Lecture 8

Learning Algorithms

Typically, the ML used for business purposes is either supervised or unsupervised in nature. Within these categories, however, there many different types of algorithms and ML routines, which can be used to accomplish different goals. Additionally, there are often different learning methods, such as "eager" and "lazy" learning methods. These methods govern how to process training data, and that governance will determine compute and storage requirements:

- ***Eager learning methods*** evaluate training data and "eagerly" begin computing (for example, classification) before receiving new (test) data. They generally depend more on upfront evaluation of training data in order to compute (that is, predict) without the need for new data. As a result, eager learning methods tend to spend more time on processing the training data.

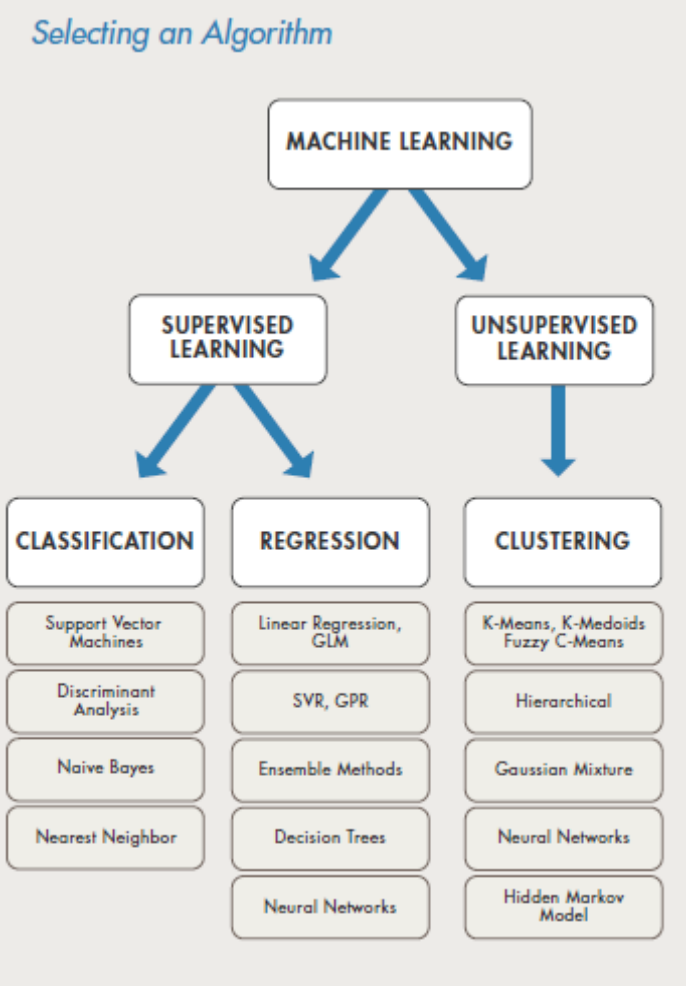
- ***Lazy learning methods*** delay processing and data evaluation until new test data is provided
— hence the term "lazy" or "lazy evaluation." As a result, lazy learning methods are often case-based, spending less time on the training data and more time on predicting.

Following table 1 provides descriptions of a few of the more common types of ML used in business, and it lists examples of the types of business applications they can be used to solve.

Table 1. Examples of Types of Machine Learning Algorithms/Problem Solving Approaches

Type	Model/ Algorithm or Task	Description	Usage Examples in Business
Supervised	Neural network	Computations are structured in terms of interconnected groups, much like the neurons in a brain. Neural networks are used to model complex relationships between inputs and outputs to find patterns in data or to capture a statistical structure among variables with unknown relationships. They may also be used to discover unknown inputs (unsupervised).	▪ Predicting financial results ▪ Fraud detection
Supervised	Classification and/or regression	Computations are structured in terms of categorized outputs or observations based on defined classifications. Classification models are used to predict new outputs based on classification rules. Regression models are generally used to predict outputs from training data.	▪ Spam filtering ▪ Fraud detection
Supervised	Decision tree	Computations are particular representations of possible solutions to a decision based on certain conditions. Decision trees are great for building classification models because they can decompose datasets into smaller, more manageable subsets.	▪ Risk assessment ▪ Threat management systems ▪ Any optimization problem where an exhaustive search is not feasible
Unsupervised	Cluster analysis	Computations are structured in terms of groups of input data (clusters) based on how similar they are to one another. Cluster analysis is heavily used to solve exploratory challenges where little is known about the data.	▪ Financial transactions ▪ Streaming analytics in IoT ▪ Underwriting in insurance
Unsupervised	Pattern recognition	Computations are used to provide a description or label to input data, such as in classification. Each input is evaluated and matched based on a pattern identified. Pattern recognition can be used for supervised learning as well.	▪ Spam detection ▪ Biometrics ▪ Identity management
Unsupervised	Association rule learning	Computations are rule-based in order to determine the relationship between different types of input or variables and to make predictions.	▪ Security and intrusion detection ▪ Bioinformatics

Type	Model/ Algorithm or Task	Description	Usage Examples in Business
			▪ Manufacturing and assembly



Supervised Learning Algorithms:

All classification and regression algorithms come under supervised learning.

- Logistic Regression
- Decision trees
- Support vector machine (SVM)
- k-Nearest Neighbors

- Naive Bayes classification
- Random forest
- Linear regression
- polynomial regression
- SVM for regression
- Neural Network

Unsupervised Learning Algorithms:

- k-Means clustering
- Neural Network
- Principal Component Analysis

Some more examples of tasks that are best solved by using a learning algorithm

• Recognizing patterns:
  – Facial identities or facial expressions
  – Handwritten or spoken words
  – Medical images
• Generating patterns:
  – Generating images or motion sequences
• Recognizing anomalies:
  – Unusual credit card transactions
  – Unusual patterns of sensor readings in a nuclear power plant
• Prediction:
  – Future stock prices or currency exchange rates