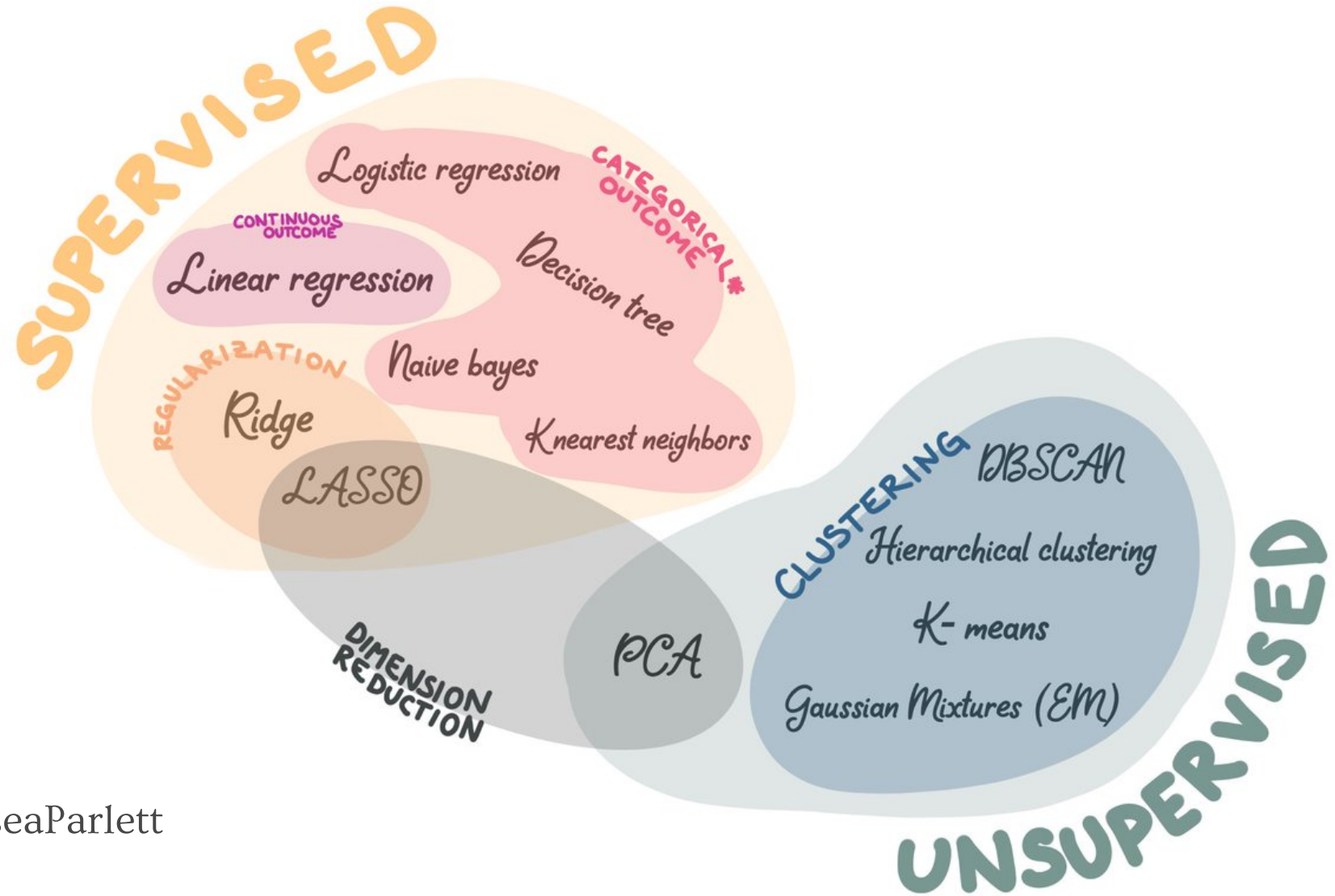
Final Review

AU STAT627

Emil Hvitfeldt

2021-11-29



By @ChelseaParlett

Unsupervised Learning

dimensionality reduction

Clustering

Dimensionality Reduction

We looked at

- Principle Component Analysis

Other methods to consider

- t-distributed Stochastic Neighbor Embedding (t-SNE)
- Autoencoder
- UMAP
- NNMF

Dimensionality Reduction

We looked at

- Principle Component Analysis
- Lasso

Not just useful in linear models

Dimensionality Reduction

We looked at

- Principle Component Analysis
- Lasso
- LDA

Clustering

We looked at

- K-means
- Hierarchical Clustering

Other models

- DBSCAN
- Gaussian Mixtures

Regularisation Methods

- Ridge
- Lasso

Serves different tasks. Can be combined in some cases

Supervised Learning

Regression

Classification

Supervised Learning

Many of the methods we looked at in this class can be used for both regression and classification

We mainly work with 2 types of trade-offs

- Flexibility / Interpretability
- Bias / Variance

Supervised Learning

The models we saw in this class lays the foundation for most models which doesn't go under the neural network umbrella

- xgboost
- lightgbm
- catboost
- stacking

Other considerations

- Implementation and run time
- What metrics are important
- Your problem statement

Thank you!