BEHAVOX

ARTIFICIAL INTELLIGENCE PRIMER

2020

~~Types of~~ Machine Learning Algorithms

Machine learning algorithm~~s could be divided into the following~~ categories:

1. Supervised learning~~;~~
    - Classification~~;~~
    - Regression~~;~~
2. Unsupervised learning
3. Semi-supervised learning~~; and~~
4. Reinforcement learning~~.~~

For **more information** ~~on the types of machine learning~~, watch a ~~3:43 minutes~~ video ~~explaining **supervised**, **unsupervised** and **reinforcement** types of~~ machine learnin~~g~~.

~~Definitions~~

Supervised learning is a process during which a computer tries to develop algorithms that connect inputs to outputs based on the labeled data it received during the training phase.

Unsupervised learning is a process during which a machine develops the algorithms in order to find underlying patterns (e.g. clusters of data) in an unlabeled data.

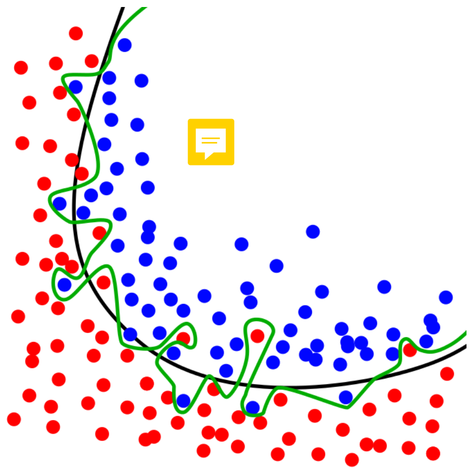
Reinforcement ~~L~~earning ~~–~~ a process ~~in a course of which~~ a computer model ~~is focused on achieving a long term goal through a series of decisions. Typically, the best next step towards achieving the goal is not clear. The model~~ implements feedback from the end results to advance its algorithms. ~~For instance, it repeat and improve the types of decisions it made if they were successful, and alters them if they were not. A~~ machine combines a reward/punishment process with a trial-and-error system to improve the results over time.

Machine Learning Challenges

The elimination of **underfitting** and **overfitting** ~~is~~ one of greatest challenge in machine learning.

Overfitting occurs when a statistical model or machine learning algorithm captures the noise of the data. ~~Intuitively, overfitting occurs when the~~ model or the algorithm fits the data too well~~.~~ ~~Specifically, overfitting occurs if the model or algorithm~~ shows low bias but high variance. Overfitting is ~~often a result of an~~ excessively complicated model~~, and it can be prevented by~~ fitting multiple models and using validation or cross-validation to compare their predictive accuracies on test data~~.~~

Picture 1. The green line represents an overfitted model and the black line represents a regularized model. While the green line best follows the training data, it is too dependent on that data and it is likely to have a higher error rate on new unseen data, compared to the black line.



Underfitting occurs when a statistical model or machine learning algorithm cannot capture the underlying trend of the data. ~~Intuitively, underfitting occurs when~~ the model or the algorithm does not fit the data well enough~~.~~ ~~Specifically, underfitting occurs if the model or algorithm~~ shows low variance but high bias. Underfitting is often a result of an excessively simple model.

Both overfitting and underfitting lead to ~~**poor predictions**~~ on new data sets.

Good fit ~~–~~ algorithm that fully understands the relationships between input and output and filters out the input points which are the outliers. Even though this model does not fit the training data as well as in the case of overfitting, it may be more successful in handling new data.

AI Technology Models and Techniques

~~There is a great number of~~ different methods and models to analyze the data for each category of machine learning. The best model ~~typically~~ depends on the problem ~~that needs~~ to ~~be~~ solve~~d~~ or the goal ~~that should be achieved.~~ In some situations, several different types of algorithms and models are mixed to address the problem.

~~The non-exhaustive list of the~~ techniques ~~used in machine learning applications is as follows:~~

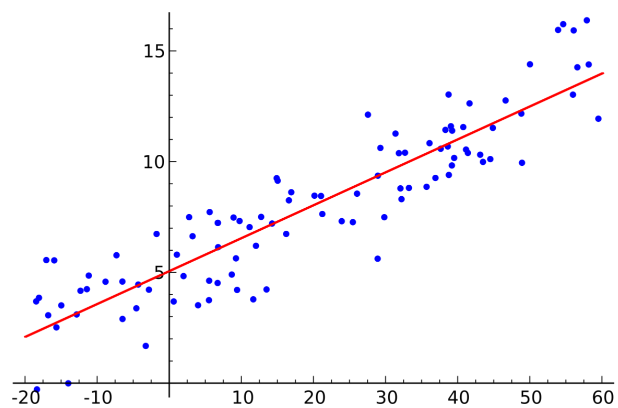
1. Linear and Polynomial Regression~~;~~
2. Bayesian/Statistical Analysis~~;~~
3. Decision Tree Learning~~;~~
4. Logistic Regression~~;~~
5. Neural Network~~; and~~
6. Classification/Clustering Algorithms: Support Vector Machines, Clustering Algorithms~~.~~

For ~~more information~~ on ~~the types of~~ machine learning techniques, watch ~~a 30:52 minutes~~ video ~~explaining the techniques listed above~~ [Resource: Udacity].

~~Definitions~~

Regression analysis: technique that ~~helps to~~ determine ~~whether~~ there is a linear relationship ~~between variables and in cases where there is one,~~ states what the relationship is. This relationship ~~between two variables~~ (developed based on the historical "training data") can then be used for forecasting the value of one variable based on the value of another.

One way to identify the relationship (or the lack of it) between the inputs in the data set is to draw a line that fits the data and then use the formula for that line to anticipate what output there will be for any of the inputs. The problem with this approach is that many of the different lines could be drawn, however, the goal is to find the optimal one that results in the least amount of error.



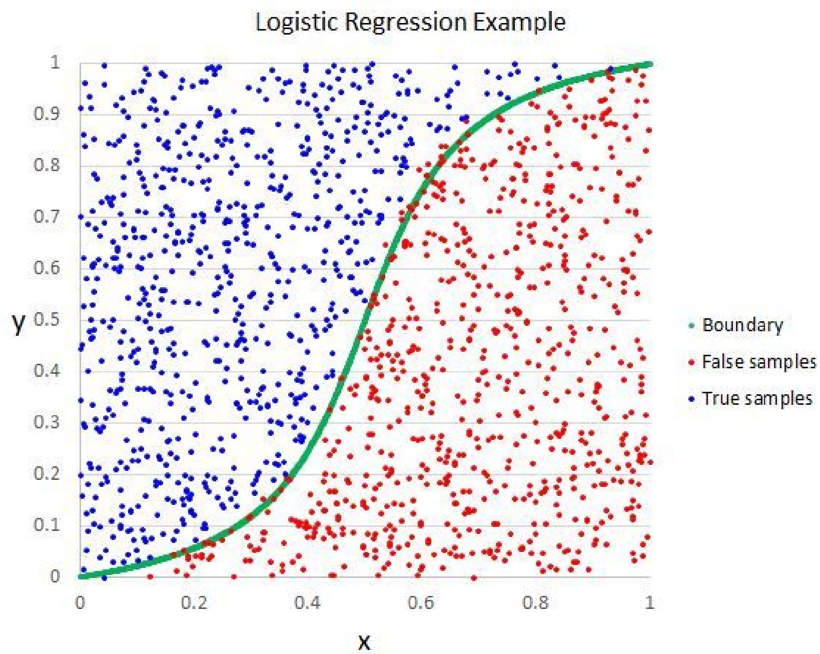
Example: the people who potentially could conspire to conduct an offense, are usually close ~~and/~~or important to each other. ~~Therefore, it could be argued that~~ the relationship between those two issues is linear: the closer and the more important a person is to a monitored employee, the greater is a potential for them to engage in a misconduct.

~~A~~ **multiple regression**: a type of a regression in which there are many inputs (several independent variables).

Example: considering the example above, it could be easily spotted that whether or not people in their communications conspire to engage in an offense depends not only on the relationship between them, but on other relevant factors, such as, for instance, a number of people in the chat (since it is less likely that a person would discuss a punishable offense in a company chat with a lot of participants).

Logistic regression: a type of a regression in which the output could be one of two or more possible discrete categories or values. Logistic regression predicts the probability of a particular output based on the inputs.

<u>Example</u>: the probability of a misconduct in a particular email could be predicted based on how the similar types of emails were dealt with previously.



Decision Tree Learning: technique that is based on the Bayes' theorem, which answers ~~“~~conditional probability~~”~~ questions. The common conditional probability questions assess the likelihood that an object belongs to a certain group based on its features and classify the object based on that. These types of questions commonly arise in machine learning applications.

Bayesian/Statistical Analysis: technique that consists of analyzing data and creating a tree structure of decisions in order to provide a predictive model for future reference.

~~“~~**Naive Bayes**: a commonly used classification technique that assigns objects to a class based on their attributes. There is a variety of these techniques based on the type of data (i.e., Gaussian Naïve Bayes is usually used for continuous data).

<u>Example</u>: an algorithm that could analyze the attributes of an incoming email based on the historical data on features of prior emails, in order to determine whether it is a blast distribution or not.

Support Vector Machines (SVMs): an algorithm that assess historical data of all of the available classes ~~and based on this~~ determines to which class a new input belongs ~~to~~. One approach to this problem is to use the line between two classes as a classifier (i.e., all data to the left is in one group, all data to the right is in another). This approach is not unimpeachable, because classifier lines could differ,~~ therefore,~~ the optimal one should be determined.

Clustering Algorithms: a technique which ~~allows~~ ~~to~~ search through data ~~in order~~ to group similar data points into clusters. This method is ~~typically~~ used for unsupervised learning or data mining applications ~~in order~~ to find relationship between data points, where there is no specific output.

For ~~**more information**~~ on Deep Learning and Neural Networks, watch ~~a 19:13 minutes~~ video ~~explaining the techniques listed above~~ [Resource: 3Blue1Brown].