Intro to ML

Created Date: 2022-07-29
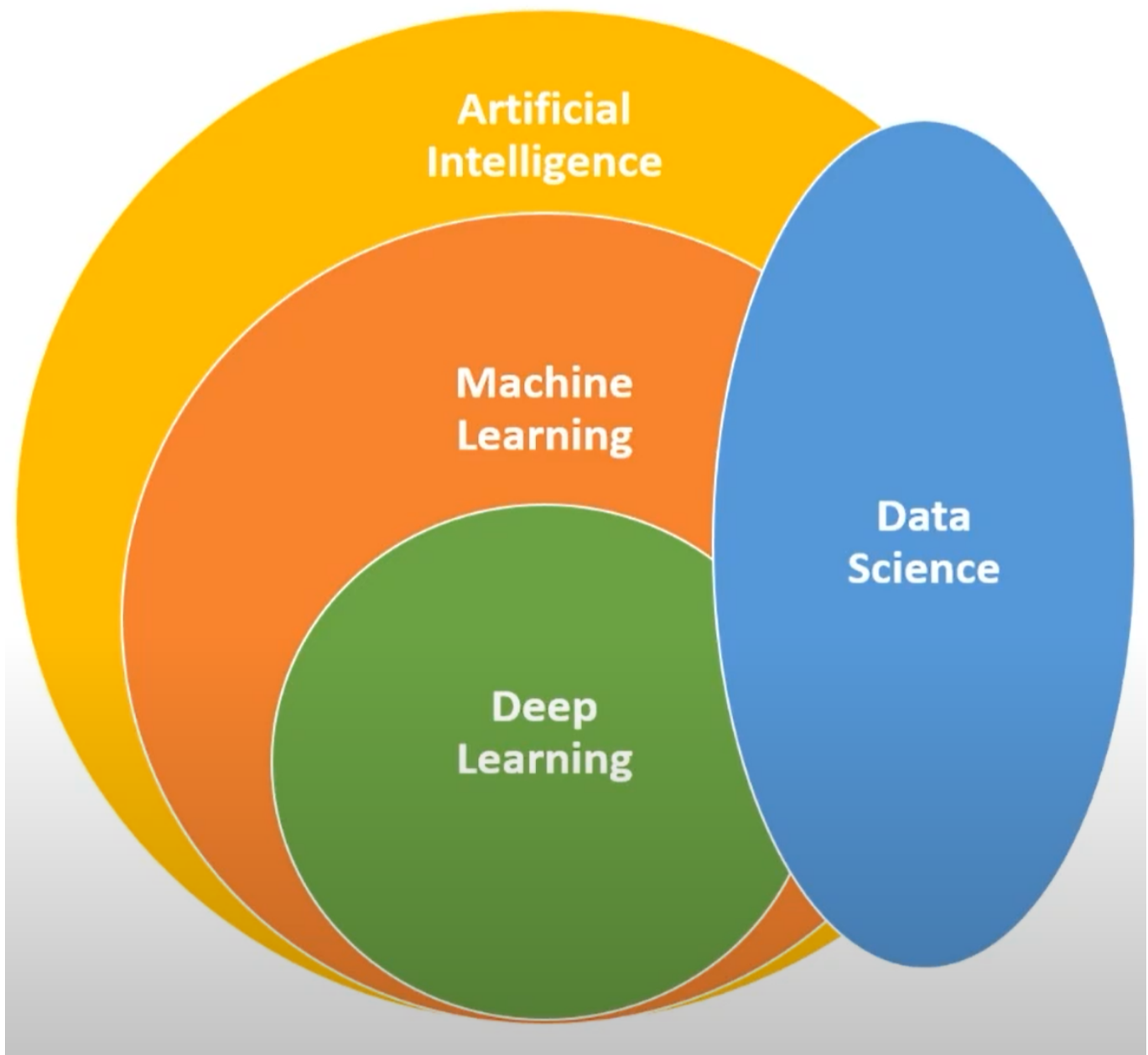

Metadata 📦

- Title: Introduction to Machine Learning
- Author: Andrew Jones
- Reference: DSI Infinity

Links & Tags 🔗

- Index: Data Science Note Index
- Atomic Tag: #datascience
- Subatomic Tags: #machinelearning
- Related Notes:

Subsets of Data Science

- Data Science
  - A broad field encompassing one or more of: Statistics, Computer science, Analytics
- Artificial Intelligence
  - The programming of machines to mimic the cognitive functions associated with humans such as learning and problem-solving.
- Machine Learning
  - Algorithms that learn from data, then apply those learnings to make informed decisions and/or predictions on new data.
- Deep Learning
  - Subset of ML algorithms specifically utilizing multi-layered Neural Networks to understand abstract and/or unstructured patterns.

Types of Machine Learning

- Supervised Learning
  - Algorithms that predicts mappings for inputs (features) and outputs (labels) based on example input-output pairs (training data)
- Unsupervised Learning
  - Algorithms that find structure and patterns from unlabeled data
- Semi-Supervised Learning
- Reinforcement Learning

- Regression
- Classification

- Clustering
- Dimensionality Reduction
- Anomaly Detection
- Association Rule Learning

Scikit-Learn

- ML library for Python
- Contains algorithms for supervised and unsupervised learning
- Contains algorithms for data preprocessing techniques
- Easy to apply but does not show underneath the hood
  - It's important to understand the theory behind any models you use

Supervised	Regression	Linear Regression, Regression Tree, Random Forest + more
	Classification	Logistic Regression, Classification Tree, Random Forest + more

Unsupervised	Clustering	k-means, hierarchical clustering + more
	Dimensionality Reduction	PCA + more

Other	Data Preprocessing	Feature Selection, Feature Scaling, One Hot Encoding + more
	Model Refinement	Validation Splitting, Parameter Tuning, Model Selection