



Major Steps of Machine Learning Process in Data Science Lifecycle

Leon James*

Department of Advanced Sciences, Uppsala University, Uppsala, Sweden

DESCRIPTION

Machine Learning (ML) is a branch of research concerned with understanding and developing methods that 'learn,' that is, methods that use data to improve performance on a set of tasks. It is one of the important components of artificial intelligence. Machine learning algorithms construct a model from sample data, referred to as training data, in order to generate predictions or judgments without being explicitly programmed. Machine learning algorithms are utilized in a wide range of applications, including medicine, email filtering, speech recognition, and computer vision, where developing traditional algorithms to execute the required tasks would be difficult or impossible.

However, not all machine learning is statistical learning. A subset of machine learning is strongly related to computational statistics, which focuses on making predictions using computers, but not all machine learning is statistical learning. Mathematical optimization research provides methodology, theory, and application domains to the subject of machine learning. A related topic of research is data mining, which focuses on exploratory data analysis *via* unsupervised learning. Some machine learning implementations use data and neural networks to replicate the operation of a biological brain. Machine learning is sometimes known as predictive analytics when used to commercial concerns.

MACHINE LEARNING PROCESS

Data collection

Data collection is the first step in machine learning. Collecting relevant and accurate data becomes critical since the quality and quantity of data directly influence the outcome of the machine learning model. This dataset is then utilized to train the data model, as stated in the preceding section.

Data preparation

The first step in the total data preparation process is data cleaning. This is a critical step in preparing the data for analysis.

Data preparation guarantees that there are no erroneous or corrupt data points in the dataset. It also entails converting the data to a single format. The dataset is also divided into two portions that will be utilized for training the data model and evaluating the trained model's performance, respectively.

Training the model

This is where the learning process begins. To forecast the output value, the training dataset is used. However, practice makes perfect. The training data is utilized to incrementally enhance the model's prediction accuracy.

Model evaluation

It is now time to assess its performance. The dataset set aside during the Data Preparation procedure is used in the evaluation process. This data was never utilized to train the model. So, evaluating a data model against a new dataset will provide insight into how the model will perform in real-world applications.

Prediction

Although the model has been trained and tested, this does not imply that it is perfect and ready for deployment. The model is improved further by adjusting the parameters. The ultimate phase in machine learning is prediction.

CONCLUSION

Machine learning programs involve computers learning from available data in order to do specific tasks. For simple jobs handed to computers, it is possible to build algorithms that instruct the machine on how to execute all steps required to solve the problem at hand; no learning is required on the computer's behalf. It can be difficult for a human to manually build the necessary algorithms for more complicated tasks. In fact, it may be more useful to assist the computer in developing its own algorithm rather than having human programmers specify every required step.

Correspondence to: Leon James, Department of Advanced Sciences, Uppsala University, Uppsala, Sweden, E-mail: leon.james@uppsala.se

Received: 17-Aug-2022, Manuscript No. SIEC-22-18412; **Editor assigned:** 19-Aug-2022, Pre QC No. SIEC-22-18412 (PQ); **Reviewed:** 05-Sep-2022, QC No SIEC-22-18412; **Revised:** 13-Sep-2022, Manuscript No. SIEC-22-18412 (R); **Published:** 22-Sep-2022, DOI: 10.35248/2090-4908.22.11.272.

Citation: James L (2022) Major Steps of Machine Learning Process in Data Science Lifecycle. Int J Swarm Evol Comput. 11:272.

Copyright: © 2022 James L. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.