

Machine Learning

CSCI 567

Dani Yogatama

About this course

Modern machine learning methods used in real-world AI applications.

Focus on conceptual understanding of these methods.

About this course

Modern machine learning methods used in real-world AI applications.

Focus on conceptual understanding of these methods.

Objectives

Develop skills to grasp abstract ML concepts and think critically.

Practice with hands on programming tasks.

Preparation for studying more advanced machine learning techniques.

Prerequisites

Undergraduate level training in probability and statistics, linear algebra, (multivariate) calculus.

Important: attend today's discussion session to see if you have the required background.

Programming: Python.

Not an intro-level CS course, no training of basic programming skills.

Logistics

Lectures: Fridays, 1–3.20pm (SGM 123)

Discussions: Fridays, 3.30–4.20pm

Course website

<https://usc-tamagotchi.github.io/csci-567/23f/>

Teaching Staff

TAs: Ting-Rui Chiang, Samuel Griesemer, Josh Robinson, Oliver Liu,
Robby Costales, Tenghao Huang, Tejas Srinivasan

CPs/Graders: Aman Bansal, Wenda Zhou, Sanying Yi, Sneha Bandi

Slides and Reading

Lecture slides will be posted before class (possibly updated after).

No required textbooks.

Grade

25%: Quiz 1 (9/29). Open book, no collaboration.

25%: Quiz 2 (11/17). Open book, no collaboration.

50%: Course Project

Grade

25%: Quiz 1 (9/29). Open book, no collaboration.

25%: Quiz 2 (11/17). Open book, no collaboration.

50%: Course Project

Initial cut-offs

$B^- = [70, 75)$, $B = [75, 80)$, $B^+ = [80, 86)$

$A^- = [86, 92)$, $A = [92, 100)$

Final cut-offs will not be released.

Grade

25%: Quiz 1 (9/29). Open book, no collaboration.

25%: Quiz 2 (11/17). Open book, no collaboration.

50%: Course Project

Initial cut-offs

$B^- = [70, 75)$, $B = [75, 80)$, $B^+ = [80, 86)$

$A^- = [86, 92)$, $A = [92, 100)$

Final cut-offs will not be released.

Academic integrity

Course Project

Done in groups (~3 students).

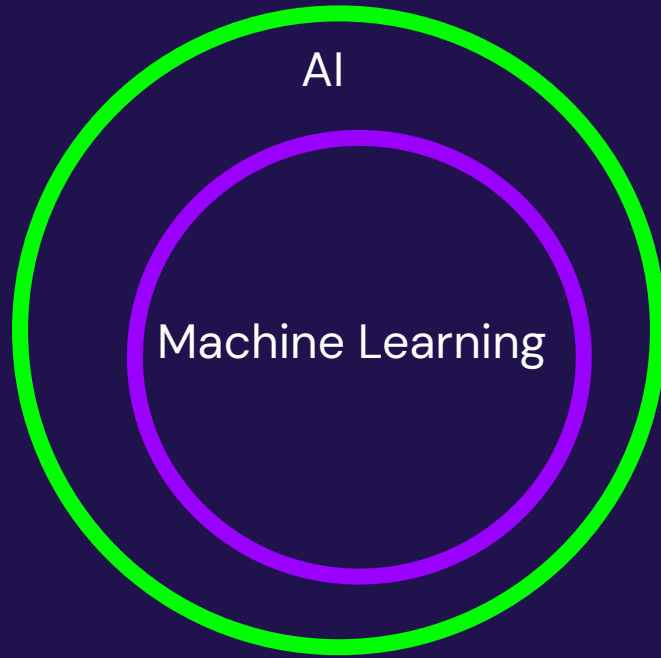
Any machine learning topic and any domain is fine.

It must include an implementation of a machine learning algorithm.

The implemented model has to work to a reasonable level (don't pick a problem that is too difficult).

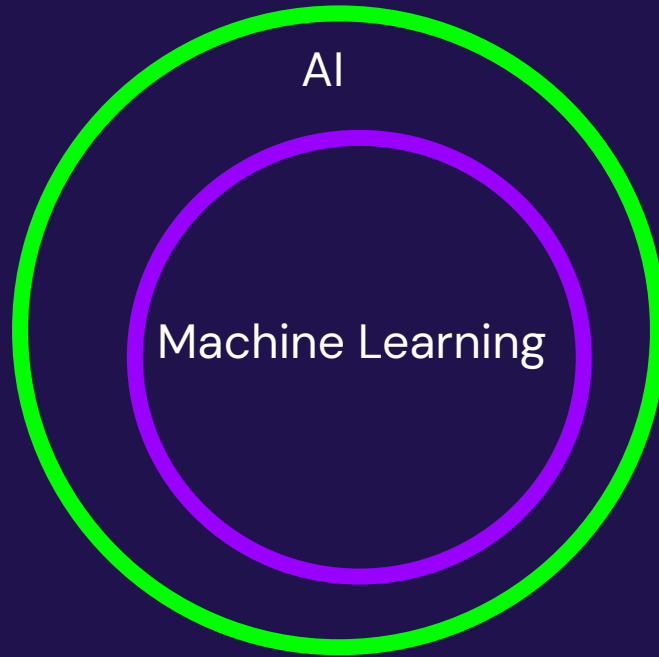
We will share more details in a separate document early next week.

What is machine learning?



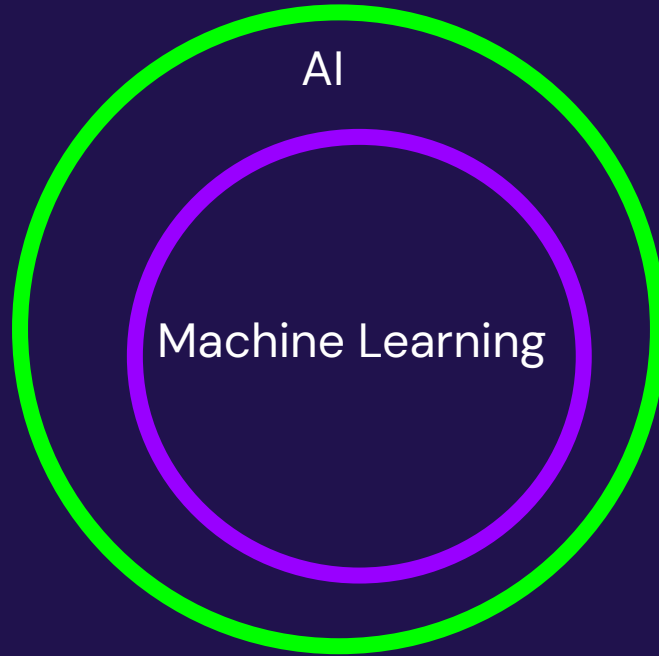
AI

Machine Learning



Machine learning is the fuel that powers state-of-the-art AI agents.

Personal opinion: AI is a civilization-altering technology that is going to transform the way we live.



Machine learning is the fuel that powers state-of-the-art AI agents.

Consumer applications

Speech recognition, information retrieval and search, text classification, stock price prediction, object recognition, ...

Scientific applications

Protein structure prediction, social network analysis, business intelligence, nuclear fusion, ...

The frontier of AI

You Create an itinerary for a two day trip to London. Please answer concisely, not longer than 100 words.

Yasa Day 1:



- Start at Buckingham Palace and watch the Changing of the Guard ceremony
- Walk to Big Ben and the Houses of Parliament
- Visit the British Museum
- OPT: Enjoy a traditional English pub dinner

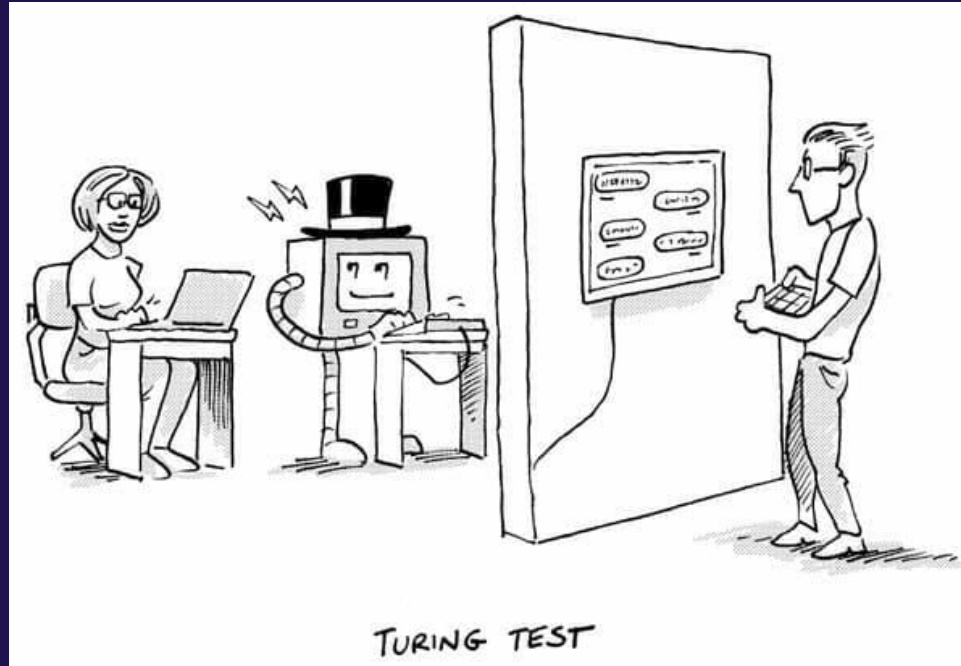
Day 2:

- Walk across Tower Bridge
- Visit the Tower of London and see the Crown Jewels
- Take a boat ride along the River Thames
- Visit St. Paul's Cathedral
- OPT: Visit the Tate Modern or the National Gallery

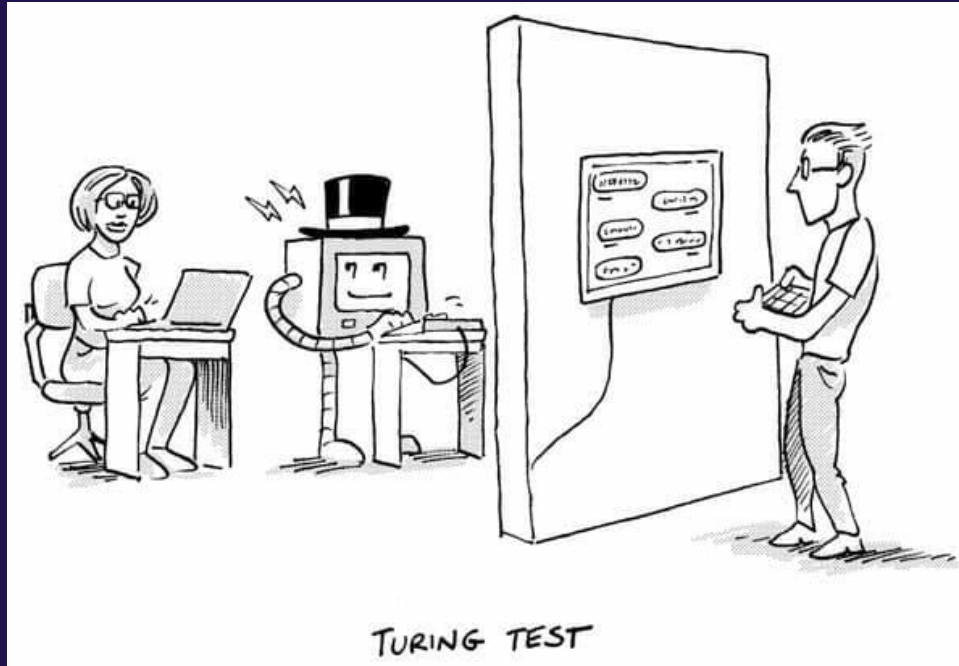
A poisonous giant green rat holding a red rose, digital art



Passing the Turing test used to be the benchmark for machine intelligence.



We have many agents that arguably have passed the test.



We can download the models and run them.

Why do we need to study ML?

How to contribute to the rapid pace of progress in AI

Understanding of fundamental concepts of how ML models work

Strong engineering skills

How to train a model

Learning a *model* from *training data* (*training examples*) by optimizing a *loss function* to minimize *generalization errors*.

Data

Model

Loss function

Generalization errors

Data

University of Southern California

Article [Talk](#)

From Wikipedia, the free encyclopedia

For other universities also known as USC, see [USC \(disambiguation\)](#).

The **University of Southern California** (**USC**, **SC**, **Southern Cal**^[a] or **SoCal**) is a [private research university](#) in [Los Angeles, California](#). Founded in 1880 by [Robert Maclay Widhey](#), it is the oldest private research university in California.^{[11][12]} The university is composed of one [liberal arts school](#), the [Dornsife College of Letters, Arts and Sciences](#), and 22 [undergraduate, graduate](#), and professional schools, enrolling roughly 21,000 undergraduate and 28,500 [post-graduate](#) students from all fifty [U.S. states](#) and more than 115 countries.^{[13][14][15][16]} It is a member of the [Association of American Universities](#), which it joined in 1969, and is also one of the [wealthiest academic institutions](#) in America.

USC sponsors a variety of intercollegiate sports and competes in the [National Collegiate Athletic Association](#) (NCAA) as a member of the [Pac-12 Conference](#). Members of USC's sports teams, the [Trojans](#), have won 107 NCAA team championships and 412 NCAA individual championships.^[17] As of 2021, Trojan athletes have won 326 medals at the [Olympic Games](#) (153 golds, 96 silvers, and 77 bronzes), more than any other university in the United States or in the world.^[18] USC has had 537 football players drafted to the [National Football League](#), the second-highest number of draftees in the country.^[19]

USC has graduated more alumni who have gone on to win [Academy](#) and [Emmy Awards](#) than any other institution, largely due to the [School of Cinematic Arts](#).^{[20][21]} USC has conferred degrees upon 29 [living billionaires](#)^[22] and is also one of the most successful universities in creating companies and attracting funding with multiple company founders as alumni or current students.^{[23][24]} USC presently has ten [Nobel Laureates](#) on staff,^[25] eleven [Rhodes Scholars](#),^{[26][27]} twelve [Marshall Scholars](#),^[28] six [MacArthur Fellows](#),^[29] 181 [Fulbright Scholars](#),^[30] one [Turing Award](#) winner,^[31] three winners of the [National Medal of Arts](#), one winner of the [National Humanities Medal](#), three winners of the [National Medal of Science](#), and three winners of the [National Medal of Technology and Innovation](#) among its alumni and faculty.^[32] USC is also the birthplace of technologies such as the [Domain Name System](#),^[33] [VoIP](#),^[34] [DNA computing](#),^[35] [transform coding](#),^[36] and [dynamic programming](#).^[37]



Data

Supervised vs. unsupervised learning

Data

Supervised vs. unsupervised learning



Data

Supervised vs. unsupervised learning



Document
Sentence
Image

Category
Sentiment
Class

Data

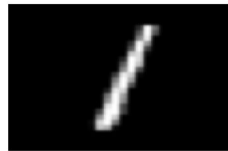
true class = 7



true class = 2



true class = 1



true class = 0



true class = 4



true class = 1



true class = 4



true class = 9



true class = 5



Data

"Just had the most amazing experience at USC! The campus is beautiful, the people are friendly, and the opportunities are endless. I'm so grateful to be a Trojan! #USC #ProudTrojan"

I hate the traffic in LA, it's always so frustrating and makes me stressed out. #LATraffic

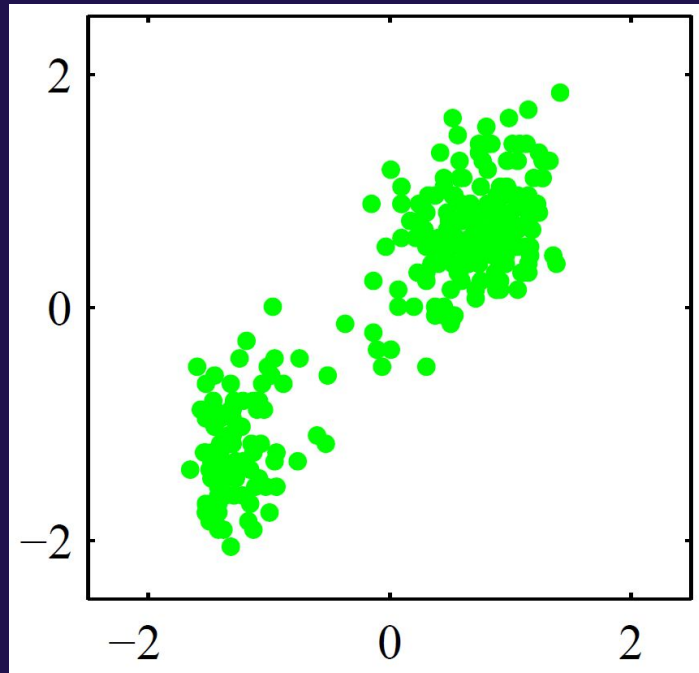
Data

Input data is often represented as a feature vector

Fisher's <i>Iris</i> Data				
Sepal length ↕	Sepal width ↕	Petal length ↕	Petal width ↕	Species ↕
5.1	3.5	1.4	0.2	<i>I. setosa</i>
4.9	3.0	1.4	0.2	<i>I. setosa</i>
4.7	3.2	1.3	0.2	<i>I. setosa</i>
4.6	3.1	1.5	0.2	<i>I. setosa</i>
5.0	3.6	1.4	0.2	<i>I. setosa</i>
5.4	3.9	1.7	0.4	<i>I. setosa</i>
4.6	3.4	1.4	0.3	<i>I. setosa</i>
5.0	3.4	1.5	0.2	<i>I. setosa</i>
4.4	2.9	1.4	0.2	<i>I. setosa</i>
4.9	3.1	1.5	0.1	<i>I. setosa</i>

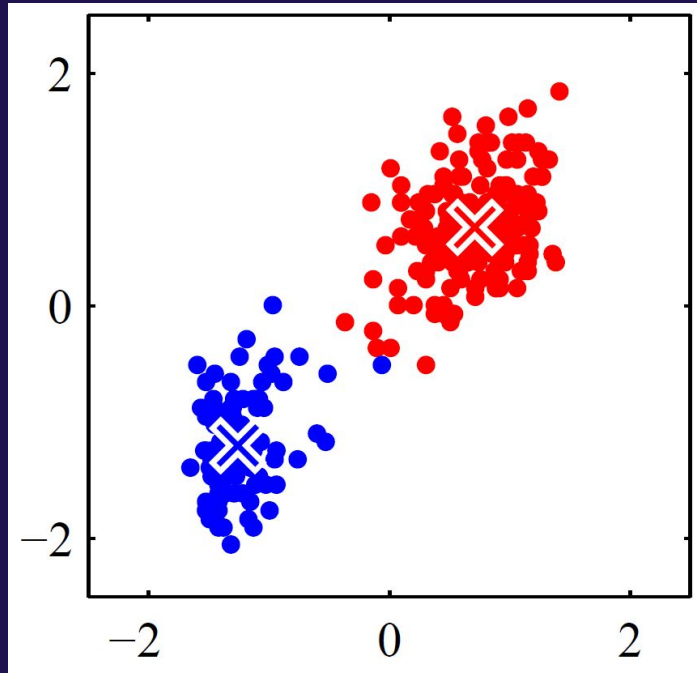
Data

Supervised vs. unsupervised learning



Data

Supervised vs. unsupervised learning



Model

Model

In supervised learning, the goal is to learn a function (based on the available training data) that maps a new input to a predicted output.

$$f(\mathbf{x}; \theta) = \mathbf{y}$$

Model

In supervised learning, the goal is to learn a function (based on the available training data) that maps a new input to a predicted output.

$$f(x; \theta) = y$$

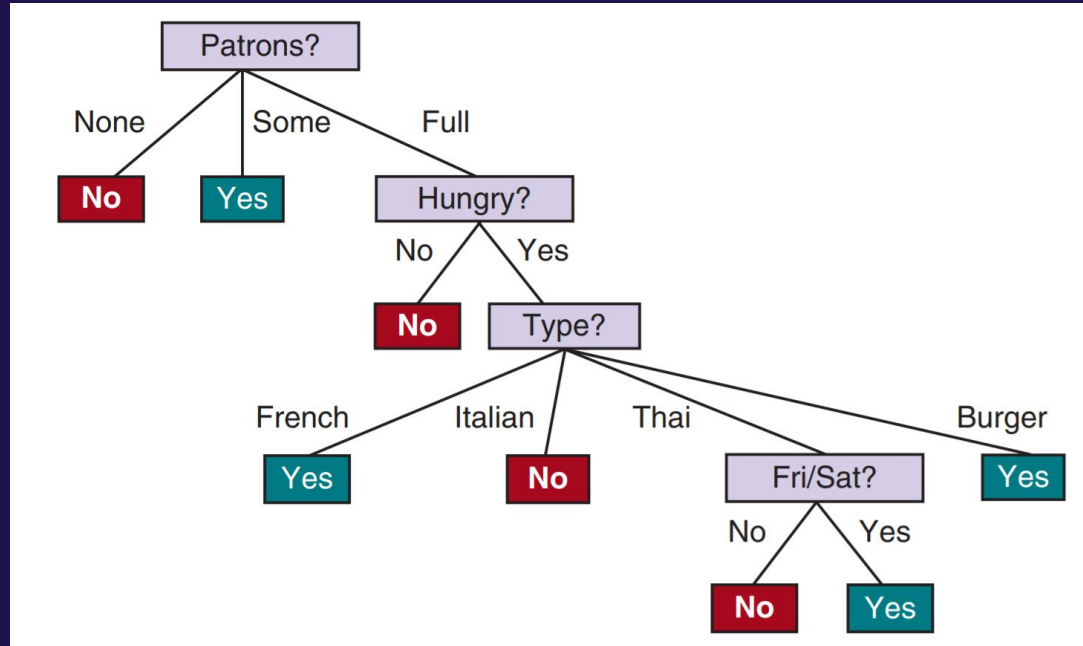
model parameters

Model

What function to choose?

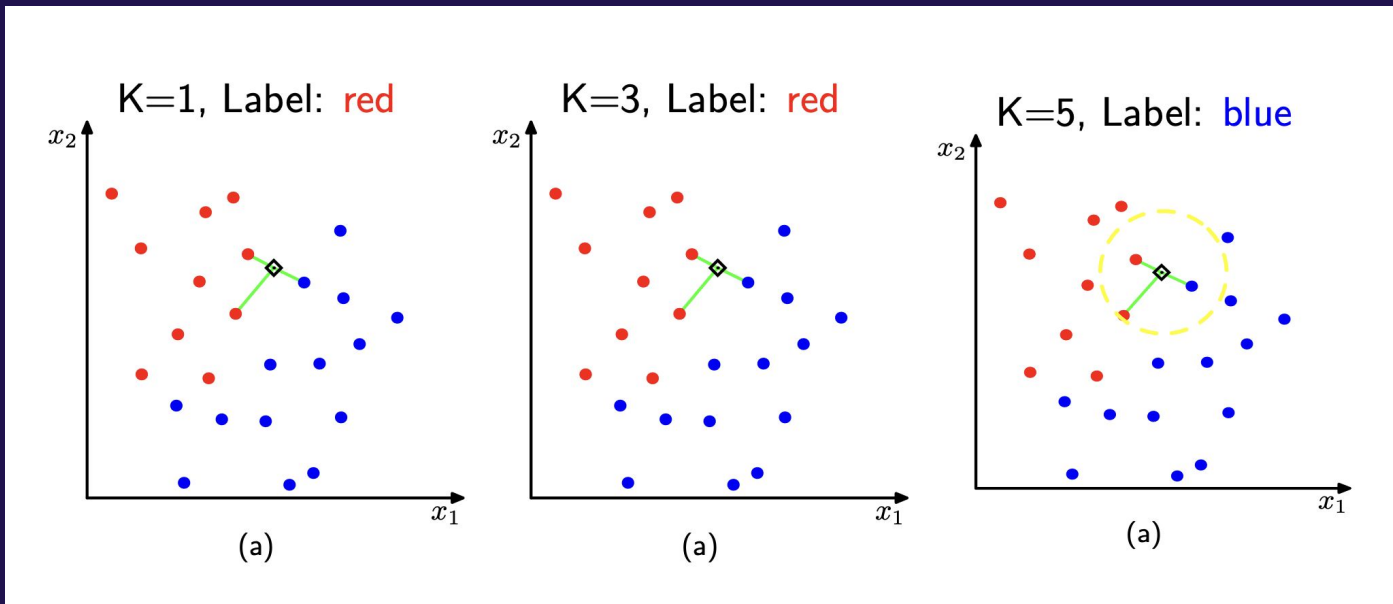
Model

What function to choose? Decision tree



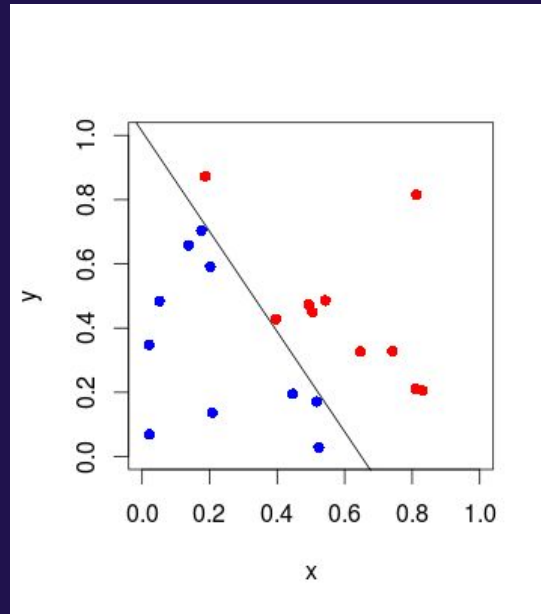
Model

What function to choose? K -nearest neighbors



Model

What function to choose? Linear classifier



Model

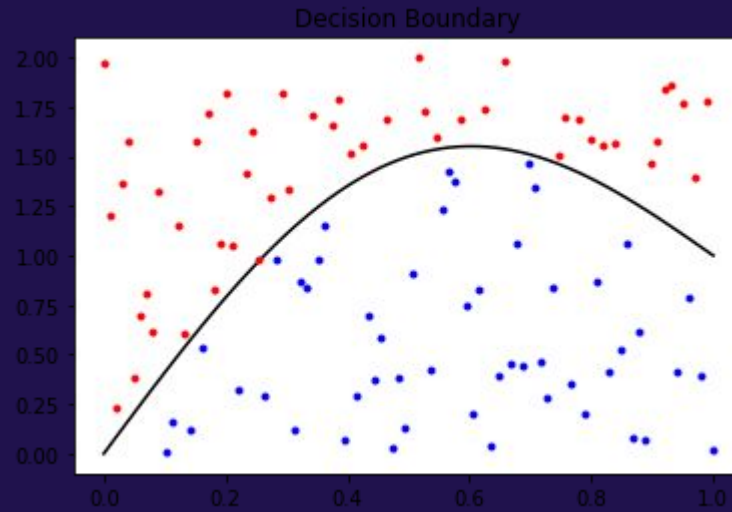
Logistic regression, a probabilistic model

$$p(\mathbf{y} = 0 \mid \mathbf{x}; \theta) = \frac{1}{1 + \exp(\theta^\top \mathbf{x})}$$

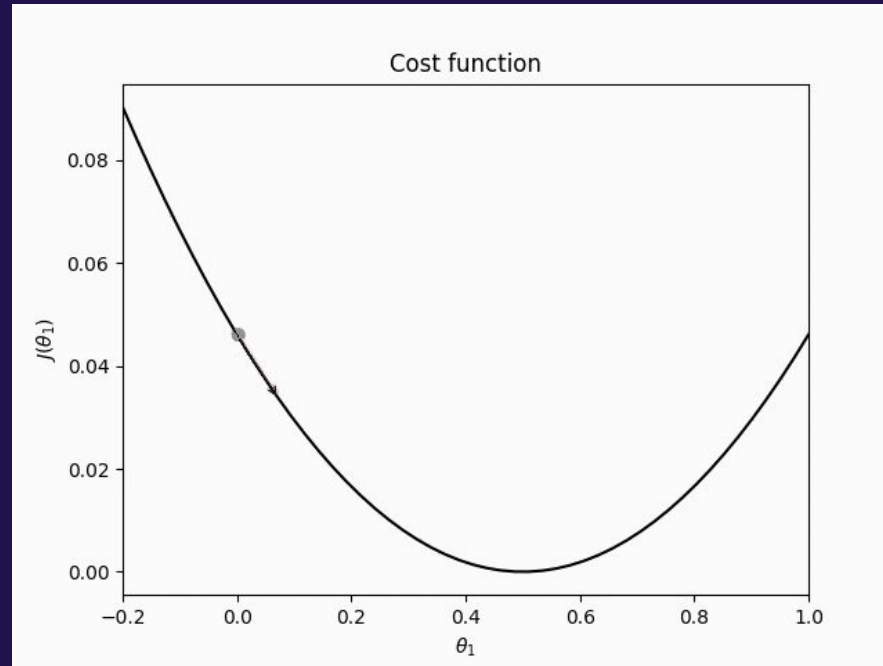
$$p(\mathbf{y} = 1 \mid \mathbf{x}; \theta) = \frac{\exp(\theta^\top \mathbf{x})}{1 + \exp(\theta^\top \mathbf{x})}$$

Model

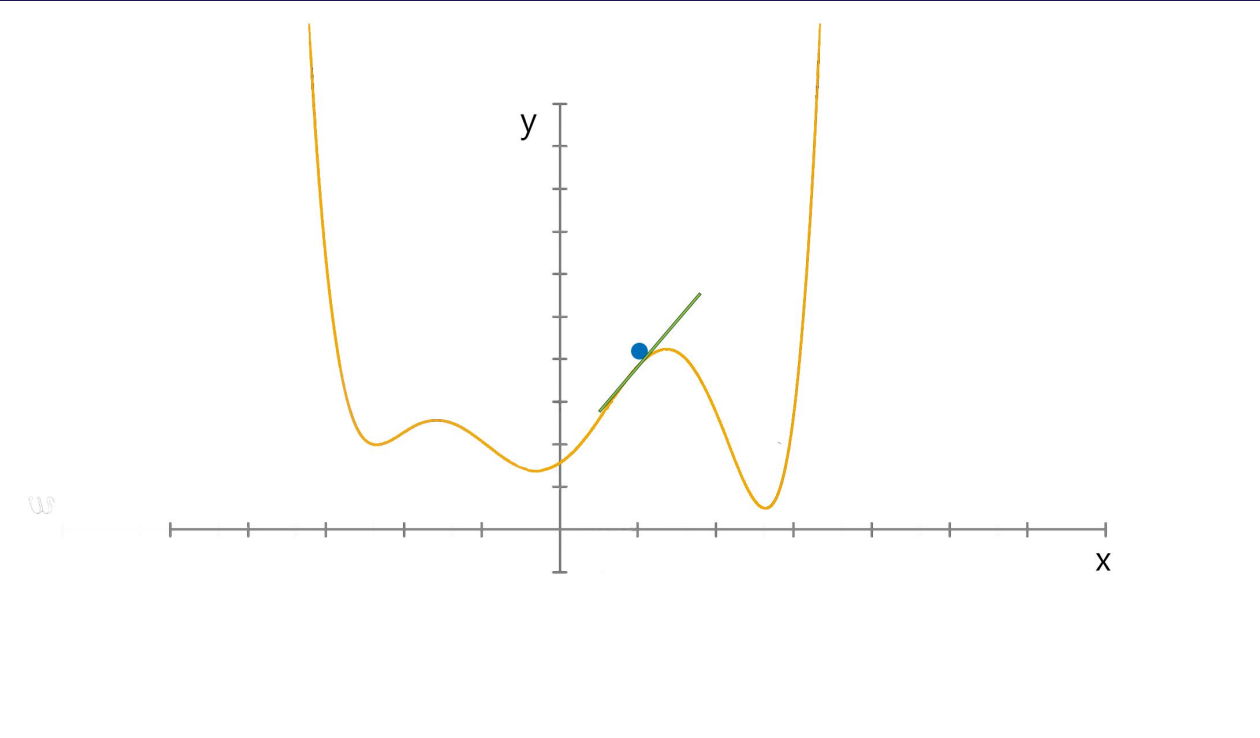
What function to choose? Non-linear classifier



Loss function



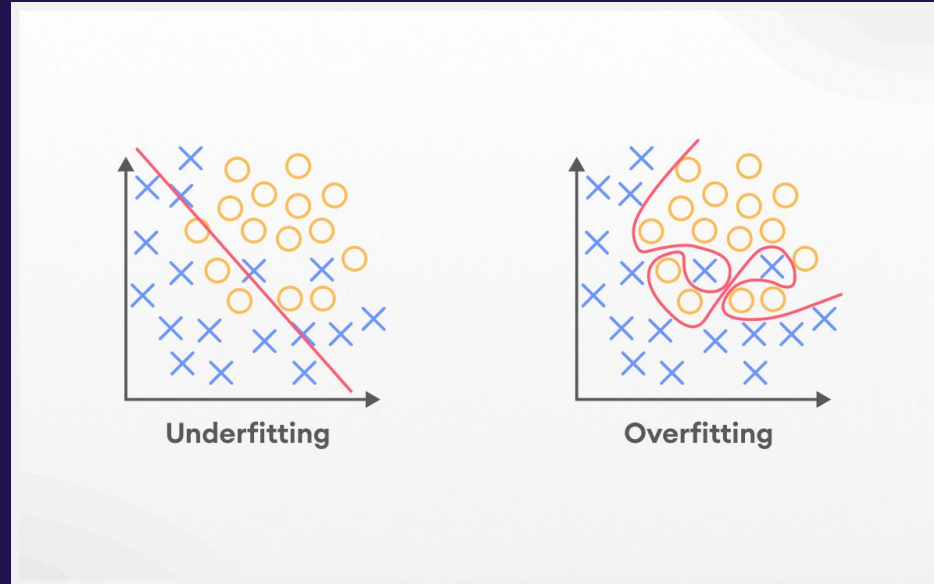
Loss function



Generalization errors

Robust to noise in the training data

Generalizes to new examples



tack շնորհակալութիւն Danke
ありがとうございました Salamat
grazie **Thank you** mulțumesc
धन्यवाद நன்றி
Terima kasih Dankie 감사합니다 Merci
Спасибо شكرا جزىلا σας ευχαριστώ
teşekkür ederim 谢谢 cảm ơn bạn