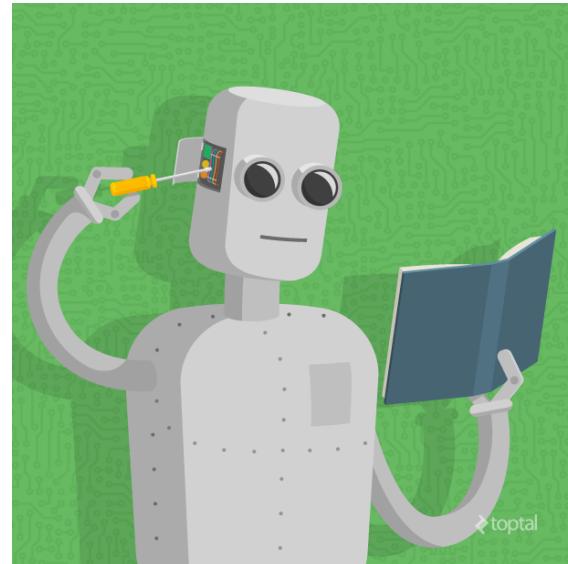


Introduction to Machine Learning



 Representently

Joseph

Agenda

- **What is Machine Learning**
- **How is Machine Learning Used**
- **Prediction Problem Types**
- **Model Types**
- **Machine Learning in Represently**
- **Basic Coding Implementation**



What is Machine Learning?

- **Machine Learning** is "giving computers the ability to learn without being explicitly programmed".
- (Thanks Wikipedia but also Arthur Samuel in 1959)

- What does that mean (to you)?



What is Machine Learning?

- **Machine Learning** is “giving computers the ability to learn without being explicitly programmed”.
- (Thanks Wikipedia but also Arthur Samuel in 1959)

- What does that mean (to you)?
- Coding the *conditions* for something to occur rather than explicitly saying that given thing is going to occur

Machine Learning Examples

- You've encountered machine learning. Give an example of machine learning you've encountered.
- For the example you've provided, explicitly state what is “being learned,” and what data is being used to uncover that insight.

Machine Learning Examples



- Netflix Recommendation Engine
- What is being predicted?

NETFLIX

Machine Learning Examples

□ Netflix Recommendation Engine

- What is being predicted?
- *Movies a user may like.*
- What data is being used?



Machine Learning Examples

□ Netflix Recommendation Engine

- What is being predicted?
 - *Movies a user may like.*
- What data is being used?
 - *Trickier! Data about movies **and** users.*
- Netflix uses collaborative filtering (rather than content only) for their algo:
the recommendations are based on other user preferences.



Machine Learning Examples

- The Veteran Affairs Website for Claims
- What is being predicted?



Machine Learning Examples

- The Veteran Affairs Website for Claims

- What is being predicted?
- *Potentially, the time it takes to process a patient claim*
- What data is being used?



Machine Learning Examples

- The Veteran Affairs Website for Claims

- What is being predicted?
- *Potentially, the time it takes to process a patient claim*
- What data is being used?
- *Previous claims data, segmented by claim type*



Machine Learning Examples

□ The Veteran Affairs Website for Claims

- What
- *Potential* to
- *process*
- BUT WAIT: How does this help the end user?
- What data is being used?
- *Previous claims data, segmented by claim type*



Machine Learning Examples

- **Takeaways:**
- 1. ML is ubiquitous, and only going to increase in prevalence through all industries
- “What we have seen is that once you implement AI successfully into an industry and into a sector it’s impossible to compete without it.”
- e.g. web search



Machine Learning Examples

- **Takeaways:**
- 2. ML relies on a high amount of robust training data
- Thus, successful implementation of ML relies on smart, early data collection.
- When you are building a service, you should ask what data you want from day zero for your ML algos of the future.

Problem Types

- Machine learning problems can be neatly segmented into a few types:

- Classification

- Regression

} Mutually exclusive

- Supervised

- Unsupervised

} Mutually exclusive

Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
- **Classification** – What are classification problems? Class guesses?

Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
- **Classification** – What are classification problems? Class guesses? (This was a pun)
- **Classification** is when we are predicting a categorical-based output.
- e.g. What color of shirt will Mark Zuckerberg wear tomorrow?; What issues will a user want to receive an email alert about



Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
- **Regression** – What are regression problems?

Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
 - **Regression** – What are regression problems?
-
- **Regression** is when we are predicting a numeric, continuous output.
 - e.g. What is the price of this home? How many impressions do I predict this given issue campaign ad spend will reach?

Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
- **Regression vs Classification**

Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
- **Supervised** – What is a supervised model?

Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
- **Supervised** – What is a supervised model?

- **Supervised machine learning** is when you utilize labeled target data for the purposes of training your model.
- e.g. These issues describe current Democrat vs Republican users. When I see these attributes in the future, are they Democratic or Republican?



Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
- **Unsupervised**– What is an unsupervised model?

- **Unsupervised machine learning** is when you do not use labeled data for the purposes of prediction.
- e.g. These attributes describe this group of users, while this other set of attributes describes this group of users. (Democrats/Republicans MAY cluster together based on similarity.)

Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
- **Supervised vs Unsupervised**

Problem Types

- Machine learning problems can be neatly segmented into a few types:
 - Classification
 - Regression
 - Supervised
 - Unsupervised
-
- Describe a situation that is a **supervised, regression** problem.

Problem Types

- Machine learning problems can be neatly segmented into a few types:
- Describe a situation that is a **supervised, regression problem**.

The screenshot shows a Zillow listing for a condo at 3101 New Mexico Ave NW APT 826, Washington, DC 20016. The listing includes a large image of the exterior, interior photos, and a map showing nearby landmarks and price points. Key details listed are:

- FOR SALE \$660,000**
- Zestimate*: \$659,120
- EST. MORTGAGE \$2,465/mo**
- Get pre-qualified**
- 2 beds • 1 bath • 1,653 sqft**
- Description: Stunning two level contemporary condo in Wesley Heights. This 2 bed, 2 bath contemporary home was built in 1979 and features double height ceiling living room, view from private balcony, and hardwood floors. A huge master BR with a large walk-in closet and bathroom are complimented by a 2nd BR with 2nd balcony and walk-in closet. W/D located in unit. The pet-friendly complex also has 2 gar pk sps.

A blue arrow points from the right side of the listing towards a contact form on the right, which is part of a sidebar for "Saved Homes". The sidebar includes fields for "Your Name", "Phone", "Email", and a message box stating "I am interested in 3101 New Mexico Ave NW APT 826, Washington, DC 20016.". It also features a photo of a real estate agent named Donald Gay and a checkbox for "I want financing information".

Problem Types

- Machine learning problems can be neatly segmented into a few types:
- Classification
- Regression
- Supervised
- Unsupervised
- Describe a situation that is an **unsupervised, classification** problem.

Problem Types

- **Machine learning problems** can be neatly segmented into a few types:
- Describe a situation that is an **unsupervised, classification** problem.



Inspired by your shopping trends



Model Types

- **There are a high number of models available,** and each is worthy of its own lecture.
- Demonstrating the vast number of models is meant to expose the number of options available, rather than be a comprehensive discussion of each.

Model Types

- To name and describe a few:
 - Linear Regression
 - Logistic Regression
 - K-Nearest Neighbors
 - Support Vector Machines
 - Naïve Bayes
 - Decision Trees + Random Forests
 - Ensembling Models
 - Neural Networks: recurrent, convolutional, ...



Machine Learning at Representently

- How can (and will) Representently benefit from machine learning?
- Consider: what data is required for your suggestion; what type of problem are you solving in your suggestion; what is the value of this feature to the end user
- This is a strategist *and* developer question.





Represently

Let's do some basic coding.

