AI Fundamentals:

Classical ML Models

By: Tiffany Le

Agenda

Introduction to Data Science

- Supervised vs. Unsupervised ML
- Coding Time!

Introduction to Data Science

It's a major at Chapman, but what really is Data Science?

–

Data Science: interdisciplinary study of combining mathematics, statistics, and algorithms to get actionable information and insights from data (paraphrased from IBM)

Machine Learning: a specific field of Artificial Intelligence where we train algorithms (models) to learn patterns from data and gain insights

Key Components

- Data Collection
- Data Cleaning
- Exploratory Data Analysis (EDA)
- Data Pre-processing
- Modeling & Algorithms
- Data Interpretation

Key Components

- **Data Collection**
  - Gathering relevant data from sources
  - **Sources**:
    - Kaggle
    - Census/Gov’t Public Data
    - Web Scraping

Image From: https://kadence.com/en-us/an-introduction-to-data-collection-in-marketing-research/

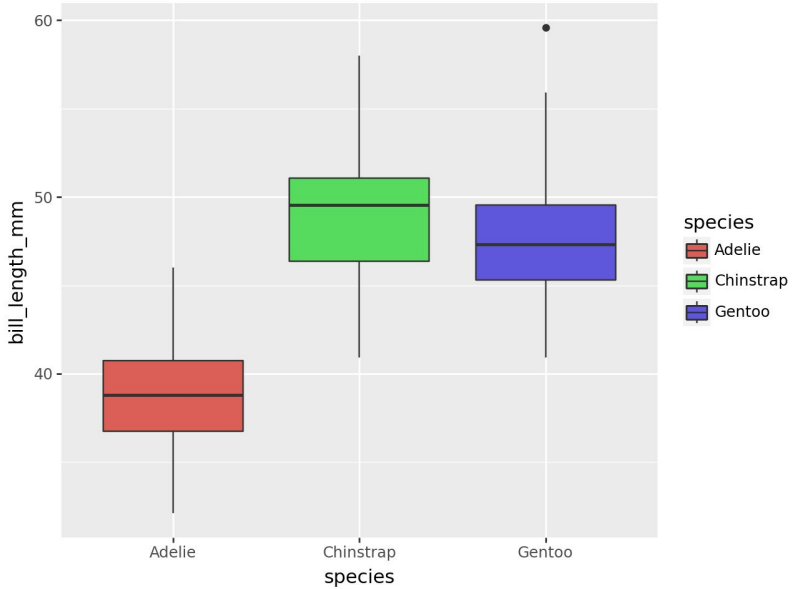
Key Components

- Data Collection
- **Data Cleaning**
  - Preparing data for analysis by:
    - Dropping Missing Data
    - Imputing Missing Data (e.g. mean, 0, etc.)
    - Checking for outliers

Image From: https://www.the-cleaning-company.co.uk/a-professional-or-regular-cleaning-service-whats-the-difference/

Key Components

- Data Collection
- Data Cleaning
- **Exploratory Data Analysis (EDA)**
  - Data Visualization
  - Summary Statistics



Key Components

- Data Collection
- Data Cleaning
- Exploratory Data Analysis (EDA)
- **Data Pre-Processing**
  - Standardizing/Normalizing Data (Z-Score)
  - Encoding (e.g. One Hot Encode)

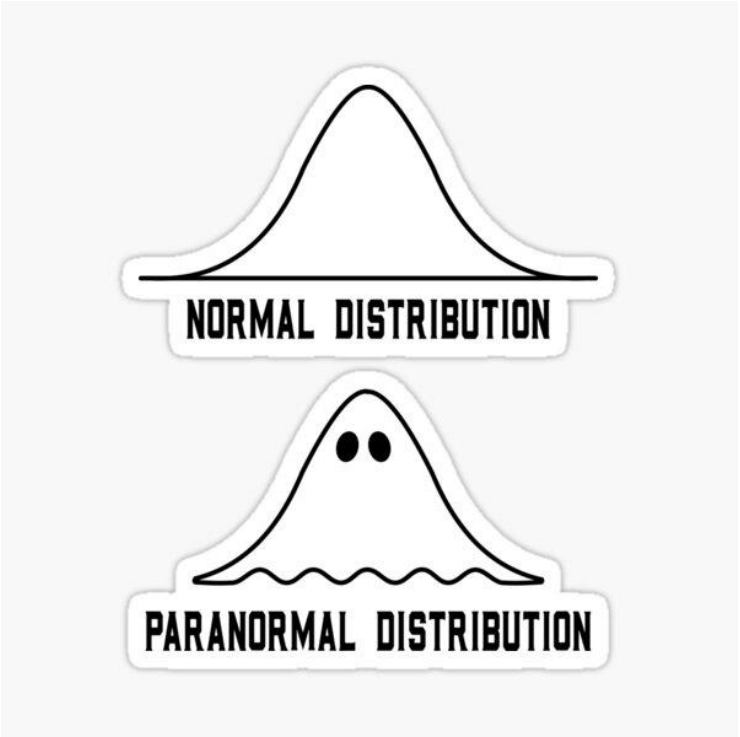


Image From: https://www.redbubble.com/shop/paranormal+distribution+stickers

Key Components

- Data Collection
- Data Cleaning
- Exploratory Data Analysis (EDA)
- Data Pre-processing
- **Modeling & Algorithms**
  - Building & Training your desired machine learning model
  - **ML Model:** Program that uses algorithms to learn patterns from the data and complete tasks
    - Supervised vs. unsupervised

Key Components

- Data Collection
- Data Cleaning
- Exploratory Data Analysis (EDA)
- Data Pre-processing
- Modeling & Algorithms
- **Data Interpretation**
  - Present actionable insights from the data
  - Business Plans & Presentation

Agenda

- Introduction to Data Science
- **Supervised vs. Unsupervised ML**
- Coding Time!

Supervised & Unsupervised Machine Learning

- **Two (2) main approaches for Machine Learning**
  - Supervised
  - Unsupervised
- **Supervised**
  - **Prediction-Based**
  - Using labeled data to make predictions
- **Unsupervised**
  - **Cluster or Grouping Based**
  - Analyze & cluster unlabeled data

Supervised Machine Learning

- **Prediction Based!**
- Using labeled data to train algorithms to recognize patterns and make predictions
- **Main Types of Supervised ML**
  - **Regression**: predicting continuous values
  - **Classification**: predicting classes/groups (cat or dog)
- **Examples:**
  - Predicting weather (temperature, hot or cold, etc.)
  - Predicting house prices

Supervised Machine Learning Example

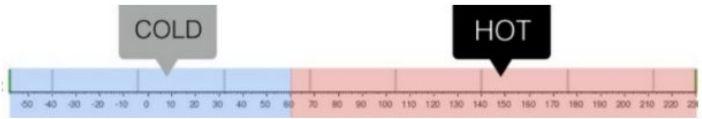
Regression

What will be the temperature tomorrow?

Fahrenheit

Classification

Will it be hot or cold tomorrow?



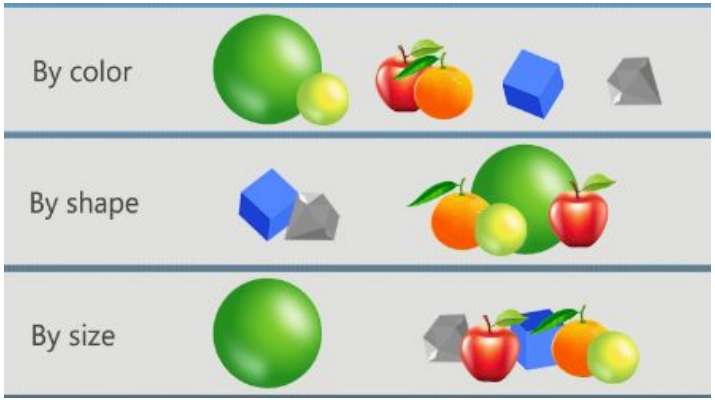
Fahrenheit

Unsupervised Machine Learning

- **Grouping or Clustering Based!**
- Analyze & cluster unlabeled data (no output data)
- Discover hidden patterns within data
- **Main Types of Unsupervised ML**
  - **Clustering**: grouping data points by on similarities or differences
  - **Dimensionality Reduction:** reducing # of features
- **Examples:**
  - Song recommendations
  - Customer purchasing patterns

Unsupervised Machine Learning Examples

Clustering:
Find patterns and similar groups



Dimensionality Reduction:
Getting rid of unnecessary variables

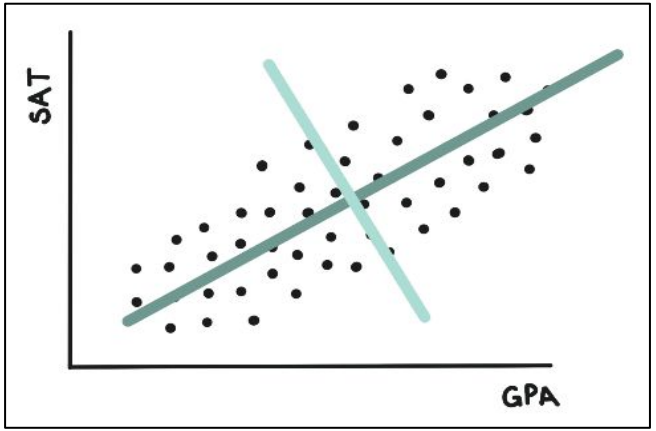


Image From: Dr. Chelsea Parlett-Pelleriti
(https://github.com/chelseaparlett/CPSC392ParlettPelleriti/blob/master/Lectures/LectureSlides/20_PrincipalComponentAnalysis.pdf)

Agenda

- Introduction to Data Science
- Supervised vs. Unsupervised ML
- **Coding Time!**

Access this Lesson's Jupyter Notebook Here!