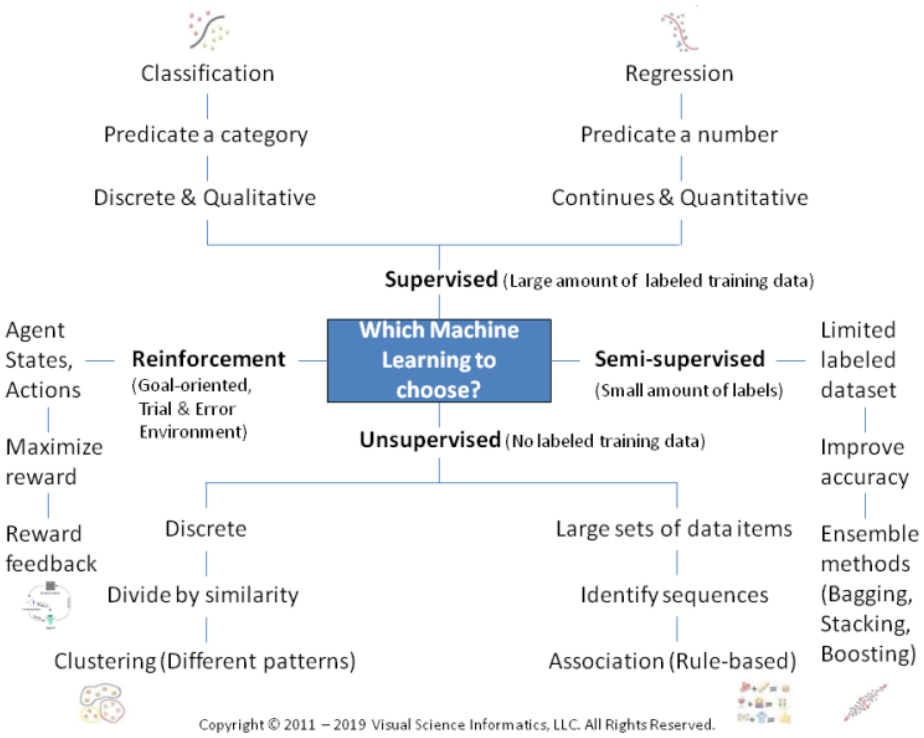
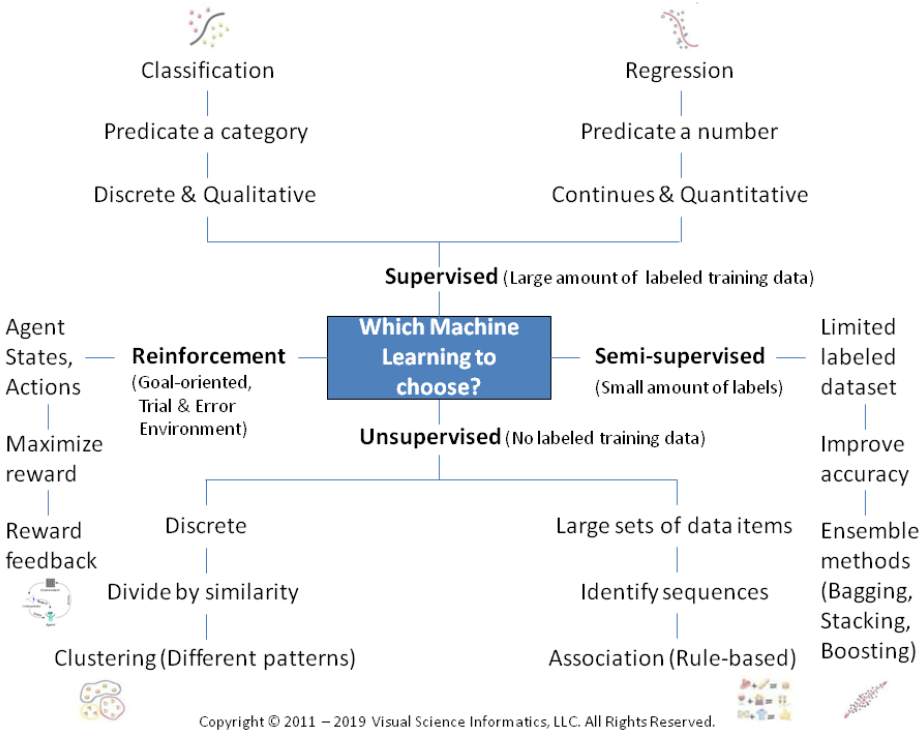
LinkedIn



Machine Learning 101 – Which Machine Learning (ML) to choose?

Add alt text



Based on the inspiration from “Which chart to choose[1],” which helps you to choose the right chart for your data, we developed the idea to chart “Which Machine Learning (ML) to choose?”

Before we present the flowchart of “Which Machine Learning (ML) to choose?” let us take a look at the big picture and zoom in on the steps that this flowchart can guide you in the selection of a machine learning to solve a business problem.

Solving a problem and finding its solution, you can follow these steps:

1. Your strategy can be to select Artificial Intelligent (AI) as your conceptual framework.
2. One of the viable approaches within AI is Machine Learning (ML).
3. After formulating the problem and exploring feasible data acquisition, part of your methodology is choosing a logical learning paradigm.
4. Then you can identify the available data type and define an objective. The logical learning paradigm, data type, and objectives are the criteria for selecting a physical learning method.
5. The next step is to follow a workflow procedure.
6. This workflow procedure can be customized for specific techniques.
7. Finally, you can select a machine learning algorithm.

Your processing pipeline should include at least the following stages:

a.) Data preprocessing and preparation

b.) Datasets sampling for training and validation

c.) Model training, validation, and evaluation

d.) Predication model deployment

e.) Production model monitor, feedback, and retrain

Selecting a logical learning paradigm and method has four major categories, four major algorithm types, and two major techniques.

The four major categories are supervised, semi-supervised, unsupervised, and reinforcement. The four major algorithm types are classification, regression, associations, and clustering. The two major techniques are ensembles methods and reward feedback. The chart at the top “Which Machine Learning (ML) to choose?” guides you through the major categories, data types, and objectives to which algorithm types or techniques to choose.

Choosing an applicable metric, for evaluating machine learning models, depends on the problem and objectives. From a business perspective, two of the most significant measurements are accuracy and interpretability. Accuracy degree measures – how reliable is the conclusion while interpretability (reasoning) measures – how well the model enables understanding the justification and reasoning to the decision conclusion.

Evaluating the accuracy of a machine learning model is critical in selecting and deploying a machine learning model. Choosing the right accuracy metric for evaluating your machine learning model depends on your problem solution objectives and datasets. Before choosing one, it is important to understand the business problem context, the pros and cons, and the usefulness of each error metric.

Add alt text

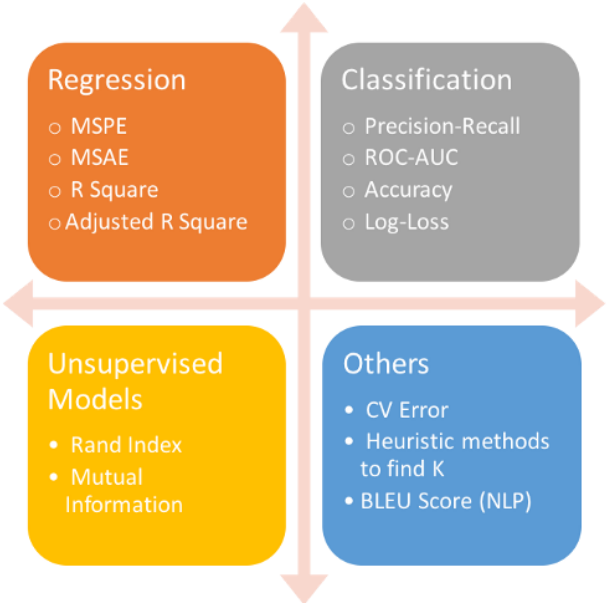


Image by Alvira Swalin via “Choosing the Right Metric for Evaluating Machine Learning Models — Part 1[2]” and Choosing the Right Metric for Evaluating Machine Learning Models — Part 2[3]”

The chart above captures and categorizes useful metrics for evaluating machine learning models for a variety of machine learning algorithms, methods, and techniques.

On the other hand, measuring interpretability (reasoning) is a more complex task because there is neither a universal agreeable definition nor an objective quantitative measure. In general, opaque methods obtain higher accuracies than transparent ones. There are methods to produce an interpretable predictive model such as a post-hoc interpretable model or an intrinsically interpretable algorithm. One measure of interpretability based on “triptych predictivity, stability, and simplicity” is proposed by Vincent Margot in “How to measure interpretability?[4]”

Add alt text

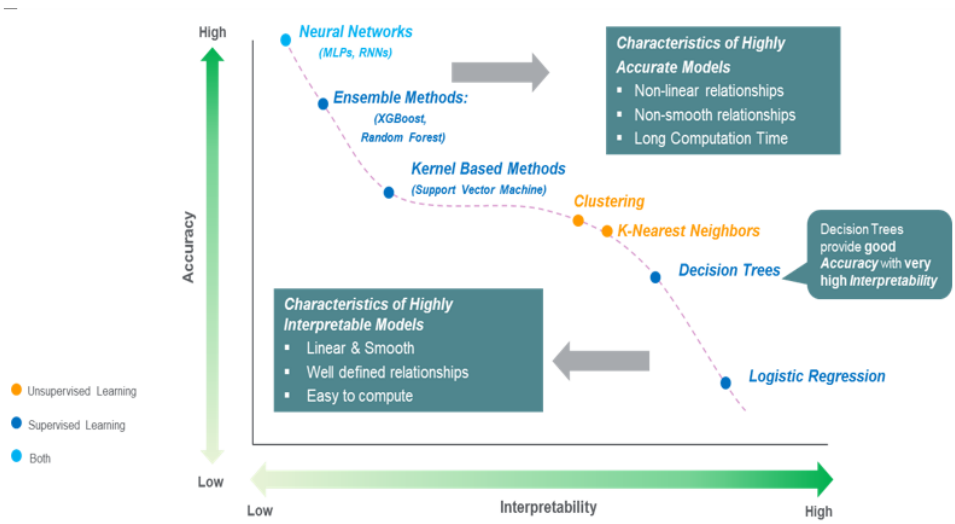


Image by Sharayu Rane via “The balance: Accuracy vs. Interpretability[5]”

The chart “The balance: Accuracy vs. Interpretability” sorts out the trade-off between accuracy and interpretability (reasoning) for a variety of machine learning algorithms, methods, and techniques.

In conclusion, choosing a Machine Learning (ML) depends on multiple complex factors and challenging trade-offs. Selecting machine learning, which balances all decision factors, is important. Because the capital investment, in the processing pipeline stages, is costly and requires considerable time and effort. Therefore, it is highly valuable to employ a rigorous process in choosing machine learning.
