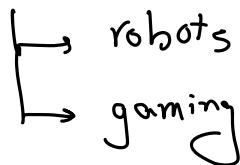
- ↳ when you show the model what you want it to do
 - ↳ classify data into categories
 - ↳ predict trends

⚠ unsupervised learning

- ↳ look through all the "unlabeled" client data, and find commonalities between customer groups
 - ↳ clustering common things together
 - ↳ detecting anomalies

⚠ reinforcement learning

↳ where you tell the machine what the goal is, and you let the machine figure out how to achieve the goal



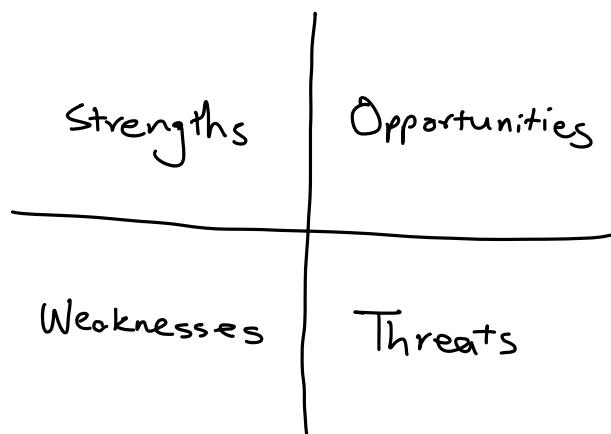
⚠ AI business model innovations

- insight generation
- operational efficiencies
- adding to an existing product
- developing a new product

⚠ Problems that AI is well suited for

- when human expert could perform the task in a few seconds
- when it's difficult or impossible to write down the rules
- when it's easy to get examples of the desired behavior

⚠ SWOT analysis



⚠ Hypothesis

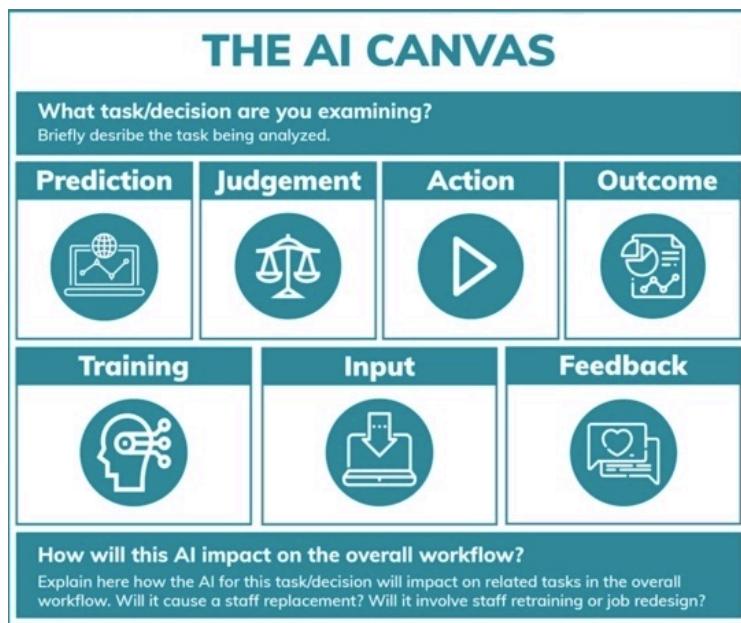
- ↳ a testable statement of what you believe to be true
- ↳ should be specific as possible

⚠ Hypothesis ensures that you clearly state five things in relation to your data

- ↳ your target market
- ↳ your assumption
- ↳ your goal
- ↳ the expected impact
- ↳ how will you measure success

⚠ AI business Canvas

- ↳ a lean startup method used to make strategic management decisions when you develop a new business model, or if you want to evaluate your existing business model



⚠ 3 core user types a product manager might build

- consumer product users
- internal product users
- software as a service users

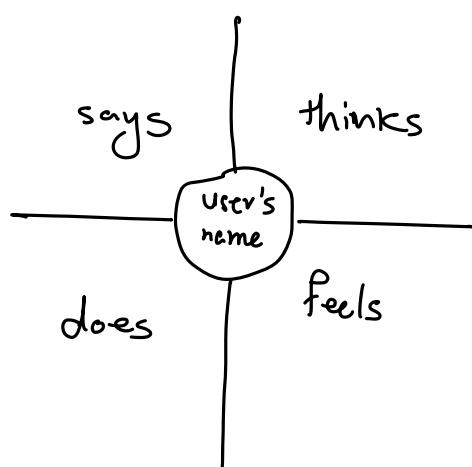
⚠ '3 whys method' is used to get the core problem

⚠ User research methods

- market research & usage data
- exploratory interviews
- surveys

⚠ building out a user persona is like creating a fictional character profile that is rooted in observed facts about your user base

⚠ user empathy map



⚠ Prototype

↳ a preliminary model or design of a product, from which other versions of the product are developed

⚠ Wireframe

↳ a simple visual layout of what a product interface might look like

⚠ Suitable experiments for AI products

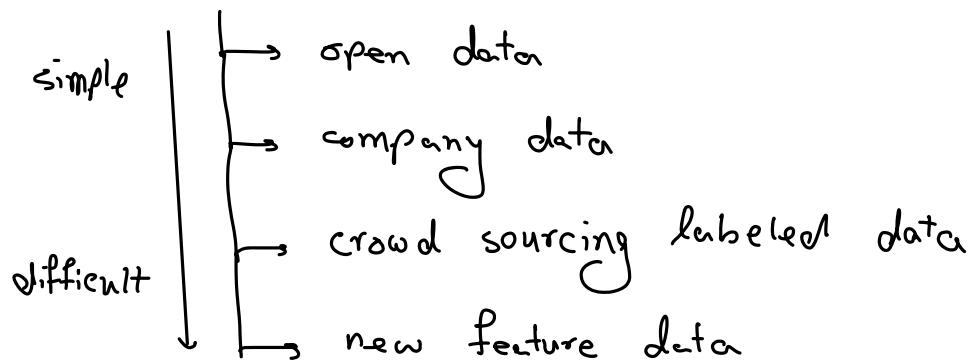
→ AI Personality design experiment

→ WOZ design experiments

 ↳ wizard of oz

↳ users interact with a computer system they believe is autonomous, but in reality it is being partially or fully operated by a human being

⚠ Data collection methods



⚠ Open data

↳ Data that is freely available for anyone to use & republish it as they wish without any copyright or Patent restrictions

⚠ Data annotation

↳ is the process of labeling data to make it usable for machine learning

⚠ Consider the following criteria when choosing a data annotation vendor

- specialization
- cost
- speed
- training

$$\Delta \text{ROI} = \frac{\text{Investment Gains} - \text{Investment costs}}{\text{Investment costs}}$$

* Investment Cost = the cost of acquiring, licensing or purchasing the data

⚠ A database contains realtime information about your product. It is used to process the transactions made in the product, what actions user takes, and more

⚠ A data warehouse is a system that brings together data from multiple sources in the organization so that it can be analyzed

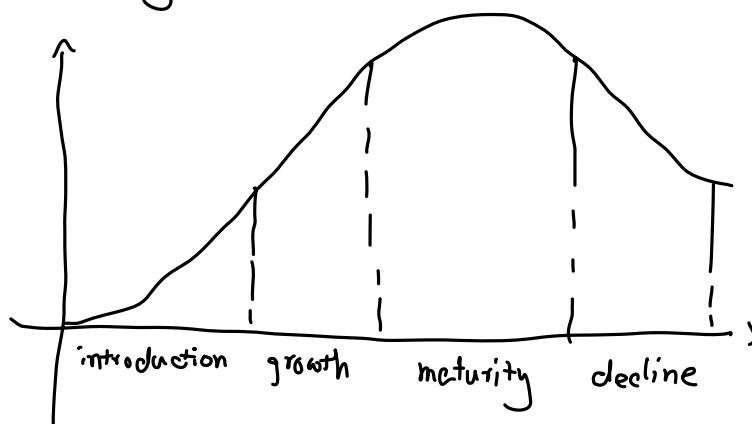
⚠ ETL systems

- extract: extract data from databases
- transform: transform data from raw data to ready to digest data
- load: load it into the data warehouse

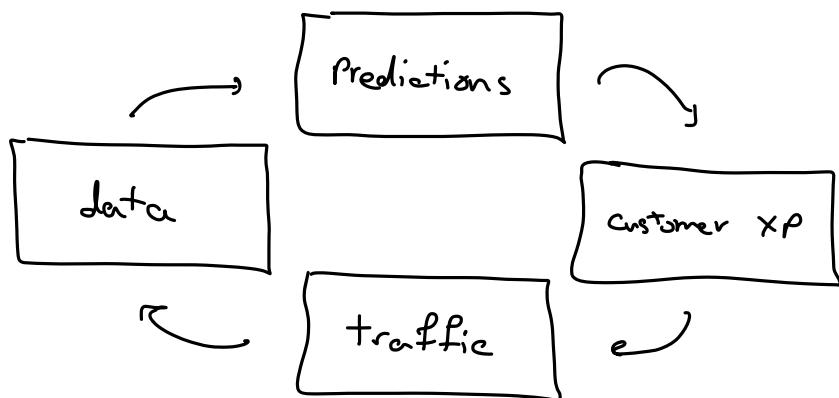
⚠ Data Lake

- It centralizes data in one place instead of in multiple databases, but unlike with a data warehouse, it is not centralized for one particular purpose

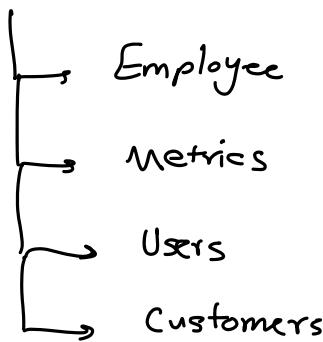
⚠ Product life cycle



⚠ AI flywheel effect



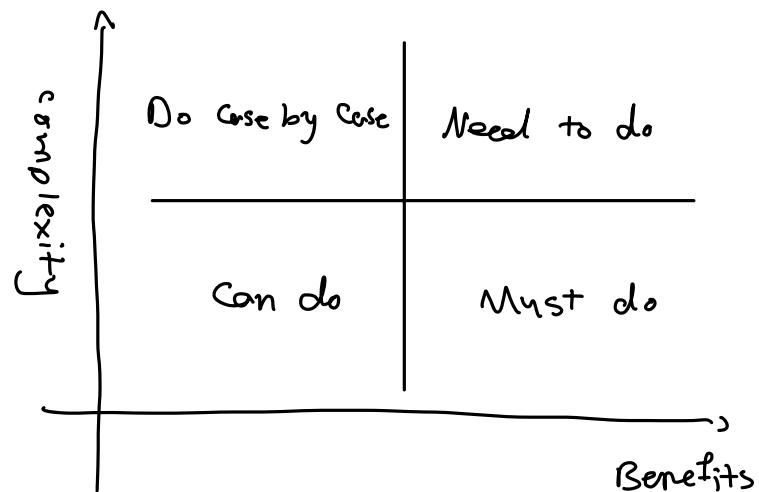
⚠ Data gathering sources (EMUC)



⚠ Product ideation techniques

- crazy eights : a brainstorm method
- round robin
- the fake press release

⚠ Benefit vs Complexity



⚠ MVP

- simplest pared down version of your product solution
- don't make as much sense for AI & Data products

⚠ AI MVD

- the smallest amount of data required to develop a potential AI product

△ Agile Software Development

↳ a framework based on developing a product iteratively instead of all at once

△ Agile methodologies like scrum are not often applicable to AI and data products

△ Kanban

↳ a methodology that was designed to help teams visualize their works and manage the number of items they were working on simultaneously

△ AI product board

DATA BACKLOG	PROCESSING	MODELING	TRAINING	TESTING	DONE

△ who should build your model are

- ↳ Is this problem core to my business?
- ↳ Can I realistically get the data set needed?
- ↳ Do we have an experienced machine learning team?

△ Enterprise AI

↳ AI and data products developed by third party companies that are sold to organizations looking to quickly leverage AI solutions

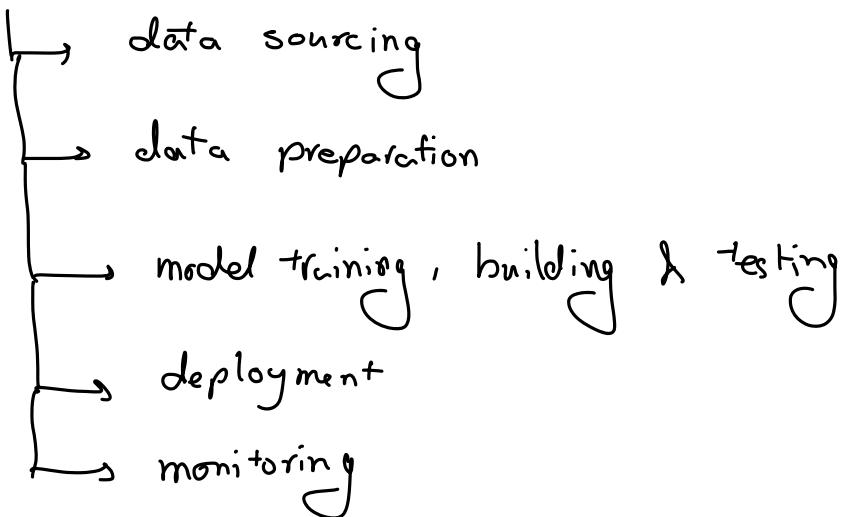
⚠ Evaluate enterprise AI vendors on these 6 criteria

- ① Data
 - what does this enterprise AI learn from?
 - is this data unique to them or not?
 - is the source of data similar enough to what you would need the model to do for your organization?
 - does this enterprise AI also enable customization of their models by including data from your own organization?
- ② specialization
- ③ Integration
- ④ customization
- ⑤ security
- ⑥ price

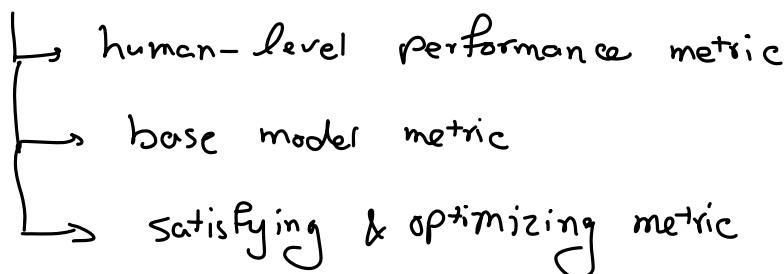
⚠ 3 ways to leverage MLaaS

- testing models
- outsourcing data infrastructure needs
- building models

⚠ In-house machine learning development takes 5 unique steps



⚠ How to set a goal for the AI



⚠ training data

↳ allows your ML model to learn

⚠ Validation data

↳ fine-tune your model's performance

⚠ testing data

↳ for the finished model

⚠ Confusion Matrix		Predicted : NO	Predicted : Yes
Actual : NO	#	TN	FP
		FN	TP
Actual : Yes			

⚠ Precision = $\frac{TP}{TP + FP}$

⚠ recall = $\frac{TP}{TP + FN}$

⚠ accuracy = $\frac{TP}{\text{all data}}$

⚠ F1 score = $\frac{\text{Precision} \times \text{recall}}{\text{Precision} + \text{recall}} \times 2$

⚠ MLOps

↳ a practice for collaboration and communication between data scientists and operations professionals to help manage the deployment of a ML model

⚠ Deploying a model based on engineering effort

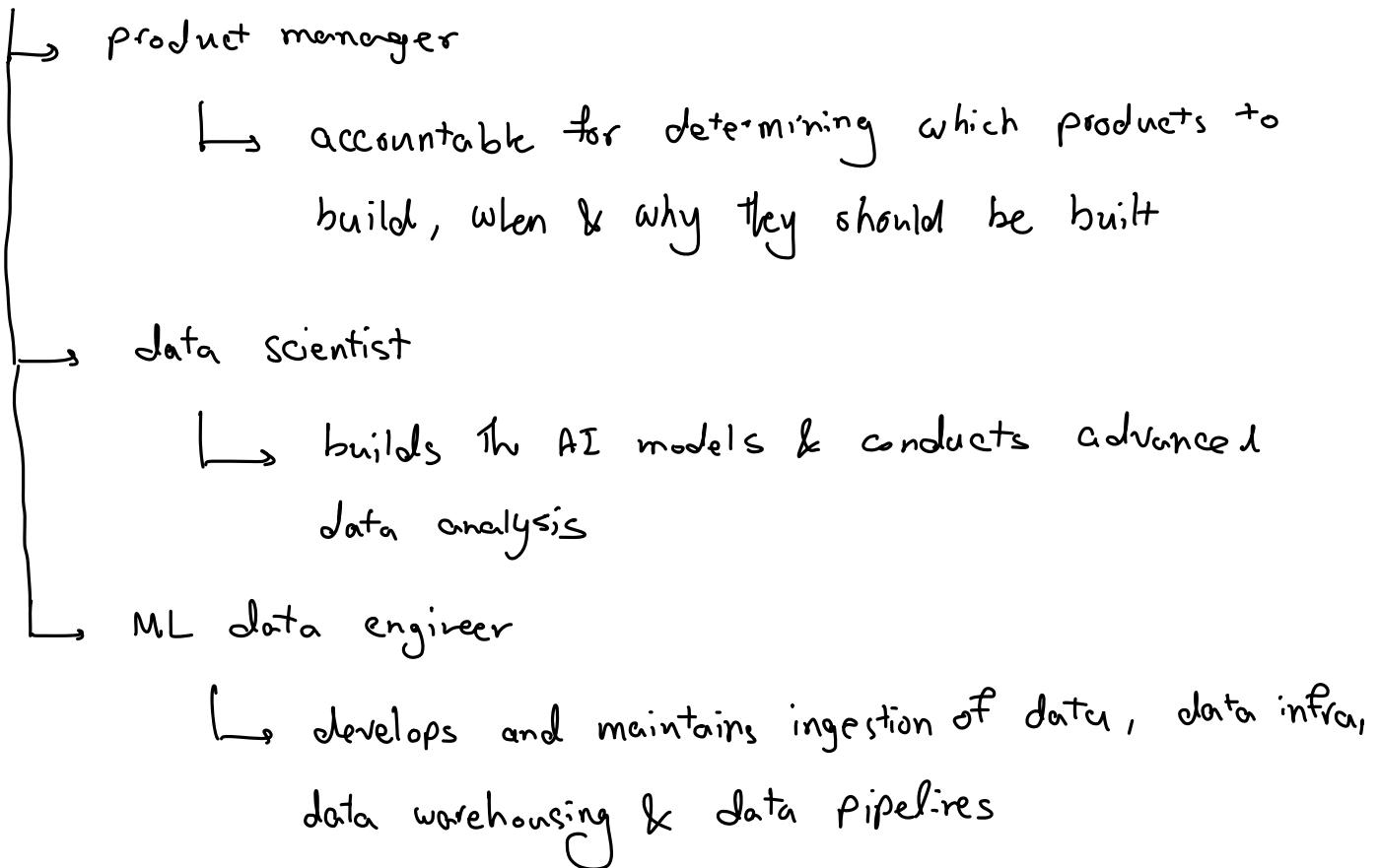
- Ad-Hoc predictions via SQL
- Batch predictions
- real-time predictions

⚠ Monitoring models

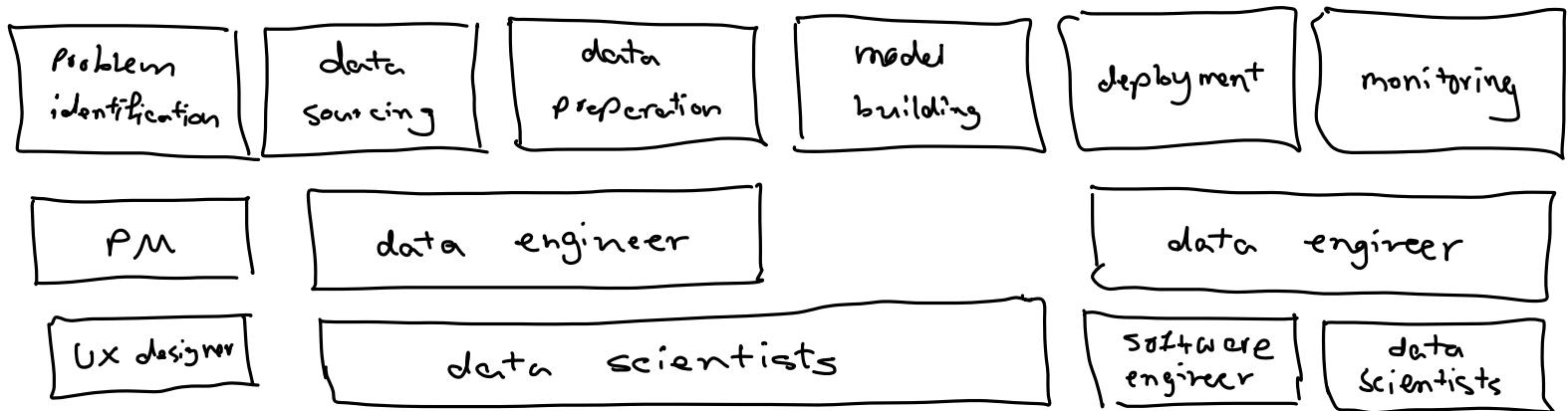
- proactive
 - ↳ identifies general outlier patterns in The real world compared to The test or training Patterns
- reactive
 - ↳ identifies bad model performance and The cause of it

⚠ shadow deployment is a way of testing a new model in production with very minimal risk

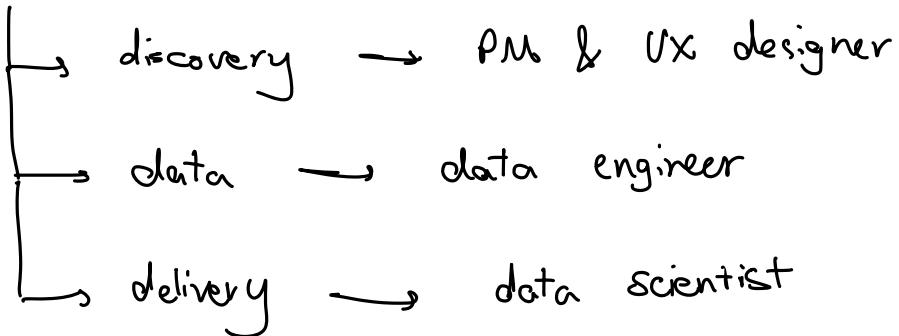
⚠ AI & Data team certain roles



⚠ A simplified workflow of an AI / Data product



⚠ triple-track agile methodology



⚠ 5 steps for successful stakeholder management

- build a relationship
- listen to their perspective
- educate them on AI and data
- keep them informed
- set boundaries

⚠ set the following expectations

- AI and data products often fail to make it to launch
- AI and data are not magic!
- AI and data products take time

⚠ active listening steps

- pay close attention
- show that you're listening
- confirm what you heard
- ask clarifying questions

A 4 steps for building user trust

- ask for permission
- explain the results
- give the user control
- differentiate between confidence

A Avoid bias with these steps

- consider any potential bias in the training data
- work hard to diversify your data
- evaluate your labeled data

A Data privacy Laws

- GDPR
- COPPA
- FOIP