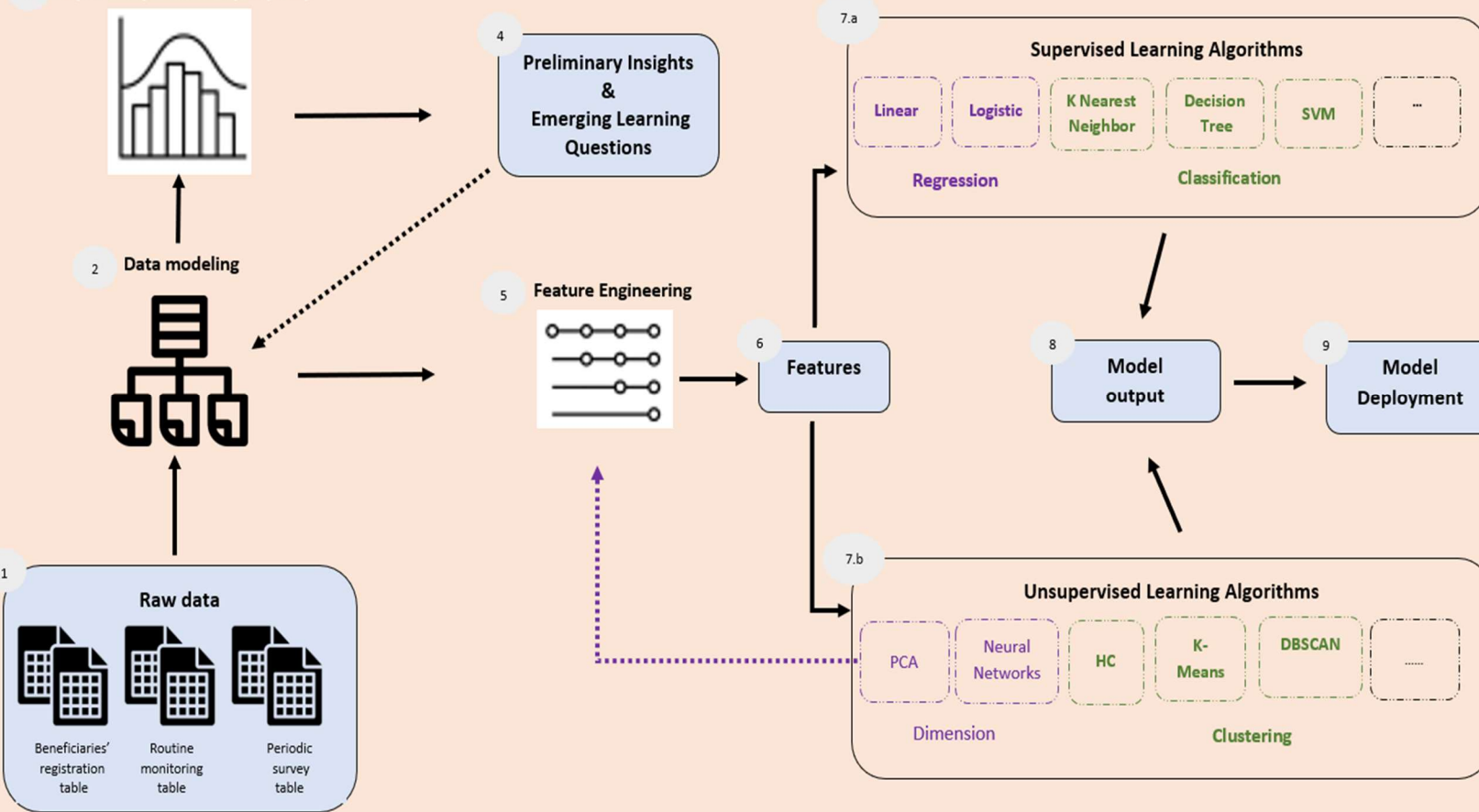


Infographic for integrating Machine Learning (ML) into Monitoring Evaluation and Learning

Machine Learning (ML) is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that human learn, gradually improving its accuracy. It is an important component of the growing field of data science. Using statistical methods, algorithms are trained to make classification or prediction, an to uncover key insight in data mining. These insights subsequently drive decision making within application and businesses, ideally impacting growth metrics. As data is at the heart of Monitoring, Evaluation and Learning (MEL) and the availability of big data continues to expand and grow, demand for ML tasks will increase. This will help identifying the most relevant business questions and the data to answer them.

This infographic suggests a path to integrate ML into MEL in the context of the implementation of a project (development, Emergency, or nexus). It is an adaptation from the IBM Machine Learning course capstone project on a recommender system.

3 Exploratory Data Analysis (EDA)



1 Generally, an M&E system will generate three types of data tables. Your program often starts by registering the beneficiaries and collecting their demographic and identification data. During implementation, goods and services provided to the beneficiaries are tracked via monitoring forms to later inform on the achievement of output indicators. Periodic survey is conducted on the beneficiaries to gather data to estimate a series of outcome indicators.

2 With an assigned unique ID to each beneficiary, you can create a connection and relationship between the tables to generate a new one that stores the variable and columns of interest.

3 Exploratory Data Analysis (EDA) allows having an initial feel for the data. You will analyze the data to summarize their main characteristics through summary statistics and visual methods. EDA will also let you determine if the data make sense or if further cleaning or more data is needed.

4 Identify patterns and trends in the data. Getting a preliminary insight can eventually trigger new learning questions or problem statement types (regression, classification, and clustering), which may result in revisiting your data modeling.

5 Once the question or problem is defined, perform features engineering by selecting, manipulating, and transforming raw data model into features.

6 A feature is an individual measurable property or characteristic of a phenomenon. Proper feature creation, transformation, and extraction are essential to make Machine Learning (ML) algorithms work properly. Once features are set up, algorithms can be run.

7.a

Supervised ML uses of labeled datasets to train algorithms. The purpose of regression is to respond to either prediction or interpretation questions on numerical data. Classification has the same purpose but deals with categorical data. Illustrative questions may include: Can I predict and how should I set the outcome targets given the beneficiaries' characteristics? What characteristics count the most in the outcome achievement? What is the weight and the effect of being exposed (or not) to a service? What are the factors explaining the adoption of a behavior?

7.b

Unsupervised learning analyzes and clusters unlabeled datasets. These algorithms discover hidden patterns or data groupings without the need for human intervention. Its ability to find similarities and differences in information makes it the ideal solution for exploratory data analysis and beneficiaries' segmentation. It can be very helpful in designing a cluster of targeted activities. Illustrative questions may be: Is there a group of beneficiaries with similar characteristics that yield a low outcome? Are there beneficiary outliers for success stories telling? Dimension reduction can be applied to feature engineering to improve supervised ML model output.

8

Once you have trained the model, you can use it to reason over data that it hasn't seen before and make prediction about those data.

9

Finally, you place a finished ML model into a live environment where it can be available to other users or projects implementer, so they can make business decision based on data, interact with the application.