# Machine Learning

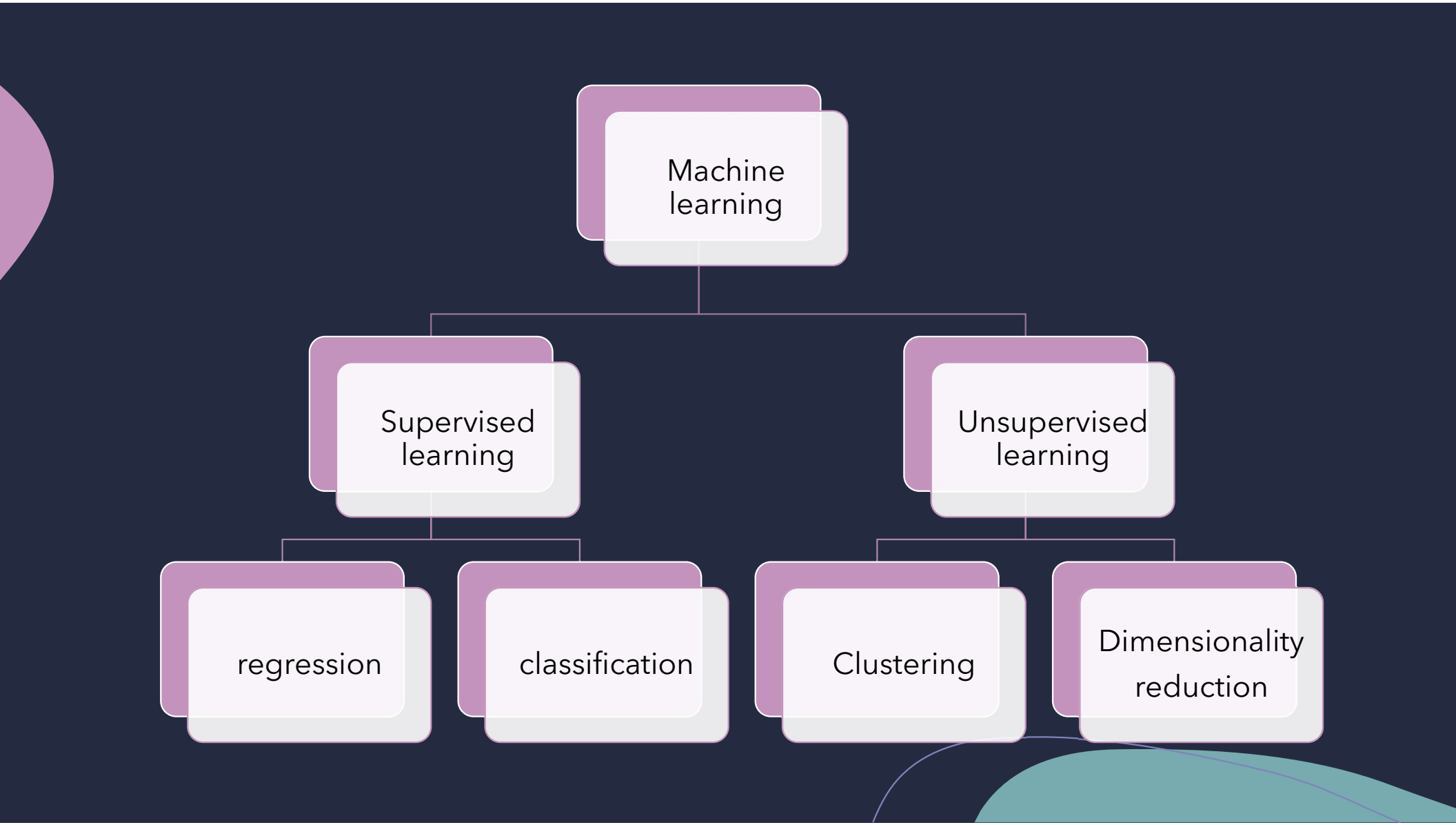
Mind-map and learnings

# Machine learning

Machine learning is a subset of artificial intelligence that enables computers to learn and improve from experience without being explicitly programmed. It involves using algorithms to analyze and find patterns in data, allowing the system to make predictions or decisions based on that data.

Machine learning is generally classified into two types

- Supervised Learning
- Unsupervised learning


Machine learning
Supervised learning
Unsupervised learning
regression
classification
Clustering
Dimensionality reduction

# Regression ( Supervised learning)

### Linear Regression

- Predicting values by creating a best fit line form the training data and this bestfit line is created by using a cost function

### Ridge and lasso Regression

- Created so that could over come few of the drawbacks of linear regression and does not lead to overfitting

### KNN algorithm (Regression)

- Can be used for both classification and regression problems

### Decision tree regressor

- the same as decision tree used for classification but there is change in the cost function used (MSE )

### XGBoost regressor

- An algorithm created by using an ensemble technique called boosting of decision trees for regression

# Classification ( Supervised learning)

## Logistic Regression

- A classification technique in machine learning which used  sigmoid function to classify samples

## Naïve bayes

- This algorithm is used when basically there are categorical values present in data and works based on the bayes probability

## KNN Algorithm

- K Nearest neighbors is classification algorithm which classifies a sample based on the k nearest samples which surround it

## Decision Trees

- This algorithm is inspired from the basic if else concept and classifies samples by forming puresets of the data

## Random forest

- Is an algorithm created by using bagging technique of decision trees

## AdaBoost

- A complex black box classification algo which classifies samples by using many weak learners by boosting them together

## XGBoost classifier

- A algorithm created by boosting binary decision trees and considering similarity weight and information gain factors to classify samples

## SVM

- The Support Vector Machine (SVM) algorithm is **a powerful machine learning method used for classification**
- It works by finding the optimal hyperplane that maximally separates data points

# Clustering (unsupervised learning)

## K Means Clustering

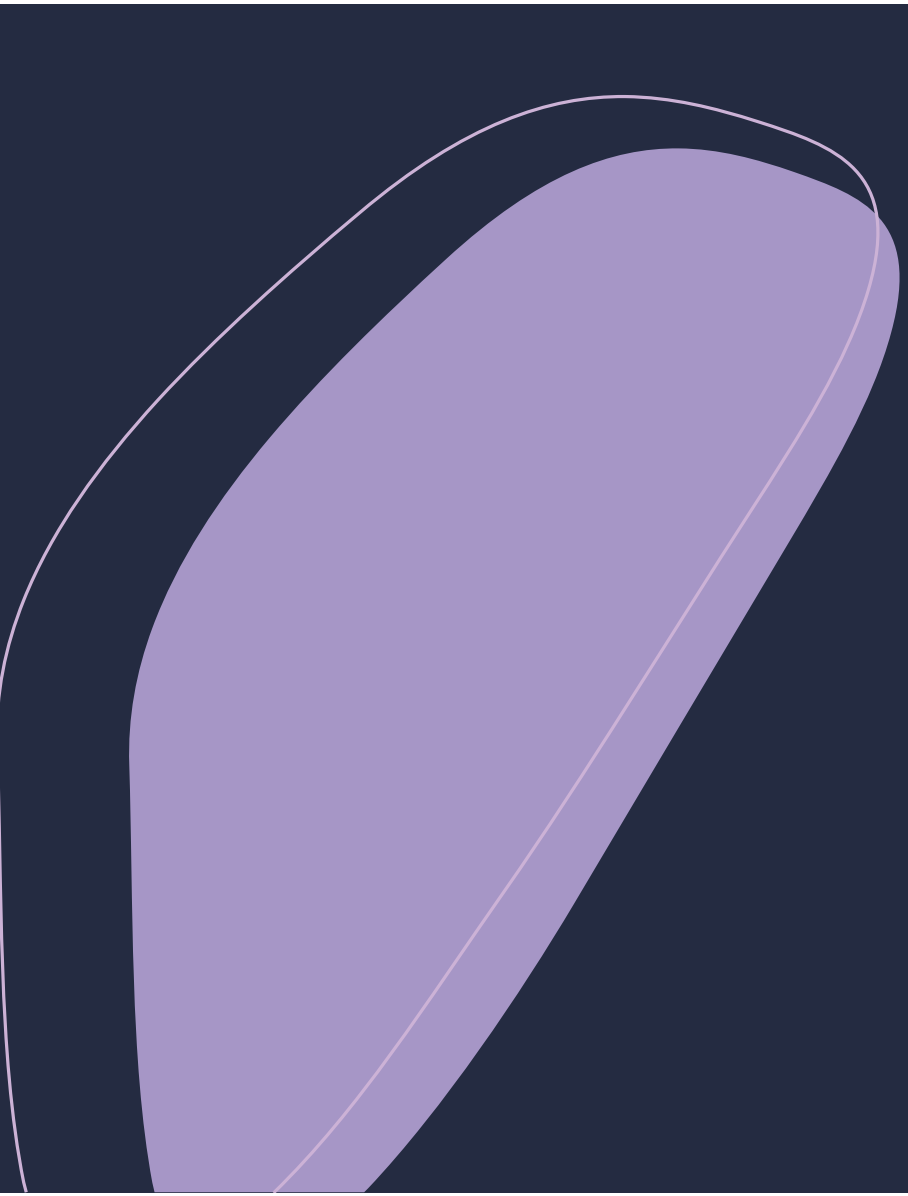
- K-means clustering is **an unsupervised machine learning algorithm used for data clustering**. It partitions a dataset into K clusters based on the distance between data points and cluster centroids

## Hierarchical Clustering

- It creates a hierarchy of clusters by repeatedly merging or splitting them based on their similarity, which is measured using distance metrics. The result is often visualized as a dendrogram to cluster datasets

## DBSCAN Clustering

- DBSCAN, which stands for Density-Based Spatial Clustering of Applications with Noise used in the clustering incase of presence of outliers



THANK YOU.....