Sl No	Terminologies	Classification (Supervised)				
		Decision Tree	Random Forest	Support Vector Machines	Naïve Bayes Classifier	K-Nearest Neighbor
	Key Concepts	**Gini, Entropy**	**Ensemble:** **Bagging (random sampling with replacement)** **Pasting (without replacement)** **Boosting (Adaptive, Gradient)** **Stacking** **OOB, Variable Importance, n_estimators**	**C (high) = Penalty for Misclassification** **Gamma(Low) = Large Margin (preferred)** **Low Gamma (~0.01) + High C (~1) = High Accuracy** **Kernels: rbf (radial basis func), Poly, Linear**	**Assumes independence of all X Variables** **Easy & Fast**	**Eucledian (Minkowski) distance with power 2** **Lazy Learners**
1	**Category**	Multi Class Classifier & Regression	Multi Class Classifier & Regression	Multi Class Classifier & Regression	Multi Class Classifier	Multi Class Classifier & Regression
2	**Linear/Non-Linear**	Linear & Non Linear	Linear & Non Linear	Linear & Non Linear	Linear & Non Linear	Linear Data Only
3	**Method**	Not Distance Based	Probablistic, Not Distance Based	Distance Based (due to SV Margin)	Probablisic, Not Distance Based	Distance (proximity) Based
4	**Scaling**	Not Required	Not Required	Required	Not Required	Required
5	**Correlation Check**	Not Required	Not Required	Not Required	Required	Required
6	**Outlier Treatment**	Not Required	Not Required	Not Required	Required	Required
7	**H/W Requirement**	Low	High	Very High	Low	High
8	**Reliability**	Low Reliability as varying the proportion of 0 and 1 in Target Variable will vary the splits. This leads to Biasing	High Reliability due to Randomness in variables and Values	High Reliability	Algorithm automatically Replace 0's with a very low non 0 value, called Laplace Correction	High Accuracy
9	**Comments**	# Greedy Algorithm # Re-Training does not happen everytime # Explainability is available # Pruning possible: max depth, sample leaf, sample split etc	# Greedy Algorithm # Re-Training does not happen everytime # Black Box - No Explainability # Pruning possible only at Tree Level # Mostly used in Industry. Best Algorithm. Cant tell in Interview	# Good when No of variables > No of Rows # Poly, RBF, Linear Kernels converts data to higher dim for segregation, # Classification boundary is supported by 2 opposite class vectors, hence the name # Mostly used in Industry as it can handle both Linear and Non Linear	# Efficient for Text analytics and not for Numerical Data due to biasing # Good for higher dim data as proba of 2 features being close reduces # Very fast and simple model with no hyper parameters # Used when we cannot drop correlated variables	# Good for small datasets # Lazy Learners: Re-Training happens for every prediction # Very efficient for Text analytics
10	**Use Cases**	Capital Budgeting	Credit Card Fraud Detection	Handwriting Recognition, Facial Expression Classify, Genes classify	Sentiment Analysis, Text Analysis	Text Mining, Forest Inventory