EEP 596: Adv Intro ML || Lecture 1

Dr. Karthik Mohan

Univ. of Washington, Seattle

Jan 3, 2022

Instruction Team

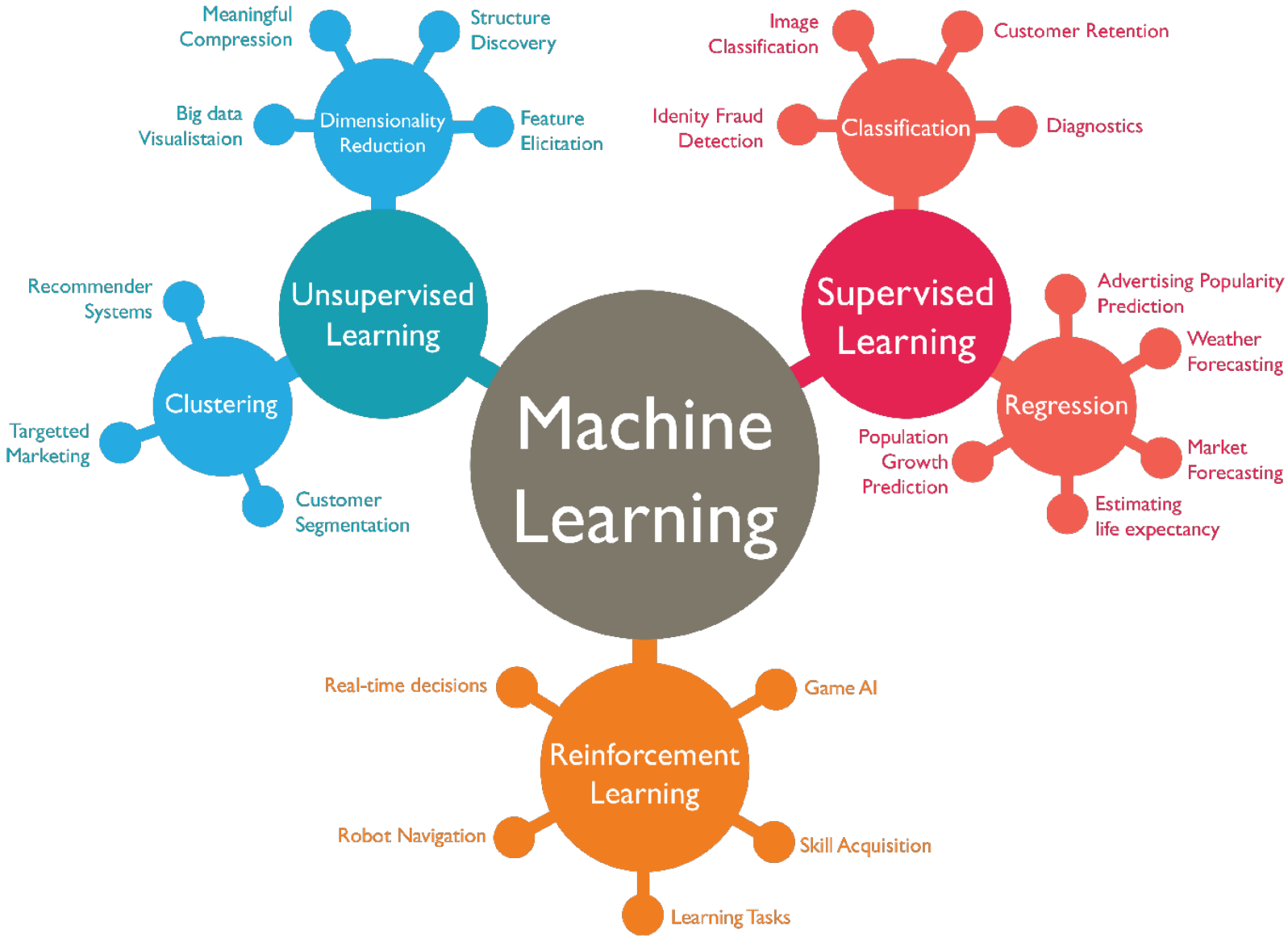
Karthik

Ayush (TA)

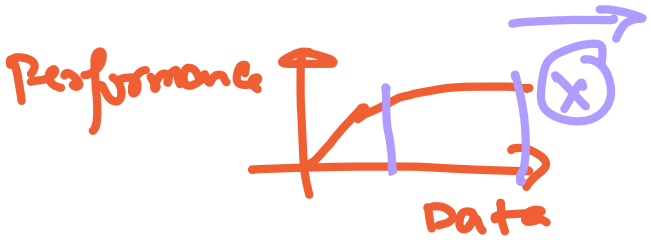
Feature (Coadv)

What is Machine Learning?

What is Machine Learning?



What is Machine Learning?



Definitions - Which ones are right?

1. Machine learning is code that improves itself with data and over time!
2. Machine learning is helping machines learn to be smarter (e.g. Tesla)
3. Machine learning relies on big data. More the data, the better the performance of the ML model.
4. Machine learning makes lives of humans easier

What is Machine Learning?

More perspectives

Have you noticed how a kid learns?

What is Machine Learning?

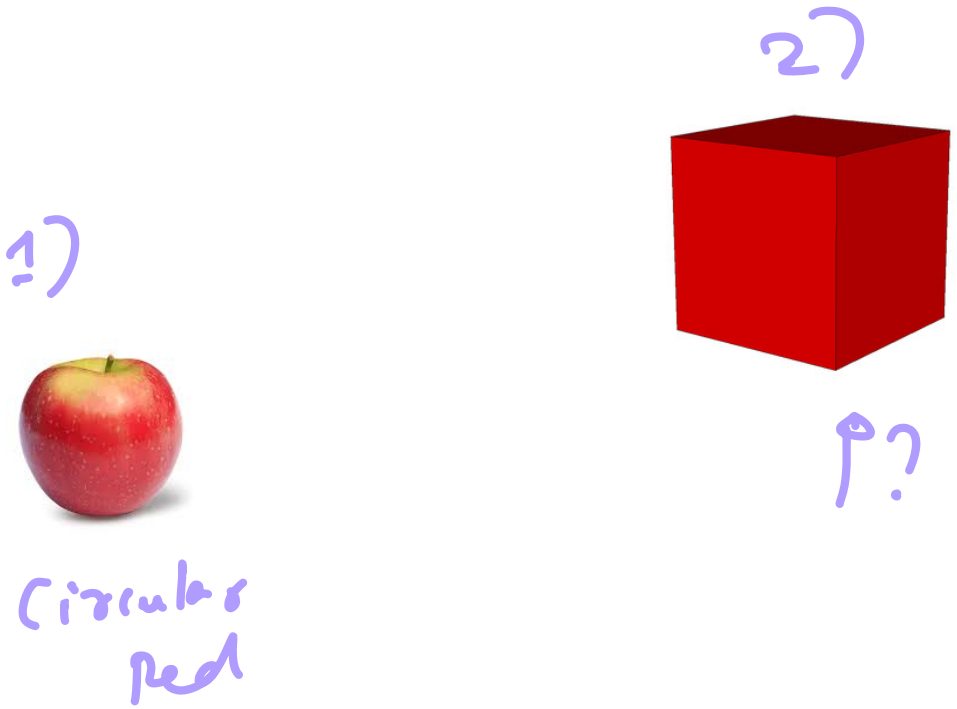
More perspectives

Have you noticed how a kid learns?

What is Machine Learning?

More perspectives

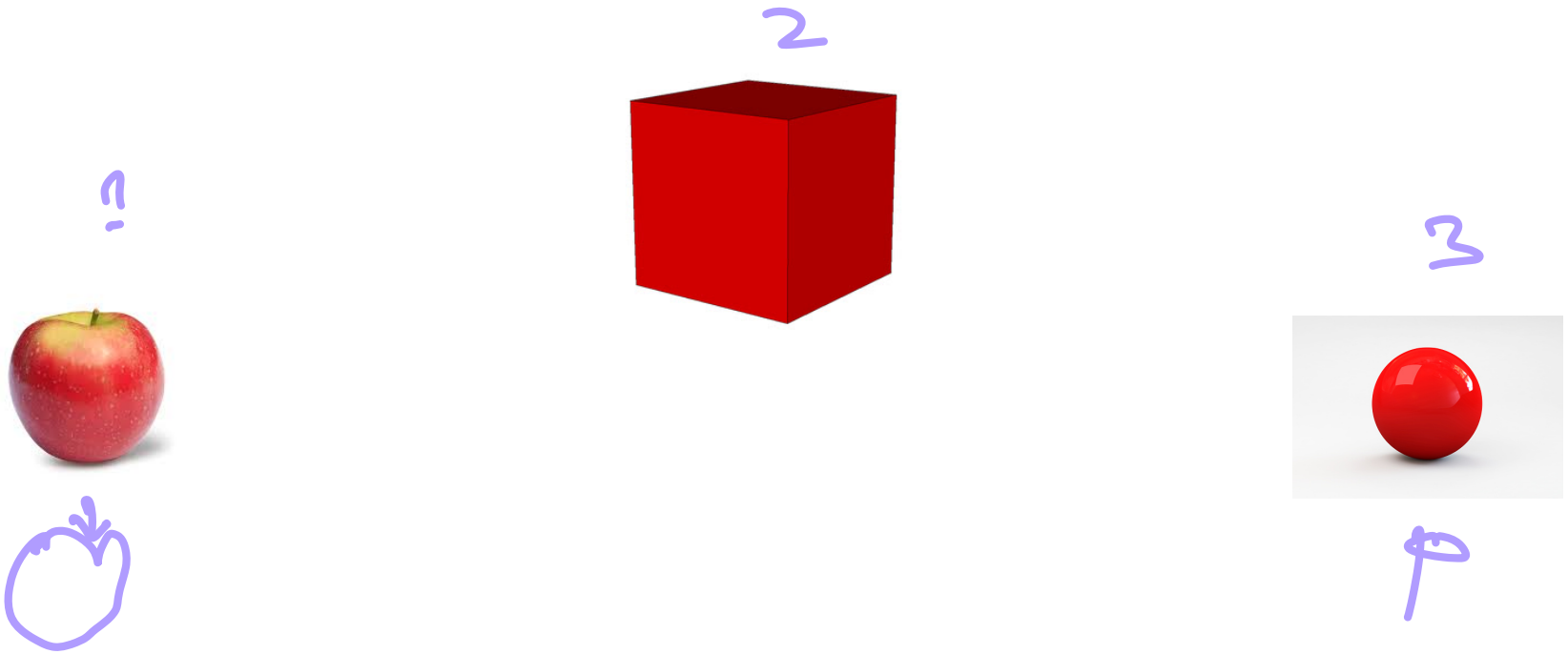
Have you noticed how a kid learns?



What is Machine Learning?

More perspectives

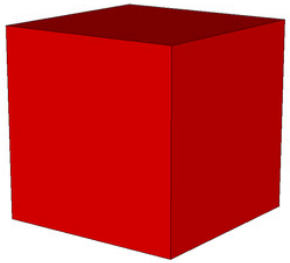
Have you noticed how a kid learns?

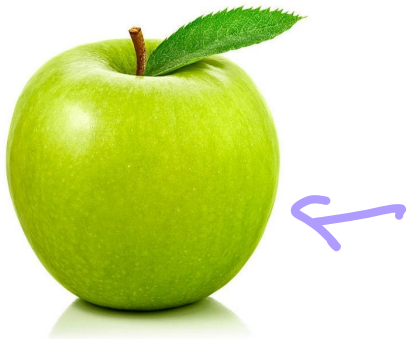


What is Machine Learning?

More perspectives

Have you noticed how a kid learns?

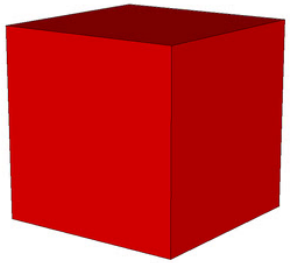




What is Machine Learning?

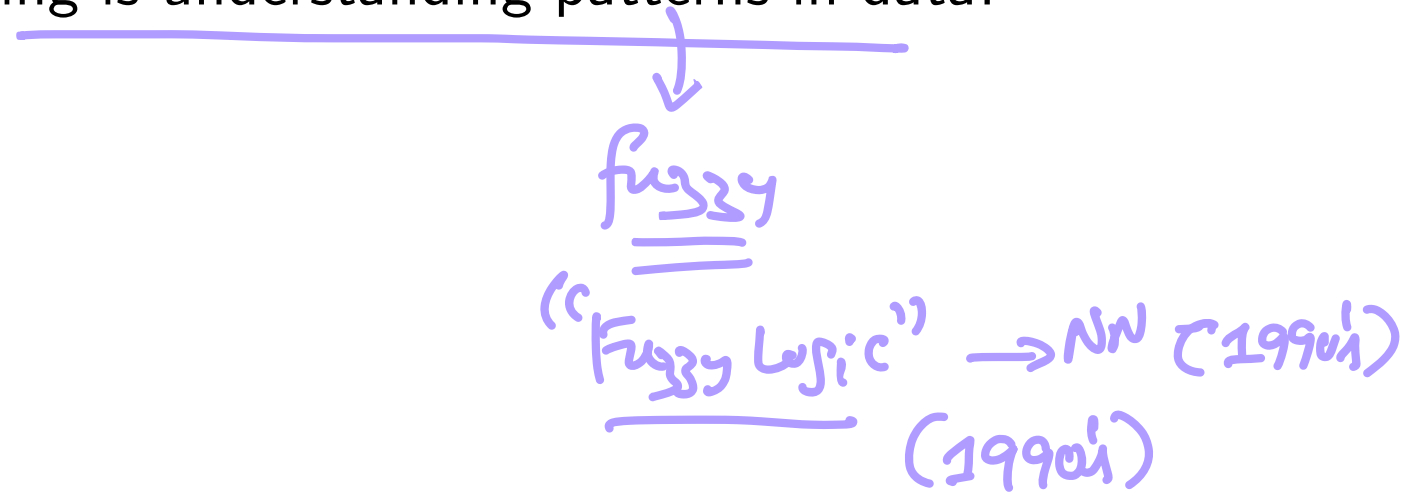
More perspectives

Have you noticed how a kid learns?



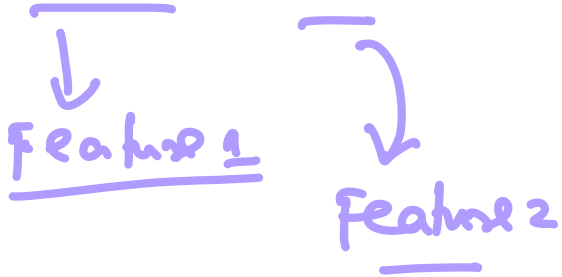
What is Machine Learning?

- Machine Learning is understanding patterns in data!



What is Machine Learning?

- Machine Learning is understanding patterns in data!
- It's knowing what combinations of features or factors in the data contribute to a decision? (e.g. shape and color for recognizing an apple)



What is Machine Learning?

- Machine Learning is understanding patterns in data!
- It's knowing what combinations of features or factors in the data contribute to a decision? (e.g. shape and color for recognizing an apple)
- Machine Learning helps you appreciate human learning! Our brains are so complex and smart - Even a simple act of driving requires tons of intelligence (some electric cars still make mistakes)!

When do you stop learning?

↳ "Learning Theory"

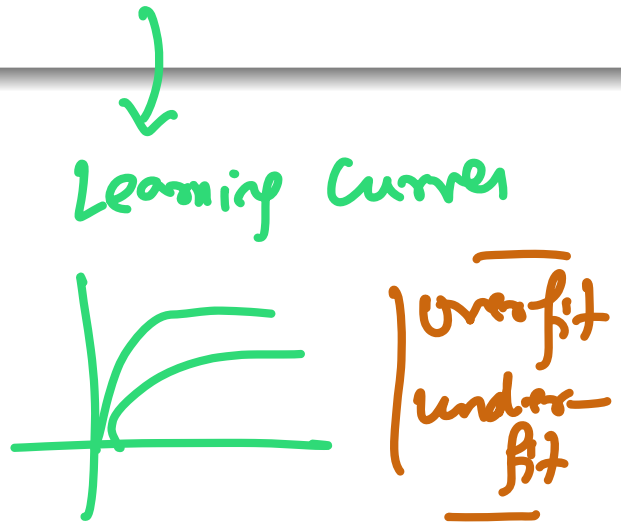
Human vs Machine

- For humans, learning doesn't stop - Isn't it?

When do you stop learning?

Human vs Machine

- For humans, learning doesn't stop - Isn't it?
- What about machines. Would you say "learning" could stop at some point in the machine learning process ? And if so, how do you check ?



When do you stop learning?

Human vs Machine

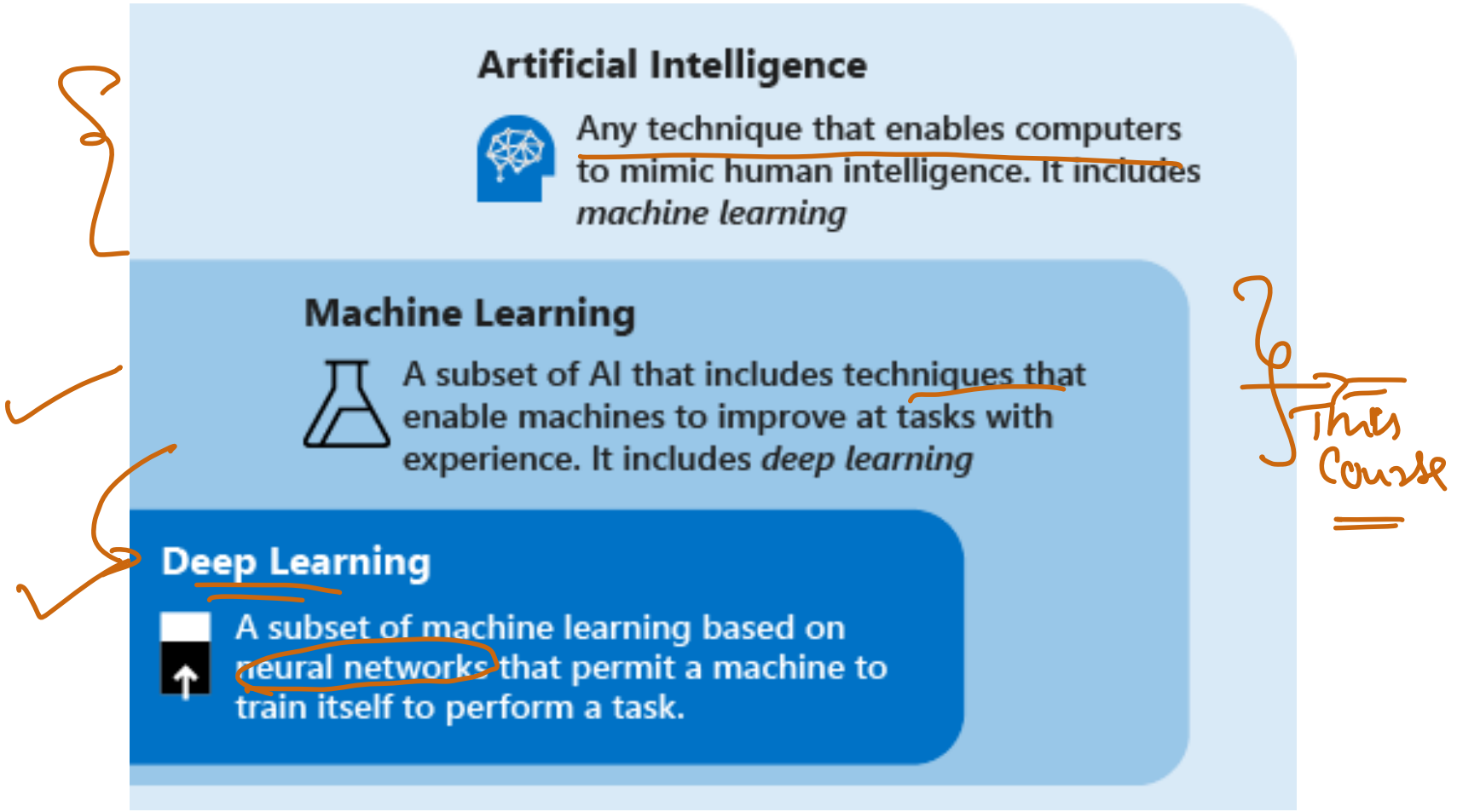
- For humans, learning doesn't stop - Isn't it?
- What about machines. Would you say "learning" could stop at some point in the machine learning process ? And if so, how do you check ?
- What exactly is "learning" in Machine Learning ?

ML vs AI: What's the difference?

ML vs AI: What's the difference?

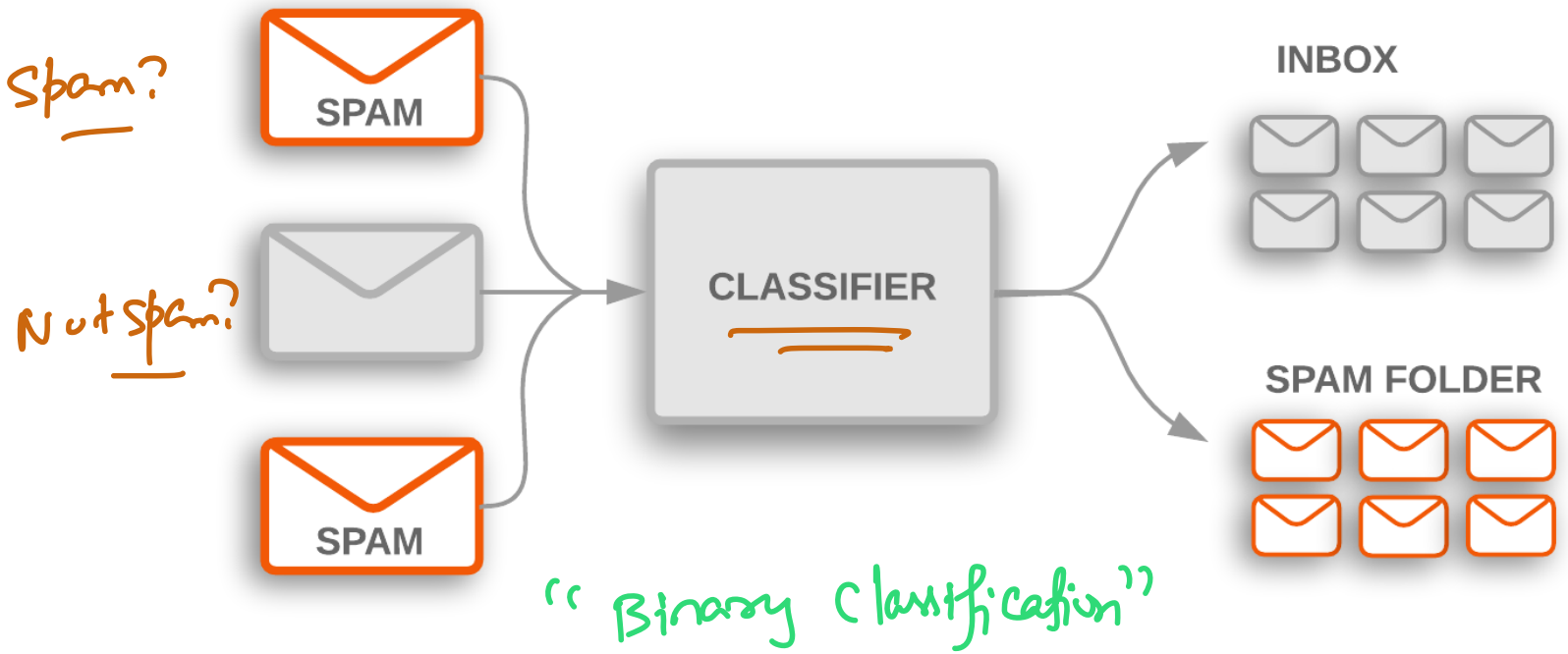
One take on this

ML vs AI: What's the difference?

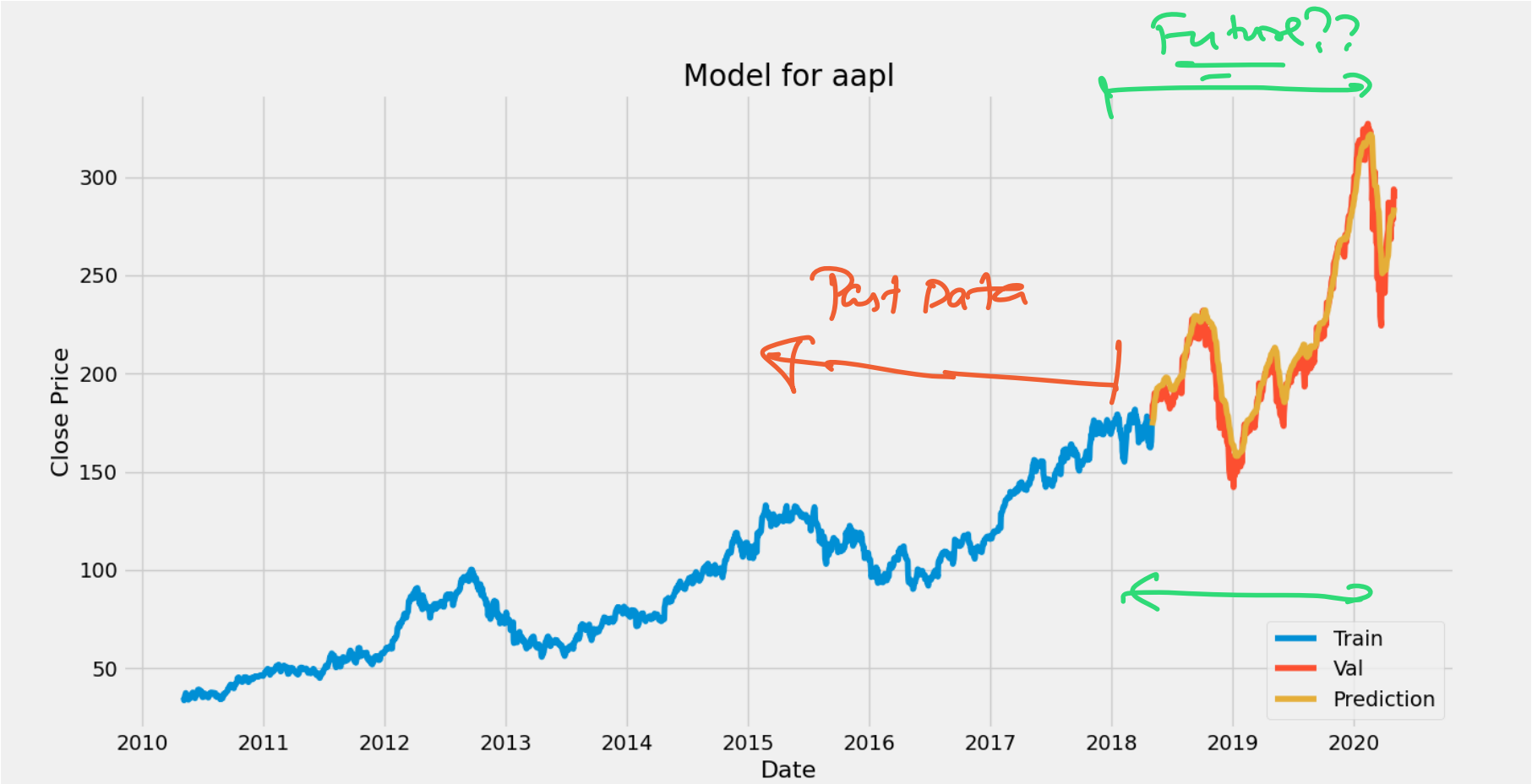


ML application: Housing price prediction

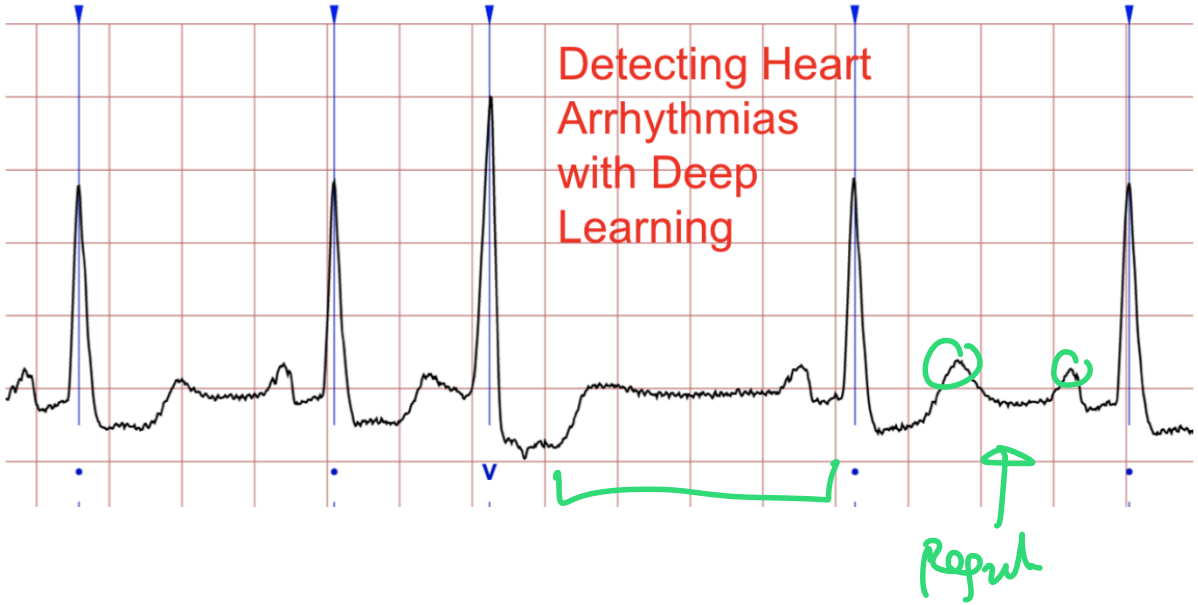
ML application: Spam detection



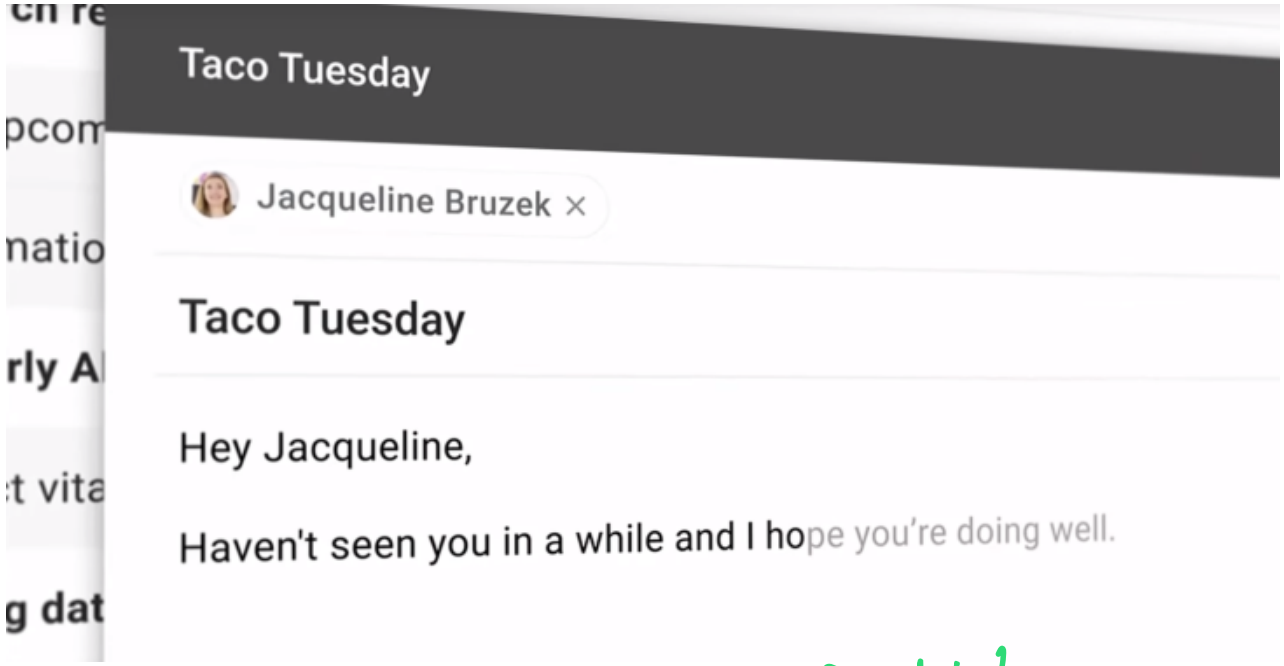
ML application: Stock Price Prediction



ML application: Arrhythmia detection



ML application: Email auto-complete



NLP → Auto Completion
→ Deep Learning Model

ML Application: Chat GPT

What's hot right now!

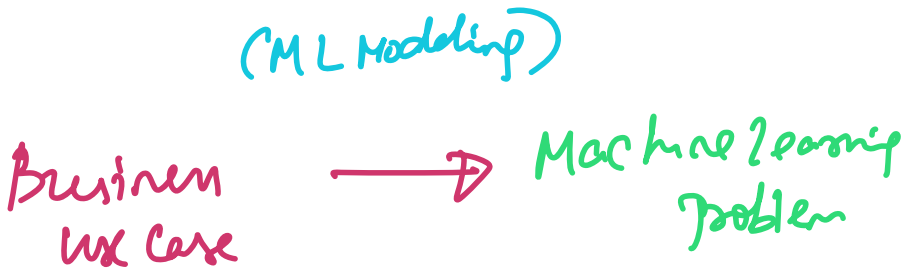
An application in the area of NLP (Natural Language Processing):
Chat GPT
A unprecedented bot that has been trained on billions of documents on the web!

Breakout #1

Discuss an application!

With your group, discuss an application of Machine Learning you have encountered in the past. Conversely, what application of Machine Learning are you excited to try out yourself either in this course or through a project?

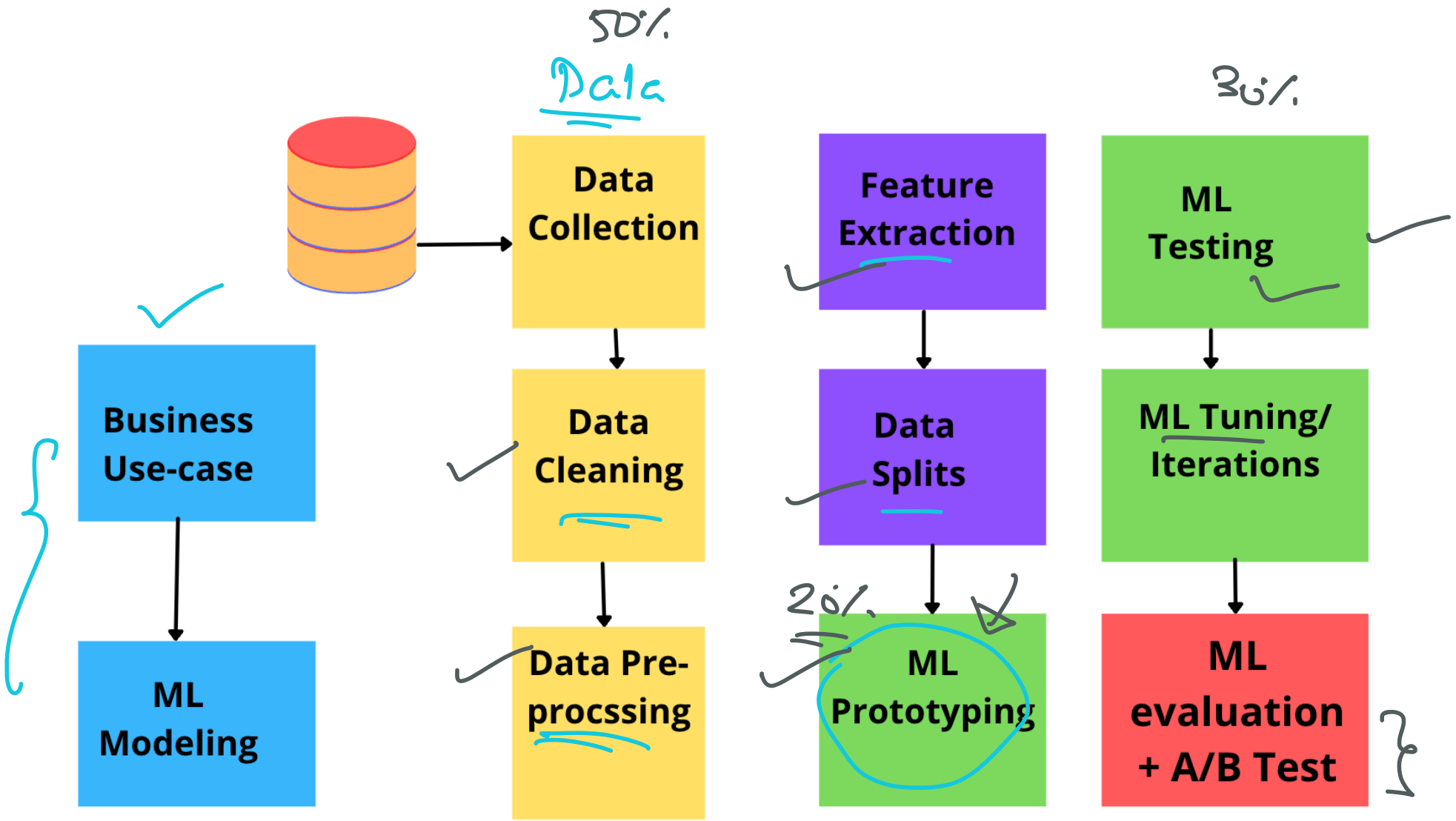
In Practice though!!



ML Modeling

No one is going to hand you a binary classification problem!!

Life of an ML scientist/ML engineer/Data Scientist!

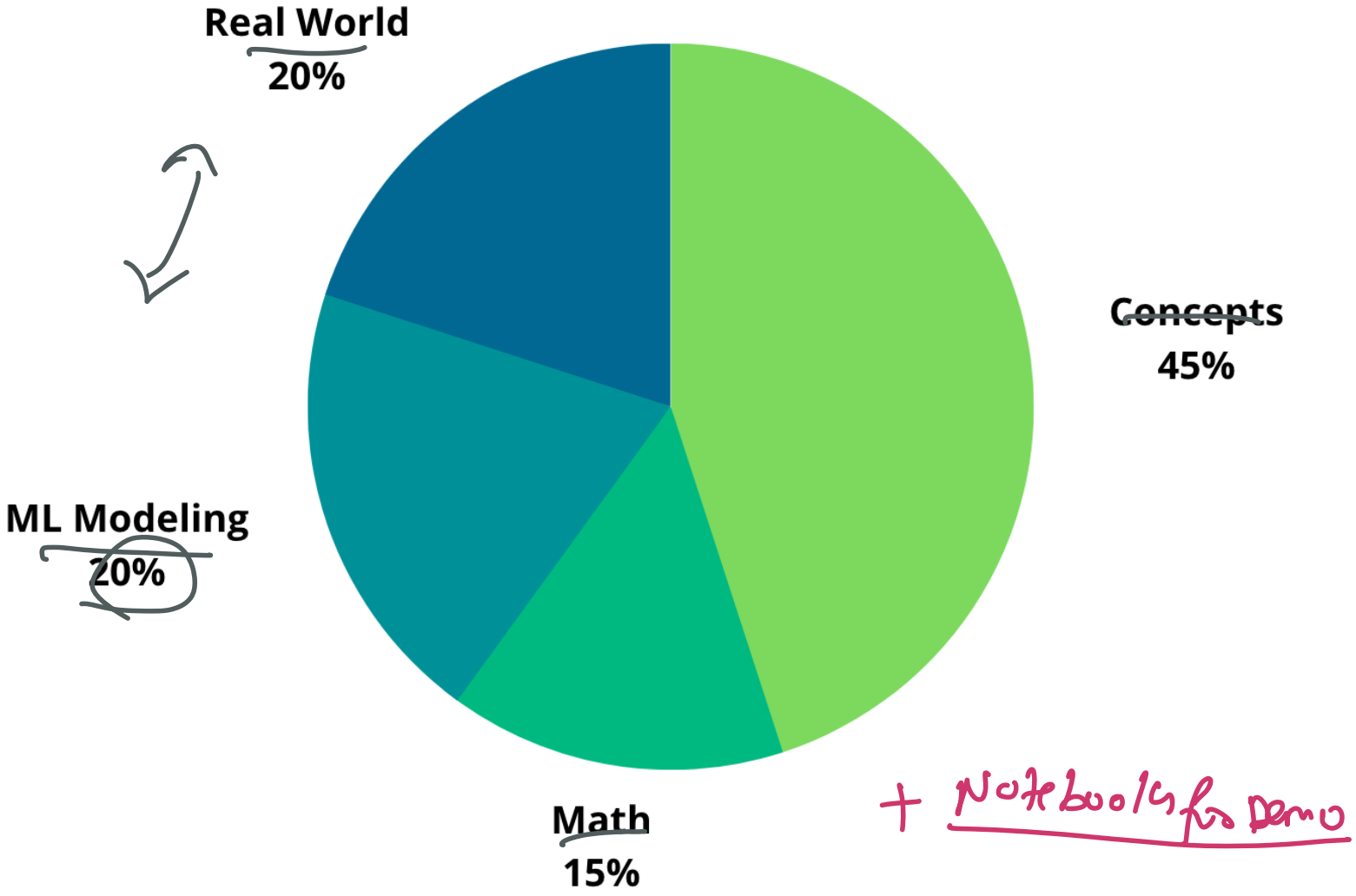


Course and Logistics

Pre-course survey results

Pre-course Survey Results

Make up of Lectures



Course Outline

Classical Machine Learning

- **Regression:** Linear, Non-linear regression, overfitting and underfitting
- **Classification:** Binary, multi-class, Naive Bayes, Logistic Regression, Random Forests
- **Unsupervised Learning:** Clustering, Anomaly Detection, Data Visualization and Dimensionality Reduction.

Course Outline

Modern and Special Topics in ML

- **Deep Learning:** Feed-forward neural networks and applications
- **NLP:** LSTM networks, BERT, Sentiment Analysis and Summarization
- **Computer Vision:** CNNs, State of the art methods

ML Foundations

Dimensions touched upon in this course

- Fundamentals of Machine Learning

ML Foundations

Dimensions touched upon in this course

- Fundamentals of Machine Learning
- Methodology behind Machine Learning

ML Foundations

Dimensions touched upon in this course

- Fundamentals of Machine Learning
- Methodology behind Machine Learning
- Mechanics behind Machine Learning

ML Foundations

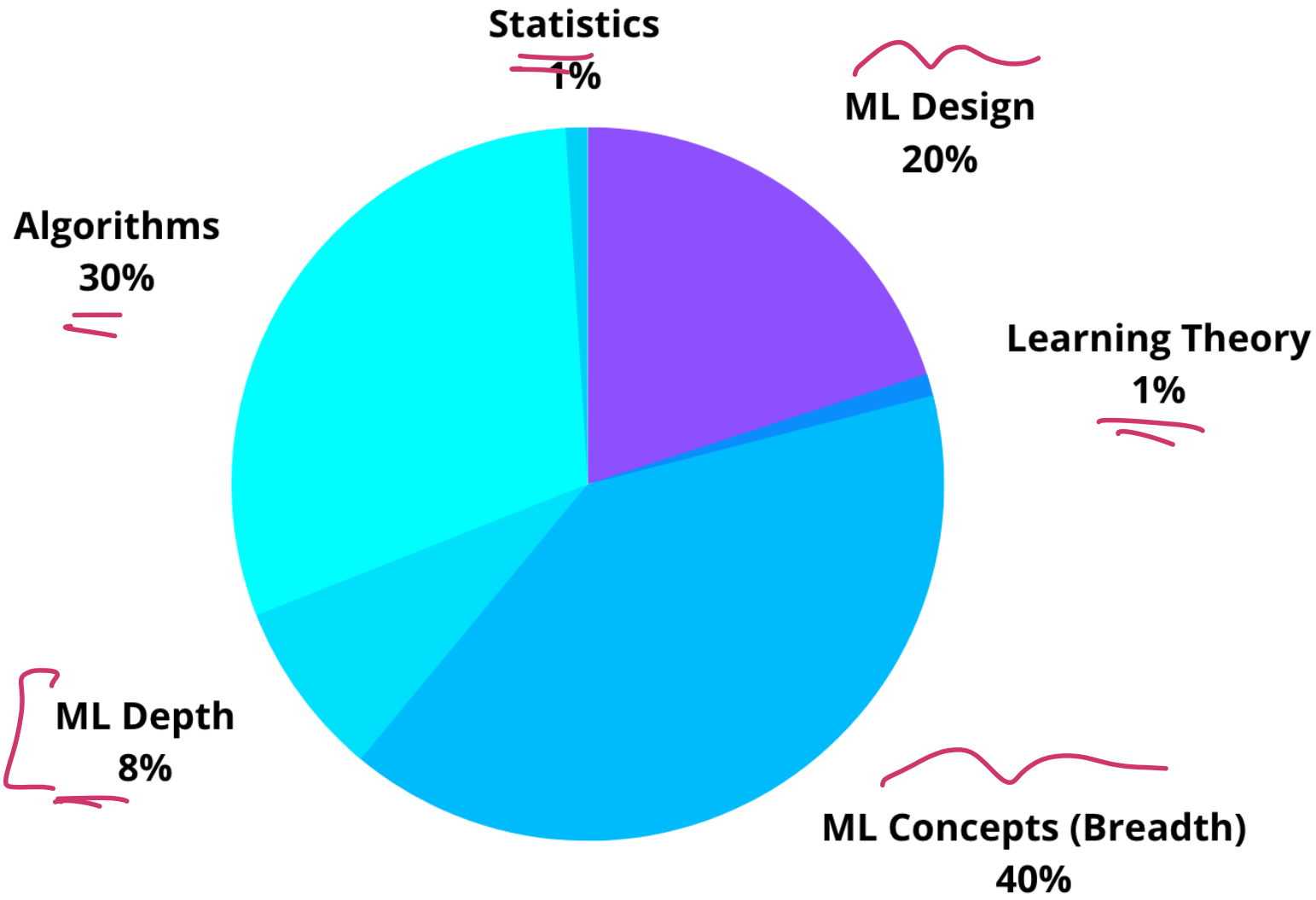
Dimensions touched upon in this course

- Fundamentals of Machine Learning
- Methodology behind Machine Learning
- Mechanics behind Machine Learning
- Both Classical and Modern Machine Learning

Lectures and Programming Assignments (Tentatively)

Week	Lecture Material	Assignment
1	Linear Regression	Housing Price Prediction
2	Classification	Spam classification (Kaggle)
3	Classification	Flower/Leaf classification
4	Clustering	MNIST digits clustering
5	Anomaly Detection	Stock price Prediction (Kaggle + P)
6	Data Visualization	Stock price Prediction (Kaggle + P)
7	Deep Learning	Visualizing 1000 images
8	Deep Learning (DL)	ECG Arrythmia Detection
9	DL in NLP	TwitterSentiment Analysis (Kaggle + P)
10	DLs in Vision	TwitterSentiment Analysis (Kaggle + P)

ML breakdown in our course



Textbook(s)

Classic ML

- Too many to name! But one good one I recommend is by Christopher Bishop.
- Note that we will not be following a textbook as such!
- However, Lecture notes will be posted for each Lecture and will serve as a reference to go through

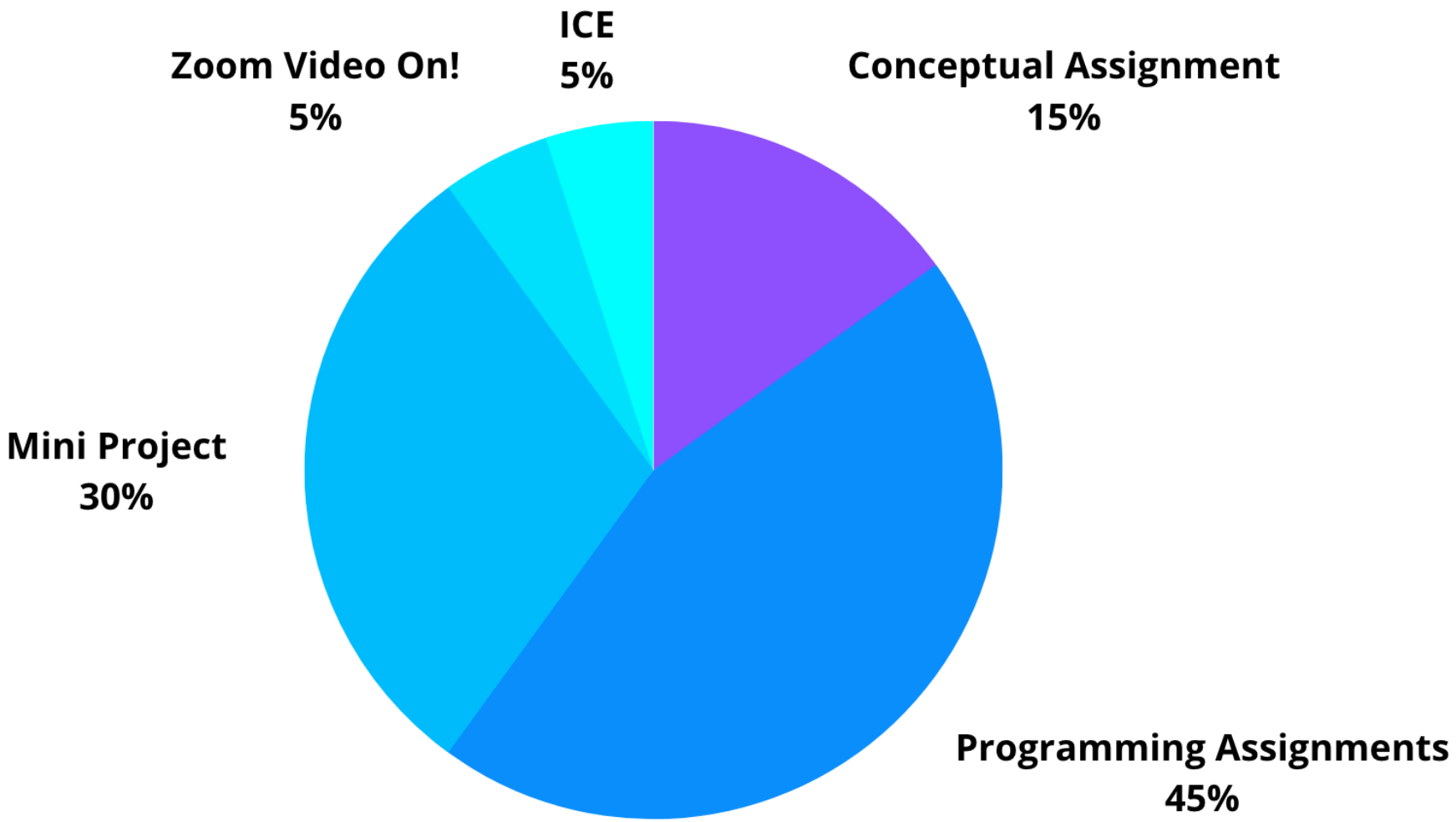
Deep Learning

Deep Learning by Yoshua Bengio et al

Assessments

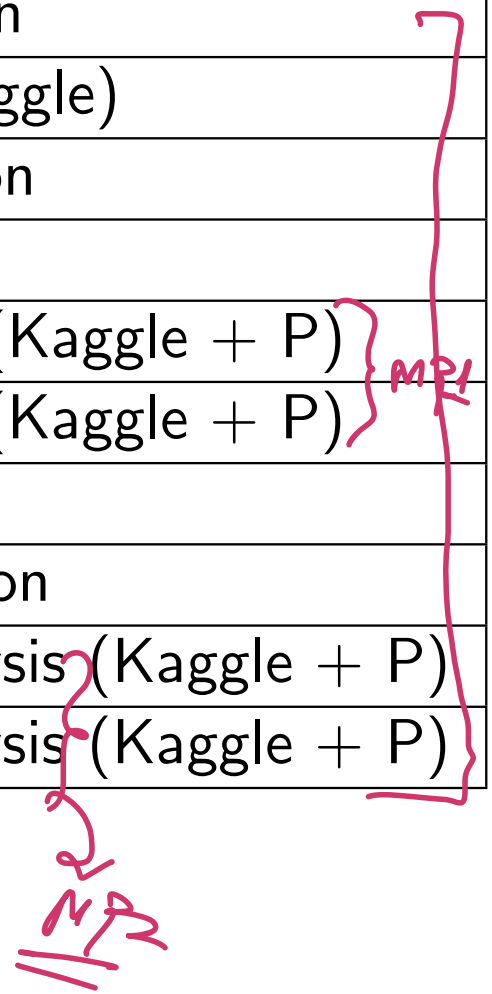
	Assigned	Due	Grade Percentage
In-class	In-Lect	In-Lect	5%
Conceptual	Wed	Next Tue	15%
Programming	Thu	Next Wed	45%
Mini-Projects	-	-	30%
Zoom Attendance			5%

Assessments



Lectures and Programming Assignments (Tentatively)

Week	Lecture Material	Assignment
1	Linear Regression	Housing Price Prediction
2	Classification	Spam classification (Kaggle)
3	Classification	Flower/Leaf classification
4	Clustering	MNIST digits clustering
5	Anomaly Detection	Stock Price Prediction (Kaggle + P)
6	Data Visualization	Stock Price Prediction (Kaggle + P)
7	Deep Learning	Visualizing 1000 images
8	Deep Learning (DL)	ECG Arrythmia Detection
9	DL in NLP	TwitterSentiment Analysis (Kaggle + P)
10	DLs in Vision	TwitterSentiment Analysis (Kaggle + P)



Coding pointers

- Assignments assume python as the main language (e.g. for hints and modules, etc)
- Coding environment set-up will be one of the problems on HW 1
- Prototyping can be done on notebooks and submitted as such for smaller assignments.
- For mini-projects and kaggle assignments - Please keep your code modular and organized.

Coding Environment

- Pointers below if you want to get set up on Google Colab for both prototyping, running machine-intensive ML experiments and working with code through IDEs
- Prototype Coding work in Notebooks recommended on Google Colab
- For terminal access on Google Colab, sign up for pro
- `pip3 install colabcode` on termainal
- ColabCode enables you to have a VSCode IDE port into Google Colab - So you can work on the IDE from your laptop but run experiments on Google Colab!

Maximizing Your Learning of Machine Learning!

- Ask questions during lectures - Clarify things as they happen!

Maximizing Your Learning of Machine Learning!

- Ask questions during lectures - Clarify things as they happen!
- Make use of office hours and quiz section!

Maximizing Your Learning of Machine Learning!

- Ask questions during lectures - Clarify things as they happen!
- Make use of office hours and quiz section!
- Collaborative learning - Discord is a great place to brainstorm concepts outside class and unblock yourself.

Maximizing Your Learning of Machine Learning!

- Ask questions during lectures - Clarify things as they happen!
- Make use of office hours and quiz section!
- Collaborative learning - Discord is a great place to brainstorm concepts outside class and unblock yourself.
- 30% of your learning happens in class and office hours - The remaining 70% happen when you work on the assignments. (You ofcourse need the 30 to get to the 70 :D)

Maximizing Your Learning of Machine Learning!

- Ask questions during lectures - Clarify things as they happen!
- Make use of office hours and quiz section!
- Collaborative learning - Discord is a great place to brainstorm concepts outside class and unblock yourself.
- 30% of your learning happens in class and office hours - The remaining 70% happen when you work on the assignments. (You ofcourse need the 30 to get to the 70 :D)
- What you put in is what you get out!

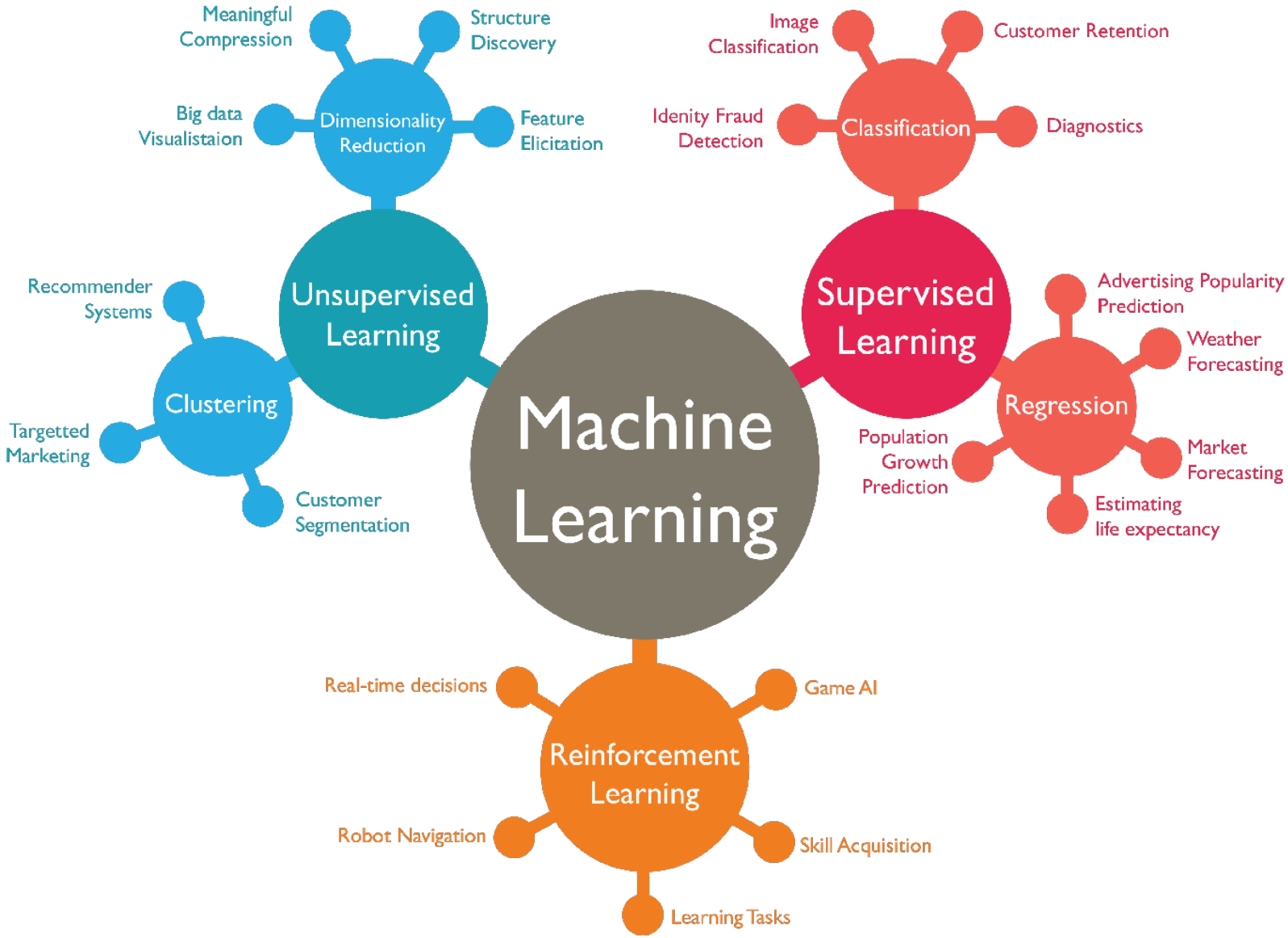
Maximizing Your Learning of Machine Learning!

- Ask questions during lectures - Clarify things as they happen!
- Make use of office hours and quiz section!
- Collaborative learning - Discord is a great place to brainstorm concepts outside class and unblock yourself.
- 30% of your learning happens in class and office hours - The remaining 70% happen when you work on the assignments. (You ofcourse need the 30 to get to the 70 :D)
- What you put in is what you get out!
- Excitement + Smart work + Inquisitiveness = Maximized learning!

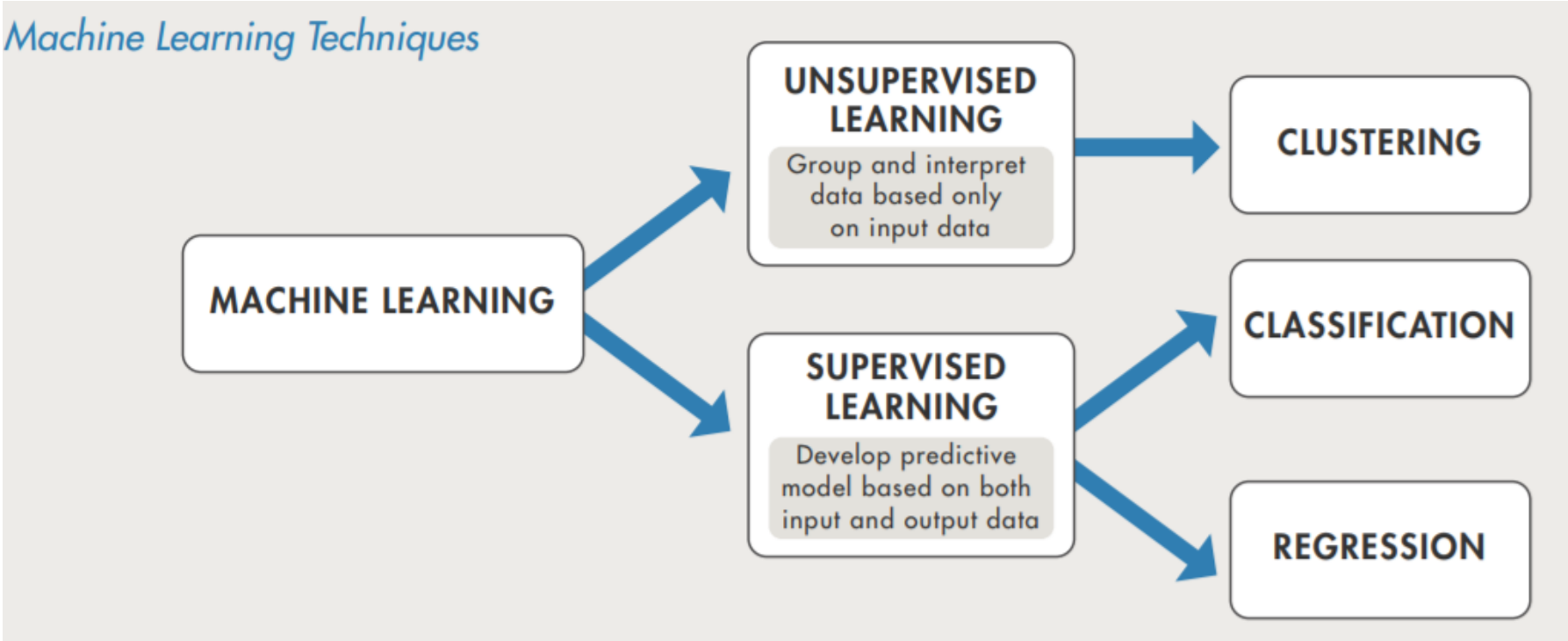
Weekly Logistics

	Day	Timings	Class type
Lecture 1 (In-person)	T	4 pm - 6 pm	In-person
Lecture 2	Th	4 pm - 6 pm	Zoom
Office Hours Karthik	T	6 - 6:30 pm	In-person
Office Hours Ayush	TBD	TBD	Zoom
Quiz Section Ayush	TBD	TBD	Zoom

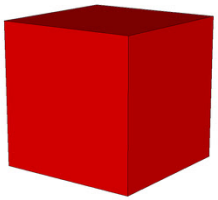
What is Machine Learning?



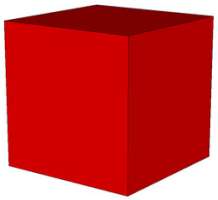
Supervised vs Unsupervised Learning



Supervised Learning



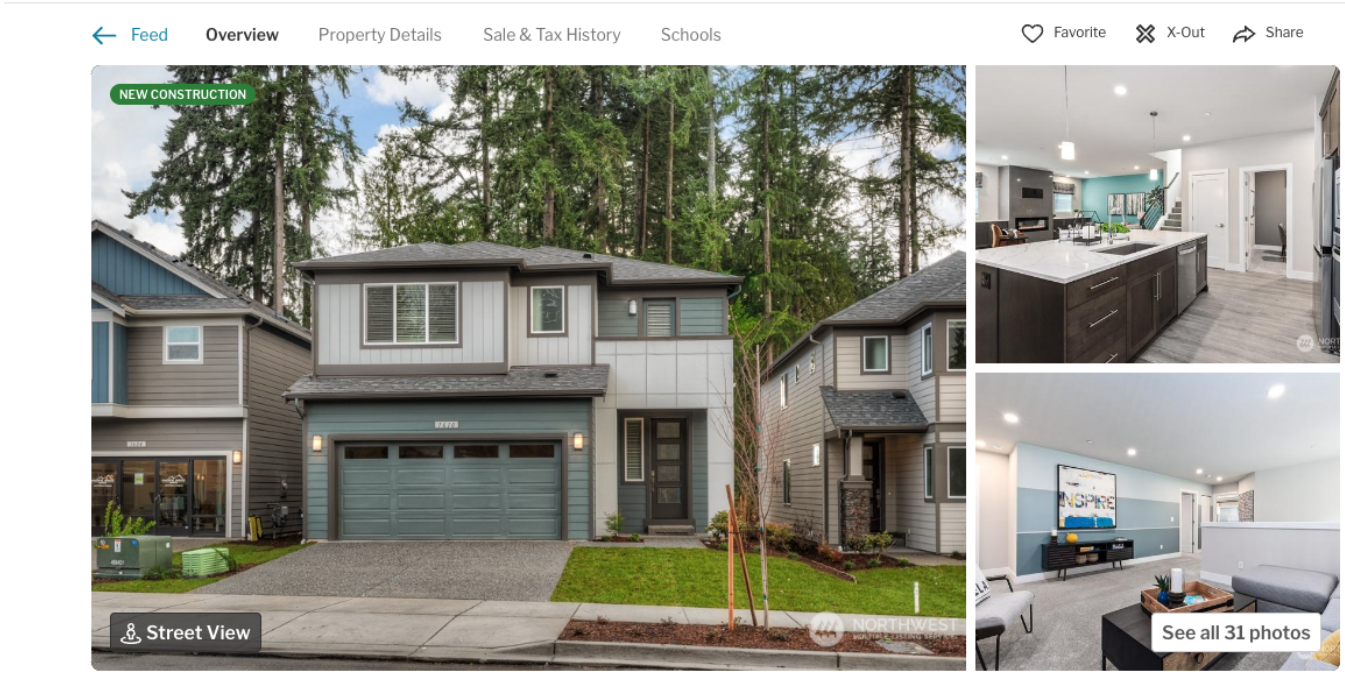
Un-Supervised Learning



Our first ML method: Linear Regression

Application: Housing Prices

Redfin



Listed by Mari Riksheim · Pacific Ridge - DRH, LLC.

17817 2nd Ave W Unit IW-42, Bothell, WA 98012

$1,134,995
Est. $7,420/mo Get pre-approved

5 Beds

3 Baths

2,703 Sq Ft

This home is popular
It's been viewed 2,022 times. Tour it in person or via video chat before it's gone!

Today: 6:00 pm · 7:00 pm · 8:00 pm · More times

About This Home

Pacific Ridge presents Ironwood! Gorgeous new home community centrally located between Bothell, Mill Creek & Lynnwood. Perched just off North Road with panoramic views to the East, this neighborhood offers a quiet place to call home with community parks & convenient access to

Go tour this home

TUESDAY 3 JAN | WEDNESDAY 4 JAN | THURSDAY 5 JAN

Tour in person | Tour via video chat

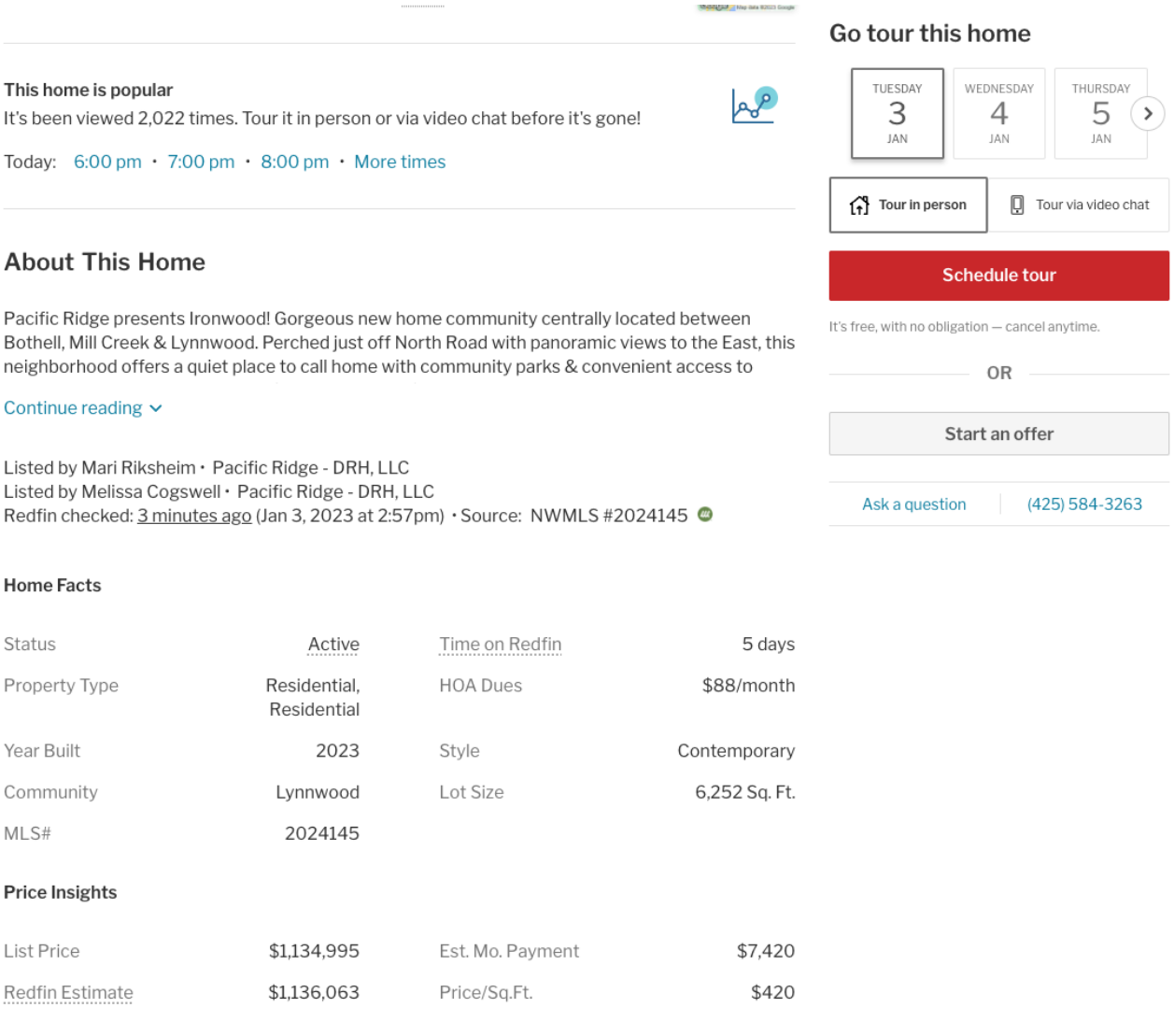
Schedule tour

It's free, with no obligation — cancel anytime.

OR

Start an offer

Redfin Estimate



This home is popular
It's been viewed 2,022 times. Tour it in person or via video chat before it's gone!

Today: 6:00 pm · 7:00 pm · 8:00 pm · More times

About This Home

Pacific Ridge presents Ironwood! Gorgeous new home community centrally located between Bothell, Mill Creek & Lynnwood. Perched just off North Road with panoramic views to the East, this neighborhood offers a quiet place to call home with community parks & convenient access to

Continue reading

Listed by Mari Riksheim · Pacific Ridge - DRH, LLC
Listed by Melissa Cogswell · Pacific Ridge - DRH, LLC
Redfin checked: 3 minutes ago (Jan 3, 2023 at 2:57pm) · Source: NWMLS #2024145

Home Facts

Status	Active	Time on Redfin	5 days
Property Type	Residential, Residential	HOA Dues	$88/month
Year Built	2023	Style	Contemporary
Community	Lynnwood	Lot Size	6,252 Sq. Ft.
MLS#	2024145		

Price Insights

List Price	$1,134,995	Est. Mo. Payment	$7,420
Redfin Estimate	$1,136,063	Price/Sq.Ft.	$420

Go tour this home

TUESDAY 3 JAN | WEDNESDAY 4 JAN | THURSDAY 5 JAN

Tour in person | Tour via video chat

Schedule tour

It's free, with no obligation — cancel anytime.

OR

Start an offer

Ask a question | (425) 584-3263

Breakout #2

Zillow Estimate/RedFin Estimate

If you are on the market to buy a house, you would perhaps be looking at "Zestimates" or "RedFin Estimates" to filter out houses in your budget range. Discuss in your group, what are the factors that influence the price of a home and what are the factors (also called features in ML) that may have been used to construct these estimates. Once you have a set of factors identified, how do you combine them to produce the final house price estimate?

Next Lecture

1. Linear Regression Applications
2. Linear Regression Models
3. Under-fitting and Over-fitting in ML