



FDP Kultivations

Spark MLlib Workshop



Workshop Agenda

- **Day 1**
 - Intro Big data Scenario
 - Spark Overview
 - Rdd's processing with hands on
- **Day 2**
 - Intro SQL and Big Data motivation
 - Spark – Dataframe and SQL API
 - Hand's on
- **Day 3**
 - Full Exercises Day 1 and Day 2
- **Day 4**
 - Intro to Spark MLlib (Plus)



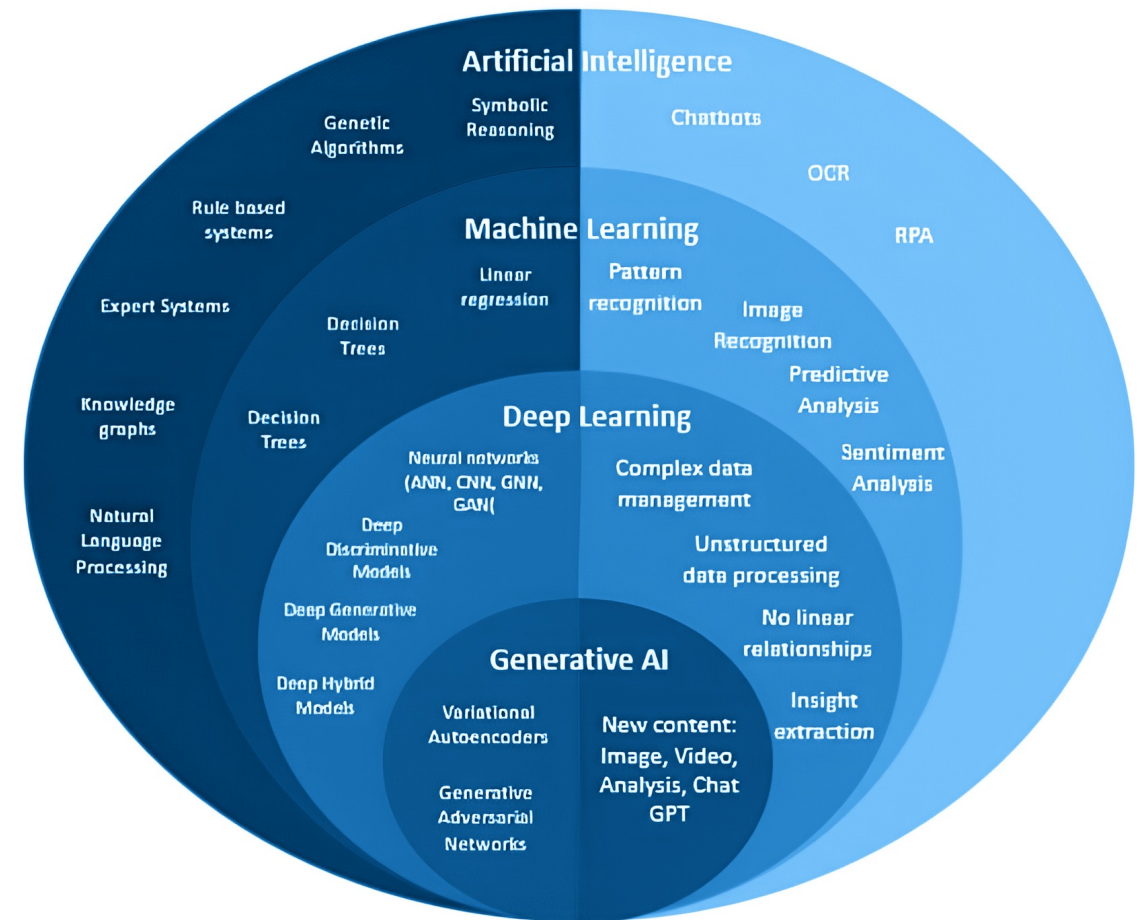
Day 4 Spark MLlib

- What is Machine Learning
- Machine Learning Lifecycle
- Spark MLlib Algorithms
- Use Cases for Spark MLlib
- Basket Analysis Lab
- Customer Churn Lab

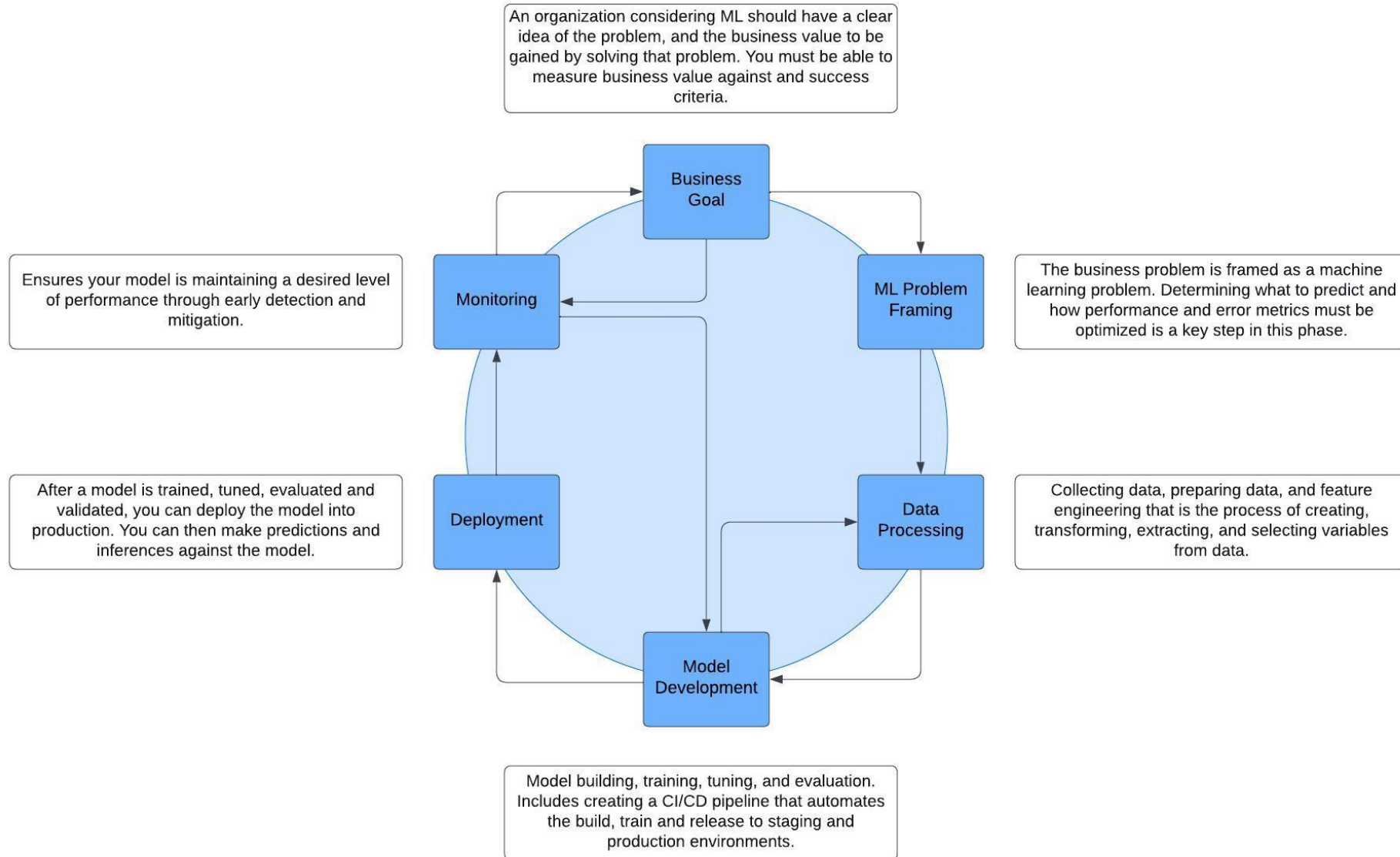
What is Machine Learning?

While many ML algorithms can be deterministic, some exhibit non-deterministic behavior due to inherent randomness.

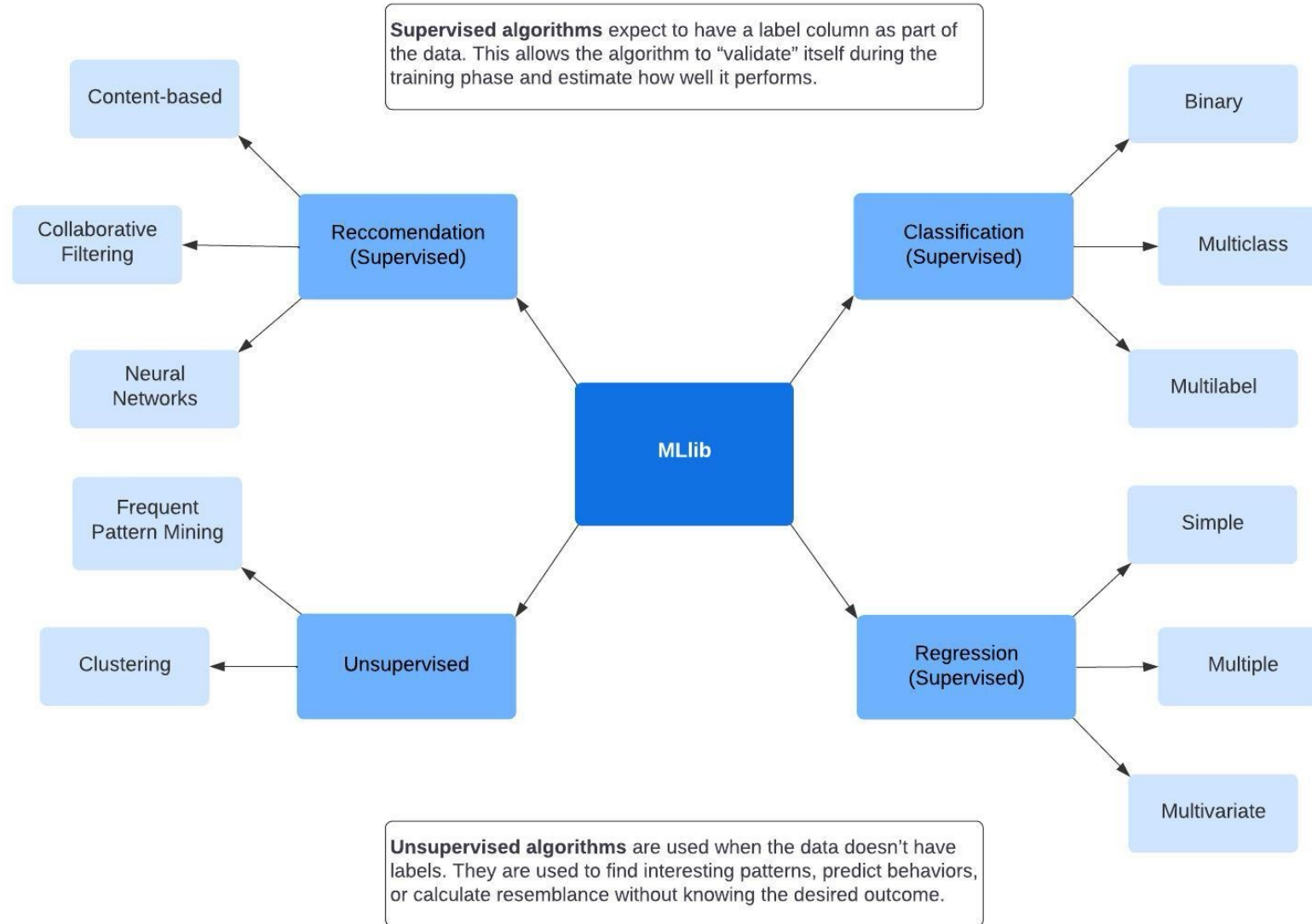
Machine Learning is a branch of AI that uses algorithms to learn from data and make predictions or decisions without explicit programming.



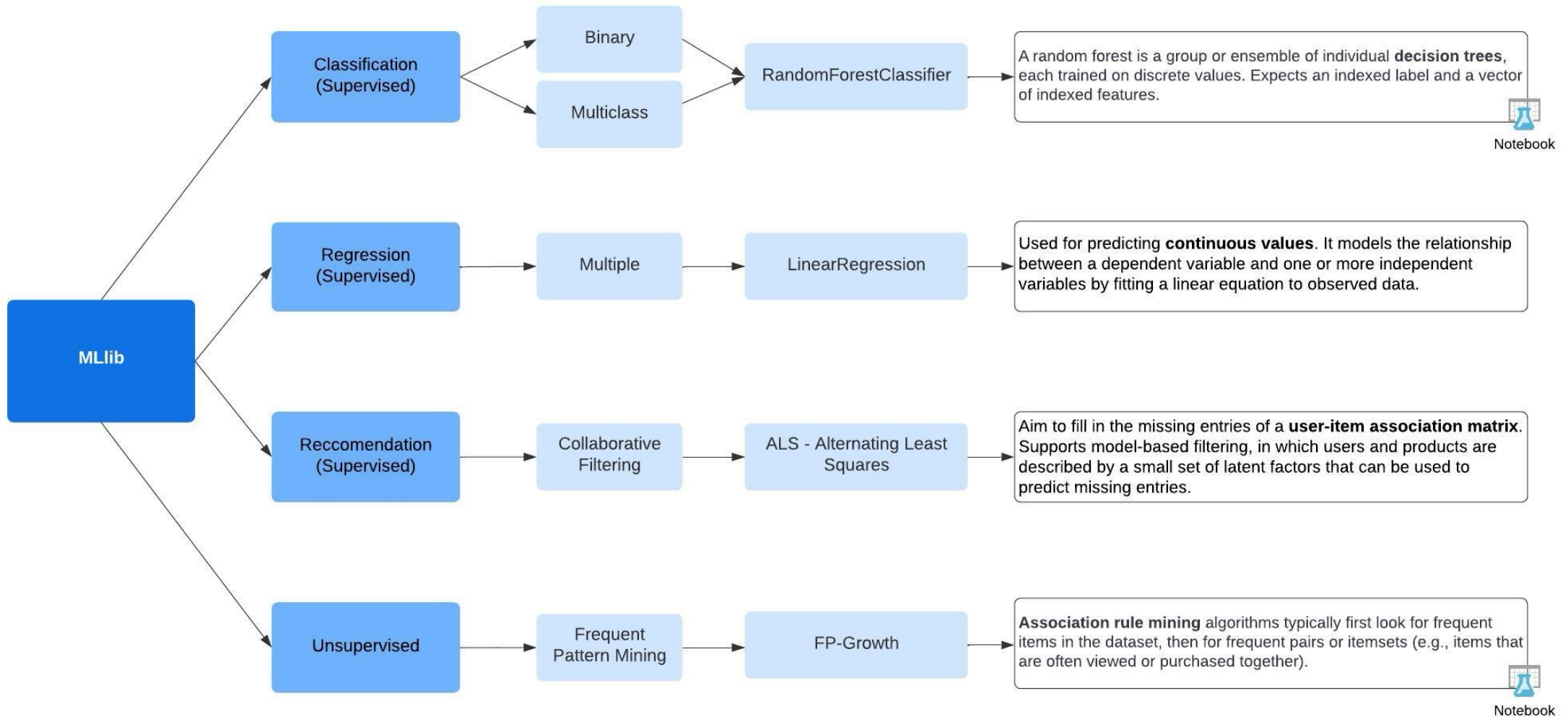
Machine Learning Lifecycle



Spark MLlib Algorithms



Use Cases for Spark MLlib



Do you have a
quick question?

rodrigo.mattos@kroger.com

Data Engineer - Team Friends From Work

