Method	Uses	Concerns
Decision Trees	• Trees handle outliers and missing observations well. • High Interpretability. • Interactions considered automatically, but implicitly. • Ensemble trees e.g. random forest and gradient boosting have the ability to increase prediction accuracy and decreases overfitting to some extent.	• Overfitting. • Unstable with a small dataset. • Unstable with outliers and noisy data. • Careful parameter tuning required.
Penalized Regression	• Supervised regression or classification. • Modeling linear or non- linear occurrence by specifying interactions terms. • Parsimonious model. • When interpretability is important.	• Standardization needed. • Careful parameter tuning required. • Treat missing and outliers values prior to algorithm implementation.
K-Mean	• Unsupervised clustering. • Finding similar observations to form the groups in a dataset without labels.	• Standardization needed. • Finding an optimal number of K. • Sensitive to missing values and outliers.
Hierarchical Clustering	• Unsupervised clustering. • Create a known number of overlapping clusters of different sizes.	• Standardization needed. • An optimal number of clusters. • Sensitive to missing values and outliers. • Curse of dimensionality.
Support Vector Machines (SVM)	• Modeling linear and non- linear occurrence by using linear and non-linear kernels. • Capture much more complex relationships between observations.	• Low interpretability. • Computationally intensive. • Missing values and outliers. • Standardization needed. • Parameter tuning.
Neural Networks	• Pattern recognition in images, videos etc. • Unsupervised feature extraction. • Supervised regression or classification. • Anomaly detection with autoencoder networks.	• Low interpretability. • Missing values and outliers. • Standardization needed. • Parameter tuning. • Computationally intensive.