Introduction to Machine Learning

CMSC 176

Rubelito Abella

Machine Learning Milestones

1950

Alan Turing devises the Turing test. A computer passes the Turing test if it is indistinguishable from a human

1952

Arthur Samuel wrote a program that plays checkers, learning and improving the more it plays

1957

Frank Rosenblatt designed the first neural network. The architecture is based on the neurons of the brain

Machine Learning Milestones

1967

The algorithm, “nearest neighnour” was written, allowing computers to solve problems based on learned patterns

1985

Terry Sejnowski builds the “Net Talk” a computer that learns to pronounce words similar to a baby

1997

IBM’s Deep blue defeats chess grandmaster Kasparov

Marr 2016

Machine Learning Milestones

2011

IBM's Watson defeats human competitors at the game show "Jeopardy"

2014

Facebook develops DeepFace, an algorithm that is capable of recognizing faces on photos on the same level of a human

2016

AlphaGo defeats 9-dan Lee Sedol in a professional game without handicaps

Marr 2016

the field of study that gives computers the ability to learn without being explicitly programmed

Samuel cited by Ng Retrieved 2017

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E

Mitchell cited by Ng Retrieved 2017

Formal Definition of Machine Learning

Playing chess

T = playing a game of chess vs an opponent

E = the experience of playing chess with an opponent

P = probability that the next game will be won

Predicting the amount of rainfall tomorrow

T = predicting the amount of rainfall tomorrow

E = historical weather records

P = variance of predicted rainfall from actual rainfall

Formal Definition of Machine Learning

Detecting Spam Emails

- **T** = ?
- **E** = ?
- **P** = ?

Investing on the most profitable stocks

- **T** = ?
- **E** = ?
- **P** = ?

Formal Definition of Machine Learning

Detecting Spam Emails

T = marking an email as spam

E = records of emails classified as spam and not classified as spam by users

P = accuracy of marking an email as spam

Investing on the most profitable stocks

T = selecting the stock that would yield the most profit

E = stock records and trends

P = amount of profit gained from investing a stock

Supervised Learning

There is an idea for the output of a given database

The relationship between the dataset and the output is learned

Unsupervised Learning

The dataset doesn’t come with the output

The similarities, differences, and structure of the data is learned

How many strains of a new virus exists

Given a set of medical records for each patient, how many strains of virus X exists?

Unsupervised Learning

How many strains of a new virus exists

Given a set of medical records for each patient, how many strains of virus X exists?

Unsupervised Learning

- How many strains of a new virus exists
  - There is no idea on how many strains exists based on the dataset but we can learn how many strains there are by studying the differences and similarities in symptoms

Predicting the amount of rainfall tomorrow

The dataset is the set of past daily records (temperature, cloud cover, humidity and etc. from the previous date) with corresponding amount of rainfall

Supervised Learning

Predicting the amount of rainfall tomorrow

- The dataset is the set of past daily records (temperature, cloud cover, humidity and etc. from the previous date) with corresponding amount of rainfall

Supervised Learning

Predicting the amount of rainfall tomorrow

The relationship between yesterdays temperature, cloud cover, humidity and etc. and the amount of rainfall tomorrow is learned

Supervised Learning

Natural Selection (Which species is the fittest)

The dataset is the generations of species with different genes and mutations. The output is which generation produces the most surviving offspring

Regression Problem

The output value of the supervised predictor is a continuous value

Classification Problem

The output value of the supervised predictor is a discrete value