**Comparison of Different Algorithms**

**1.Linear Regression:**

- As the linear regression is a regression algorithm, we will compare it with other regression algorithms

- One basic difference of linear regression is, LR can only support linear solutions.

***LR vs Decision Tree****:*

* Decision trees supports non linearity, where LR supports only linear solutions.
* When there are large number of features with less data-sets(with low noise), linear regressions may outperform Decision trees/random forests. In general cases, Decision trees will be having better average accuracy.
* For categorical independent variables, decision trees are better than linear regression.
* Decision trees handles colinearity better than LR.

***LR vs SVM:***

* **SVM supports both linear and non-linear solutions using kernel trick.**
* **SVM handles outliers better than LR.**
* **Both perform well when the training data is less, and there are large number of features.**

***LR vs KNN :***

* **KNN is a non -parametric model, whereas LR is a parametric model.**
* **KNN is slow in real time as it have to keep track of all training data and find the neighbor nodes, whereas LR can easily extract output from the tuned θ coefficients.**

**2. KNN (K-Nearest Neighbors):**

**-** A general difference between KNN and other models is the large real time computation needed by KNN compared to others.

***KNN vs naive bayes:***

* Naive bayes is much faster than KNN due to KNN’s real-time execution.
* Naive bayes is parametric whereas KNN is non-parametric.

***KNN vs linear regression :***

* KNN is better than linear regression when the data have high SNR.

***KNN vs SVM :***

* SVM take cares of outliers better than KNN.
* If training data is much larger than no. of features(m>>n), KNN is better than SVM. SVM outperforms KNN when there are large features and lesser training data.

**3.** **Decision Tree:**

***Decision tree vs Random Forest :***

* Random Forest is a collection of decision trees and average/majority vote of the forest is selected as the predicted output.
* Random Forest model will be less prone to overfitting than Decision tree, and gives a more generalized solution.
* Random Forest is more robust and accurate than decision trees.

***Decision tree vs KNN :***

* Both are non-parametric methods.
* Decision tree supports automatic feature interaction, whereas KNN cant.
* Decision tree is faster due to KNN’s expensive real time execution.

***Decision tree vs naive Bayes :***

* Decision tree is a discriminative model, whereas Naive bayes is a generative model.
* Decision trees are more flexible and easy.
* Decision tree pruning may neglect some key values in training data, which can lead the accuracy for a toss.

***Decision tree vs SVM :***

* SVM uses kernel trick to solve non-linear problems whereas decision trees derive hyper-rectangles in input space to solve the problem.
* Decision trees are better for categorical data and it deals colinearity better than SVM.

**4. Support Vector Machine**

***SVM vs Random Forest :***

* Random Forest supports multiclass classification,whereas SVM needs multiple models for the same.
* Random Forest can give a probability over the prediction, whereas SVM cannot give.
* Random Forest deals categorical data better than SVM.

***SVM vs Naive Bayes :***

* Both performs better with low amount of training data and large features.
* If features are mutually dependent, SVM outperforms Naive Bayes.
* SVM is a discriminative model whereas NB is generative model.

References:

[Comparative Study on Classic Machine learning Algorithms | by Danny Varghese | Towards Data Science](https://towardsdatascience.com/comparative-study-on-classic-machine-learning-algorithms-24f9ff6ab222)

[Comparative Study on Classic Machine learning Algorithms , Part-2 | by Danny Varghese | Medium](https://medium.com/@dannymvarghese/comparative-study-on-classic-machine-learning-algorithms-part-2-5ab58b683ec0)