Jordan Black

Introduction to Machine Learning 14802 CGS2020 202501

Dr. Lakshmi Prayaga,

1/31/2025

**Assignment 2**

1. **From the readings and references provided list at least two differences between supervised and unsupervised learning**

Supervised Learning: The algorithm is trained on labeled data, where the correct output is known. An example is predicting the house process based on historical data**.**

Unsupervised Learning: The algorithm is trained on unlabeled data and identifies patterns or clusters. An example of clustering would be Customer segmentation in marketing**.**

1. **What is an example algorithm of supervised and unsupervised learning.**
2. Supervised Learning Algorithm: Linear Regression, Decision Trees, or Support Vector Machines (SVM).

Supervised learning algorithms: Random forest example:

Random forest employs ensemble learning to build multiple decision trees during training before delivering the most common class result (classification) or the average prediction (Regression) from these trees. This method improves prediction accuracy while simultaneously managing the issue of overfitting. Medical diagnostics utilizes Random Forest to evaluate patient information, including age, blood pressure, and cholesterol levels, to determine the risk of heart conditions. Combining multiple decision tree outputs creates a robust model that identifies intricate feature interactions, thus delivering precise and dependable predictions.

1. Unsupervised Learning algorithm: K- Means Clustering or Principal Component Analysis (PCA).

Unsupervised Learning algorithm example:

The algorithm analyzes large datasets to extract frequent item sets and create association rules. Market basket analysis benefits from this tool because it reveals product combinations that appear together often in transactions. The algorithm allows retailers to analyze transaction data, demonstrating that bread-and-butter purchases are commonly followed by milk purchases. Businesses can improve sales and customer satisfaction by optimizing product placements and designing targeted marketing strategies using insights to create effective cross-selling strategies.

1. **Create a drill down for the notion of machine learning. Think of it as a tree or concept map or flowchart where you start with the main topics and break it down further. It should contain what is machine learning, what are the main types of machine learning, (supervised and unsupervised learning methods). What are the main features of each type and example case scenario for each. Draw text boxes and include text in the boxes and connect them by arrows or lines. Here are examples of creating a Concept map for a given topic**

**A diagram of a company

Description automatically generated**

1. **You want to be able to determine the salary of a new employee. You have a group of employees. What attributes would you need from this set of employees to build a model that can predict the salary of the new employee. (Hint: think of the fruits example we discussed). Build a small table with 10 entries. Is this a supervised learning model? If so why, if not why not.**

**Sample table (10 Entries):**

**A table with numbers and text

Description automatically generated**

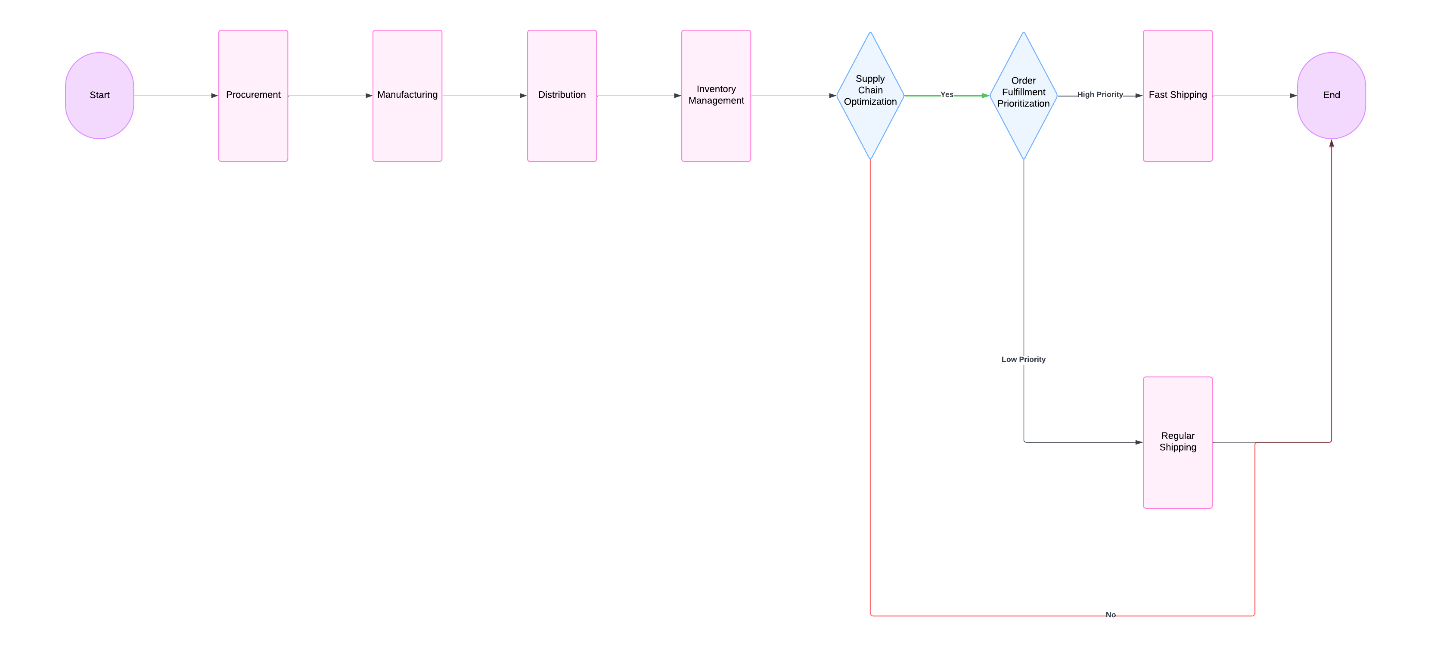
* **Yes, this is supervised by the learning model:**
* **Why?** The data is labeled with both input features (like experience and education) and output labels (salary). The algorithm learns the mapping between inputs to predict future salaries.

1. **You want to launch a new product. It is imperative to understand your customer population. What is the best type of machine learning technique to explore and understand your customer population? Use a diagram to help you think through the process.**

**Unsupervised Learning, particularly clustering algorithms like k-Means.**

**Explanation:**

Through unsupervised learning techniques, we can categorize customers using various features such as age and gender alongside location and historical purchases without needing predefined labels.



Work Cited

Brownlee, Jason. "What Is Data Preparation in a Machine Learning Project." *Machine Learning Mastery*, 2 Jan. 2019, [www.machinelearningmastery.com/what-is-data-preparation-in-machine-learning/](http://www.machinelearningmastery.com/what-is-data-preparation-in-machine-learning/).

"Introduction to Unsupervised Learning: Types, Applications and Techniques." *DataCamp*, 1.8 years ago, <https://www.datacamp.com/blog/introduction-to-unsupervised-learning>.