Introduction to Machine Learning

TIET, PATIALA

Introduction

- Machine Learning (ML) techniques are widely used nowadays in number of applications.

- Social media such as Facebook, Twitter, Instagram, etc. use machine learning algorithms for custom feeds, recommendations, auto photo tagging, etc.

- Various Email providers classify Emails as spam or ham using machine learning algorithms.

- Self customization programs for applications like Amazon, Netflix, YouTube uses machine learning for recommendations.

- Search engines rank pages using various learning algorithms.

Why ML is so prevalent?

- ML grew out of Artificial Intelligence that is as intelligent as human mind.

- New capability for computers that has touched many aspects of industry and basic science.

- Includes set of algorithms that mimic the learning process of human brain.

ML-DEFINITIONS

- Arthur Samuel (1959) coined the term machine learning and gave an informal definition of machine learning.

- Arthur defined machine learning as:

  Machine Learning is the field of the study that gives computers

  the ability to learn without being explicitly programmed.

ML-DEFINITIONS

- In 1950s Arthur wrote checkers playing program (though he himself was not a good player)
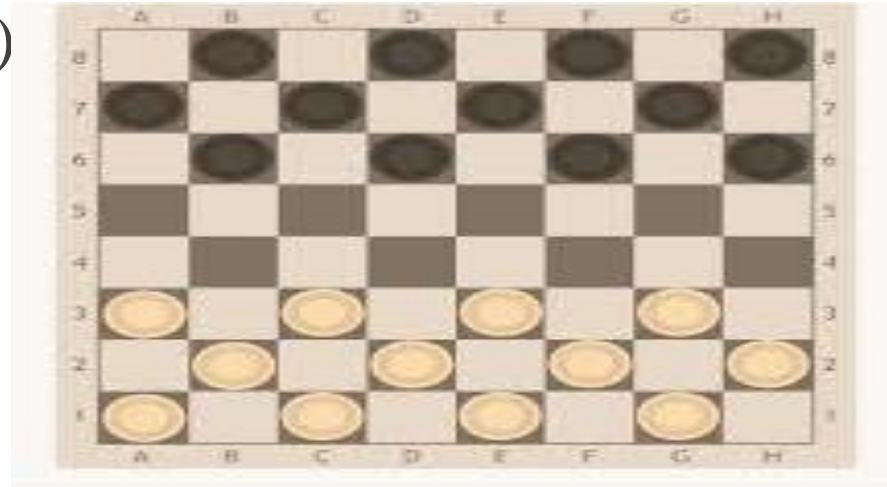
- He played tons of thousands of games against himself.

- The program learnt over time what are good and bad board positions, and what are the wining or losing positions.

ML-DEFINITIONS

- ***Tom Mitchell (1998),*** redefined the concept of Machine Learning as a Well Posed Learning Problem.

- He defined Well-Posed Learning Problem as:

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

Understanding T, E, and P

- For checkers player example:

    Task T: playing checkers,

    Experience E: having played tens of thousands of games,

    Performance measure P: probability to win the game against a new opponent.

Understanding T, E, and P

- Task (T): Stock Price Prediction / Housing Prices Prediction / Bitcoin price Predictor

- Experience (E): Tens of thousands of examples of previous years stock prices or housing prices based on certain factors (number of rooms, area, number of floors, crime rate, etc.)

- Performance (P): how close are the predicted prices to the actual prices measured in terms of number of metrics such as mean error, root mean square error, etc.

Understanding T, E, and P

- Task (T): Image Segmentation / Breast Cancer Detection / Cartoonify or Emojify an image

- Experience (E): Tens of thousands of images with labeling of different objects (for image segmentation), ROI area (for breast cancer detection), emojis or cartoons according to different moods in image (for cartooning or Emojify).

- Performance (P): whether or not the image is properly segmented or classified measured in terms of number of metrics such as precision, recall etc.

Designing a Learning System

In order to design a learning system, some of the basic design principles and approaches must be followed:

➤ Choosing the Training Experience

➤ Choosing the Target Function

➤ Choosing a Representation for the Target Function

➤ Choosing a Function Approximation Algorithm

➤ The Final Design

Choosing the Training Experience

➢ Whether the training experience provides direct or indirect feedback.

❑ Direct training examples for playing checkers may include individual checkers board states and the correct move for each.

❑ Indirect training examples may include move sequences and final outcomes of game.

➢ Degree to which the learner controls the sequence of training examples.

❑ Learner may be provided training experience by a random process or learner collects training examples by exploring its environment.

➢ How well the training experience represents distribution of examples over the final system performance.

Choosing the Target Function

- Reduce the learning task to the problem of discovering an ideal target function (function approximation) that performs the desired task.

- For instance, for a checkers player program, we need to define a function:

$$V: B \rightarrow R$$

- V maps any legal board state from the set B to some real value.

- The target function V is intended to assign better score to better board states.

Choosing a Representation of Target Function

- We must choose a representation that the learning system will use to describe target function.

- We may allow the program to represent V as a large table with a distinct entry specifying the value for each distinct board state.

- We can also specify it as linear, quadratic or polynomial function of pre-defined board features.

- For example, $V(b) = w_0 + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + w_6x_6$

where wi represents weights and xi corresponds to board positions.

Choosing a Function Approximation Algorithm

In order to learn the target (approximation function), we need methods for the following:

1. Estimating Training Values:

   ➢ Each training example, is an ordered pair of the form $(b, V_{train}(b))$; where b is a particular board state and $V_{train}(b)$ is the corresponding training value.

   ➢ These training values need to be estimated in case of indirect training experience.

Choosing a Function Approximation Algorithm Contd….

2. Adjusting the weights

➤ The learning algorithm must choose the weights w_i to best fit the training examples(b,V_{train}(b)).

➤ One common approach is to define the best hypothesis, or set the weights, as that which minimized the squared error E between the training values and the values predicted by the hypothesis V.

➤ Several algorithms exist for finding weights of function that minimize E.

➤One such example is Least Mean squares (LMS) training rule. For each observed training example it adjusts the weight a small amount in a direction that reduces E.

Choosing a Function Approximation Algorithm Contd….

- LMS Updation Rule:

For each training example (b, V_{train}(b));

➢ Use the current weights to estimate $V^{\hat{}}(b)$.

➢ For each weight wi, update $w_i \leftarrow w_i + \alpha$ (Vtrain(b)-$V^{\hat{}}(b)$) x_i

Issues in Machine Learning

The field of machine learning is concerned with answering the questions related to design of learning system:

- What learning algorithms exist for learning general target function from specific training examples?

- In what settings will particular algorithms converge to the desired function, given sufficient data?

- Which algorithms perform best for which types of problems and representations?

- How much training data (examples) is sufficient for a particular model?

- When and how can prior knowledge held by the learner guide the learning process of generalizing from examples?

Issues in Machine Learning Contd…..

- What is the best way to reduce the learning task to one or more function approximation problems? In other words, what specific functions should system attempt to learn?

- Can the process of learning target function be automated?

- How can the learner automatically alter its representation to improve its ability to represent and learn the target function?

Applications of Machine Learning

Machine Learning is used in a number of engineering, applied sciences, life sciences, and mundane tasks. Broad applications of ML are :

- **Applications that can't be programmed by man**: Automatic driving, handwriting recognition, NLP, Computer Vision

- **Self Customization Programs**: Amazon, NetFlix recommendations.

- **Data mining :** web click data, medical records, biology, engineering

- **Understanding human behavior:** brain, real AI