Oregon State University

Advancing Transportation Data Analysis through AI and Big Data Techniques

By

Sal Hernandez

Associate Professor

School of Civil and Construction Engineering

Oregon State University

Oregon State University
College of Engineering

Breakdown

- What is Meant by AI
  - Unsupervised vs Supervised
- Recent Projects
- Learning AI
- Transportation Agency Adoption

Oregon State University
College of Engineering

What is Artificial Intelligence?

Some general definitions

- “Artificial intelligence is a computerized system that exhibits behavior that is commonly thought of as requiring intelligence.”
- “Artificial Intelligence is the science of making machines do things that would require intelligence if done by man.”
- The founding father of AI, Alan Turing, defines this discipline as: “AI is the science and engineering of making intelligent machines, especially intelligent computer programs.”

BIG DATA

Capable of processing massive amounts of **structured and unstructured data** which can change constantly

REASONING

Ability to reason (deductive or inductive) and to draw inferences based on situation. **Context driven awareness** of system.

Ability to **learn** based on historical patterns, expert input and feed-back loop

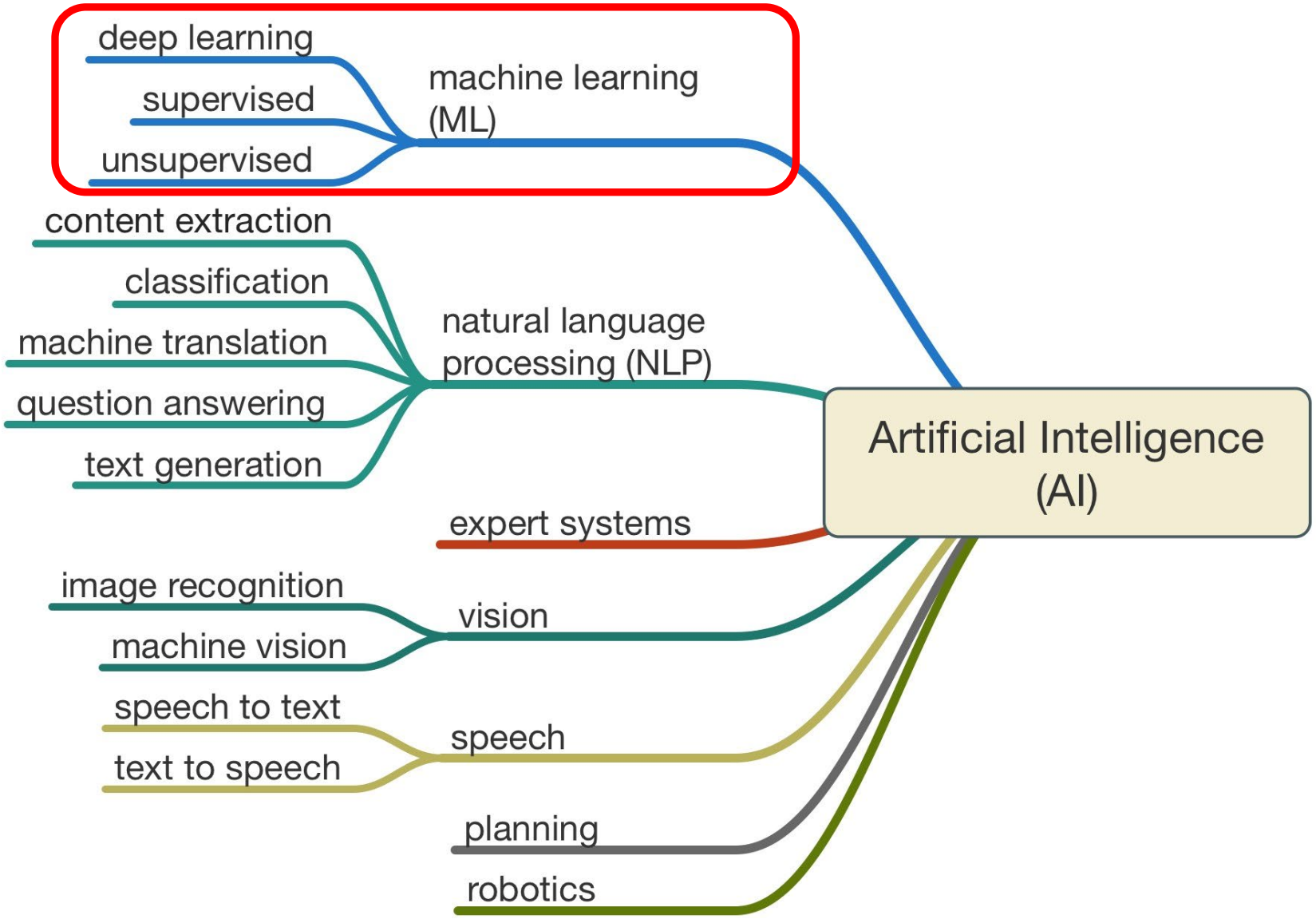
LEARNING

Capable of analyzing and **solving complex problems** in special-purpose and general-purpose domain

PROBLEM SOLVING

AI in a Nutshell

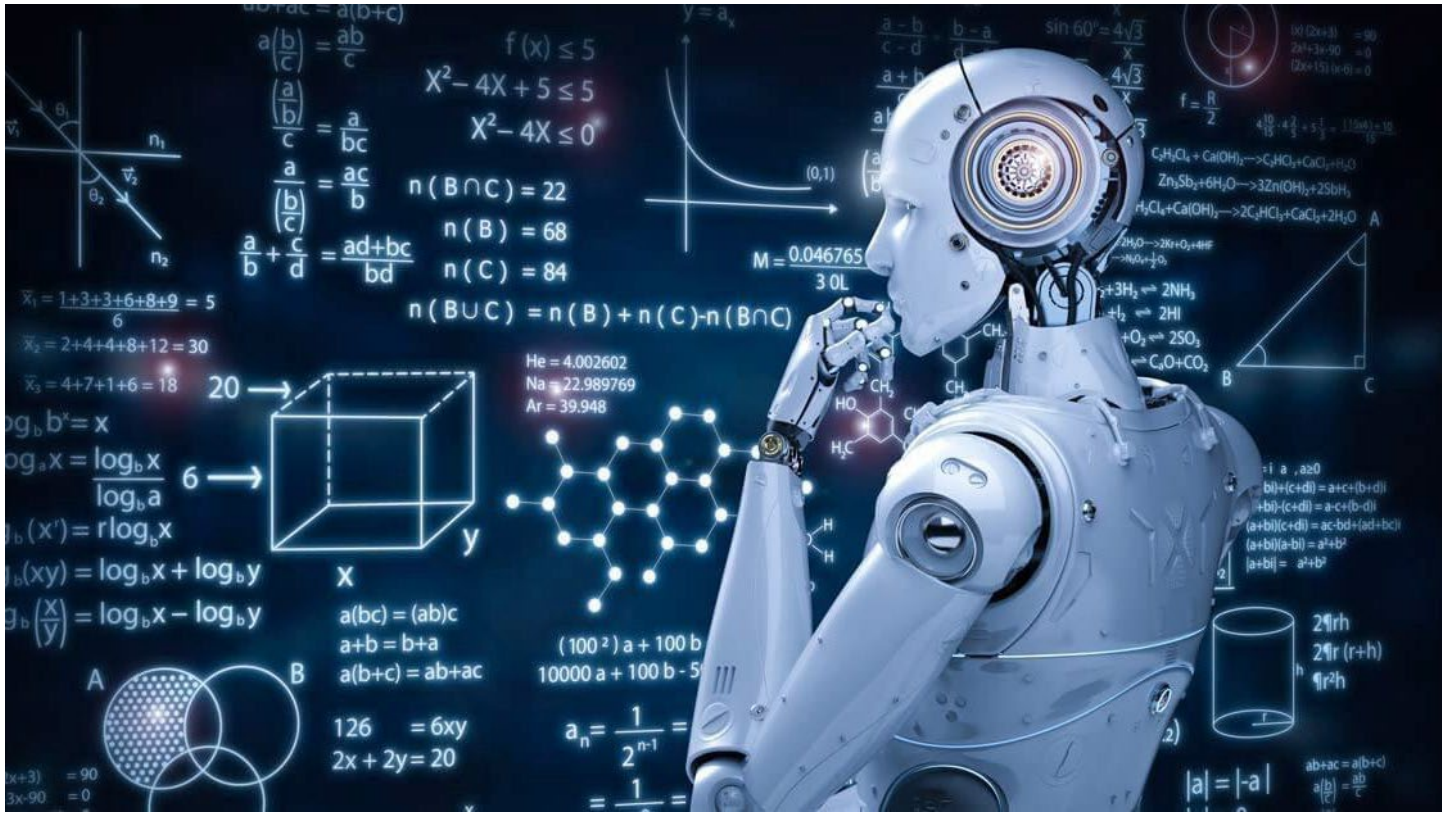
Oregon State University
College of Engineering

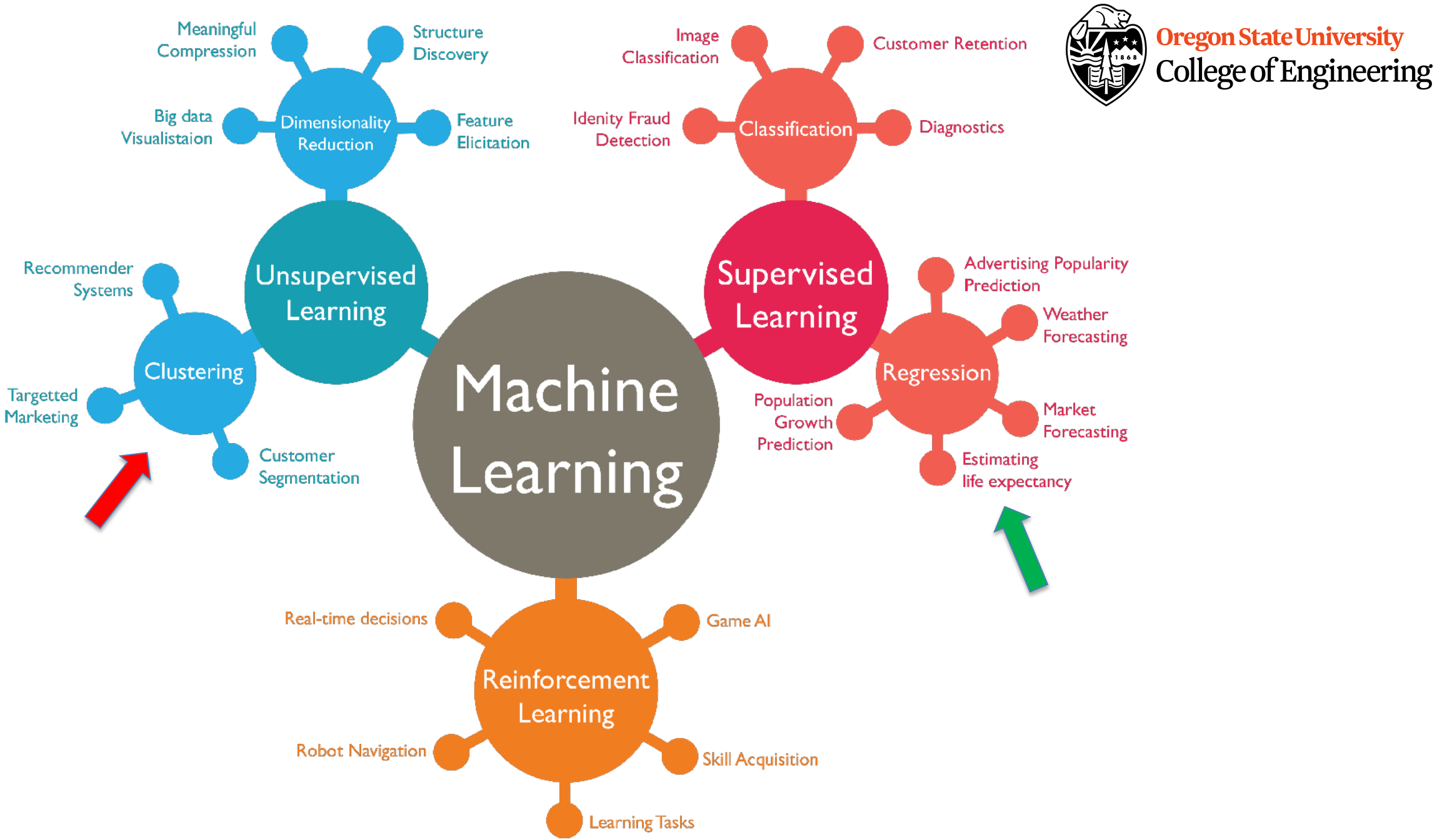


Oregon State University
College of Engineering

What is machine learning?

- It’s a branch of artificial intelligence (AI) based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.





Oregon State University
College of Engineering

Machine Learning Techniques

Supervised Learning

- Labels are provided, there is a strong learning signal.
- e.g. classification, regression.

Unsupervised learning

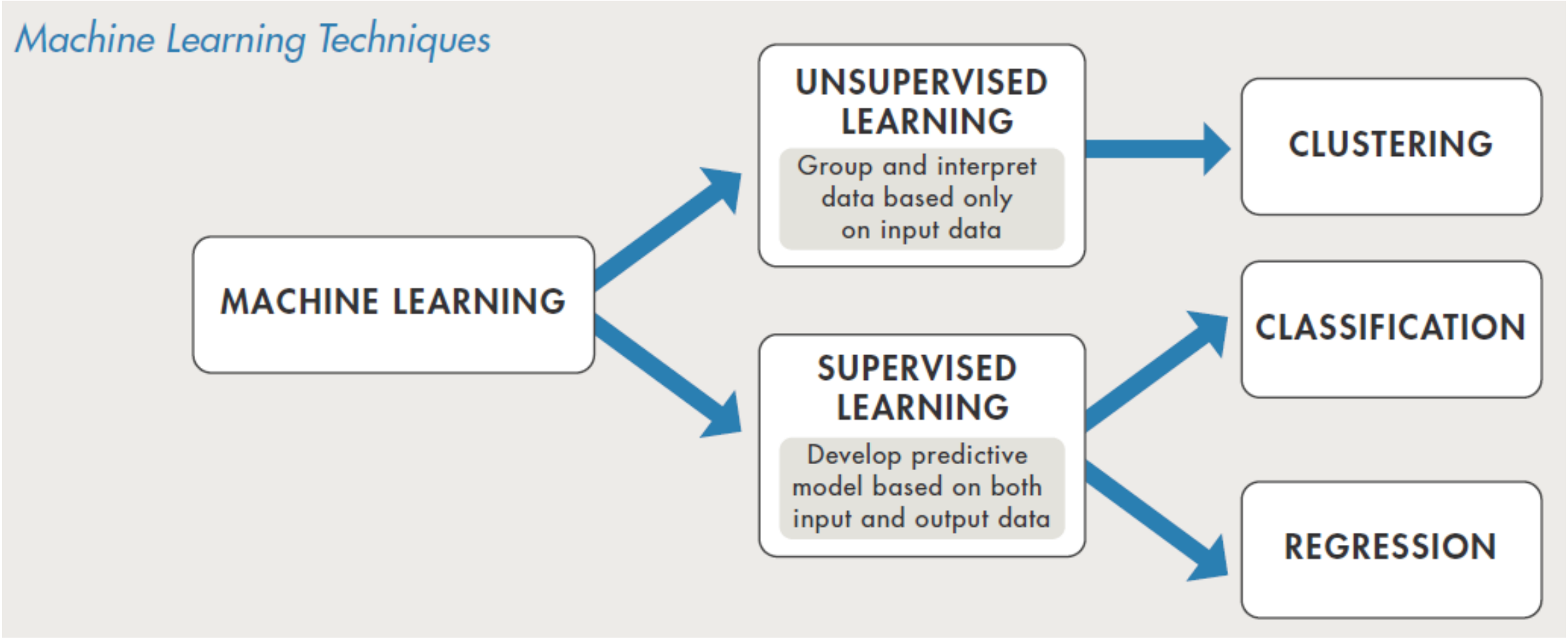
- There is no direct learning signal. We are simply trying to find structure in data.
- e.g. clustering, dimensionality reduction.

Reinforcement learning.

- The learning signal is a (scalar) reward and may come with a delay.
- e.g. trying to learn to play chess, a mouse in a maze.

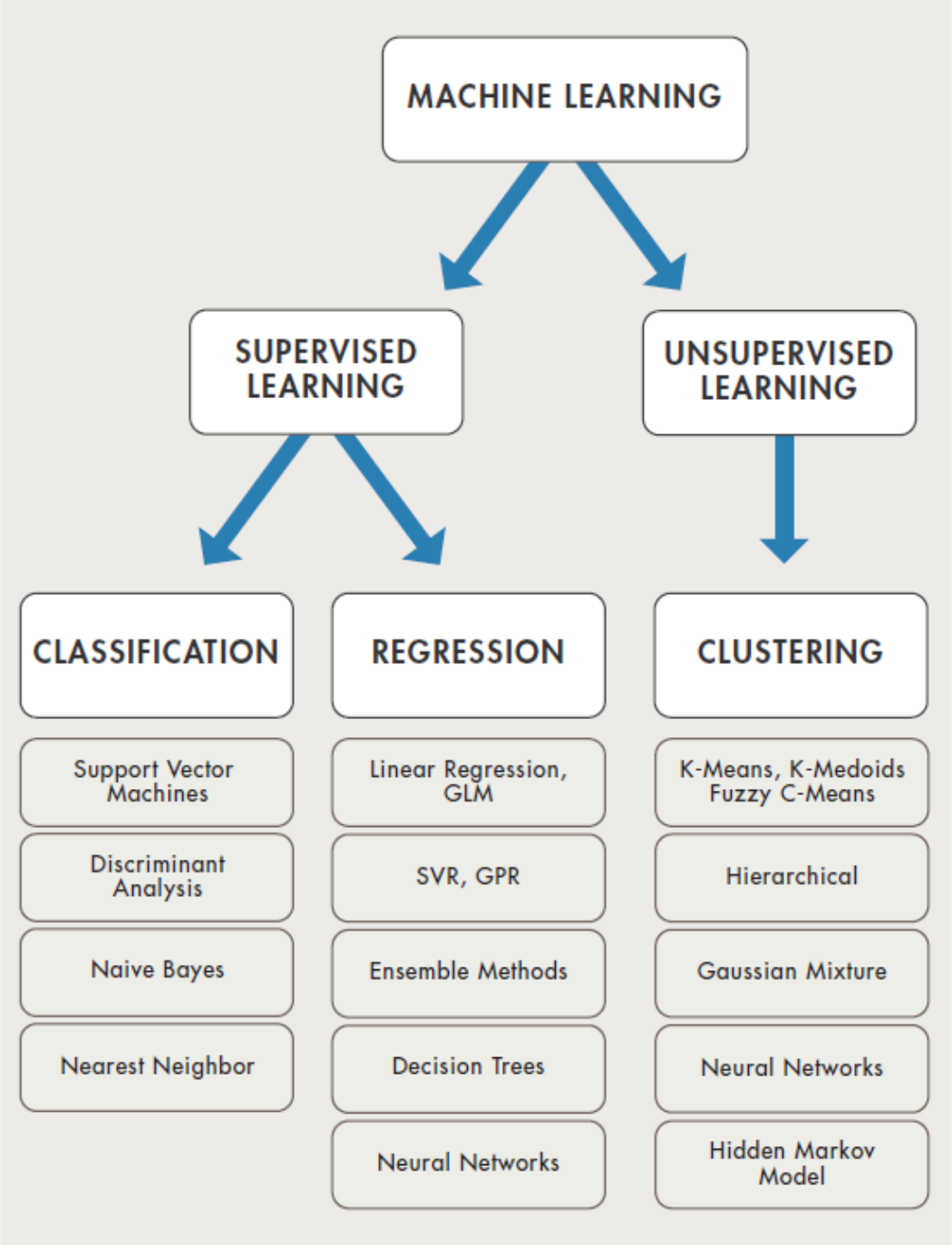
Machine learning

Oregon State University
College of Engineering



Selecting an Algorithm

- There is no best method or one size fits all.
- Finding the right algorithm is partly just trial and error.
- But algorithm selection also depends on the size and type of data you're working with, the insights you want to get from the data, and how those insights will be used.



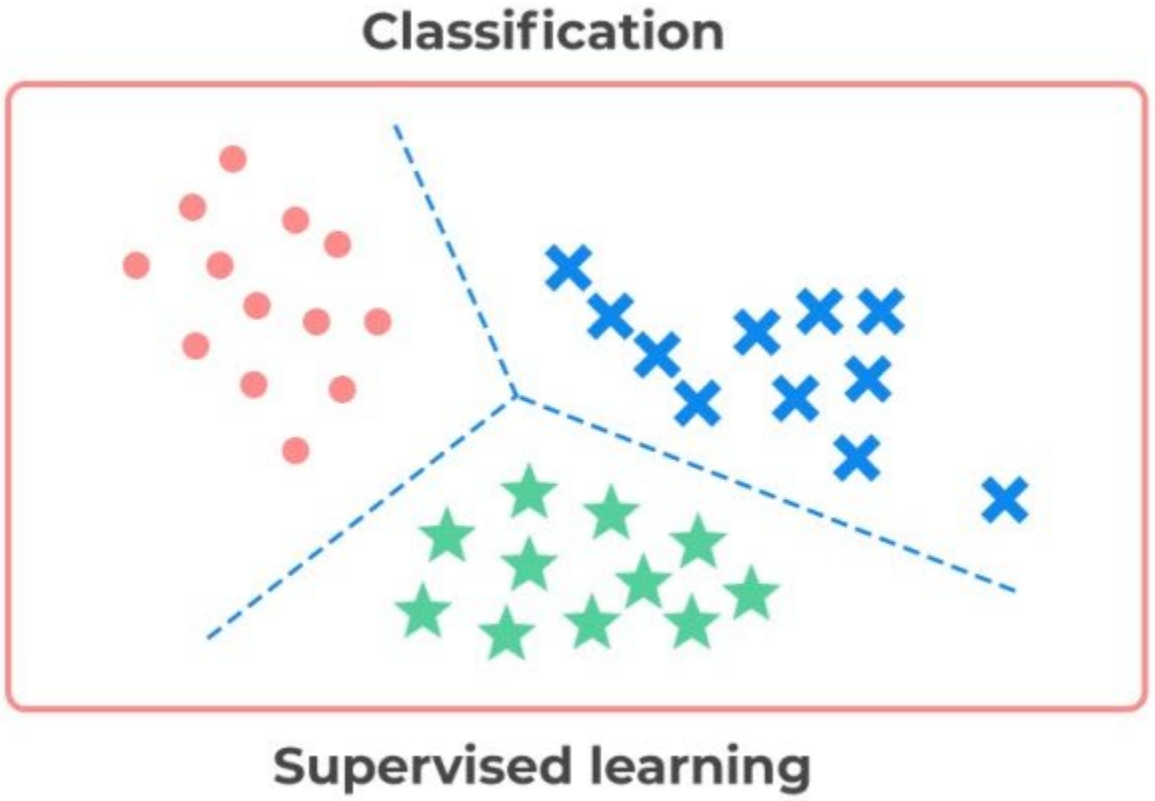
Algorithms

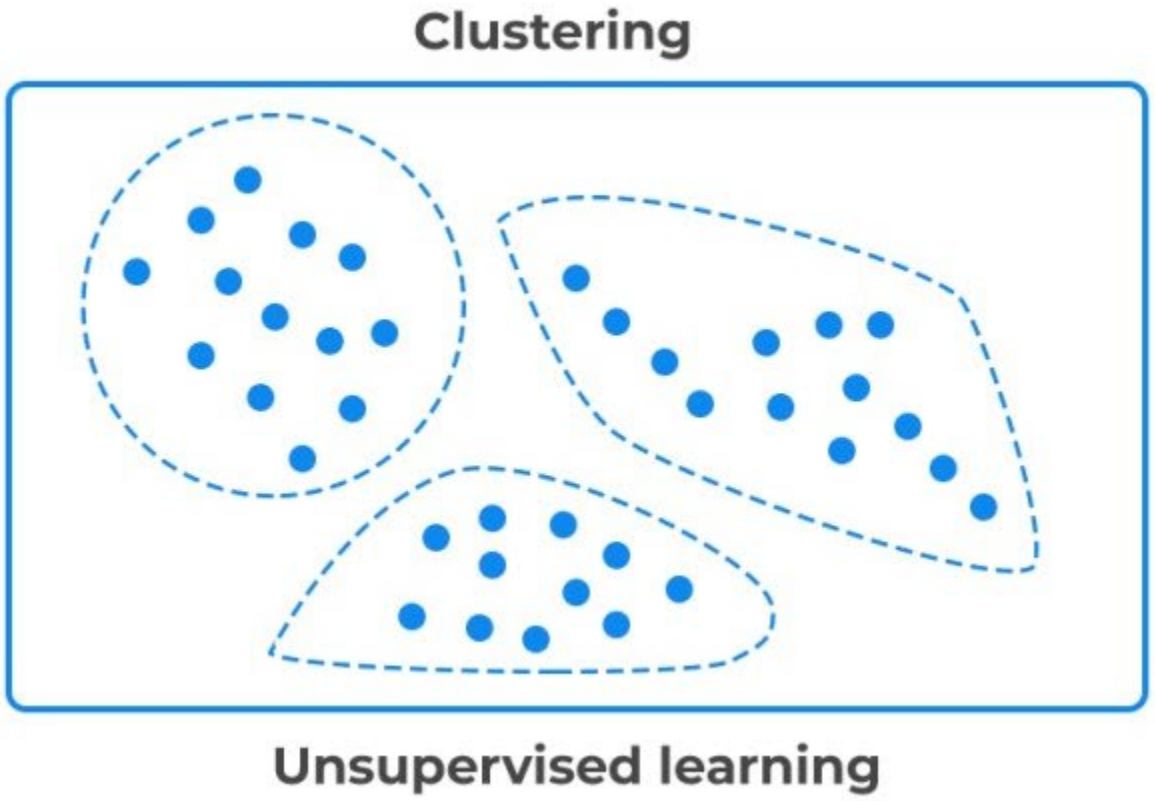
The success of machine learning system also depends on the algorithms.

The algorithms control the search to find and build the knowledge structures.

The learning algorithms should extract useful information from training examples.

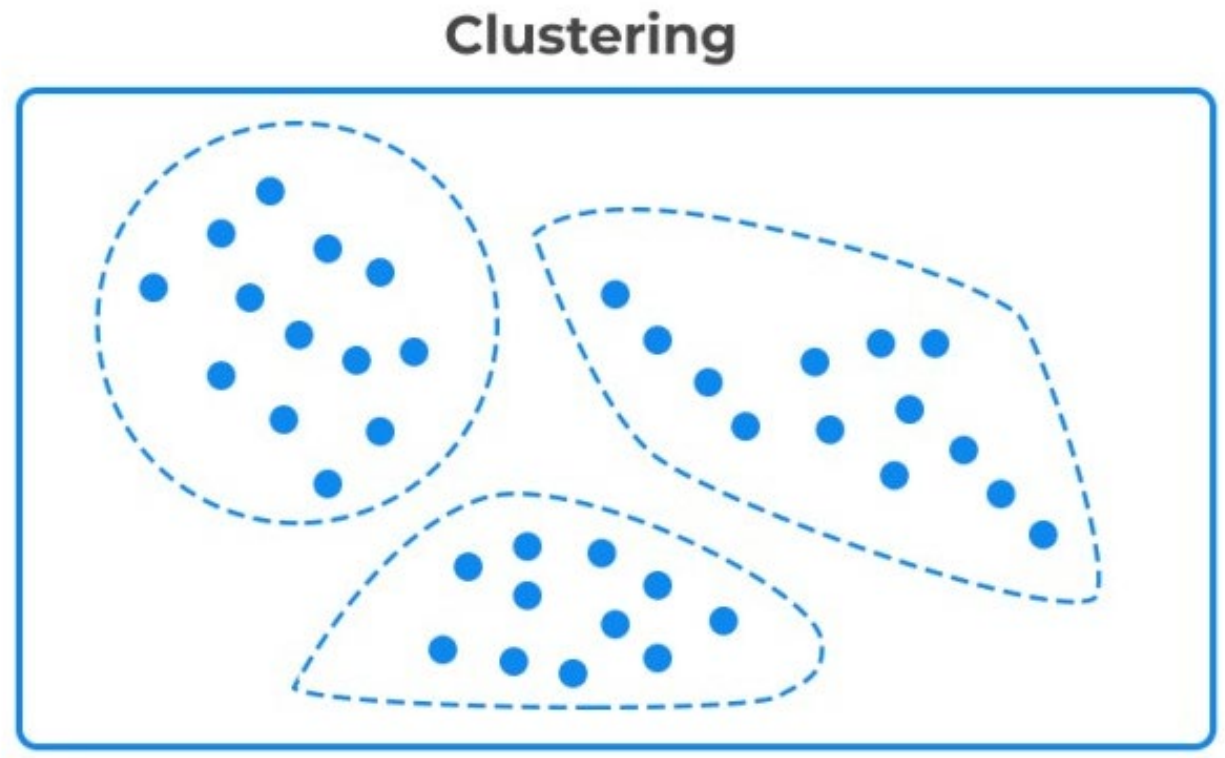
Recent Projects





Oregon State University
College of Engineering

Unsupervised Machine Learning



Oregon State University
College of Engineering

Predicting Crashes by Applying Machine Learning on New Sources of Driver Behaviour Data

- **Goal:** To identify areas on road network where heavy vehicles exhibited a high number of harsh braking events, to predict a sites crash harm potential.
- **Data:** Two data sets were considered
  - EROAD GPS Data
  - New Zealand Crash Analysis System (CAS) Data

Oregon State University
College of Engineering

Predicting Crashes by Applying Machine Learning on New Sources of Driver Behaviour Data

- **Methodology:** Unsupervised Learning
  - Machine Learning
    - DBSCAN Clustering - Density-based spatial clustering of applications with noise

RESULTS

Oregon State University
College of Engineering

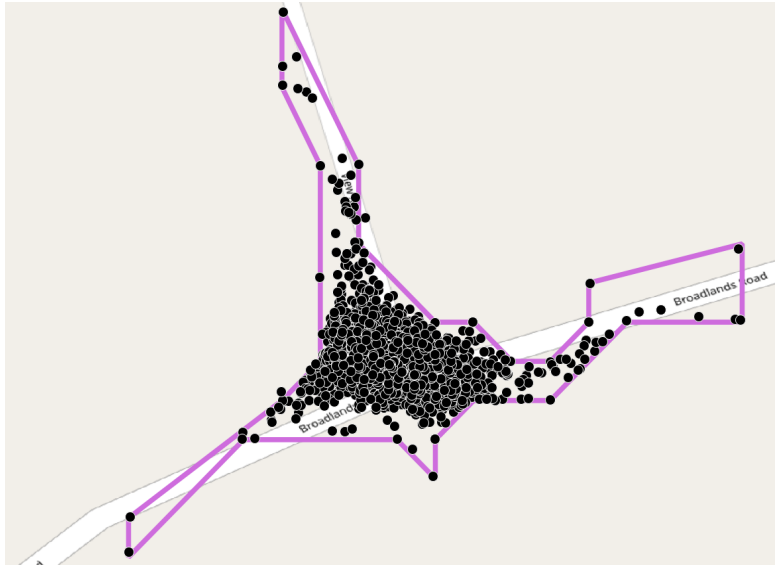


Figure 1: Harsh Braking Events at a Single Intersection Encompassed by a Concave Hull Geometry

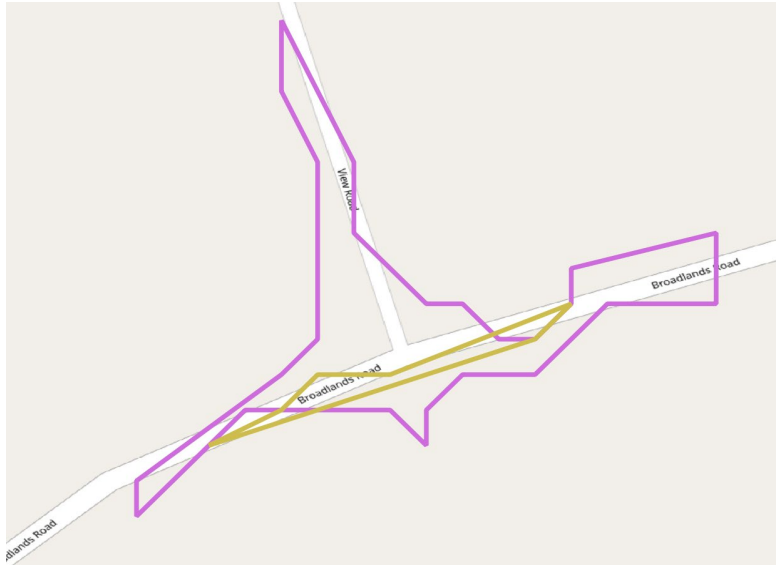


Figure 2: Concave Hull From Harsh Braking Cluster Overlapping With Concave Hull From Crash Cluster

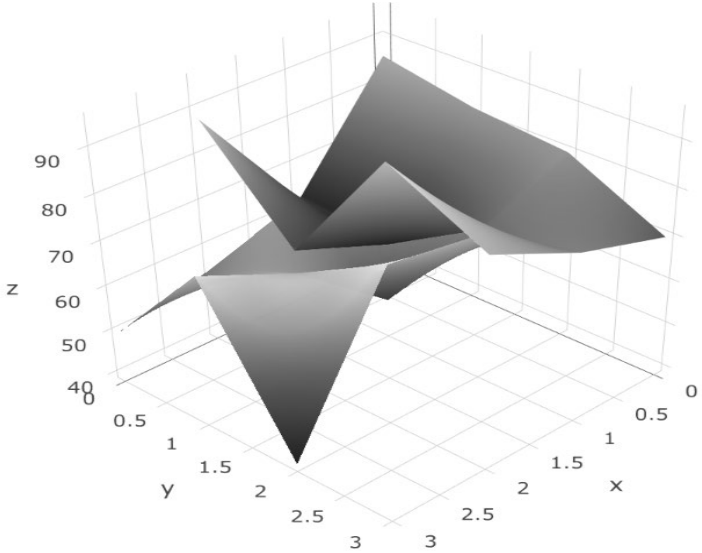


Figure 3: Illustration Representing the Overlapping Nature of Two Planes in Both 2D and 3D Space

RESULTS

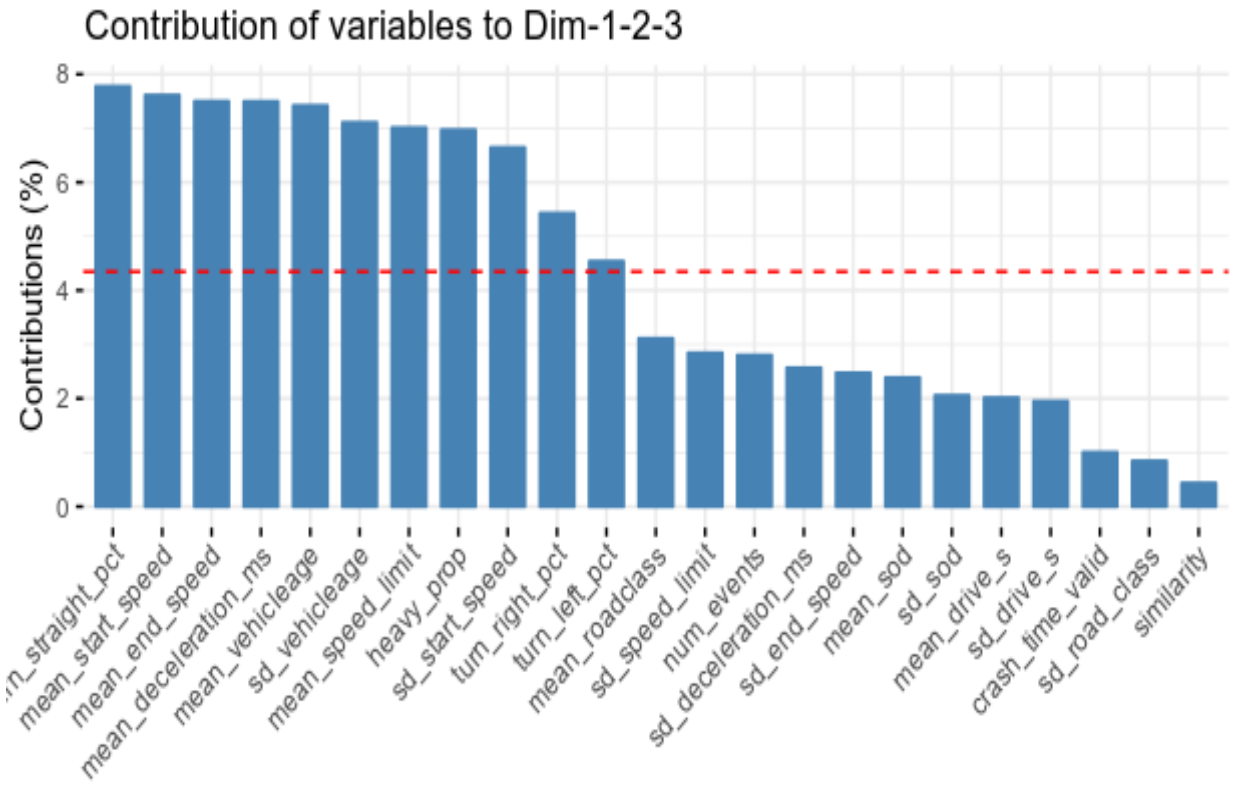


Figure 4: Results of PCA Showing Contributions of All Variables in Dimensions 1, 2, and 3

DISTRIBUTION OF PREDICTED CRASH LOCATIONS

Functional Class	***Original Clusters***	***Predicted Crash Locations***
National	215	0
Regional	402	5
Arterial	1423	26
Collector	2446	60
Access	3103	10

Oregon State University
College of Engineering

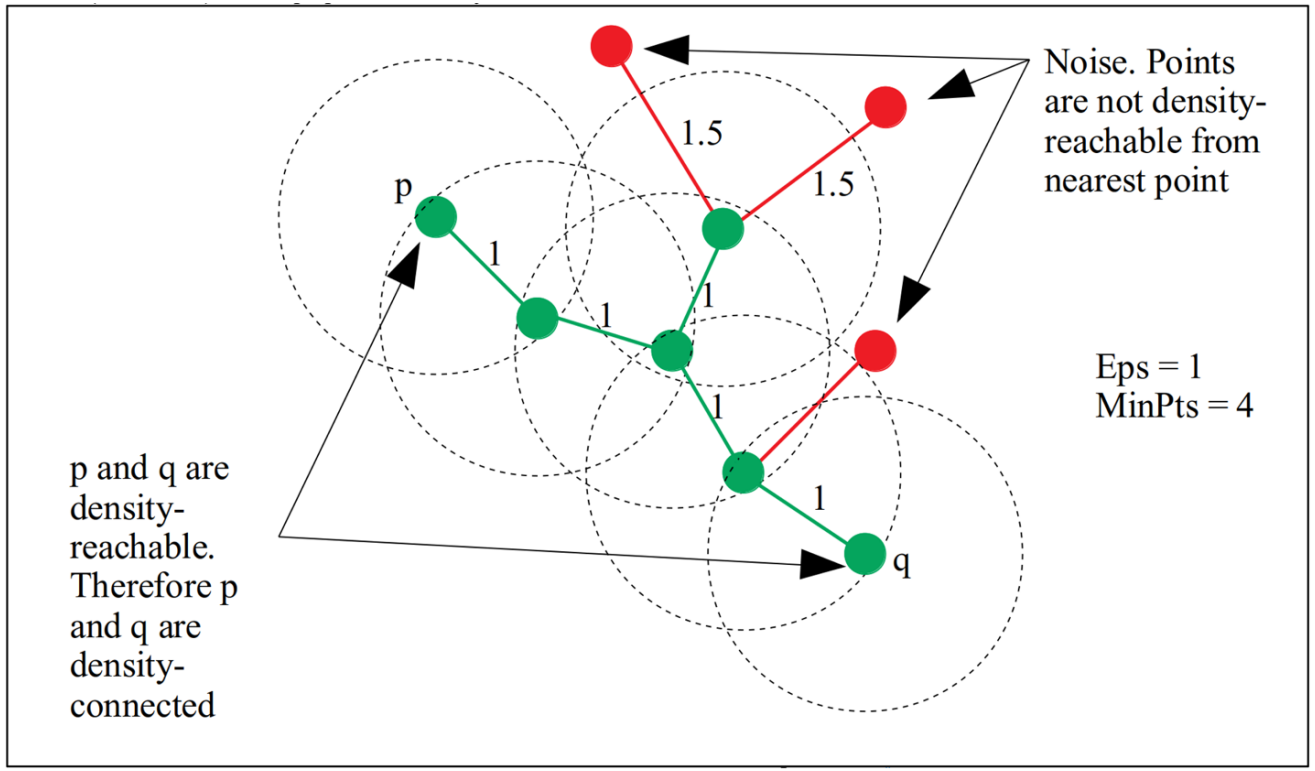
TRENDS AND CHALLENGES POSED BY MEDIUM-DUTY TRUCKS TO THE OPERATION AND SAFETY OF OREGON HIGHWAYS

- **Goal:** To attempt to estimate crash potential of Medium-duty vehicles (10,000 to 26,000 lbs.) by four potential generators
- **Data:**
  - Oregon Department of Transportation merged crash data
    - ODOT Crash Data System (CDS)
    - DMV Driver Crash Data
    - DMV Vehicle Characteristics
    - National Highway Transportation Safety Administration (NHTSA) Vehicle Data
  - Medium-duty truck generators Geospatial information

Oregon State University
College of Engineering

TRENDS AND CHALLENGES POSED BY MEDIUM-DUTY TRUCKS TO THE OPERATION AND SAFETY OF OREGON HIGHWAYS

- **Methodology:**
  Unsupervised Learning
  - –Machine Learning
    - DBSCAN Clustering - Density-based spatial clustering of applications with noise



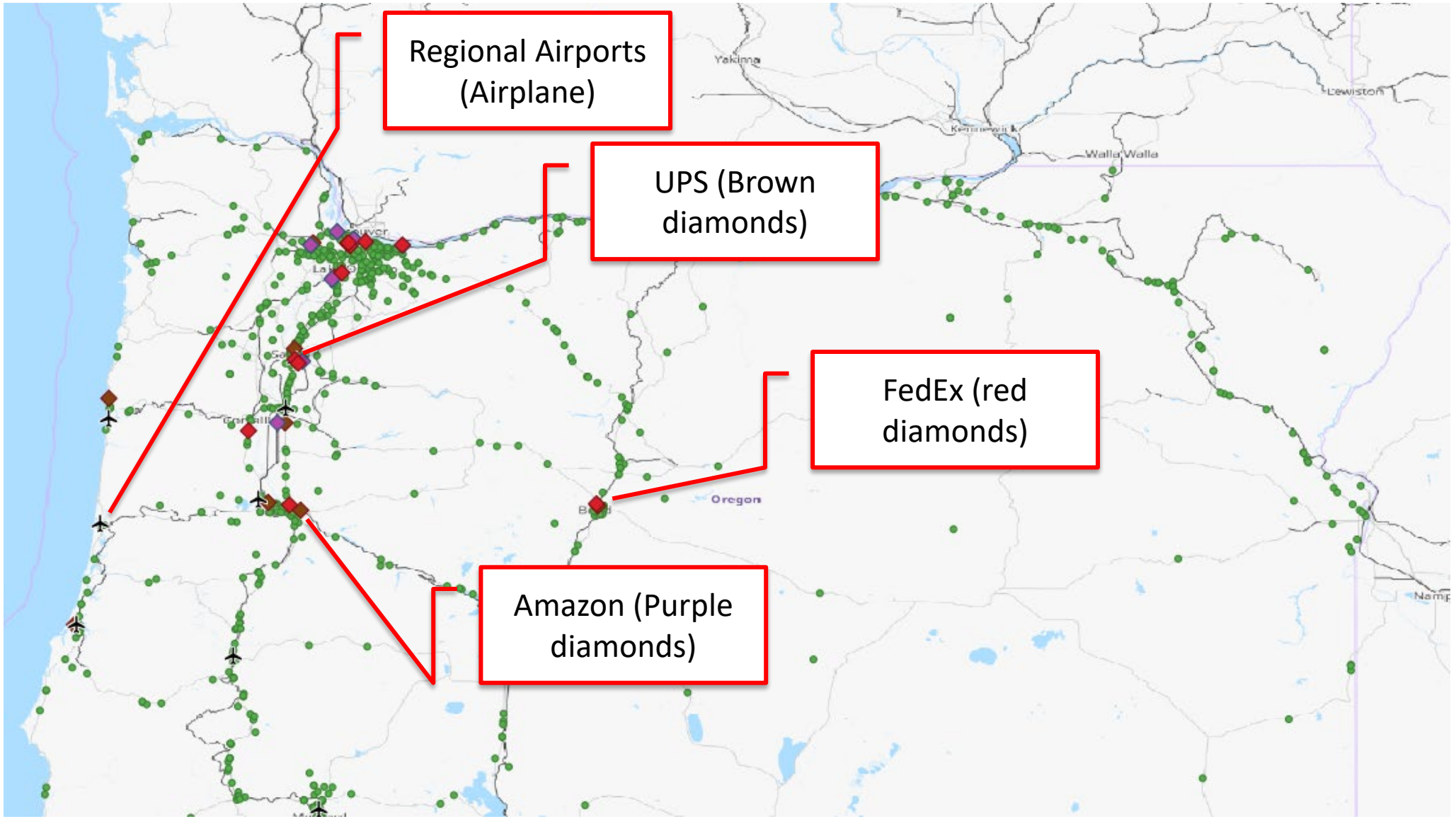


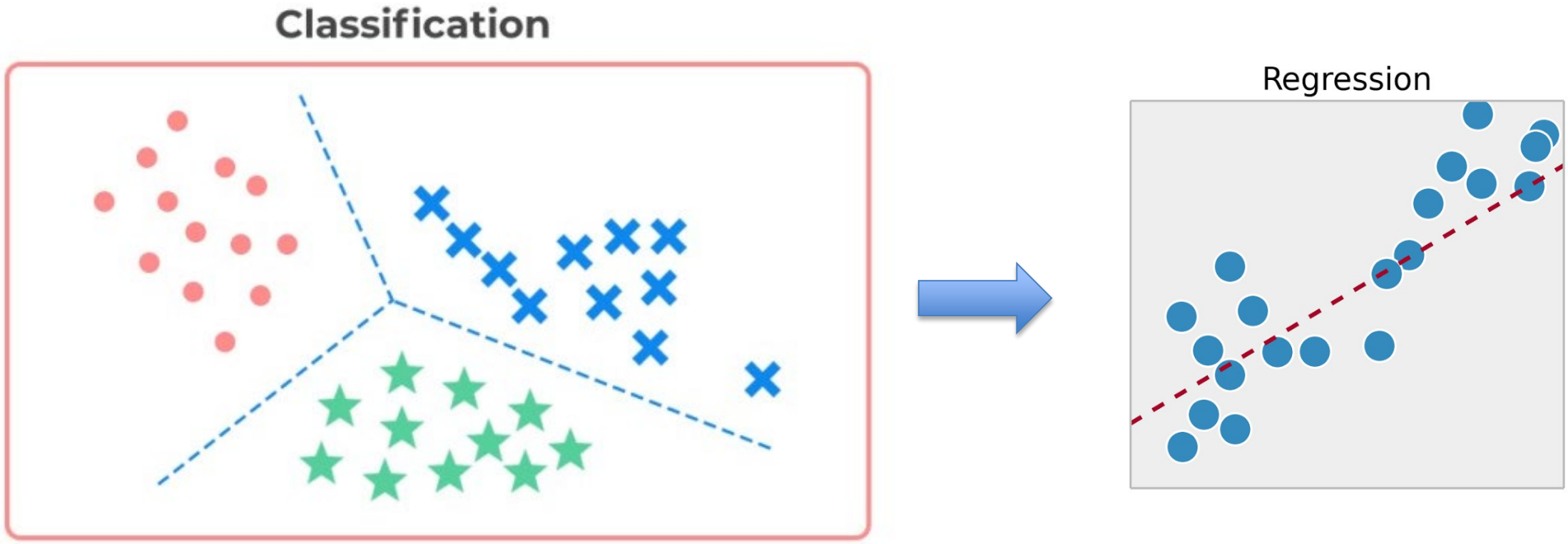
Figure 1: Spatial spread of medium duty Medium-Truck crashes and Generators in Oregon

Oregon State University
College of Engineering

Summary

- Unsupervised Machine Learning
  - Clustering
    - DBSCAN

Supervised Machine Learning



Oregon State University
College of Engineering

Roundabout Safety: Econometric and Machine Learning Models and Applications

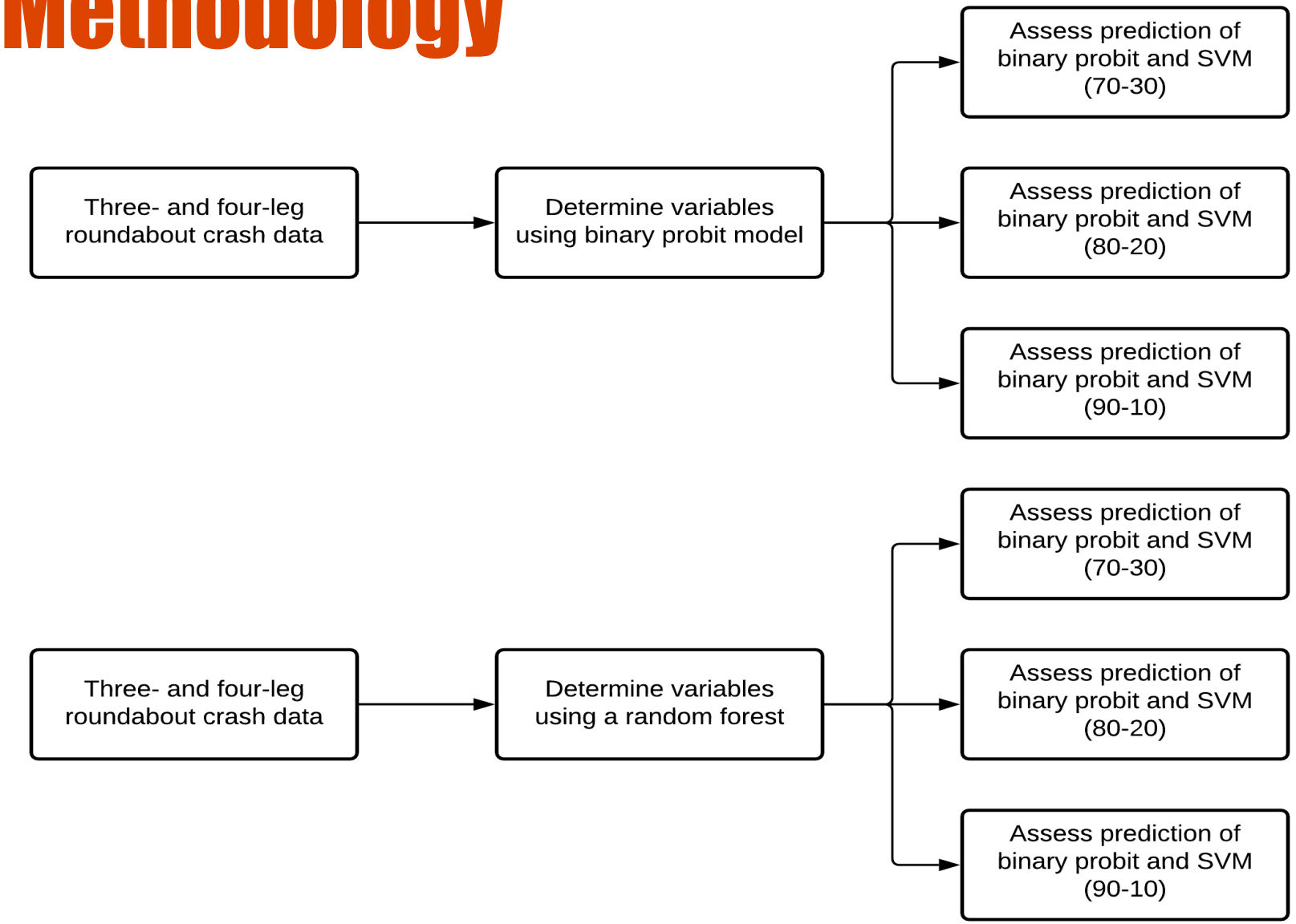
- **Goal:** To compare machine learning supervised techniques (algorithms) to econometric techniques
- **Data:** Oregon Department of Transportation Crash data at roundabouts

Oregon State University
College of Engineering

Roundabout Safety: Econometric and Machine Learning Models and Applications

- **Methodology:**
  - Econometric model
    - Random parameter binary probit model (RPBP)
  - Machine learning
    - Support vector machine (SVM)
    - Linear kernel
    - Radial (nonlinear) kernel
    - Polynomial kernel
    - Sigmoid kernal

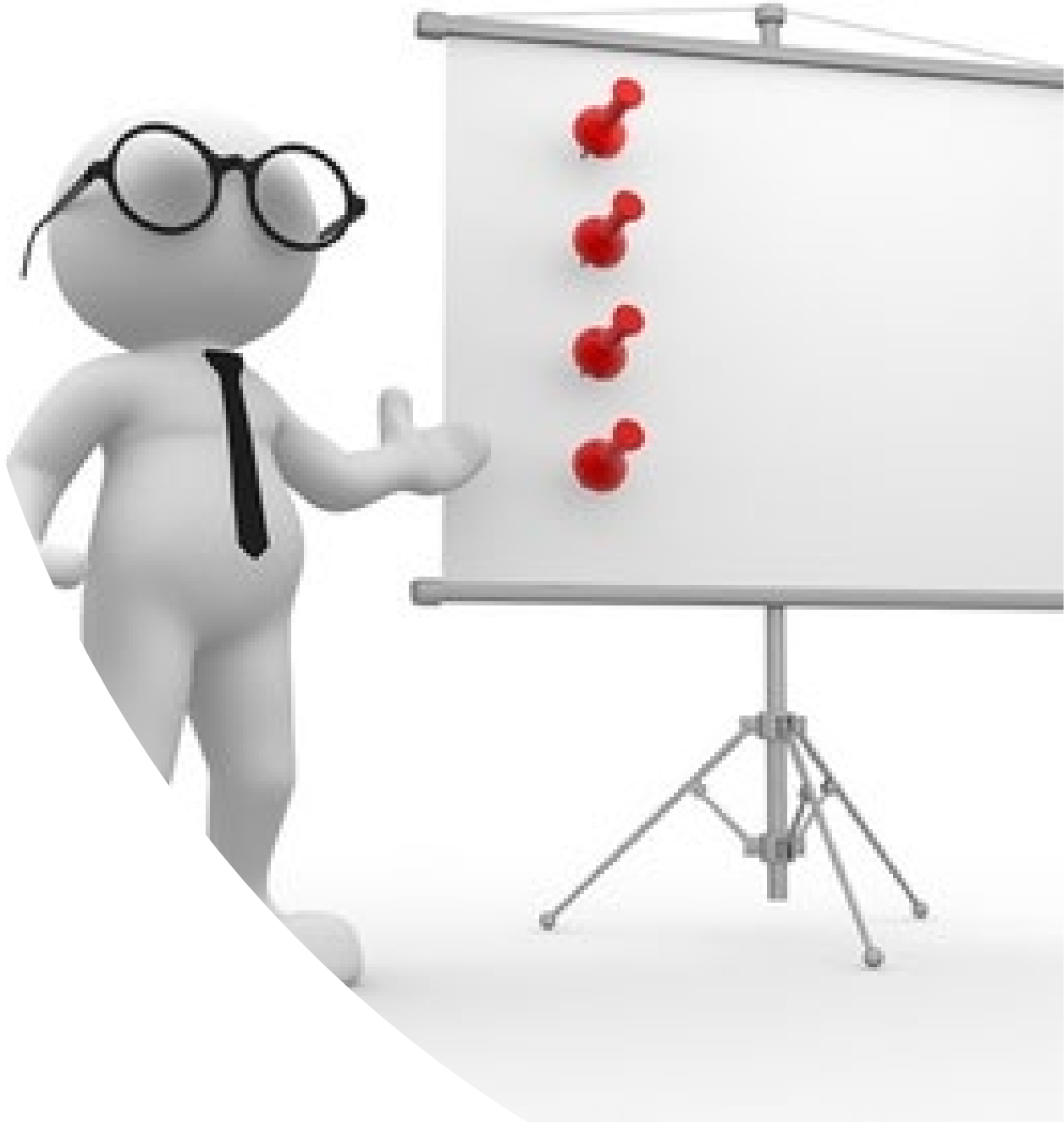
Methodology



Oregon State University
College of Engineering

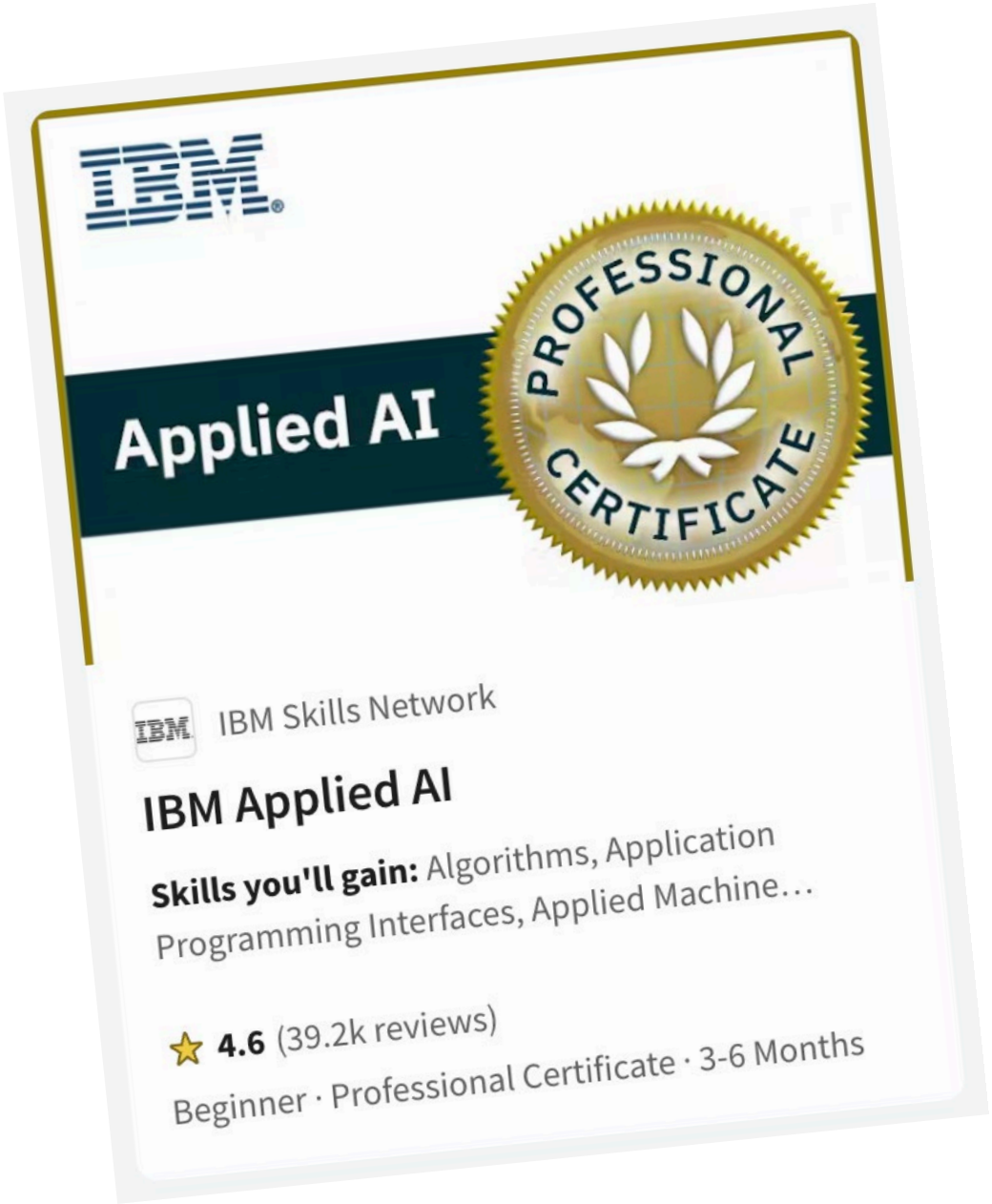
Summary

- The study compared the predictive performance of crash injury severity between various machine-learning and econometric techniques based on three-leg and four-leg roundabout crash data from 2011 to 2015 in Oregon.
- Machine-learning models outperformed the econometric model in injury severity prediction.



Learning AI

- Structured
  - Certificate Programs
- Self Taught
  - Short courses on YouTube or.,
  - free content provided by universities



Oregon State University
College of Engineering

Transportation Agency Adoption

- Collaborative Research
- Training
- Use Cases
- Data

Summary

- Given the amount of data being collected by today's transportation agencies, methods are needed to more efficient analyze and sort through the mountains of data
- Several Machine Learning techniques have already been applied, many are in process, and much more to come.

Oregon State University
College of Engineering

Thank you!

SAL HERNANDEZ, PHD,
ASSOCIATE PROFESSOR
Oregon State University
College of Engineering
School of Civil and Construction Engineering
101 Kearney Hall
Corvallis, OR 97331
Direct: 541-737-4740
Fax: 541-737-3052
Email: sal.hernandez@oregonstate.edu
www: http://research.engr.oregonstate.edu/hernandez/

Go Beavs!