Introduction to Machine Learning

Introduction

Varun Chandola

March 8, 2021

Outline

Contents

1 Making Intelligent Machines

`https://quickdraw.withgoogle.com/`

Artificial Intelligence or AI has been getting immense publicity in recent years, mostly due to advances in deep learning methods and their applications to areas such as self driving cars, robotic assistants, and language translation systems. The area of AI is more than half a century old, when computer scientists and mathematicians embarked on the quest for making intelligent machines. Now that idea pervades or threatens to pervade every aspect of our daily life. Read an interesting article on AI in "The Great A.I. Awakening" in New York Times, December 2016.

1. Talk. See. Hear.
   - Natural Language Processing, Computer Vision, Speech Recognition
2. Store. Access. Represent. (Knowledge)
   - Ontologies. Semantic Networks. Information Retrieval.
3. Reason.
   - Mathematical Logic. Bayesian Inference.
4. **Learn.**
   - Improve with Experience
     - Machine Learning

2 Human Learning

- What do we learn?
  - Concepts (this is a chair, that is not a chair)
  - Distinguishing concepts (this is a chair, that is a table)
  - Other things (language, juggling, using a remote)
- How do we learn?
  1. Teaching (Passive).
  2. Experience (Active).
     (a) Examples.
     (b) Queries.
     (c) Experimentation.

3 Definition of Machine Learning

- Computers learn without being **explicitly programmed**.
  - Arthur Samuel (1959)
- A computer program learns from experience E with respect to some task T, if its performance P while performing task T improves over E.
  - Tom Mitchell (1989)

3.1 Why Machine Learning?

One could argue as to why one would want a machine to "learn" over experience instead of programming a machine to "know" everything. But such approach is infeasible. First, because knowing everything will require a significant amount of storage. Second, the success of such a system assumes that the creator already knows everything about the problem, which often is an optimistic and false assumption.

Machine learning is implicitly based on the notion that the new experiences that a machine will encounter will be similar to the past experiences. Hence, instead of "pre-loading" the machine with knowledge about all possible experiences, it is efficient to selectively learn from these experiences. Moreoever, as long as the new experience has some structural relationship with the past experiences, the machine can perform well in unseen situations as well.

- Machines that know everything from the beginning?
  - Too bulky. Creator already knows everything. Fails with new experiences.
- Machines that learn?
  - Compact. Learn what is necessary.
  - Adapt.
  - Assumption: Future experiences are not too different from past experiences.
    - Have (structural) relationship.

4 Learning from (Past) Data

Deductive Logic

- All birds can fly
- Dodo is a bird
- $\Rightarrow$ Dodo can fly

Inductive Logic

- A stingray can swim
- Stingray is a fish
- $\Rightarrow$ All fish can swim

Core Tenet

- ~~Deduce~~ Induce from past
- Generalize for future

5 Overview of ML

Supervised Learning

- Given a finite set of $\mathbf{x}$'s and corresponding y's, **learn** $f()$
- **Infer** y for a new $\mathbf{x}$
  - y - continuous (regression)
  - y - discrete (classification)

Unsupervised Learning

- Given only $\mathbf{x}$'s, infer structure in data

– hidden (latent) relationships among the objects

- e.g., clustering, embedding, dimensionality reduction, etc.

Reinforcement Learning

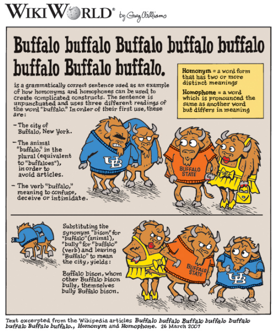
- Find the best mapping of situations to actions to maximize a numerical reward
- Agent learns to behave in an environment

6 Supervised Learning

- Data: A collection of data objects
  - A representation of the object of interest or the state of the target system, and/or
  - A label or a target value or a low-dimensional representation or an embedding or an action associated with the object/state
- Mother nature generates this data using an unknown generative process (secret recipe)
- ML problem - given part of the information, infer the other part

Primary

- Given an individual's data (criminal and otherwise), predict risk of recidivism
- Given an email, determine if it is spam or normal



- Given an image, identify the object
- For a given day, predict the number of travelers that will pass through a subway station

Secondary

- Color a black and white image
- Translate English text to French
  - `translate.google.com`

True Test for Any Language System

- Buffalo buffalo Buffalo buffalo buffalo buffalo Buffalo buffalo.

- Buffalo buffalo, Buffalo buffalo buffalo, buffalo Buffalo buffalo.

- An agent, operating in a changing environment, finds the best sequence of actions that maximize an end goal

- Key concepts: policy, reward, end-value, environment
- Examples
  1. Learning to play a game (Breakout, AlphaGo)
  2. A robotic vaccum cleaner figures out the best time to recharge
  3. Almost all robotic tasks
  4. Traffic light control

- `https://www.youtube.com/v/V1eYniJ0Rnk`

ML as a generator - faces

Generating faces

- `https://thispersondoesnotexist.com`
- `https://www.whichfaceisreal.com`

ML as a generator - text

You will rejoice to hear that no disaster has overtaken you. If your heart is grieved, and your joy is swallowed up, you have the LORD to thank. You will rejoice to hear that no disaster has accompanied the commencement of an enterprise which you have regarded with such evil forebodings.

ML as a decision maker - criminal justice

The above image is from a study done by Propublica in 2017 on the inherent bias in a widely used ML based criminal justice software called COMPAS. The study[1] showed that the COMPAS software, which is increasingly being used in the criminal justice system across the country, was biased in assigning a risk of recidivism against black defendants.

References

[1] https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing