Machine Learning Lecture 1

Machine learning is everywhere

 Search, content recommendation, image/scene analysis, machine translation, dialogue systems, automated assistants, game playing, sciences (biology, chemistry, etc), ...



Machine learning: what is it?

A brief definition

Machine learning as a discipline aims to design, understand and apply computer programs that learn from experience (i.e., data) for the purpose of modeling, prediction, or control

Prediction problems

About future events

Market value

Time

Also collision avoidance, monitoring, medical risk, etc.

Prediction problems

About properties we don't yet know



would I like this movie?



soluble in water?



what is the image about?

"ML is very cool"

what is it in Spanish?

- It is easier to express tasks in terms of examples of what you want (rather than how to solve them)
- E.g., image classification (1K categories)

Image Image

Category

mushroom

cherry

- It is easier to express tasks in terms of examples of what you want (rather than how to solve them)
- E.g., image classification (1K categories)



 Rather than specify the solution directly (hard), we automate the process of finding one based on examples

- It is easier to express tasks in terms of examples of what you want (rather than how to solve them)
- E.g., image classification (1K categories)



 Rather than specify the solution directly (hard), we automate the process of finding one based on examples

- It is easier to express tasks in terms of examples of what you want (rather than how to solve them)
- No limit to what you can learn to predict...

| English | | <u>Spanish</u> |
|---------------------------------|---|---------------------|
| $h($ Is it real? ; θ $)$ | — | ¿Es real? |
| Will it continue? | | ¿Continuará? |
| For how long? | | ¿Por cuanto tiempo? |
| | | |

Already in production for some language pairs (Google)

A concrete example

Learning to predict preferences from just a little data...

In April 1945, the Allee are making that fire outh in the Biogeon transfer. A battle haddnet Arry engages transmission for "Nacional QC Caller (that FBB, Hedrag a chemical transmission transmission and enderstates a dealey invited interfacementy lines. However, a restriction of solar Again Lemma in their risks are interfaced a solar Again Lemma in their risks and the investigations. Some cancel along others are the investigation at Machine T of Nacional Generals.



In Contragge, MEL, Service Texes, Varia Scienter Heise, Davino (Der Afflicht) and Hug Generoscieuwich Antre Reasonant: Place growent is contract of a biland at managest the poster. Everywent, when a variage set managest the coupling Mith resolution and the statistical set with a solid measure in the decooperation. The vector generation and in texp and have a solid real demonstration of a binary and the solid set of the solid posterior and the solid measure and when you and the solid means of the solid measure about when you and they solid posterior about the view and they solid posterior about the you and they solid posterior



Instantion follow, a global snaphiliphi antiferented Desi Even ano similar tradition the planet arbitrationable for dense there. Defined catego antifere status physicistic is working to plane to state tradition by transporting for planet possible in the new tradition by protocology (particular brock response) and to state of transporting (particular brock response) and a state of transported (particular brock response) and a state of transported to any state the planet scale bereast off a new lower.



Hoodboop redgement piges retter, i was in the ony of Save Transfers was to include instrume Transfer (http:/disease), somework in the one of the piges. These developes the pigesen in the one of backet, these developes the pigesen in the one of backet. The save are included in the other than the transfer for which of a develope and the other than the formation of a developed by the pigestations for which of a developed by the transfer them that the other than the transfer that the other than the formation of the other than the other than the formation of the other than the other than the other of high term thereas.



In Juan 1795 the record prophers at Electrony Yan, Biosen 1795 percent listed 12, set toring their were characterized and the set of the set of the set interference factor 1, the toring their were characterized listed 14, and the set of the s



A concrete example

Learning to predict preferences from just a little data...

encapit tiers, the alive any making their first purchase the European theater. A better hardware any mergene is exerned box "Nearing" Callie (that Picz), Heating a chemical box "Nearing the child beauty incomentation a body mixtude bodied over yields incomentation a body window bodied over yields in the second second bodies (the second over the second second second over the second over the second second second second second second over the second second second second second over the second seco



-1

In Contrapp, ME, Sentor Henry Factoreal writes Hole Durano (Dar Afflach) and Hall (gamessearcheal Anny Reasonant: Ners present a sourcest of a biologial memogene the public, however, when Anny Sentor Hall Alexan and the public sectoring and investory Hall Alexan and the public sectoring present and the depotentiation of the mean of the public media feasor planes. The Health application and the public media feasor planes the Depotent in language and the cost and Any Alexano.



-1

In Earth's future, a global implifield and second Deep Even an elevely repetition the planet printedhalo Inviews there (Derived Card), a vollare water, physican, is working or plane to seve transform by its report for plant's population to a new horizon by processoor (Lasthan Aucustican) control that the processoor (Lasthan Aucustican) control that the planet to one (Lasthan Aucustican) and a second the mean of a new lower.



+1

Hoodhologineed pright retain() nets in the one of Service Instantiages have to introduce therman Traditional Hards officered comparison to Departure Tradition (Molt, a topol embode one programme to the force of program. Therma Comparison (Instantia Control Traditional Retains Comparison (Instantia Control Traditional Retains and instantia Retain). Active Tradition (Instantia Retains and Control Retains and Ambient Control Department). Allowed (Control Traditional Ambient Department) and the Retains and Ambient Ambient Department Retains.



+1

Is based 1992 the excel process all thinking have, biocaring present lateous, set torget the were determined by the excel problem of the excel to based 1992, the excel problem all thinking have indifference face of the the torget tori were relatively there is be exceleded of the excel to problem the torget tory were and the all the torget of the excel torget to provide the torget torget torget to promote the torget torget to compose of the excel torget to provide the torget torget to the promote torget torget to the torget of the excel torget torget to the torget torget to the torget torget to the torget to the torget torget to the excellence to the torget torget to the torget to the torget torget torget torget to the torget to the torget torget torget torget to the torget torget torget to the excellence torget torget torget to the torget torget to the torget torget torget to the torget torget to the torget torget torget torget torget to the torget torget torget torget to the torget torget torget to the torget torget torget torget to the torget torget torget torget to the torget torget torget torget torget to the torget torge



?, ?,

A concrete example

Learning to predict preferences from just a little data...



Supervised learning

Learning to predict preferences from just a little data...



Supervised learning



Supervised learning: training set



Supervised learning: test set



Supervised learning: training set















Supervised learning: generalization



Supervised learning +

 $h\left(\begin{array}{c} \\ \end{array}\right) = \text{politics} \qquad h: \mathcal{X} \to \{\text{politics}, \text{sports}, \text{other}\}$

Multi-way classification (e.g., three-way classification)

Regression

$$h\left(\begin{array}{c} \blacksquare \end{array}\right) = \$1,349,000 \qquad h: \mathcal{X} \to \mathbb{R}$$

Structured prediction



A group of people = shopping at an outdoor market

$$h: \mathcal{X} \to \{\text{English sentences}\}$$

Types of machine learning

- Supervised learning
 - prediction based on examples of correct behavior
- Unsupervised learning
 - no explicit target, only data, goal to model/discover
- Semi-supervised learning
 - supplement limited annotations with unsupervised learning
- Active learning
 - learn to query the examples actually needed for learning
- Transfer learning
 - how to apply what you have learned from A to B
- Reinforcement learning
 - learning to act, not just predict; goal to optimize the consequences of actions
- Etc.

Key things to understand

- Posing supervised machine learning problems
- Supervised classification
- The role of training/test sets
- A classifier
- A set of classifiers
- Errors, generalization