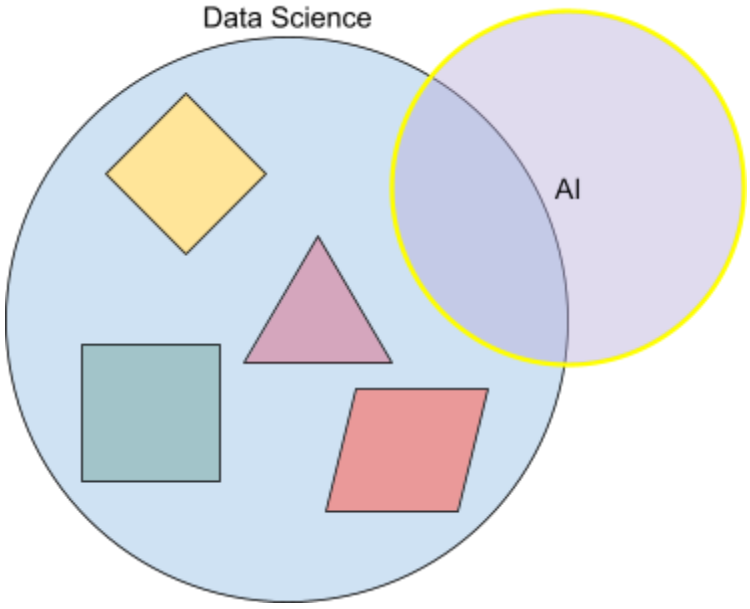
Table of Contents

1/10/24

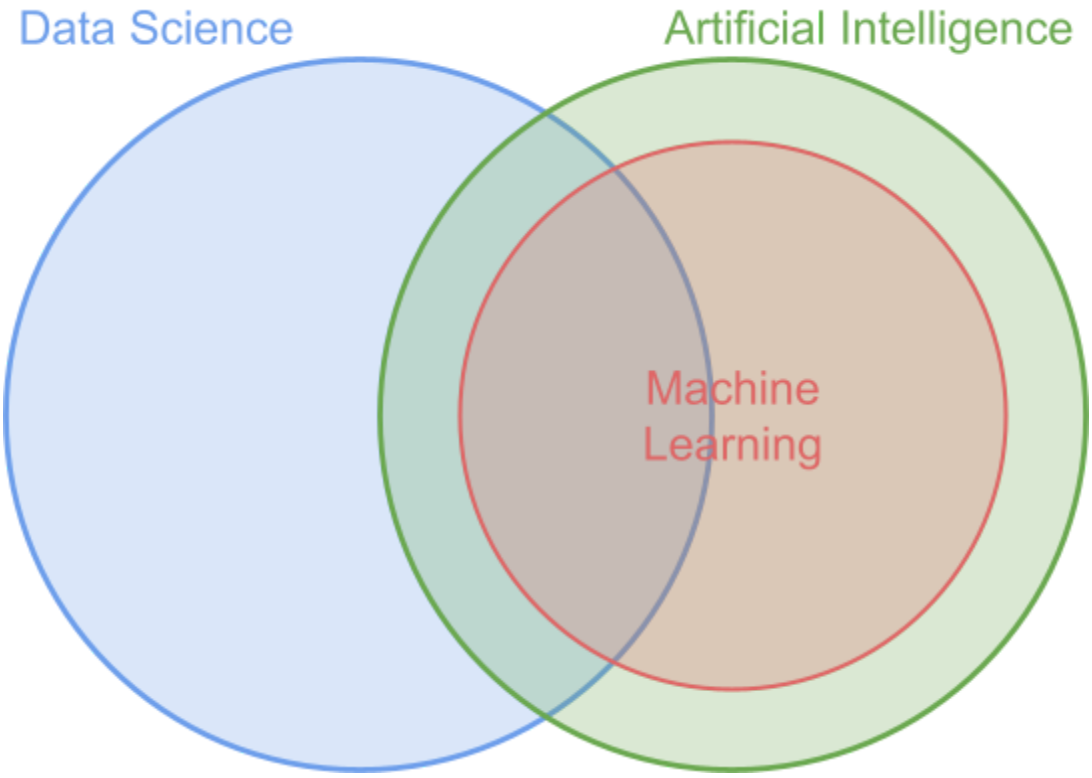
Intro to AI and ML

1. Data Science
    a. Analyze and compute data
    b. Extracting statistics from data
    c. Generates meaning from data
    d. “The use of scientific methods to obtain information from computer data”
2. Relationship between AI and Data Science
    a. 

3. Artificial Intelligence
    a. A field of study in computer science that involves the development and study of “intelligent” machines.
    b. Designed to mimic human functions
        i. Reasoning
            1. Should be able to reason through problems and use problem-solve skills
        ii. Knowledge Representation
            1. Should be able to make deductions or conclusions and apply its knowledge to problems
        iii. Planning
            1. Should be able to plan and make decisions based on logic

iv. Learning
    1. Should be able to learn from mistakes and successes
v. Language
    1. Should be able to process and identify languages.
    2. Can identify patterns in human speech
vi. Perception
    1. Can identify patterns and can make predictions based on patterns

4. Machine learning
    a. A field of study housed in the main field of Artificial Intelligence that focuses on learning from data and can perform tasks without explicit instructions
5. Relationship between AI, Data Science, and Machine Learning
    a.



1/31/24

Predicting Continuous Values

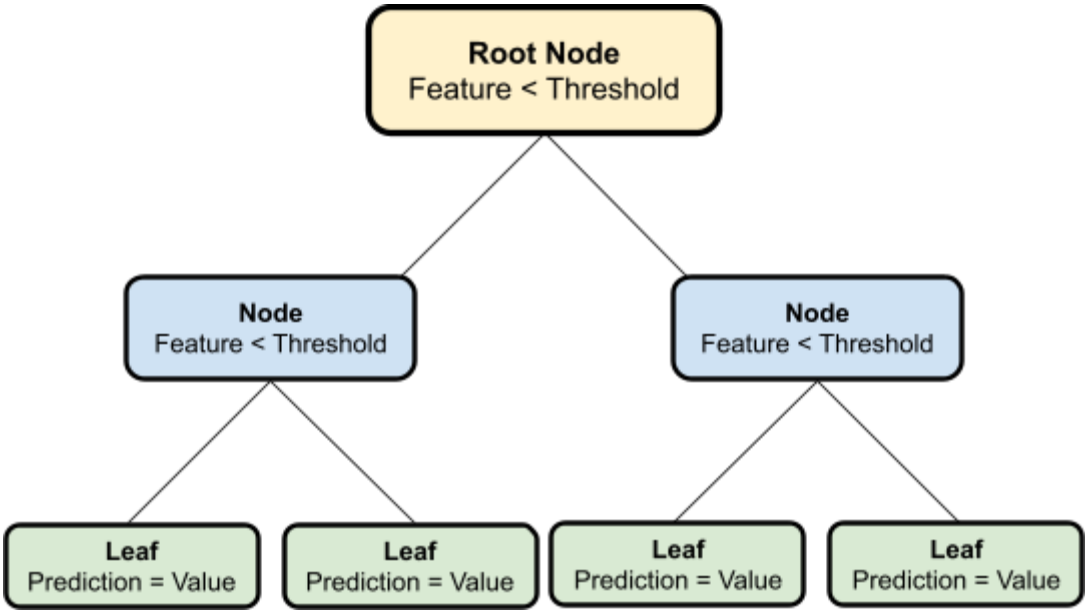
1. Polynomial Regression

    a. A model that fits the data following a polynomial form (coefficients are raised to a degree)
2. Difference between linear and polynomial regression
    a. Linear regression follows a straight line indicated by y=mx+b, while polynomial regression follows a line indicated by y=nx^2+mx+b
3. MSE
    a. The mean squared error
    b. Shows the average distance between actual and predicted value squared
    c. Lower values show a more accurate graph
4. Decision Tree Regression
    a. A "flowchart" that checks to see if values meet certain criteria, then predicts values after following the path of the criteria.
    b. Works well with a variety of data types
        i. Good with categorical and small data
5. RMSE
    a. The root mean squared error
    b. Shows the average distance between the actual and predicted values
    c. Lower values show a more accurate graph

2/16/24

Predicting Discrete Values

1. Classification
    a. Predicting categorical or discrete variables
2. Decision Tree Classification
    a. Similar to decision tree classification, but uses categorical data instead of numerical data
    b. Best suited for binary or categorical data

c. 

3. K-Nearest Neighbors
    a. Plots points on a line, checks to see distance between points to group points
4. Naive Bayes
    a. A model that calculates the conditional probability for different features and uses that to determine a prediction based on probability
    b. Best used with categorical, binary, or text data.
5. Featurization/Vectorization
    a. Converting variables and values to different features which are assigned probabilities or other classifiers
    b. Especially useful in Naive Bayes
6. Naive Bayes formula
    a. probability = (1 / (np.sqrt(2 * np.pi) * std)) * (np.exp(-((x - mean) ** 2) / (2 * std ** 2)))