Chapter 3
Explaining the Challenges of Accountability in Machine Learning Systems Beyond Technical Obstacles

Srinivas Kumar Palvadi
Koneru Lakshmaiah Education Foundation, India

ABSTRACT

The ability to make a note regarding machine learning systems decisions for people is becoming increasingly sought after, particularly in situations where decisions have significant repercussions for those affected and where capability in terms of maintaining is required. To increase comprehension based on referred to as "black box" mechanism, explaining ability is frequently cited as a technical obstacle in the design of ML systems and decision procedures. The quantities that ML systems aim to optimize must be specified by their users. This leads to the revealing of policy trade-offs that may have previously been hidden or implicit. Important decisions, as well as judgments, help what may need to be explicitly discussed in public debate as ML's use in policy expands.

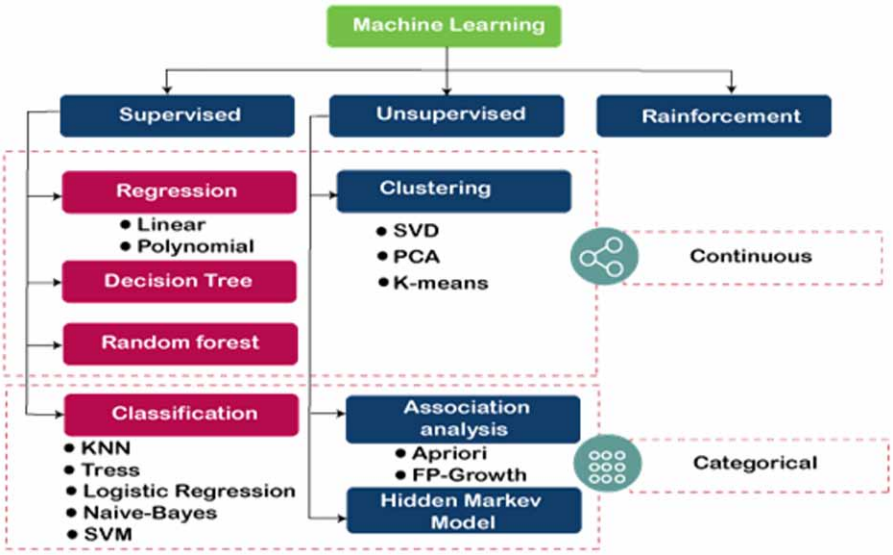
1. INTRODUCTION

The modern era is characterized by interconnectivity with data sources, and virtually every aspect of our lives is digitally documented. We are undeniably immersed in the data age. This encompasses a spectrum of numerous amount of trending technologies such as IoT, ML, DL etc., and a multitude of other data categories that define the wealth of today's digital landscape. These data variations can take on structures ranging from structured to semi-structured to unstructured, as briefly expounded upon in the section titled "Machine Learning Techniques and Real-World Data Types."

The prevalence of these data types is continually expanding, leading to the potential for insights to be derived from them and subsequently applied in intelligent applications within their respective domains. For instance, pertinent cybersecurity data can fuel the development of a data-centric, automation process

DOI: 10.4018/979-8-3693-1479-1.ch003

Figure 1. ML classification



for systems. Similarly, mobile information which was harnessed in the case of craft personalization, information regarding smart devices. In essence, practical applications in various fields rely on adept data management tools and techniques to effectively extract timely insights and valuable knowledge from the available data sources (A. Bharat, et al., 2018)

Programs known as machine learning algorithms help in training the hidden patterns of data and predict output which improves the performance in the overall experience. In ML, various techniques which utilized for various things including the KNN algorithm for classification problems and simple linear regression for stock market prediction issues.

2. TYPES OF MACHINE LEARNING (ML) ALGORITHMS

ML techniques shown in Figure 1, were broadly defined into various categories like:

1. **Supervised Learning**
2. **Unsupervised Learning**
3. **Reinforcement Learning**

Deep ML models, particularly deep neural networks, help in solving a range of complex applications. The diagnosis of a disease by models using medical images or other medical data is one demonstrative example. At a time, DL mechanisms frequently operate like black-box models, with the operation's specifics frequently remaining largely unknown (Deo, R. C. 2015). In this instance, it's hard to define how the project for the particular decision succeeded. As a consequence of this, machine learning models have a hard time fitting into a lot of important applications, like medicine, where doctors need to know what a specific diagnosis means to choose the right treatment. Many methods have been developed to