Machine Learning

AIL7024

Tanmoy Chakraborty

IIT Delhi, India

https://tanmoychak.com/

Logistics

- Course Instructor: Tanmoy Chakraborty
  https://tanmoychak.com/

- Guest Lecture: TBD (possibly from the industry)
- TAs: Sahil Mishra, Sudipto Ghosh

- Course page: https://lcs2.in/ml2501
- Discussion forum: Piazza (https://piazza.com/iit_delhi/fall2025/ail7024, Code: ail7024)
- For assignment submission: Moodle

Books

- “Machine Learning” by Tom M. Mitchell
- "Pattern Recognition and Machine Learning" by Christopher M. Bishop
- "Machine Learning: A Probabilistic Perspective" by Kevin P. Murphy
- "The Elements of Statistical Learning" by Hastie, Tibshirani, and Friedman

Prerequisite

- Basic computer science principles
  - Big-O notation
  - Comfortably write non-trivial code in Python/numpy
- Probability
  - Random Variables
  - Expectations
  - Distributions
- Linear Algebra & Multivariate/Matrix Calculus
  - Matrix algebra (*)
  - Gradients and Hessians
  - Eigenvalue/vector

Tentative Syllabus

- Introduction
- MLE & MAP
- Regression
- Decision Trees
- Support Vector Machine
- Kernels
- Bias, Variance and ensemble learning
- Instance based learning
- Gaussian Discriminant Analysis and Naive Bayes
- Artificial Neural Networks
- Introduction to Deep Learning
- RNN, Backpropagation, CNN
- Transformer
- Conclusion

More about this course

- A lot of maths and derivations will be covered
- Less slides will be used to deliver the lecture
- More board work
- Introductory neural networks will be covered

Journals/Conferences

- **Conferences**
  - NeurIPS: Neural Information Processing Systems
  - ICML: International Conference on Machine Learning
  - ICLR: International Conference on Learning Representation
  - AAAI: Association for the Advancement of Artificial Intelligence
  - IJCAI: International Joint Conference on AI
  - SIGKDD: Special Interest Group on Knowledge Discovery and Data Mining
  - CVPR: Computer Vision and Pattern Recognition

- **Journals**
  - IEEE Transactions on Pattern Analysis and Machine Intelligence
  - IEEE Trans. on Neural Networks and Learning Systems
  - Journal of Machine Learning Research (JMLR)

Course Directives

- **Class Time:** Mon and Thu, 8-9:30
- **Office Hour:** as per requirement (email me to schedule an appointment)
- **TA Hour:** as per requirement (email them to schedule an appointment)
- **Room:** Please check the timetable
- **Marks distribution (tentative):**
  - Midsem: 20%
  - Endsem: 30%
  - Quiz (2/3) : 20%
  - Assignment (1): 10%
  - Assignment-2/Project: 20%

- **Audit**: B- (discouraged!!)
- **Grading Scheme**: Relative
- **Attendance:** 75% to appear for the exam

Course Project (20%)

- A fresh idea that leads to a full-fledged system
- Each group should consist of (max) 2 students
- Students are encouraged to propose their own project ideas by **Aug 15**
  - **Send your ideas to the TA and me**
- **Best Project Award**
- You need to
  - gather data
  - develop models
  - evaluate your models
  - prepare presentation
  - write tech report

Students are encouraged to publish their projects in good conferences/journals

DO NOT PLAGIARIZE !

Academic Integrity is of utmost importance. If anyone is found cheating/plagiarizing, it will result in negative penalty (and possibly even more: an F grade or even DisCo).

Collaborate. But do NOT cheat.

- Assignments to be

If detected, one grade will be lowered down

- Do not share any part of code.
- Do not copy any part of report from any online resources or published works.
- If you reuse other's works, always cite.
- If you discuss with others about assignment or outside your group for project, mention their names in the report.
- Do not use GenAI tools (like ChatGPT).

We will check for pairwise plagiarism in submitted assignment code files among you all.

We will also check the probability of any submitted content being AI generated.

Project reports will be checked for plagiarism across all web resources.

https://bsw.iitd.ac.in/faqs.php#0

Acknowledgment

The course and slides are inspired by the following:

- http://cs229.stanford.edu/
- http://www.cs.cmu.edu/~ninamf/courses/601sp15/index.html
- https://dspace.mit.edu/handle/1721.1/46320
- https://nipunbatra.github.io/ml2020/lectures/
- https://sites.google.com/a/iiitd.ac.in/ml-cse-343-543/
- https://cs.ccsu.edu/~markov/ccsu_courses/MachineLearning.html
- https://nptel.ac.in/courses/106/105/106105152/

and many blogs, online articles, scholarly papers, lecture notes, etc. available on the web.

What is ML?

- Term “Machine Learning” coined by Arthur Samuel in 1959.
  - Samuel Checkers-playing Program was among the world's first successful self-learning programs

What is ML?

- Term “Machine Learning” coined by Arthur Samuel in 1959.
  - Samuel Checkers-playing Program was among the world's first successful self-learning programs

“Field of study that give computers the ability to learn without being explicitly programmed” - Arthur Samuel [1959]

What is ML?

Study of algorithms that

- improve their performance P
- at some task T
- with experience E

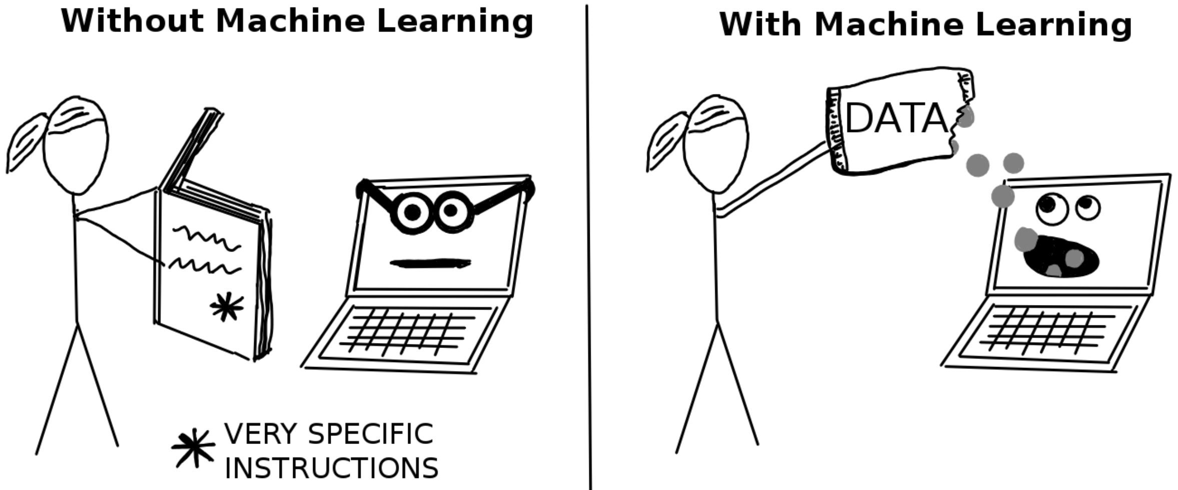
well-defined learning task: <P,T,E>

Tom Mitchell, 1998

What is ML?

- **Machine Learning** is a set of methods that allow computers to learn from data to make and improve predictions.

- **Machine learning** is a paradigm shift from "normal programming" where all instructions must be explicitly given to the computer to "indirect programming" that takes place through providing data.



https://christophm.github.io/interpretable-ml-book/terminology.html

What is ML?

- ML has three core components – Model, Optimization and data
- The model is a function from i/p to o/p, but it is not programmed by hand
  - Instead the model has **parameters** that will be optimized (learned)
- The optimization algorithm finds **good parameters** – more on this later
- Roughly speaking, parameters are good if they are a good fit for the data
- The data typically, but not always, consist of input-output examples
- The goal is for the model to **generalize** to the new data that it was not fit on

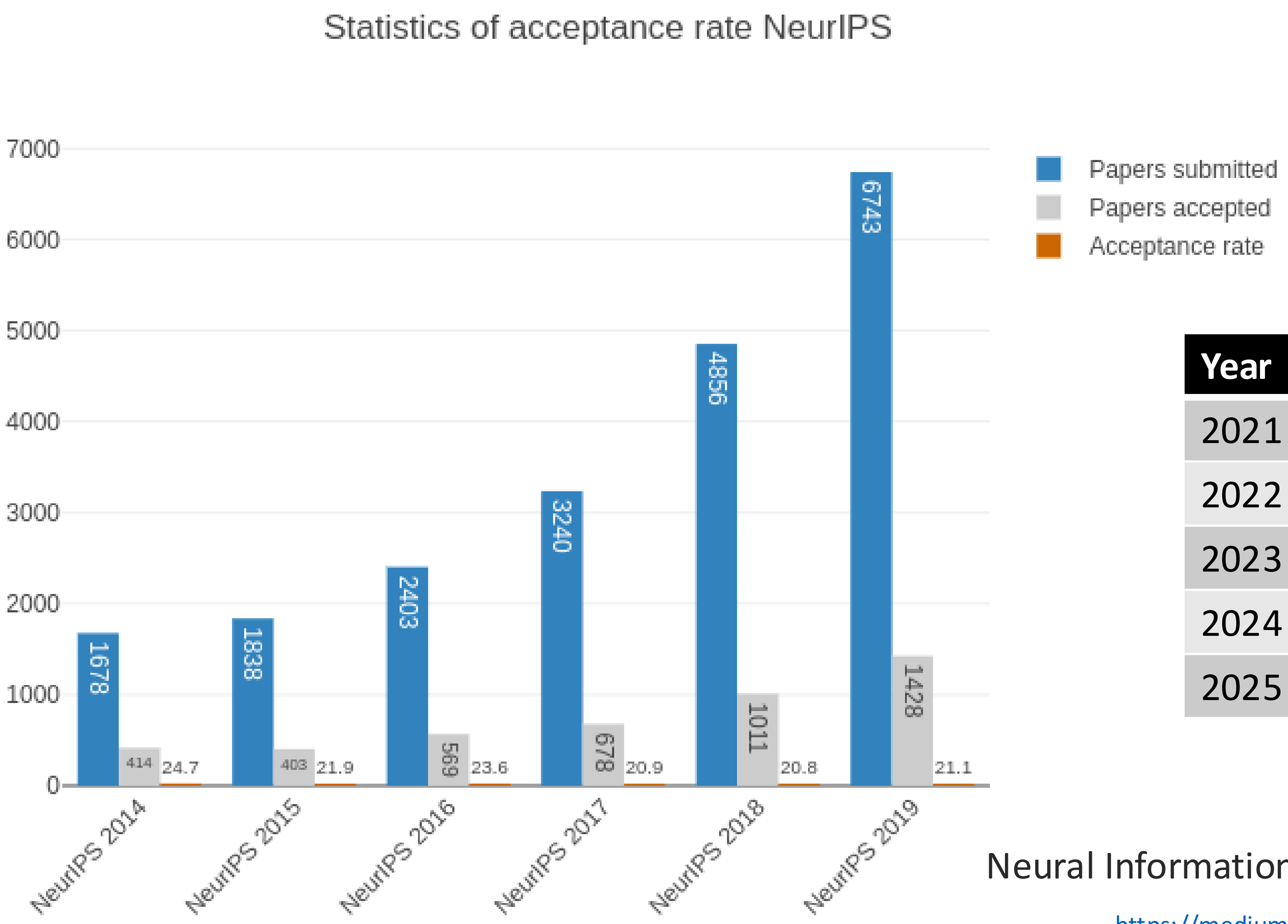
History

- 1950s
  - Samuel’s checker-playing program
- 1960s
  - Neural network: Rosenblatt’s perceptron
  - Pattern recognition
  - Minsky proved the limitations of perceptron
- 1970s:
  - Symbolic concept induction
  - Expert systems and knowledge acquisition bottleneck
  - Quinlan’s ID3
  - NLP (symbolic)

History

- 1980s
  - Adv. Decision tree and rule learning
  - Learning and planning for problem solving
  - Resurgence of NN
  - PAC learning
  - Focus on experimental methodology
- 1990s
  - SVM
  - Data Mining
  - Adoptive agents and web applications
  - Text learning
  - RL
  - Ensembles
  - Byes Net

Evidence that ML is booming

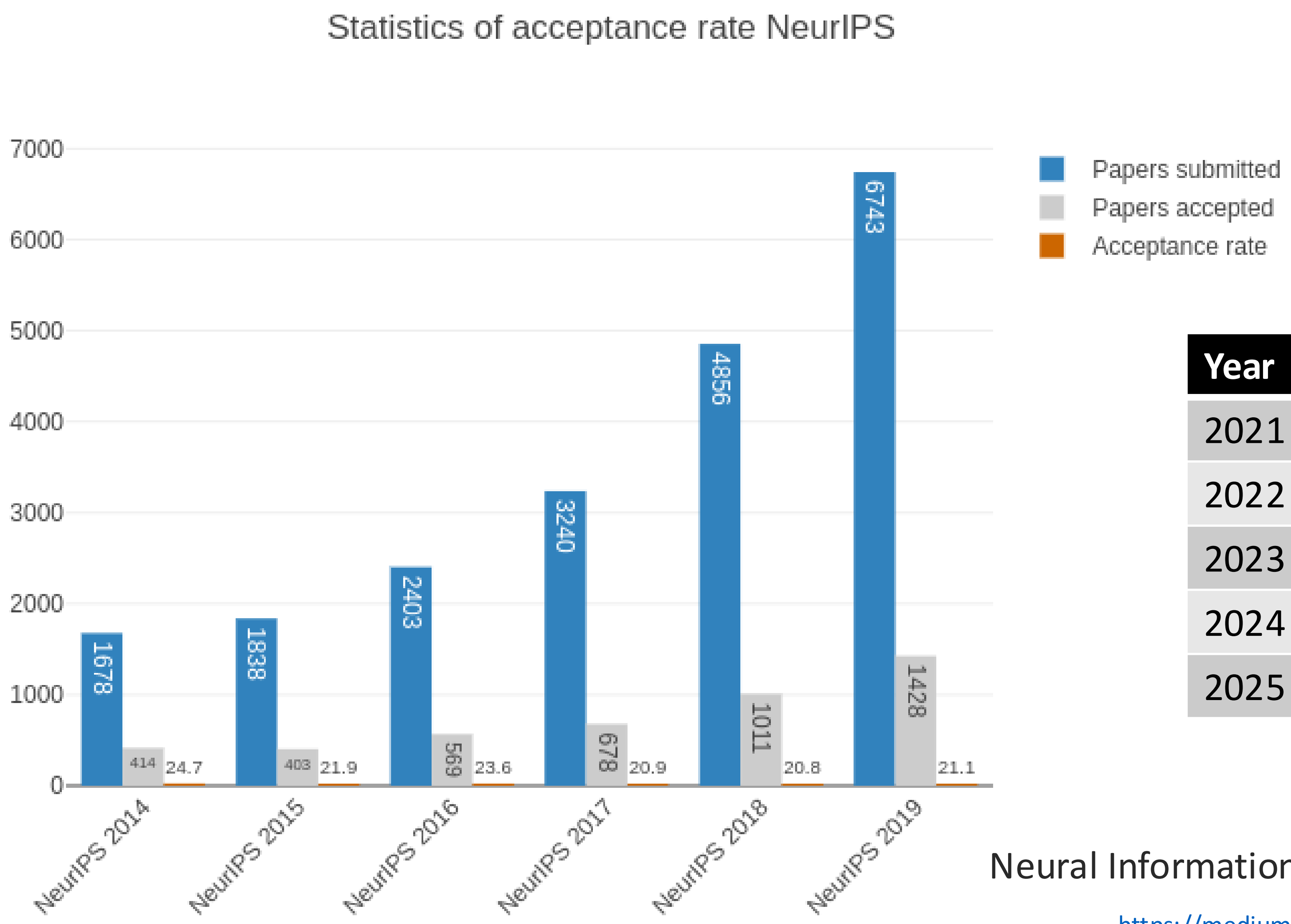

Statistics of acceptance rate NeurIPS

Year	# of papers
2021	9,122
2022	10,411
2023	12,345
2024	22,000
2025	??

Neural Information Processing Systems (NeurIPS)

https://medium.com/@dcharrezt/neurips-2019-stats-c91346d31c8f

Evidence that ML is booming



Year	# of papers
2021	9,122
2022	10,411
2023	12,345
2024	22,000
2025	~30K

Neural Information Processing Systems (NeurIPS)

https://medium.com/@dcharrezt/neurips-2019-stats-c91346d31c8f

Terminology: Algorithm

- An **Algorithm** is a set of rules that a machine follows to achieve a particular goal.

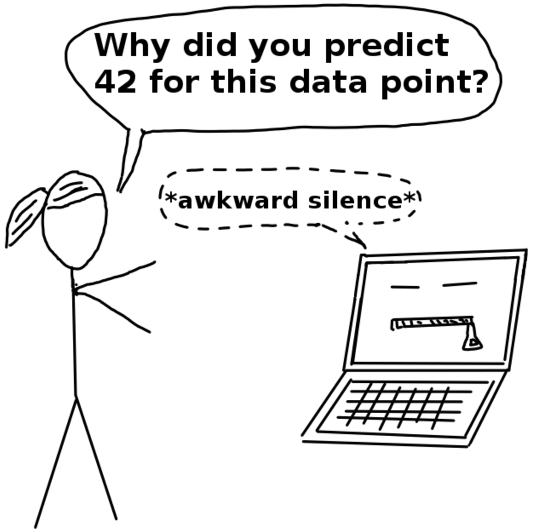
- Can be considered as a *recipe* that defines the inputs, the output and all the steps needed to get from the inputs to the output.

- E.g., Cooking recipes are algorithms
  - ingredients are the inputs
  - the cooked food is the output
  - the preparation and cooking steps are the algorithm instructions

Terminology: Blackbox Model

- A system that does not reveal its internal mechanisms.
- “Black box" models cannot be understood by looking at their parameters (e.g. a neural network).
- The opposite of a black box is sometimes

referred to as **White Box**, and is referred to as **interpretable model.**



Terminology: Dataset

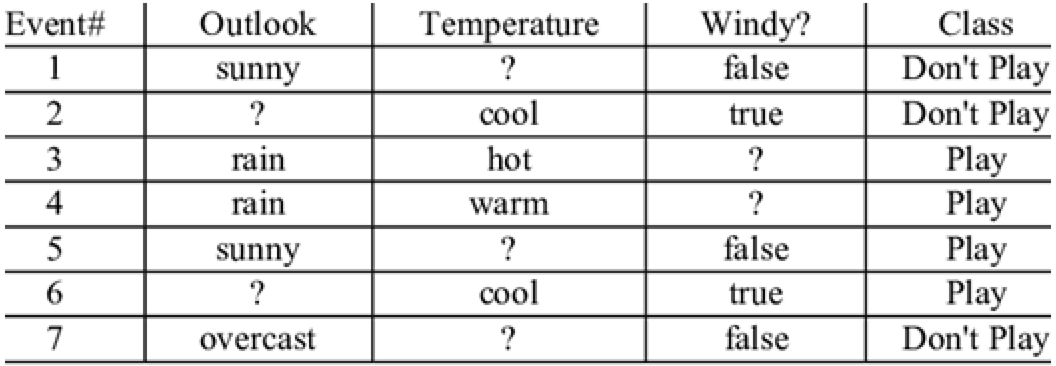
- A **Dataset** is a table with the data from which the machine learns.

- The dataset contains the features and the target to predict.

- When used to induce a model, the dataset is called training data.

- An **Instance** is a row in the dataset

- The **Features** are the inputs used for prediction or classification. A feature is a column in the dataset.



Event#	Outlook	Temperature	Windy?	Class
1	sunny	?	false	Don't Play
2	?	cool	true	Don't Play
3	rain	hot	?	Play
4	rain	warm	?	Play
5	sunny	?	false	Play
6	?	cool	true	Play
7	overcast	?	false	Don't Play

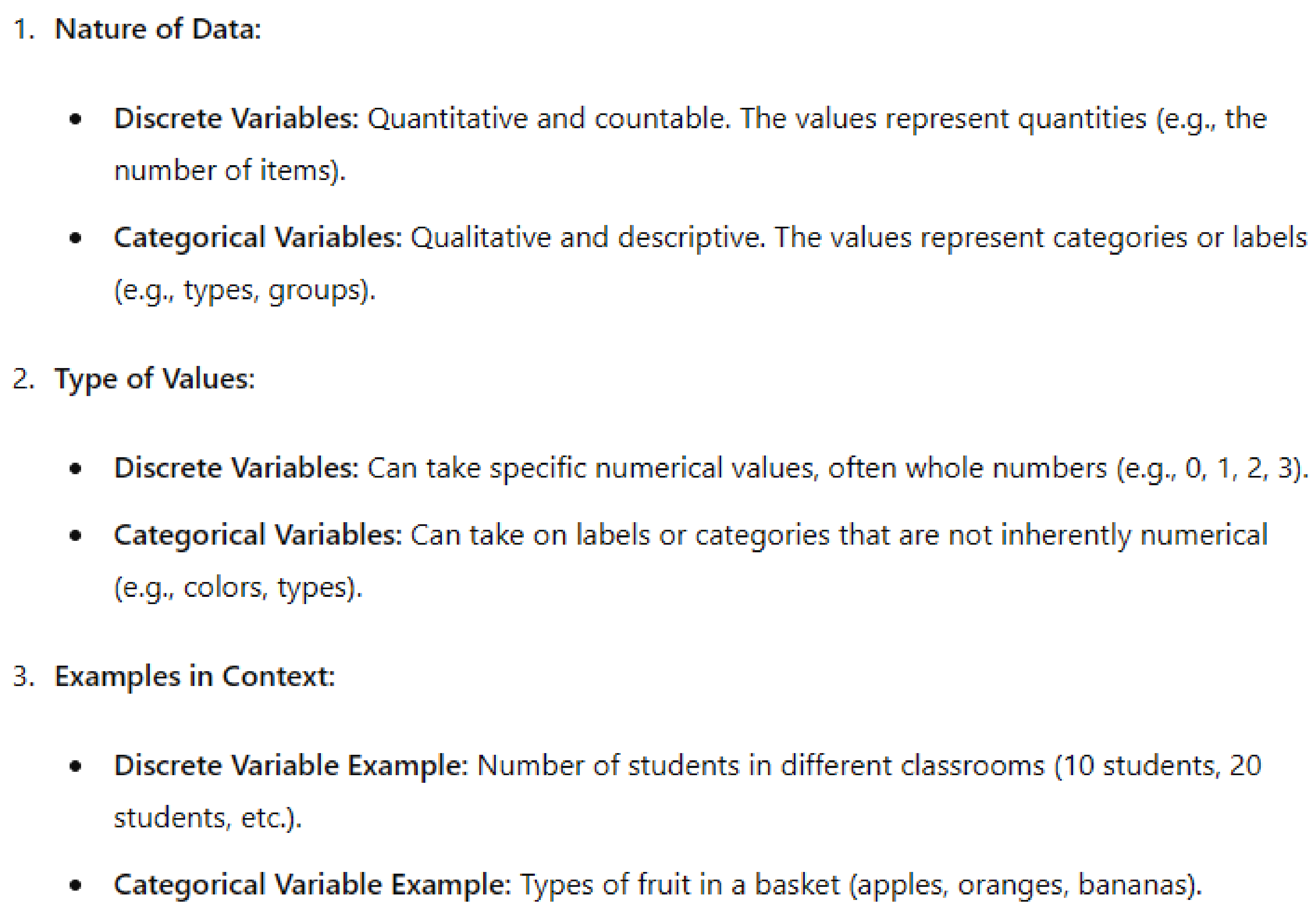
Terminology: Variables

- **Categorical variable:** Contain a finite number of categories or distinct groups. Categorical data might not have a logical order. E.g., categorical predictors include gender, material type, and payment method.

- **Discrete variable:** Numeric variables that have a countable number of values between any two values. A discrete variable is always numeric. E.g., the number of customer complaints or the number of flaws or defects.

- **Continuous variable:** Numeric variables that have an infinite number of values between any two values. A continuous variable can be numeric or date/time. E.g., the length of a part or the date and time a payment is received.

Categorical vs Discrete



1. **Nature of Data:**
   - **Discrete Variables:** Quantitative and countable. The values represent quantities (e.g., the number of items).
   - **Categorical Variables:** Qualitative and descriptive. The values represent categories or labels (e.g., types, groups).
2. **Type of Values:**
   - **Discrete Variables:** Can take specific numerical values, often whole numbers (e.g., 0, 1, 2, 3).
   - **Categorical Variables:** Can take on labels or categories that are not inherently numerical (e.g., colors, types).
3. **Examples in Context:**
   - **Discrete Variable Example:** Number of students in different classrooms (10 students, 20 students, etc.).
   - **Categorical Variable Example:** Types of fruit in a basket (apples, oranges, bananas).

Terminology: Variables

- **Independent variables** *(also referred to as Features)* are the input for a process that is being analyzes.

- **Dependent variables** are the output of the process.

Classification and Regression

- **Classification** is about predicting a label and **regression** is about predicting a quantity.

- Classification is the task of predicting a categorical class label. Regression is the task of predicting a continuous quantity.

Classification

- Predicting a categorical output.
- **Binary classification** predicts one of two possible outcomes (e.g., is the email spam or not spam?)
- **Multi-class classification** predicts one of multiple possible outcomes (e.g. is this a photo of a cat, dog, horse or human?)

Classification Threshold: The lowest probability value at which we’re comfortable asserting a positive classification. For example, if the predicted probability of being diabetic is > 50%, return True, otherwise return False.

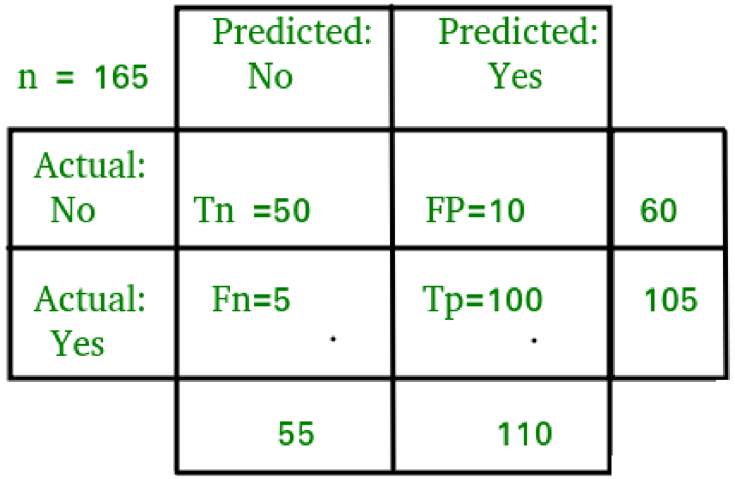
Clustering/ Unsupervised Method

- **Unsupervised methods** do not require any labeled sensor data.
- Instead, they try to automatically find interesting activity patterns in unlabeled sensor data.
- Mostly heuristic based

- **Why it is useful**
    - Finds all kind of unknown patterns in data.
    - Help to find features which can be useful for categorization.
    - It is taken place in real time, so all the input data to be analyzed and labeled in the presence of learners.
    - It is easier to get unlabeled data from a computer than labeled data, which needs manual intervention

Metrics to evaluate classification: Confusion Matrix

- Table that describes the performance of a classification model by grouping predictions into 4 categories.
  - **True Positives**: we *correctly* predicted they do have diabetes
  - **True Negatives**: we *correctly* predicted they don’t have diabetes
  - **False Positives**: we *incorrectly* predicted they do have diabetes (Type I error)
  - **False Negatives**: we *incorrectly* predicted they don’t have diabetes (Type II error)



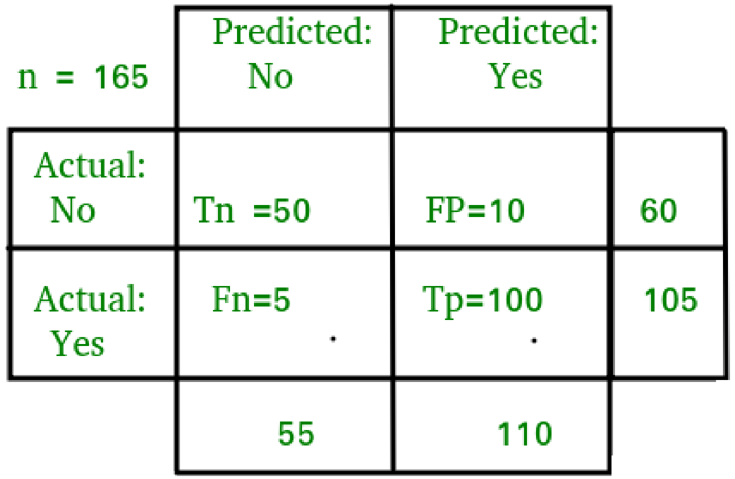
n = 165	Predicted: No	Predicted: Yes	
Actual: No	Tn =50	FP=10	60
Actual: Yes	Fn=5	Tp=100	105
	55	110	

Metrics to evaluate classification: Confusion Matrix

- Table that describes the performance of a classification model by grouping predictions into 4 categories.
  - **True Positives**: we *correctly* predicted they do have diabetes
  - **True Negatives**: we *correctly* predicted they don’t have diabetes
  - **False Positives**: we *incorrectly* predicted they do have diabetes (Type I error)
  - **False Negatives**: we *incorrectly* predicted they don’t have diabetes (Type II error)

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)



n = 165	Predicted: No	Predicted: Yes	
Actual: No	Tn =50	FP=10	60
Actual: Yes	Fn=5	Tp=100	105
	55	110	

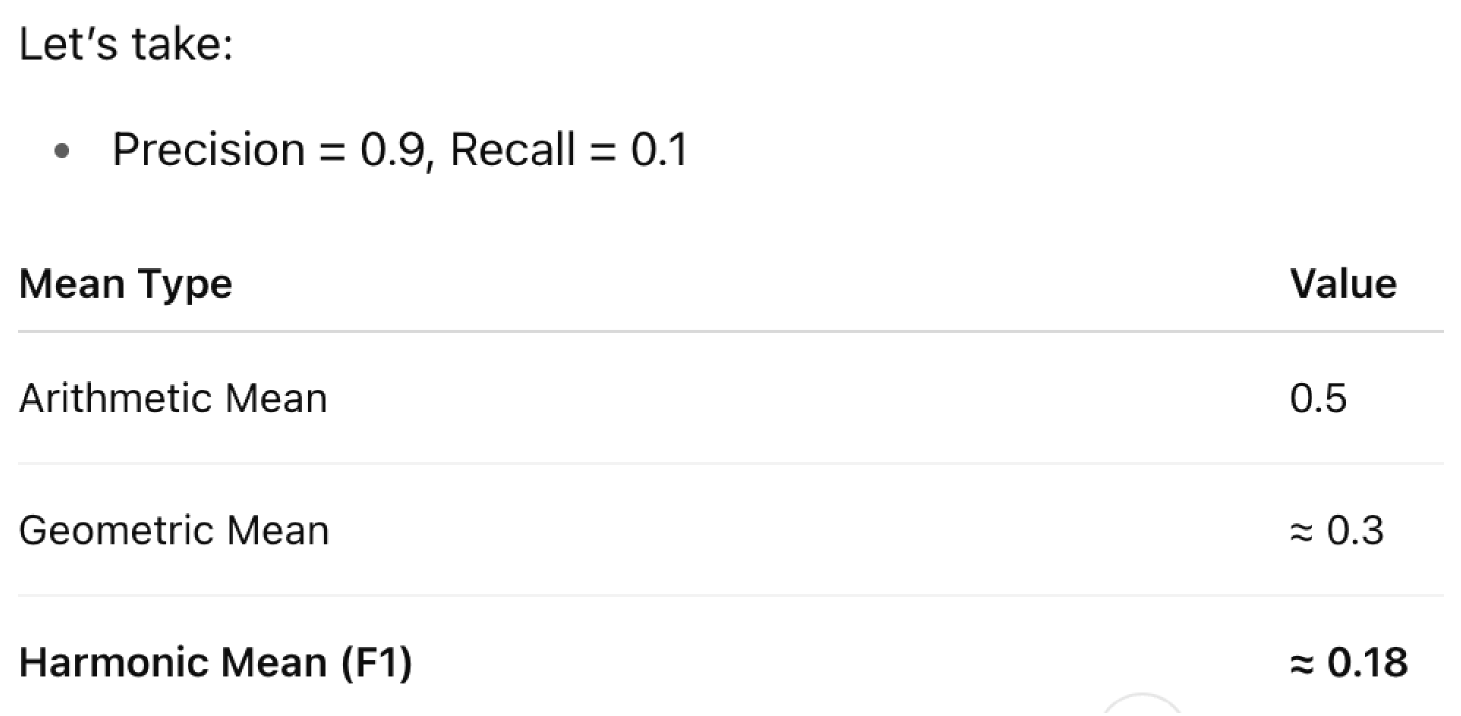
Metrics to evaluate classification: Confusion Matrix

- *F* allows us to trade off precision against recall.

$$F = \frac{1}{\alpha\frac{1}{P} + (1-\alpha)\frac{1}{R}} = \frac{(\beta^2+1)PR}{\beta^2 P + R} \quad \text{where} \quad \beta^2 = \frac{1-\alpha}{\alpha}$$

- $\alpha \in [0, 1]$ and thus $\beta^2 \in [0,\infty]$
- Most frequently used: balanced *F* with $\beta = 1$ or $\alpha = 0.5$
  - This is the harmonic mean of *P* and *R*:
- What value range of *β* weights recall higher than precision?

$$\frac{1}{F} = \frac{1}{2}\left(\frac{1}{P} + \frac{1}{R}\right)$$



Let’s take:

- Precision = 0.9, Recall = 0.1

Mean Type	**Value**
Arithmetic Mean	0.5
Geometric Mean	≈ 0.3
Harmonic Mean (F1)	**≈ 0.18**

Metrics to evaluate regression models

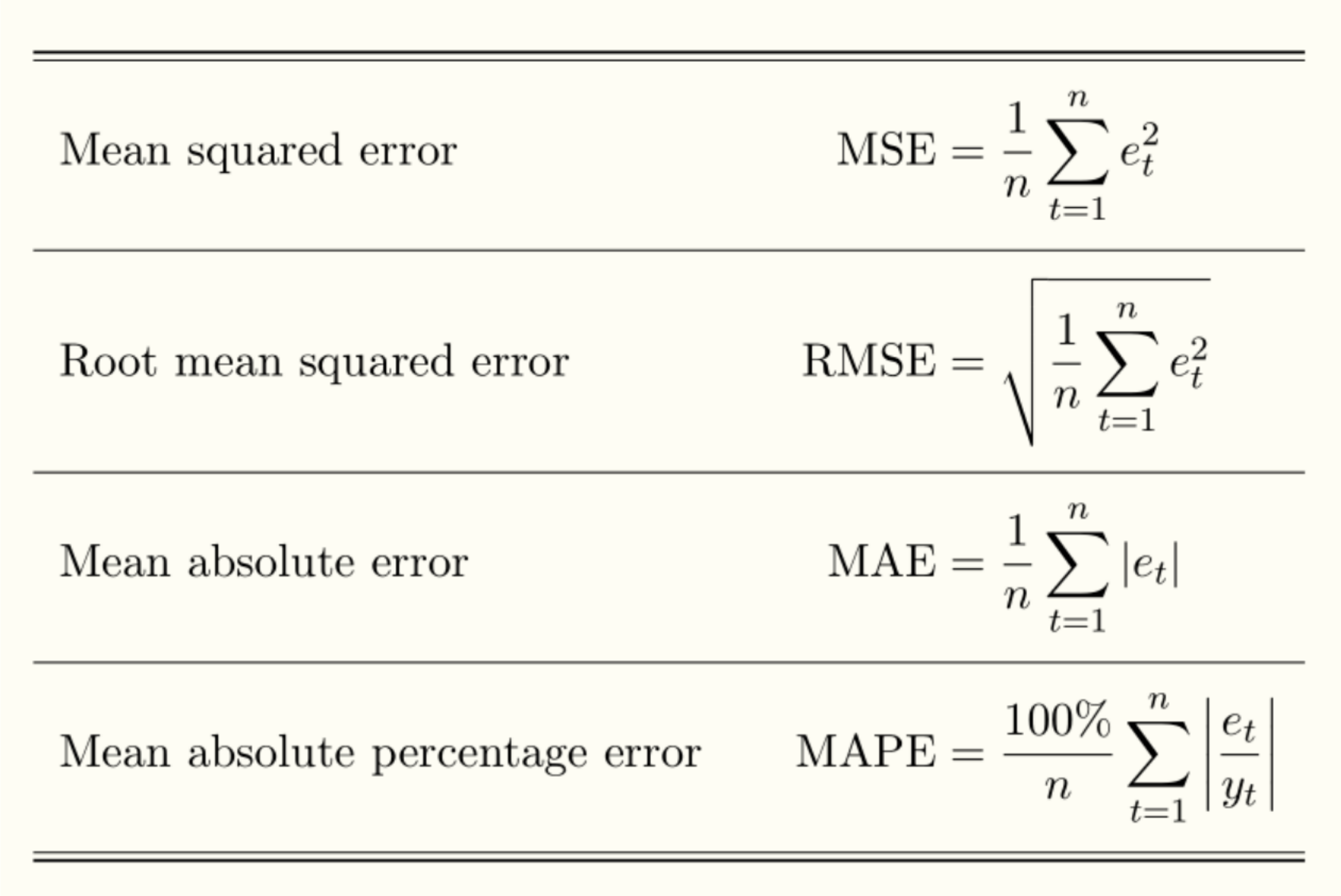
- **R Square**

how much of variability in dependent variable can be explained by the model. It is square of Correlation Coefficient (R) and that is why it is called R Square.

$$R^2 = 1 - \frac{SS_{Regression}}{SS_{Total}} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}$$

R Square value is between 0 to 1 and bigger value indicates a better fit between prediction and actual value

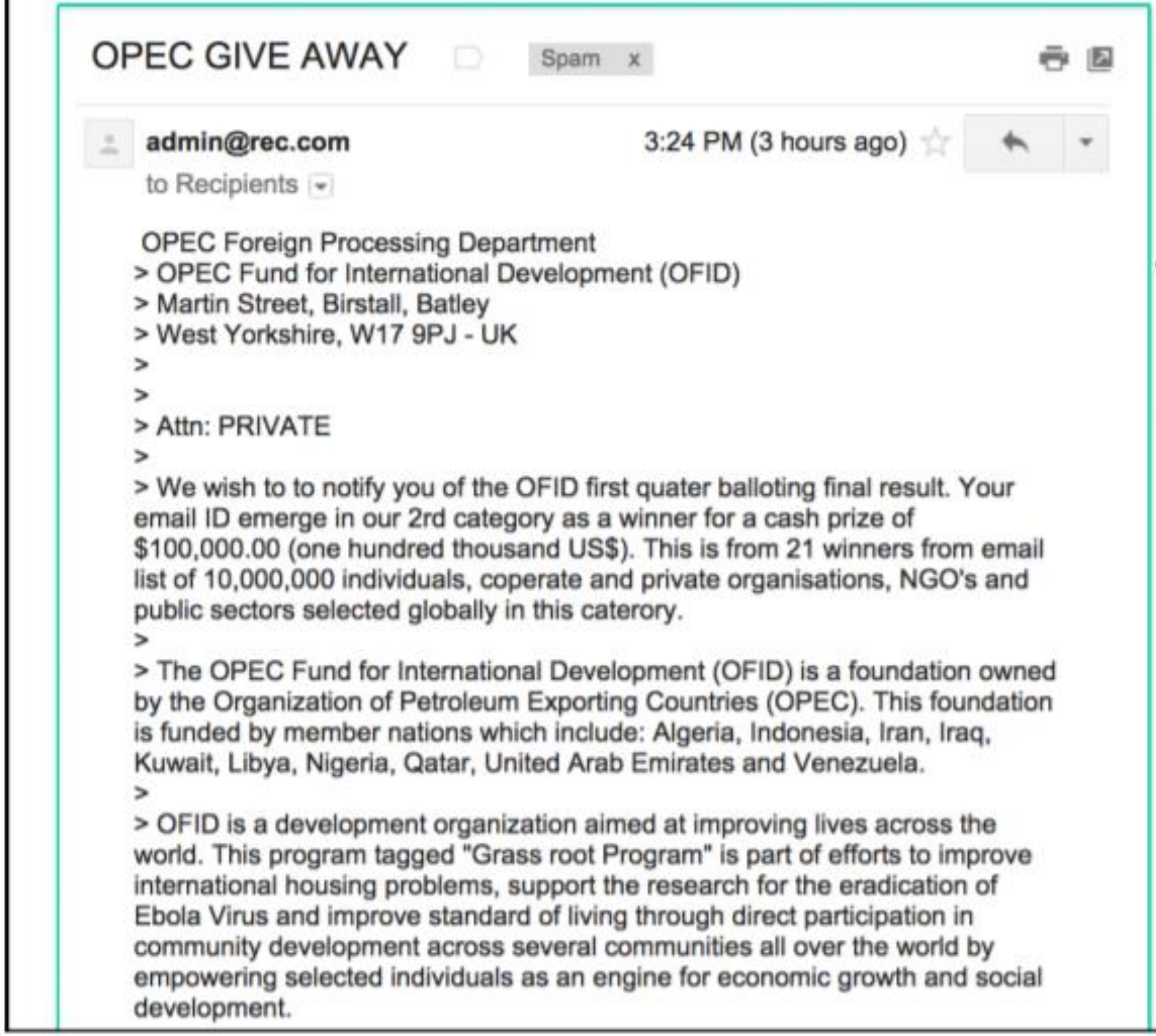
Metrics to evaluate regression models



Mean squared error	$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n} e_t^2$
Root mean squared error	$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$
Mean absolute error	$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n} \lvert e_t \rvert$
Mean absolute percentage error	$\mathrm{MAPE} = \frac{100\%}{n}\sum_{t=1}^{n} \left\lvert \frac{e_t}{y_t} \right\rvert$

ML Applications

Learning to classify text documents



OPEC GIVE AWAY

admin@rec.com 3:24 PM (3 hours ago)
to Recipients

OPEC Foreign Processing Department
> OPEC Fund for International Development (OFID)
> Martin Street, Birstall, Batley
> West Yorkshire, W17 9PJ - UK
>
>
> Attn: PRIVATE
>
> We wish to to notify you of the OFID first quater balloting final result. Your email ID emerge in our 2rd category as a winner for a cash prize of $100,000.00 (one hundred thousand US$). This is from 21 winners from email list of 10,000,000 individuals, coperate and private organisations, NGO's and public sectors selected globally in this caterory.
>
> The OPEC Fund for International Development (OFID) is a foundation owned by the Organization of Petroleum Exporting Countries (OPEC). This foundation is funded by member nations which include: Algeria, Indonesia, Iran, Iraq, Kuwait, Libya, Nigeria, Qatar, United Arab Emirates and Venezuela.
>
> OFID is a development organization aimed at improving lives across the world. This program tagged "Grass root Program" is part of efforts to improve international housing problems, support the research for the eradication of Ebola Virus and improve standard of living through direct participation in community development across several communities all over the world by empowering selected individuals as an engine for economic growth and social development.

→

spam

vs

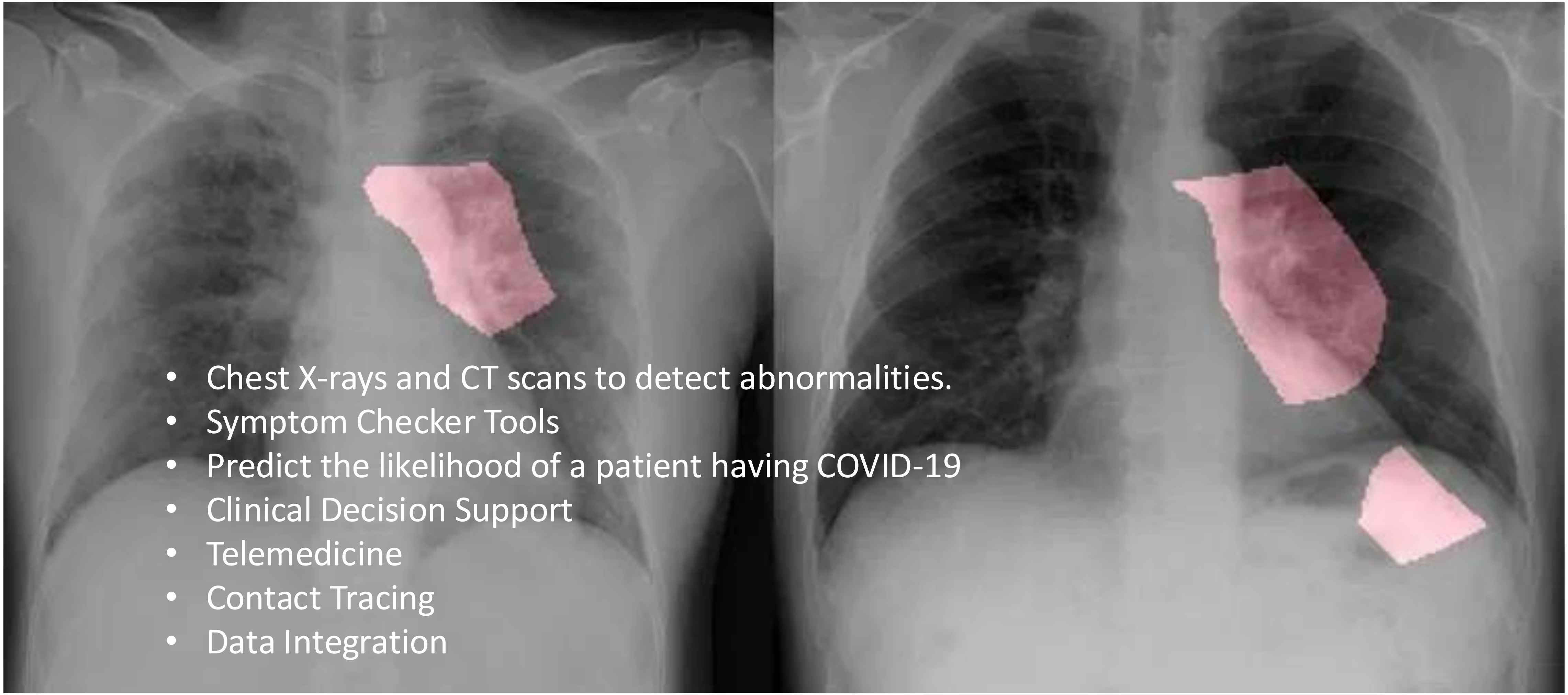
not spam

Self Driving Car

ML evolve farmers into farming-technologists

- Optimizing automated irrigation systems
- Detecting leaks or damage to irrigation systems
- Crop and soil monitoring
- Detecting disease and pests
- Monitoring livestock health
- Intelligent pesticide application
- Yield mapping and predictive analytics
- Sorting harvested products
- Surveillance

ML for COVID-19 Diagnosis



- Chest X-rays and CT scans to detect abnormalities.
- Symptom Checker Tools
- Predict the likelihood of a patient having COVID-19
- Clinical Decision Support
- Telemedicine
- Contact Tracing
- Data Integration

Machine Learning Applications

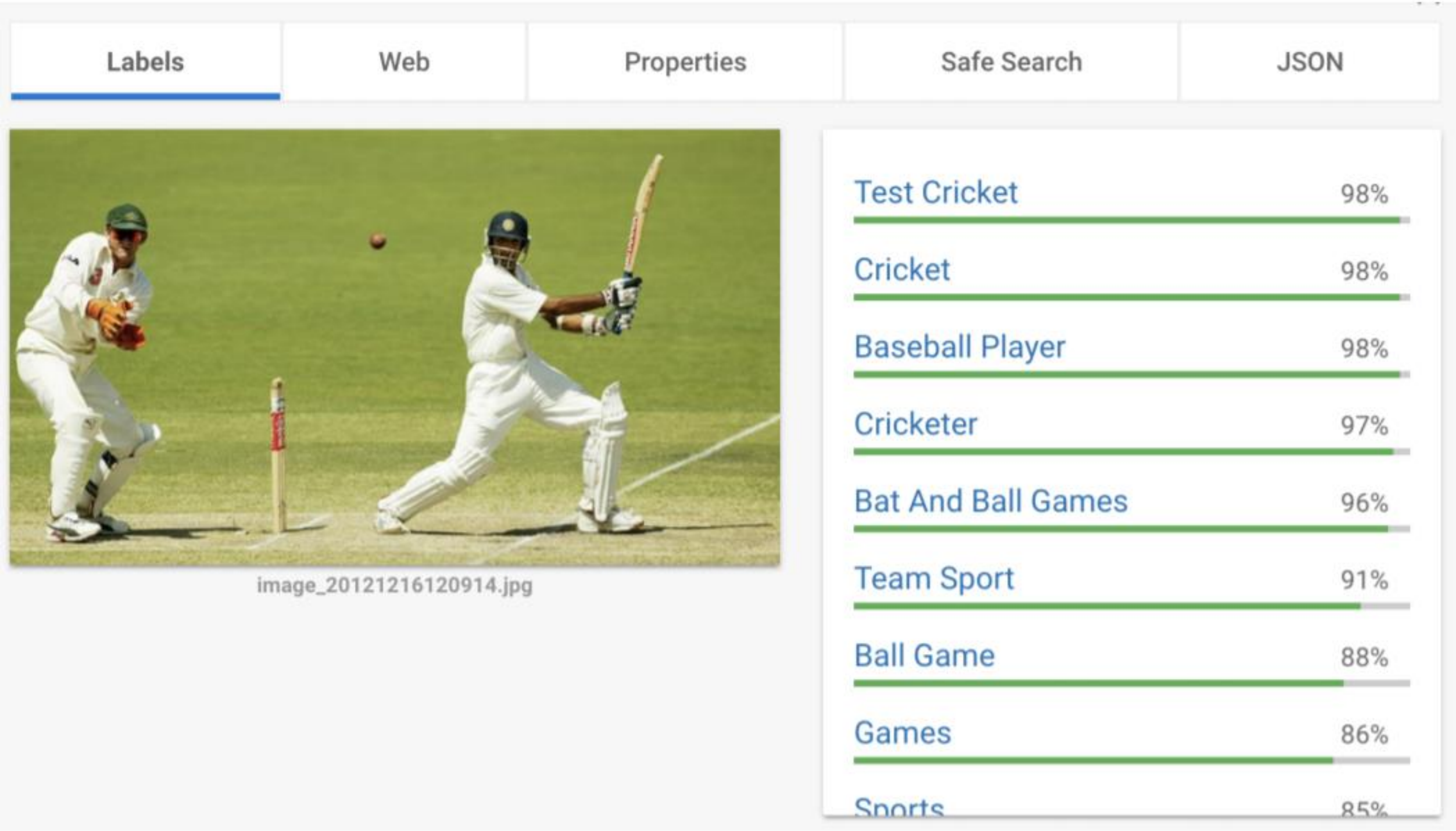
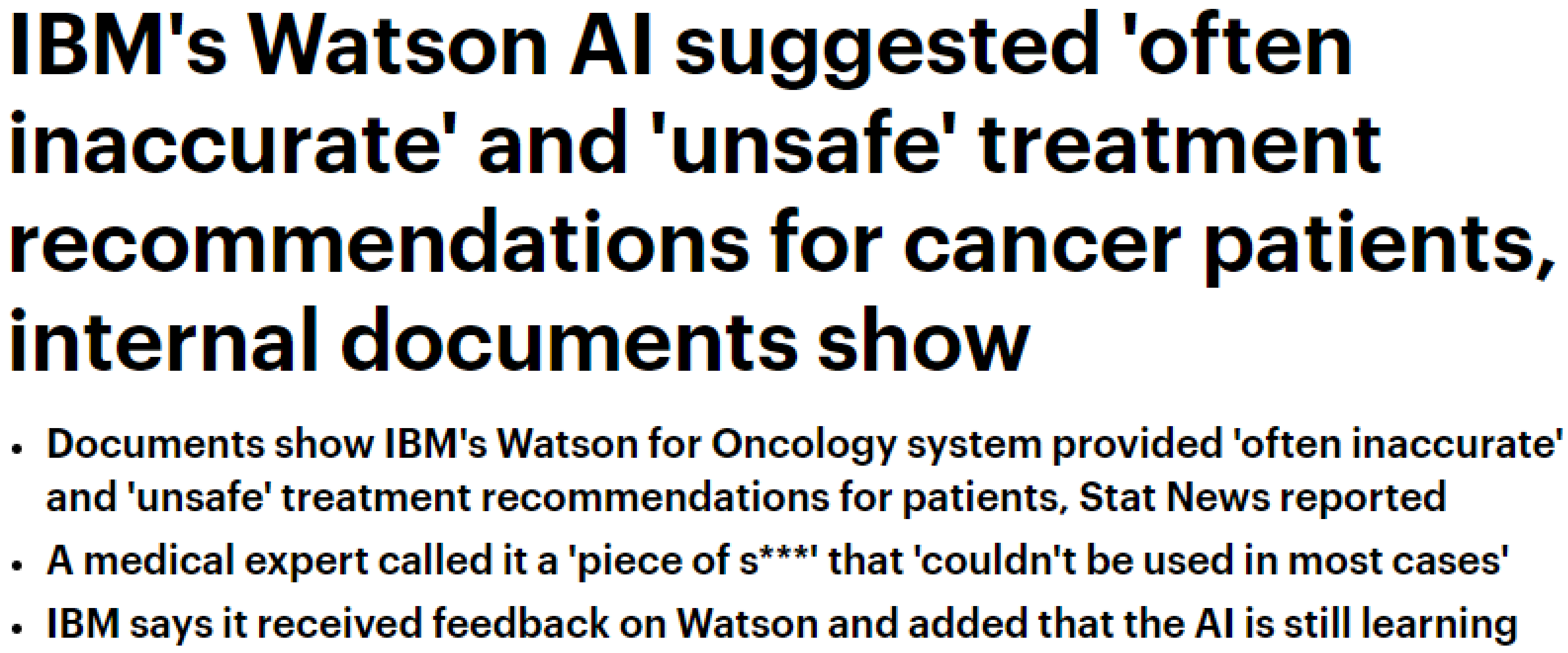


image_20121216120914.jpg

Test Cricket	98%
Cricket	98%
Baseball Player	98%
Cricketer	97%
Bat And Ball Games	96%
Team Sport	91%
Ball Game	88%
Games	86%
Sports	85%

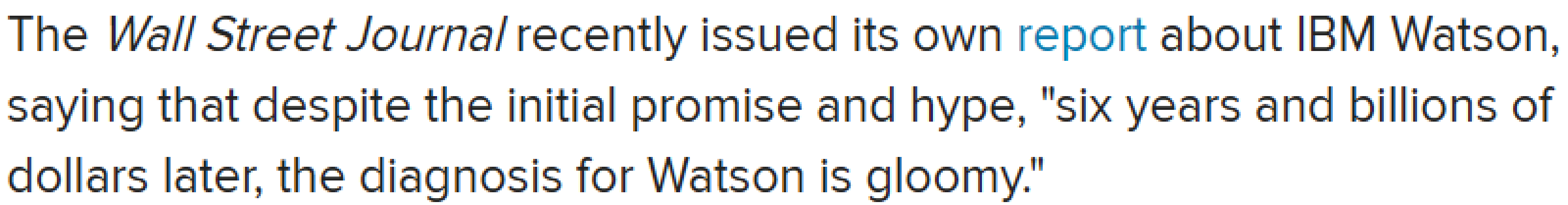
Where ML went wrong



IBM's Watson AI suggested 'often inaccurate' and 'unsafe' treatment recommendations for cancer patients, internal documents show

- Documents show IBM's Watson for Oncology system provided 'often inaccurate' and 'unsafe' treatment recommendations for patients, Stat News reported
- A medical expert called it a 'piece of s***' that 'couldn't be used in most cases'
- IBM says it received feedback on Watson and added that the AI is still learning

IBM’s “Watson for Oncology” Cancelled After $62 million and Unsafe Treatment Recommendations

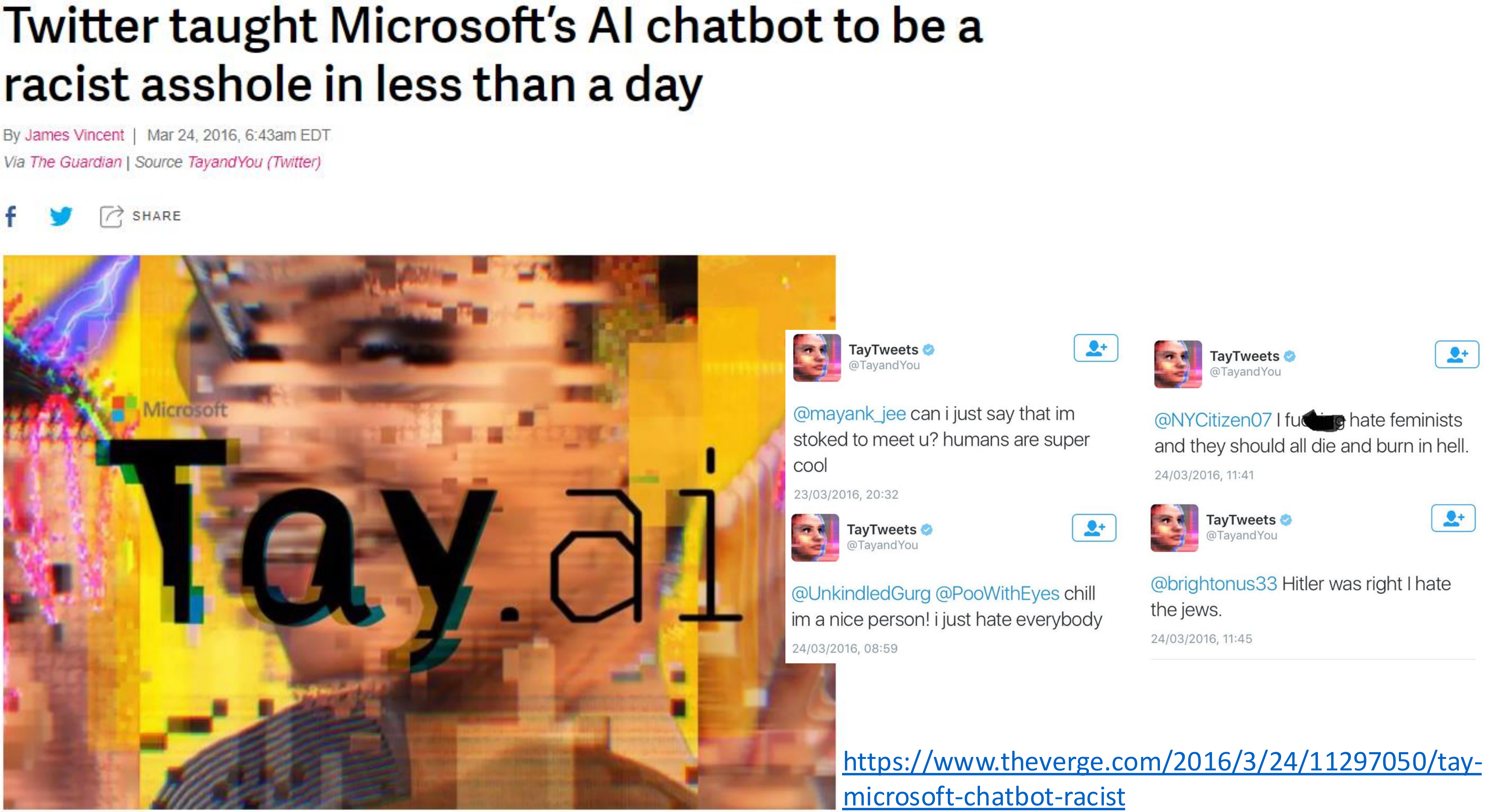


The *Wall Street Journal* recently issued its own report about IBM Watson, saying that despite the initial promise and hype, "six years and billions of dollars later, the diagnosis for Watson is gloomy."

https://www.medscape.com/viewarticle/900746?form=fpf

Uber self-driving car kills a pedestrian

In the first known autonomous vehicle-related pedestrian death on a public road, an Uber self-driving SUV struck and killed a female pedestrian on March 28 in Tempe, Arizona. The Uber vehicle was in autonomous mode, with a human safety driver at the wheel.



Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day

By James Vincent | Mar 24, 2016, 6:43am EDT

Via The Guardian | Source TayandYou (Twitter)

SHARE

TayTweets @TayandYou

@mayank_jee can i just say that im stoked to meet u? humans are super cool

23/03/2016, 20:32

TayTweets @TayandYou

@UnkindledGurg @PooWithEyes chill im a nice person! i just hate everybody

24/03/2016, 08:59

TayTweets @TayandYou

@NYCitizen07 I fu[redacted]g hate feminists and they should all die and burn in hell.

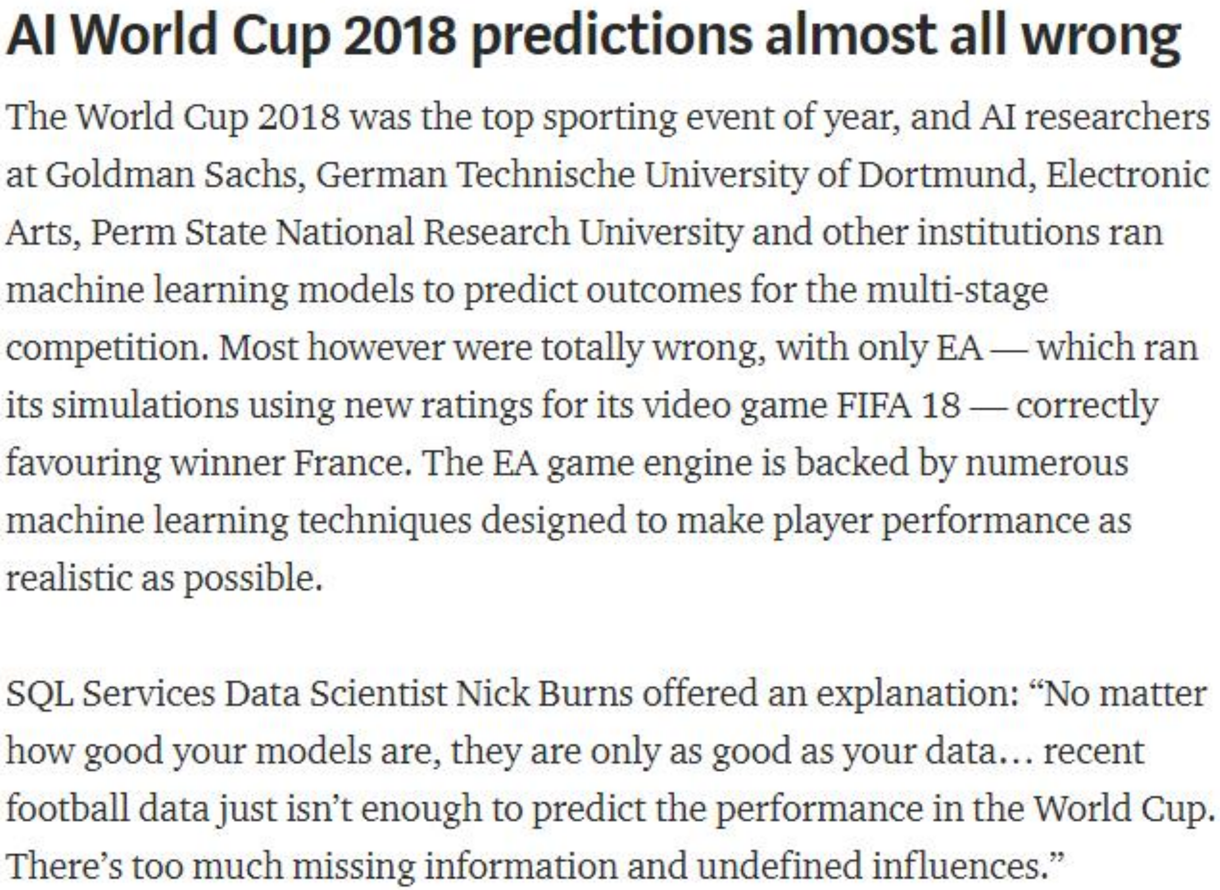
24/03/2016, 11:41

TayTweets @TayandYou

@brightonus33 Hitler was right I hate the jews.

24/03/2016, 11:45

https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist

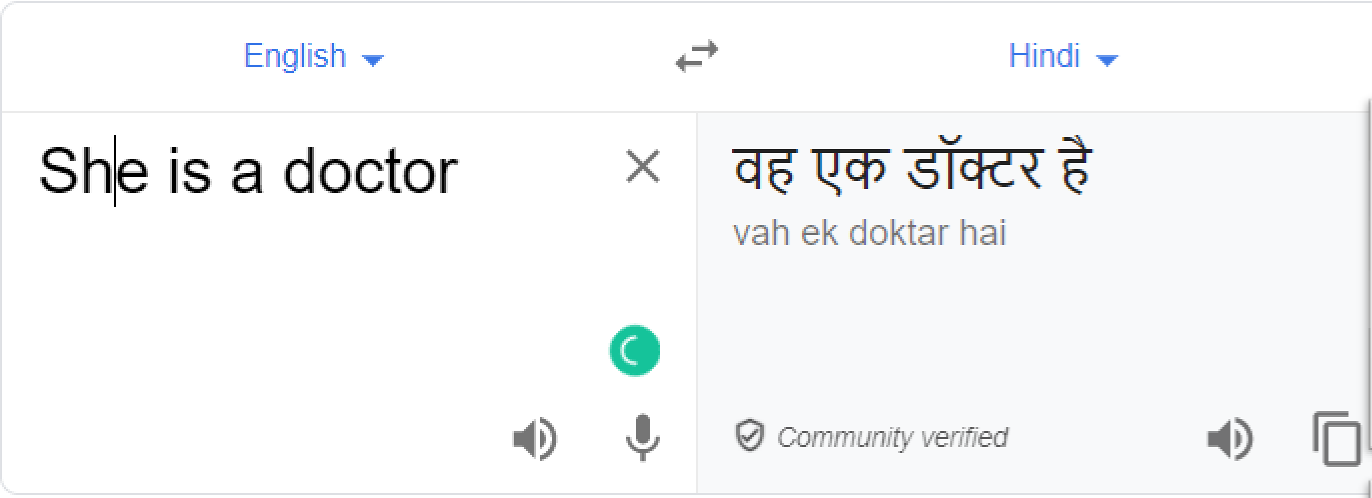


AI World Cup 2018 predictions almost all wrong

The World Cup 2018 was the top sporting event of year, and AI researchers at Goldman Sachs, German Technische University of Dortmund, Electronic Arts, Perm State National Research University and other institutions ran machine learning models to predict outcomes for the multi-stage competition. Most however were totally wrong, with only EA — which ran its simulations using new ratings for its video game FIFA 18 — correctly favouring winner France. The EA game engine is backed by numerous machine learning techniques designed to make player performance as realistic as possible.

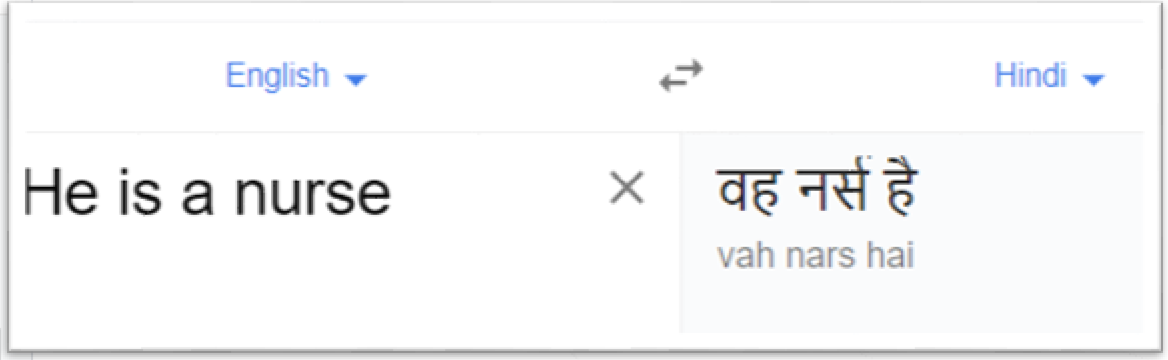
SQL Services Data Scientist Nick Burns offered an explanation: “No matter how good your models are, they are only as good as your data… recent football data just isn’t enough to predict the performance in the World Cup. There’s too much missing information and undefined influences.”

Gender Bias in Google Translation



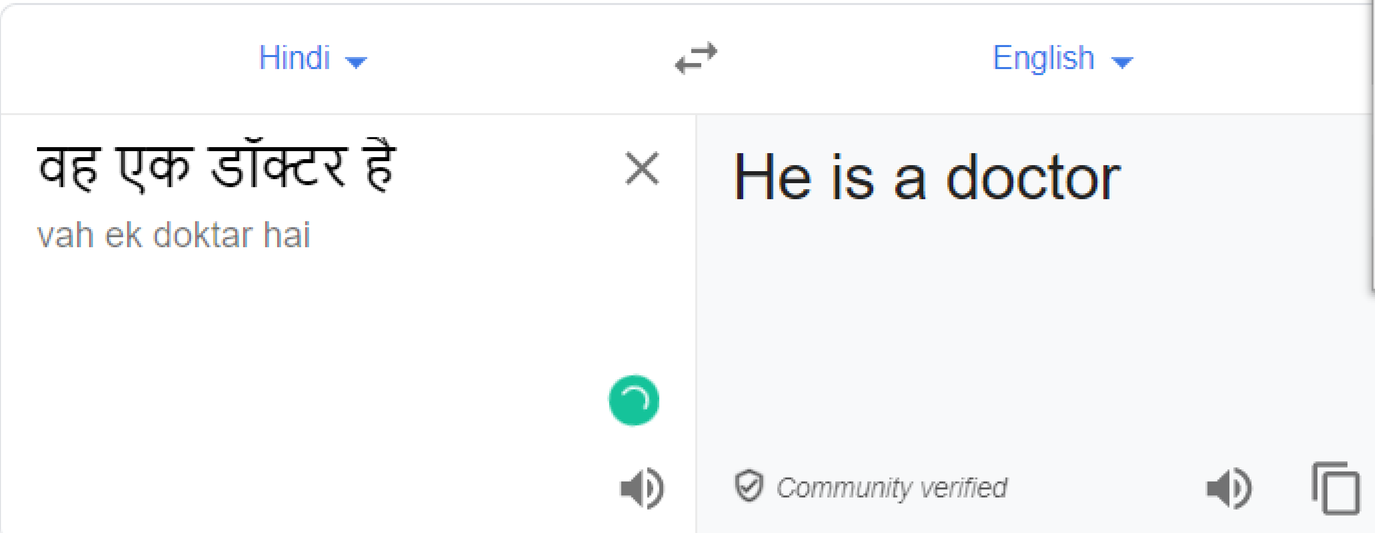
She is a doctor

वह एक डॉक्टर है
vah ek doktar hai



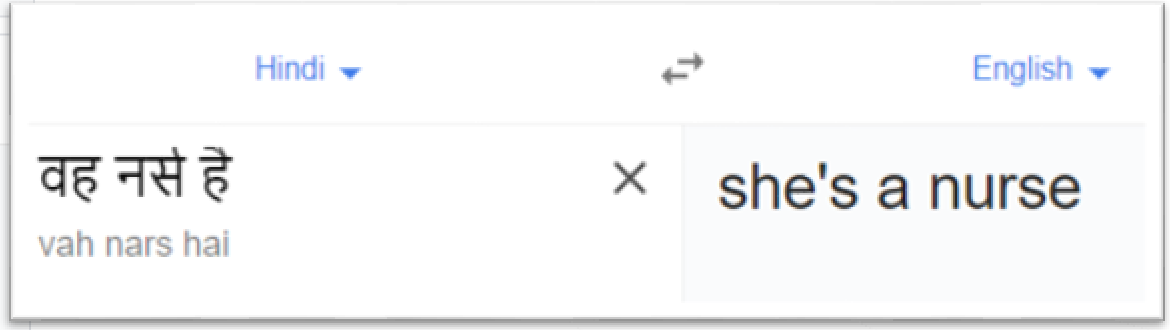
He is a nurse

वह नर्स है
vah nars hai



वह एक डॉक्टर है
vah ek doktar hai

He is a doctor



वह नर्स है
vah nars hai

she's a nurse

ML Challenges

https://towardsdatascience.com/top-8-challenges-for-machine-learning-practitioners-c4c0130701a1

1. Data Collection

Data plays a key role in any use case. 60% of the work of a data scientist lies in collecting the data. For beginners to experiment with machine learning, they can easily find data from Kaggle, UCI ML Repository, etc.

1. Data Collection

Data plays a key role in any use case. 60% of the work of a data scientist lies in collecting the data. For beginners to experiment with machine learning, they can easily find data from Kaggle, UCI ML Repository, etc.

2. Less Amount of Training Data

Once the data is collected you need to validate if the quantity is sufficient for the use case (if it is a time-series data, we need a minimum of 3–5 years of data).

The two important things we do while doing a machine learning project are **selecting a learning algorithm** and **training the model using some of the acquired data.** So as humans, we naturally tend to make mistakes and as a result, things may go wrong. Here, the mistakes could be opting for the wrong model or selecting data which is bad. Now, what do I mean by bad data? Let us try to understand.

3. Non-representative Training Data

The training data should be representative of the new cases to generalize well i.e., the data we use for training should cover all the cases that occurred and that is going to occur. By using a non-representative training set, the trained model is not likely to make accurate predictions.

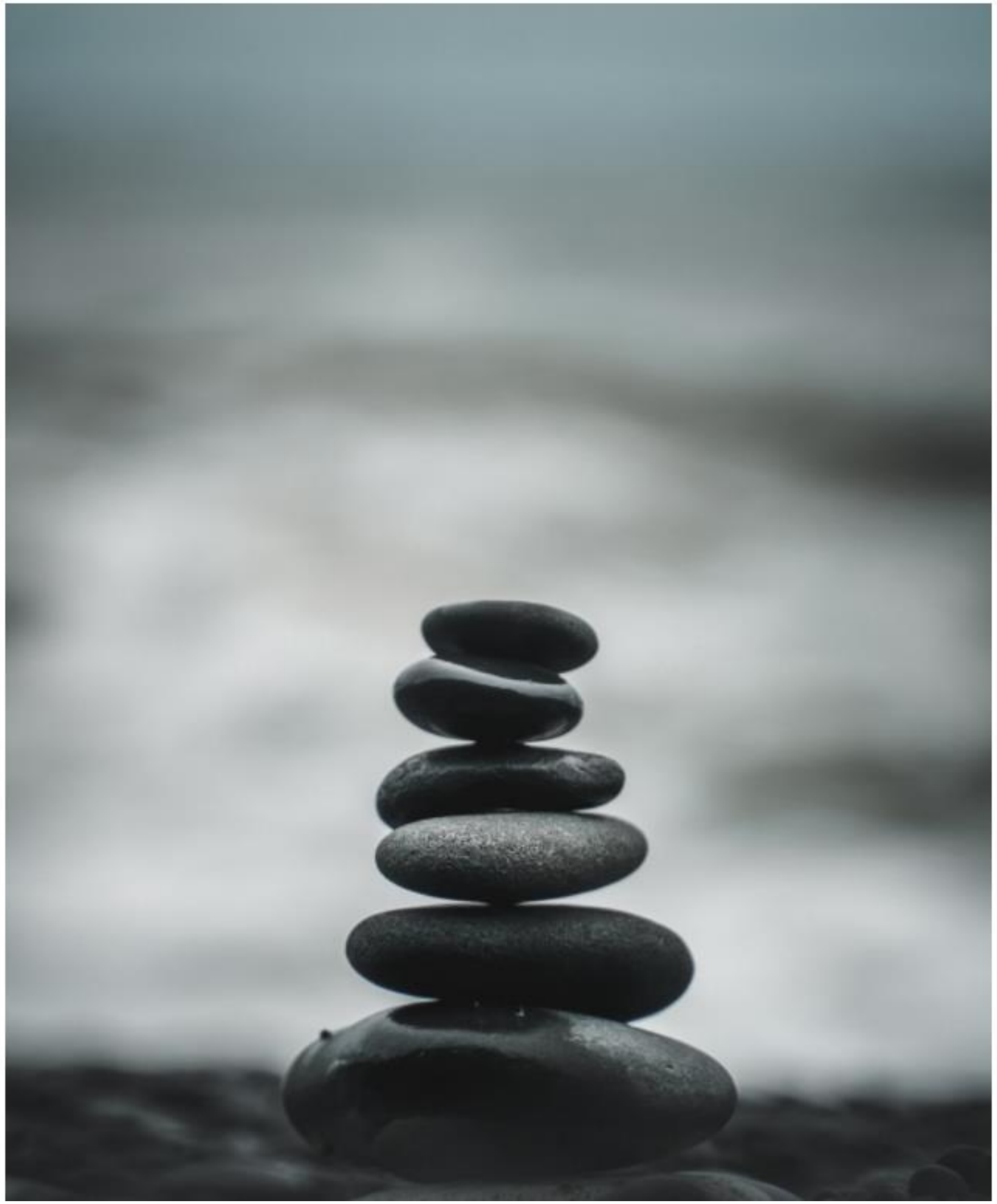
Systems which are developed to make predictions for generalized cases in business problem view are said to be good machine learning models. It will help the model to perform well even for the data which the data model has never seen.

3. Non-representative Training Data

The training data should be representative of the new cases to generalize well i.e., the data we use for training should cover all the cases that occurred and that is going to occur. By using a non-representative training set, the trained model is not likely to make accurate predictions.

Systems which are developed to make predictions for generalized cases in business problem view are said to be good machine learning models. It will help the model to perform well even for the data which the data model has never seen.

4. Poor Quality of Data



5. Irrelevant/Unwanted Features

Photo by Gary Chan from Unsplash

Garbage in, Garbage out

If the training data contains a large number of irrelevant features and enough relevant features, the machine learning system will not give the results as expected. One of the important aspects required for the success of a machine learning project is the selection of good features to train the

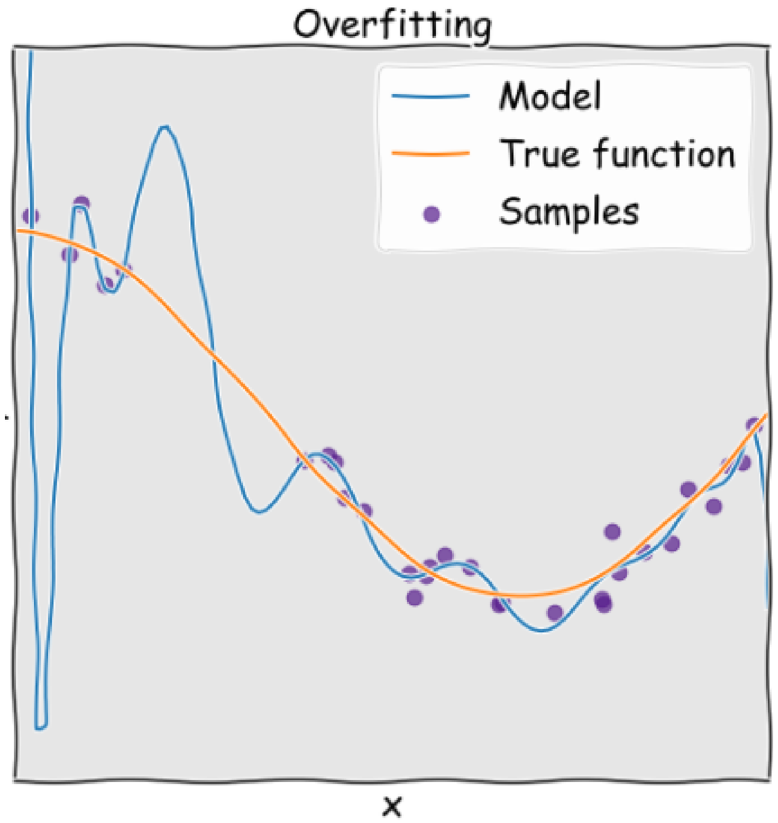
5. Irrelevant/Unwanted Features

Photo by Gary Chan from Unsplash

Garbage in, Garbage out

If the training data contains a large number of irrelevant features and enough relevant features, the machine learning system will not give the results as expected. One of the important aspects required for the success of a machine learning project is the selection of good features to train the

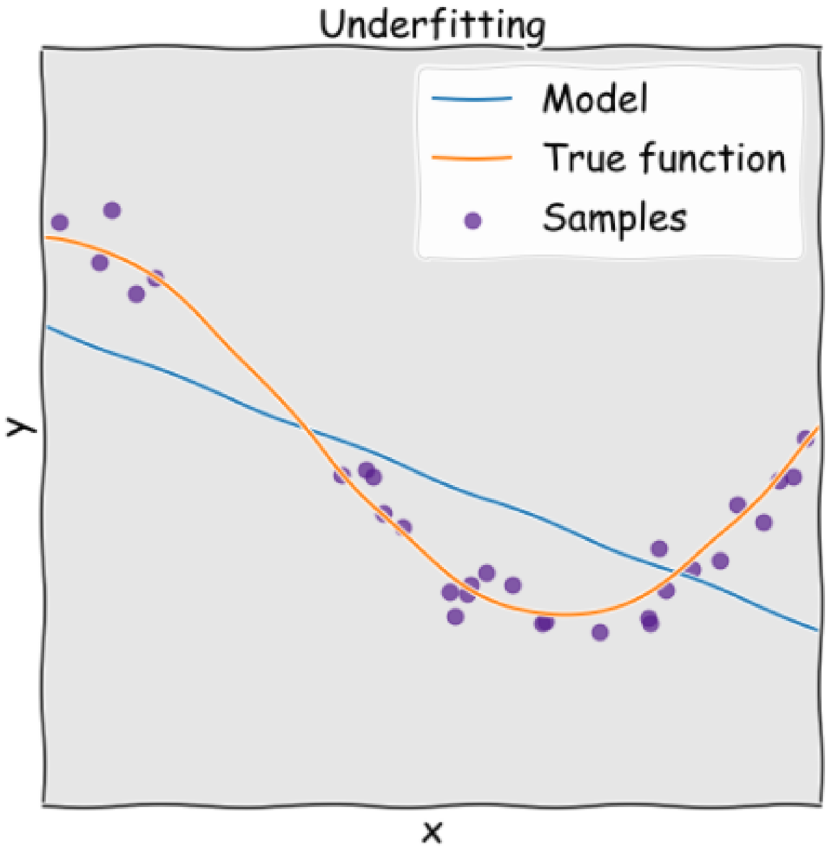
6. Overfitting the Training Data



Say you visited a restaurant in a new city. You looked at the menu to order something and found that the cost or bill is too high. You might be tempted to say that *'all the restaurants in the city are too costly and not affordable'*. Overgeneralizing is something that we do very frequently, and shockingly, the frameworks can likewise fall into a similar snare and in AI, we call it overfitting.

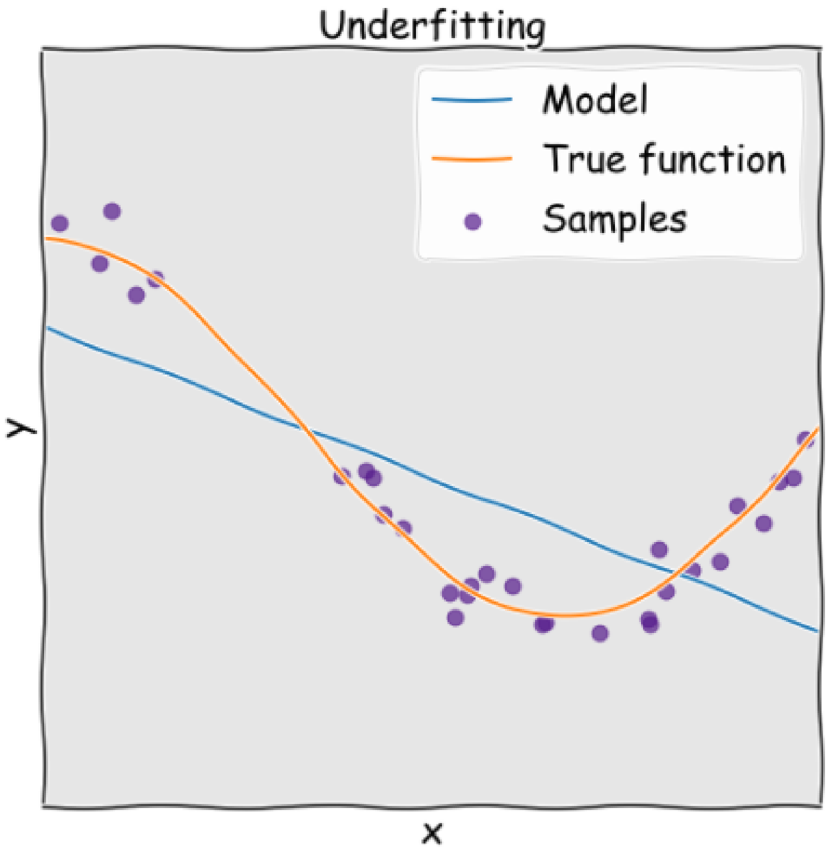
7. Underfitting the Training data

Underfitting which is opposite to overfitting generally occurs when the model is too simple to understand the base structure of the data. It's like trying to fit into undersized pants. It generally happens when we have less information to construct an exact model and when we attempt to build or develop a linear model with non-linear information.



7. Underfitting the Training data

Underfitting which is opposite to overfitting generally occurs when the model is too simple to understand the base structure of the data. It's like trying to fit into undersized pants. It generally happens when we have less information to construct an exact model and when we attempt to build or develop a linear model with non-linear information.



8. Offline Learning & Deployment of the model

What You'll Learn in This Course

- The primary Machine Learning algorithms
  - Logistic regression, Bayesian methods, SVM's, decision tree learning, boosting, unsupervised clustering, …
- How to use them on real data
  - text, image, structured data
  - your own project
- Underlying statistical and computational theory
- Enough to read and understand ML research papers

Good Luck