Intro. to Machine Learning

Abdullah Karasan
Author of ML for Financial Risk Management with Python
Senior Data Science Consultant
Lecturer at University of Maryland

March 17, 2022

Outline

Definition of Artificial Intelligence

- After 1950s, scientists focused on whether computers could be made to "think". This plants the seed of the birth of **Artificial Intelligence**
- AI is refered to as the effort to automate intellectual tasks normally performed by humans.

Definition of Artificial Intelligence-Ctd

To Nick Bostrom, there are two important revolutions in the history of mankind:

- Agricultural Revolution
- Industrial Revolution

These two revolutions have such a profound impact that any third revolution of similar magnitude would double in size of world economy in 2 weeks.

Artificial Intelligence as Third Revolution

The question now is: can AI accomplish the third revolution? Historically, there are two periods in which AI first flourished and then decline, called AI winter. These periods are:

- Dartmouth Summer Project in 1950s
- Fifth Generation Computer System in 1980s

The common setbacks bringing to a halt the progress of AI are data scarcity and hardware limitations.

Hardware Requirements

A team at DeepMind that was able to achieve breakthroughs for AI agents playing Go, Asian board game, with their AlphaGo algorithm. Here you see the hardware requirements of different Alpha algos.

Version	Year	Elo Rating[a]	Hardware	Power Consumption [TDP]
AlphaGo Fan	2015	>3,000	176 GPUs	>40,000
AlphaGo Lee	2016	>3,500	48 TPUs	10,000+
AlphaGo Master	2016	>4,500	4 TPUs	<2,000
AlphaGo Zero	2017	>5,000	4 TPUs	<2,000

What is Machine Learning?

Machine learning is programming computers to optimize a performance criterion using example data or past experience.

Less and less human guidance and more autonomous learning

Supervised Learning	Reinforcement Learning	Unsupervised Learning
Learning from the guidance: ground truths (labeled data) is given.	**Learning from the exploration: environment is given.**	**Learning with no guidance.**

Machine vs Rule-based Learning?

Rules-based AI systems exhibit "fixed intelligence", following a definite series of rules and instructions, rule-based AI systems produce pre-set outputs. Instructions are poured into rule-based systems in the form of four main components, namely-

- Facts or knowledge
- Set of rules, i.e. "IF A happens Then do B"
- Inference Engine that infers information and acts according to the interactions between input and rule base, and

When to use Rule-based Learning?

- Lack of data
- Non-complex structure such as expense approval.

Supervised Learning

- In supervised learning, we are given labeled data, i.e pairs (X, Y) and the goal is to learn the relationship between X and Y. Each observation x_i is referred to as a feature vector and y_i is the label or response.
- Linear regression, logistic regression, decision tree, RF, SVM, Boosting are some of the supervised learning models.

Unsupervised Learning

- In unsupervised learning, we are given unlabeled data, x_1, x_2, . . . , x_n and our goal is to retrieve exploratory information about the data, perhaps grouping similar observations or capturing some hidden patterns.
- K-means clustering, Gaussian mixture, and hidden Markov models are some of the unsupervised learning models.

Reinforcement Learning

- Defines a set of states and actions in response to a changing regime so as to maximize some notion of cumulative reward.
- In contrast to supervised learning, which just considers a single action at each point in time, reinforcement learning is concerned with the optimal sequence of actions. So, it requires dynamic programming.

Machine Learning-Optimization

- Model: $g(X|\theta)$
- Loss function: $\Sigma_i \mathbb{L}(r^i, g(x^i|\theta)$
- Optimization:$\theta* = \text{argmin } \mathbb{L}(\theta|X)$

$$y = a_0 + a_1 x_1 + a_2 x_2 + \epsilon \tag{1}$$

Three Pillars of Machine Learning

- Statistics
- Programming skill
- Domain Knowledge

Some Important Concepts

From this point on, we focus on AI-based financial applications. However before proceeding, it is worth mentioning the following topics:

- Learning
- Performance
- Overfitting
- Bias-Variance
- Hyperparameter Optimization
- Cross-Validation

Learning

In statistics, learning means estimation. In simpler terms, learning is nothing but finding out the best parameters from experience.

$$y = a_0 + a_1 x_1 + a_2 x_2 + \epsilon \tag{2}$$

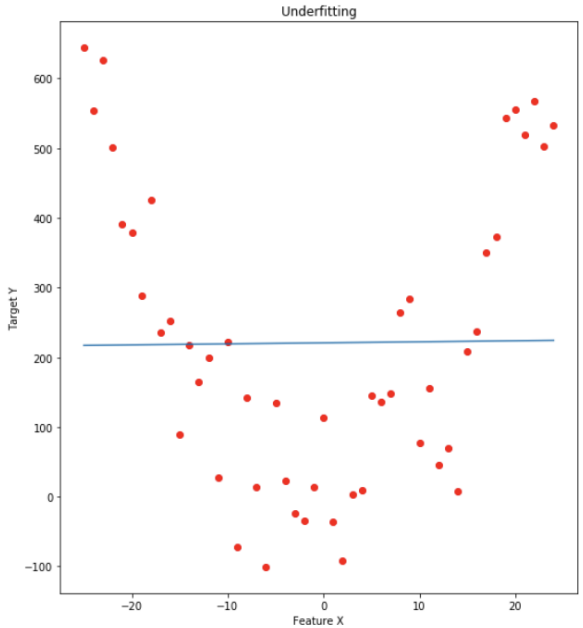
Performance

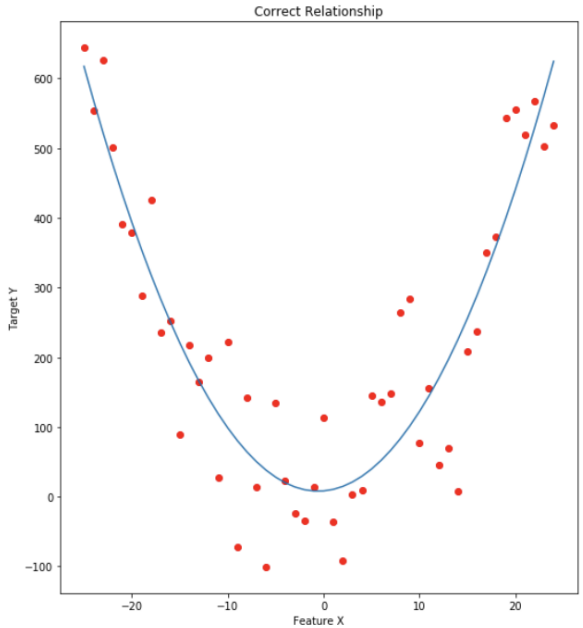
Machine learning is defined as programming computers to optimize a performance criterion using example data or past experience. So, what does performance mean?
A good fitting model is one where the difference between the actual or observed values and predicted values for the selected model is small. This can be measured by well-known metrics such as RSME, MSE, MAE.

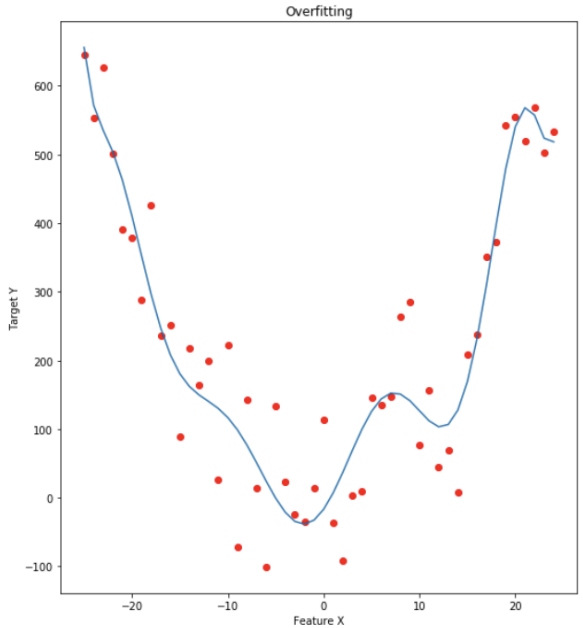
Overfitting

Overfitting and underfitting are foundational concepts in machine learning, and understanding them is crucial for designing good models. If the generalization gap of a model is larger than we want, we call this overfitting. If the training performance of a model is low, then we call this underfitting.

Overfitting







Bias-Variance Trade-Off

In the context of polynomial regression, as the order of the polynomial increases, small changes in the dataset cause a greater change in the fitted polynomials; thus variance increases. But a complex model on the average allows a better fit to the underlying function; thus bias decreases. This is called the bias/variance dilemma.

Bias-Variance-Ctd

To decrease bias, the model should be flexible, at the risk of having high variance. If the variance is kept low, we may not be able to make a good fit to data and have high bias. The optimal model is the one that has the best trade-off between the bias and the variance.

Hyperparameter Optimization

- Hyperparameter: a parameter whose values is set prior to the commencement of the learning process.
- Hyperparameter turning is an indispensable part of ML modeling in that we are looking for the optimum model parameters.

Cross Validation

- Predictive performance in ML models has the central focus. So, generalization is the ultimate aim of the machine learning model.
- Cross-validation is a tool for checking the stability of training performance.
- Cross-validation is also an essential tool for checking bias-variance trade-off.

Thank You