Class 1- Welcome to the world of Machine Learning

This is me!

Pedram Jahangiry

Professional Practice Assistant Professor, Project Mentor for the Analytics Solutions Center

Department(s):

Data Analytics and Information Systems

Contact Information

Eccles Business Building 720

435.797.2345

pedram.jahangiry@usu.edu

Personal Website

Curriculum Vitae

Education

PhD, Economics, Arizona State University, 2017

Master, Economics, Simon Fraser University, 2013

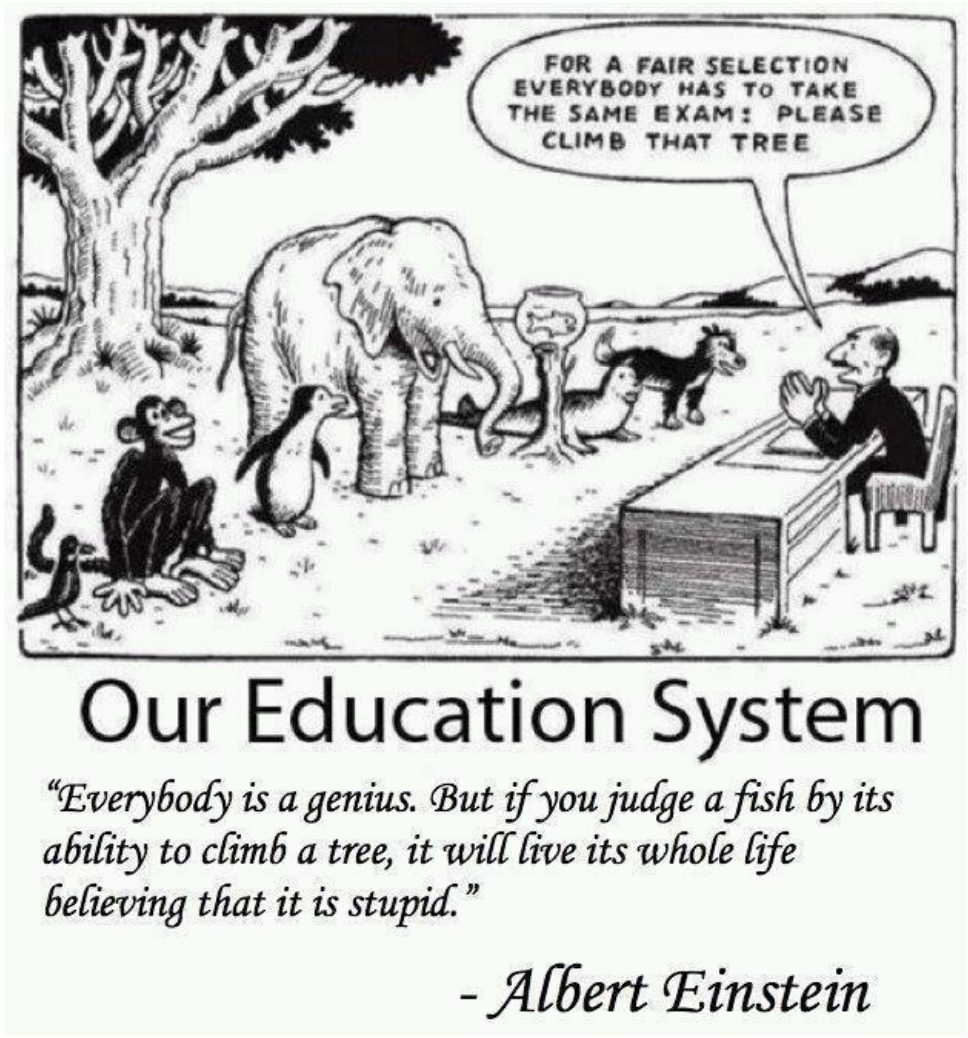
MBA, Sharif University, 2012

Industrial Engineering, IUST, 2009

Biography

Pedram Jahangiry, PhD, CFA, is an assistant professor in the Economics and Finance Department of the Jon M. Huntsman School of Business at Utah State University. Prior to joining the Huntsman School in 2018, Pedram was a research associate within Financial Modeling Group at BlackRock NYC. His research is involved in machine learning applications in finance, empirical asset pricing, and factor models.

My Teaching Philosophy



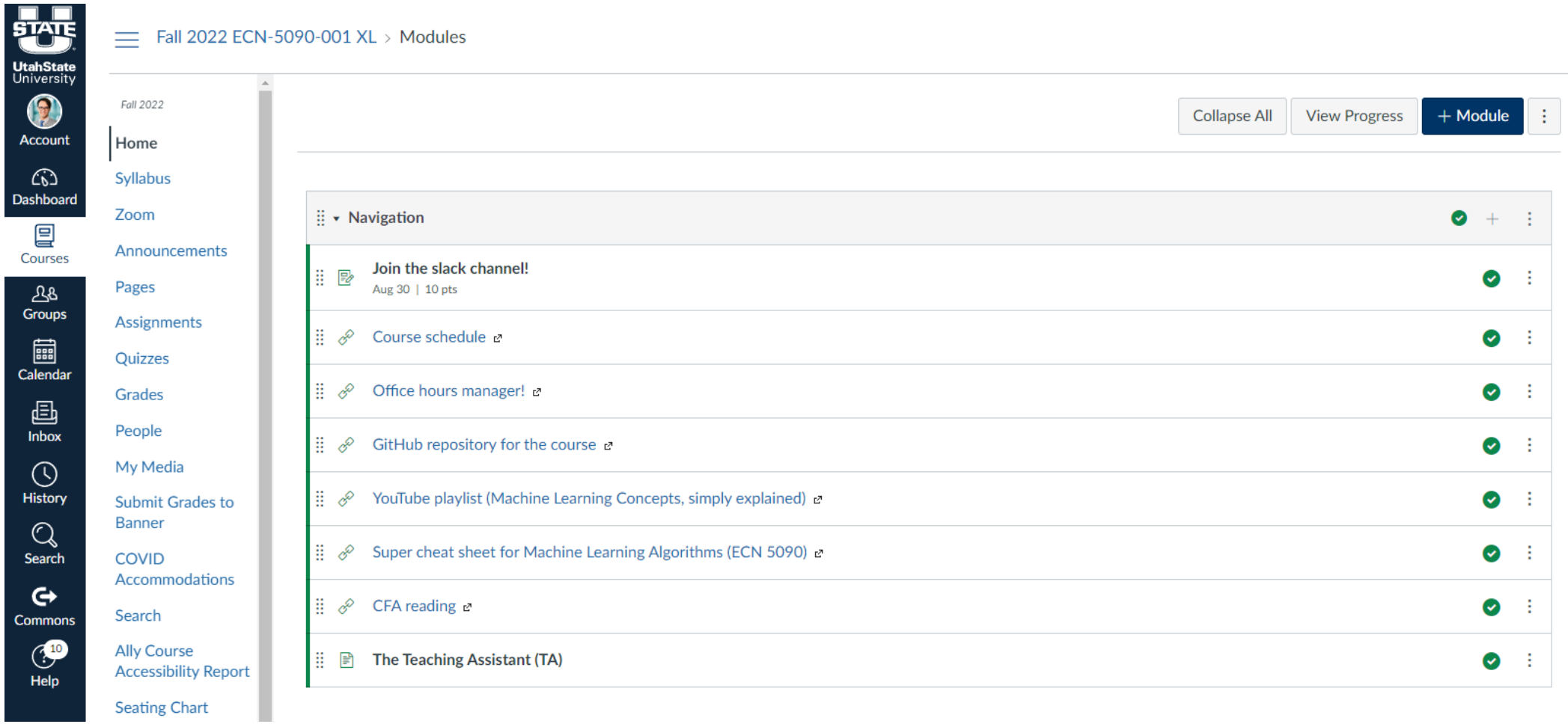
My Teaching style!

More focus on interaction and application in class while putting more lecture-like material online

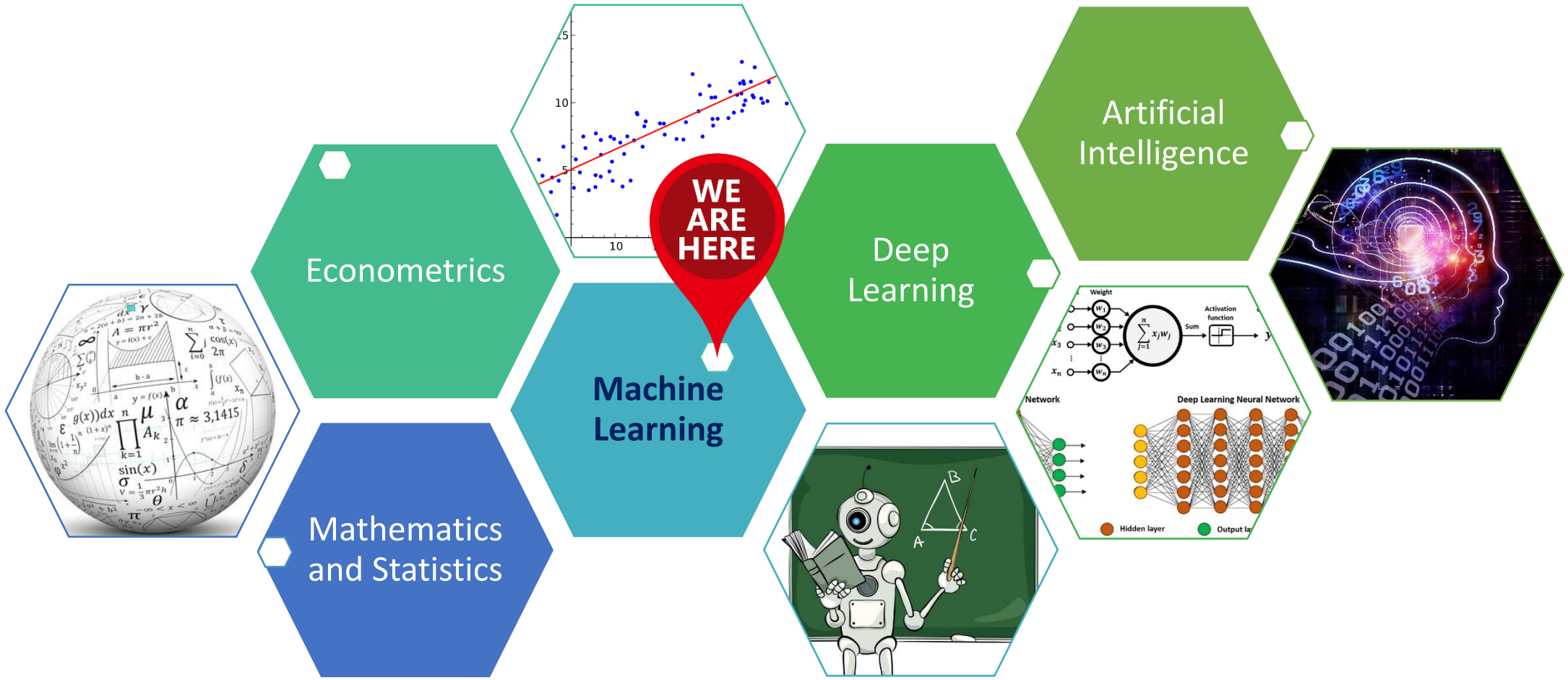
Classroom time will be used more to make mistakes in a safe environment

Balancing learning-by-studying with learning-by-doing

What's on Canvas?



Where we are?


Econometrics
WE ARE HERE
Machine Learning
Deep Learning
Artificial Intelligence
Mathematics and Statistics
Weight
Activation function
Sum
Network
Deep Learning Neural Network
Hidden layer

Big picture: Econometrics vs Machine Learning

What are we trying to do as a researcher?

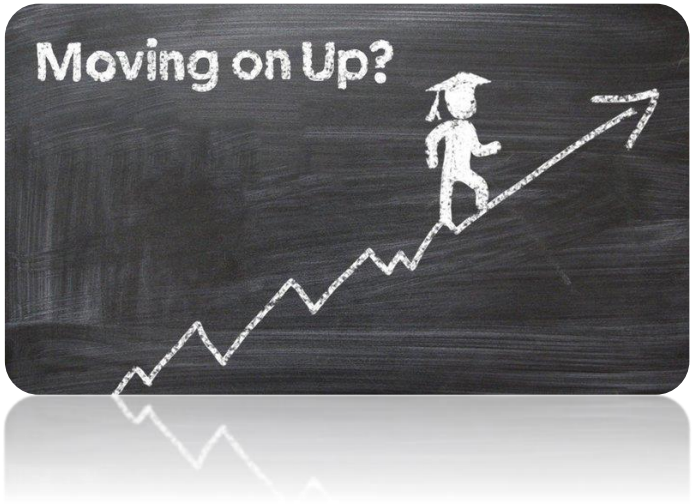
Solve real world problems, right?

Is there a theory?

What is the relationship between

- Sales and advertisement / R&D expenditure / seasonality / industry / … ?
- Quantity demanded and price / income / technology / price of competitors / … ?
- Wage and education/ age/ gender/ experience/ …?

A simple example



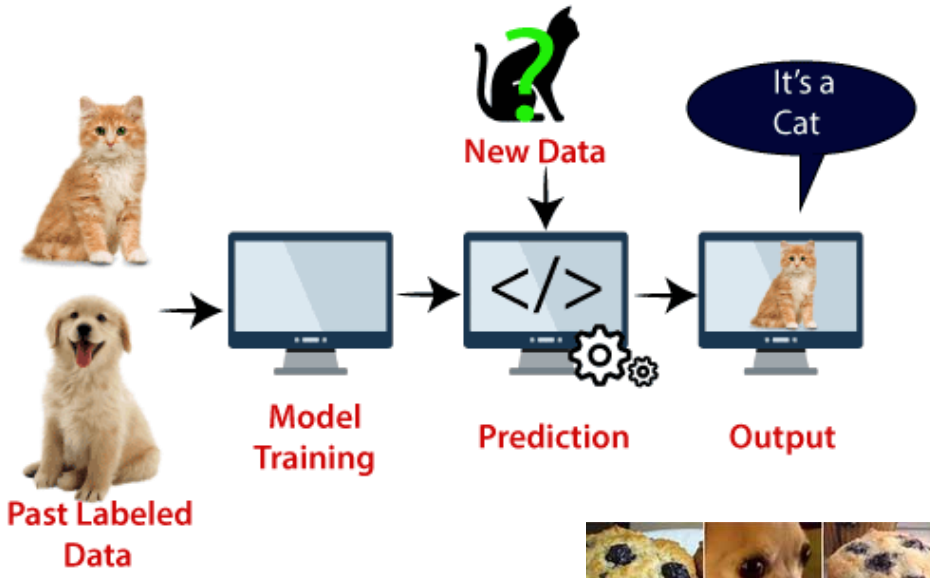
- Quantifying wage components! (is there a theory?)
- What are the drivers:
  - Education, age, experience, IQ, …
  - Ethnicity, race, gender, …
  - Industry, location, working hours, …
- Let's build a model (assuming a linear functional form!)

$$wage = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 exper + \beta_4 IQ + \cdots + \beta_k hours + u$$

- Can you **interpret** this model? Do you care about the interpretability?
- Can you make **predictions** using your model?
- Can you make this functional form more flexible? What are the caveats?

A different example

- Cat vs dog classification problem (image recognition)



- Do you really care about interpretability of the model here?
- What about accuracy of your predictions?

Statistical learning vs machine learning

	Statistical Learning	**Machine Learning**
Focus	Hypothesis testing & interpretability	Predictive accuracy
Driver	Math, theory, hypothesis	Fitting data
Data size	Any reasonable set	Big data
Data type	Structured	Structured, unstructured, semi-structured
Dimensions / scalability	Mostly low dimensional data	High dimensional data
Model choice	Parameter significance & in-sample goodness of fit	Cross-validation of predictive accuracy on partitions of data
Interpretability	High	Low
Strength	Understand **causal** relationship & behavior	Prediction (forecasting and nowcasting)

A more complex example

Apple stock price prediction

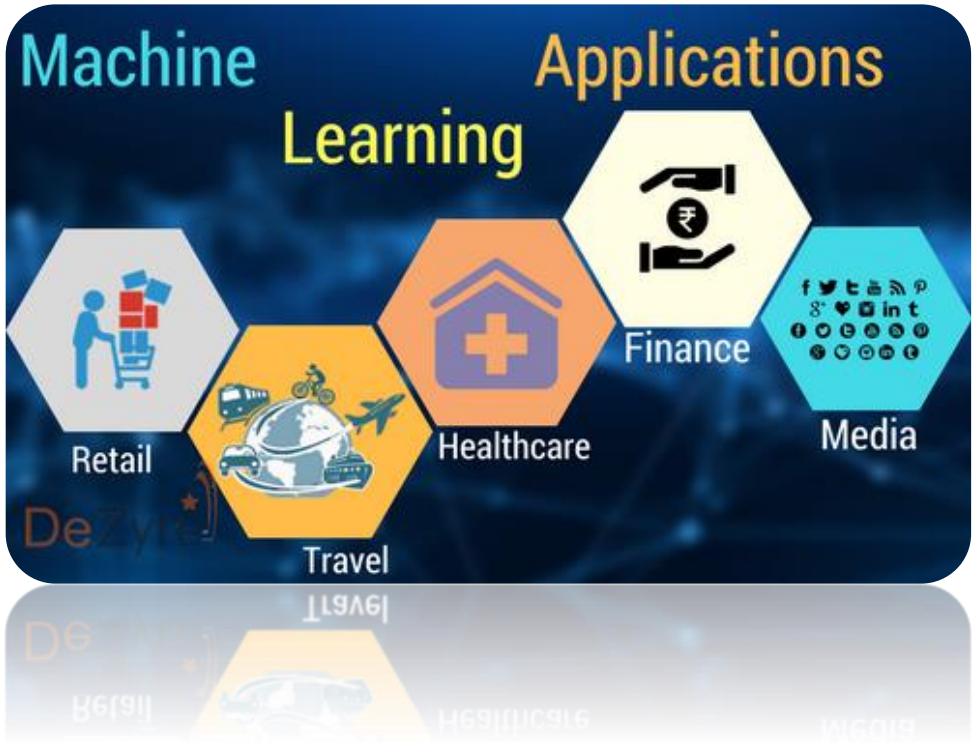
- What are the classical drivers:
  - Company's fundamentals (balance sheet, income statement, cash flow statement)
  - Competitors (comparing multiples)
  - Technical analysis!
  - Seasonality (holidays, months, days, …)

What else?

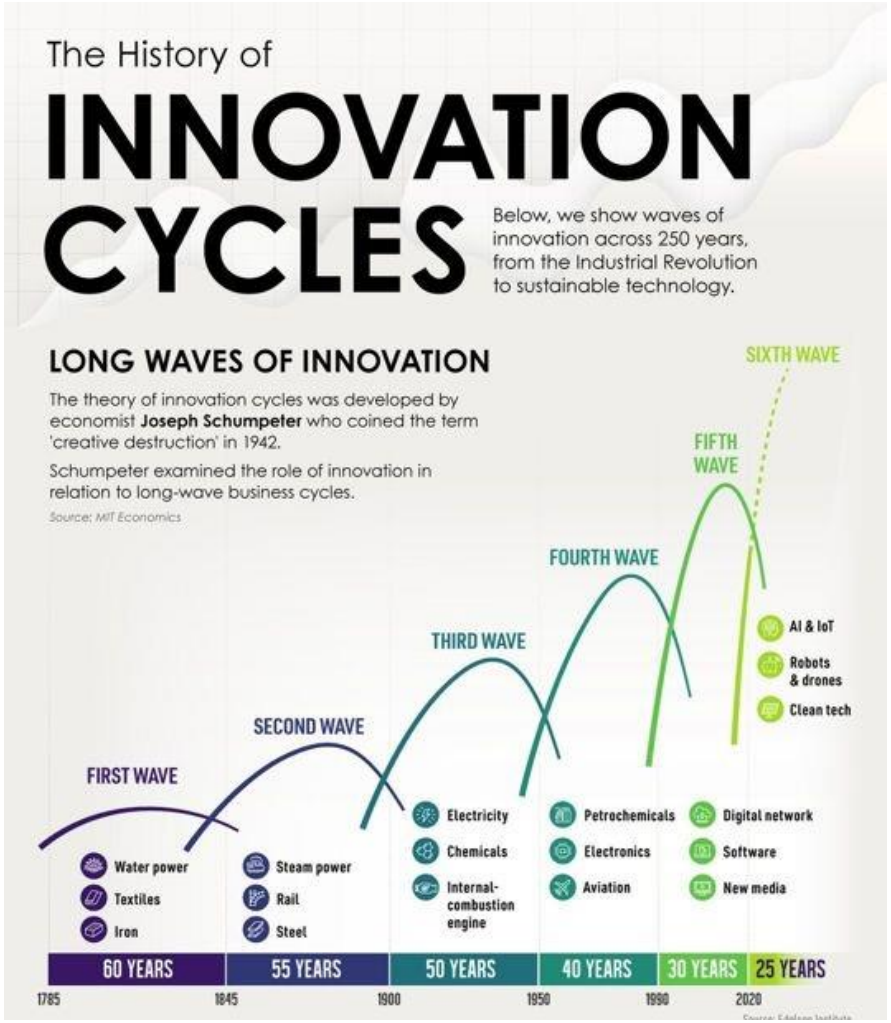
- Market sentiment (news, tweets, blogger opinions, conference calls, …)
- Satellite images from Apple store parking lots!

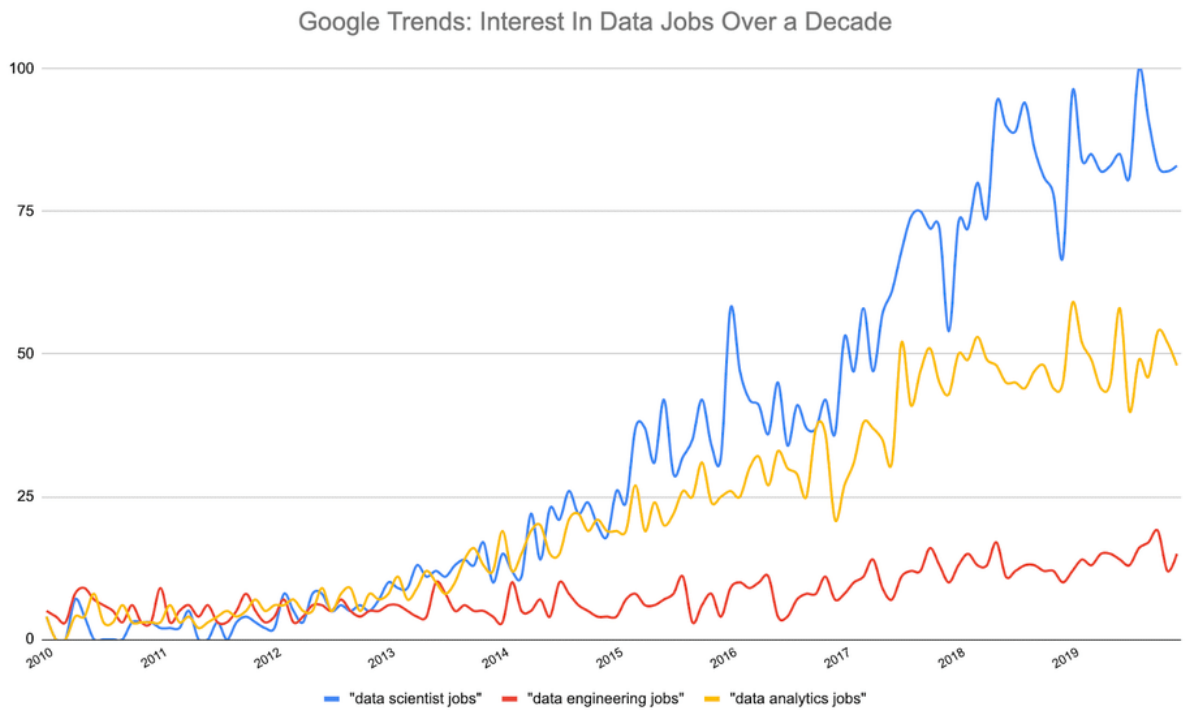
Why should I learn it?

- It’s a bid deal
- ML is closely linked to data science
- Better Career Opportunities
- Better salaries
- Hedge against next recession



Why should I learn it?





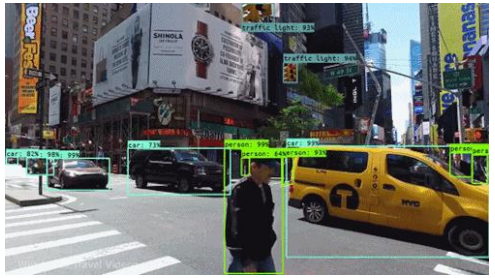
Limitations of Econometrics/Structured ML

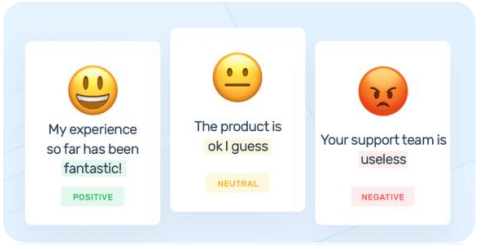
Econometrics/structured ML can only handle structured data (tabular data)!

Structured Data

	A	B	C	D
1	Date	Account	Transaction Type	Amount
2	2017-01-12	123	Credit	6089.78
3	2017-01-12	123	Fee	9.99
4	2017-01-12	456	Debit	1997
5	2017-01-12	123	Debit	20996.12
6	2017-01-13	123	Debit	17
7	2017-01-13	123	Debit	914.36
8	2017-01-14	789	Credit	11314
9	2017-01-14	789	Fee	9.99
10	2017-01-14	456	Debit	15247.89
11	2017-01-14	123	Debit	671.28
12	2017-01-15	456	Credit	5072.1
13	2017-01-15	456	Fee	9.99
14	2017-01-16	456	Debit	5109.07
15	2017-01-19	123	Credit	482.01

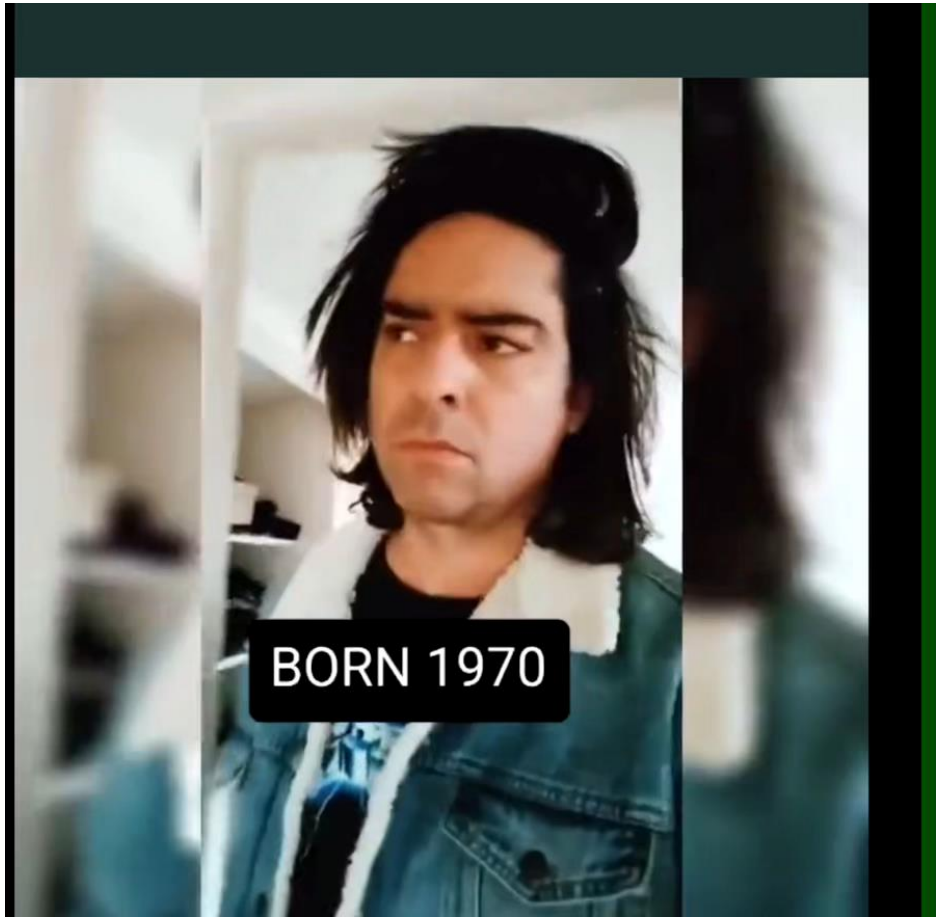
Unstructured Data (everything else!!)





You will be surprised to see what you can do when someone is counting on you!

I am counting on YOU


BORN 1970