4. Top two scores :
3.51859-- For Gradient Boosting Regressor
5.78872-- For AdaBoost Regressor

The Features that are used are:

1. Dropped the **dropoff_longitude, dropoff_latitiude, pickup_longitude, pickup_longitude** and added 2 new columns that had the difference between the dropoff_latitude and the pickup_latitude and the **dropoff_longitude** and **pickup_longitude** called as diff_latitude and diff_longitude
2. Removed the rows that had **fare_amount** lesser than 0.
3. Removed the rows that had **no_of_passengers** lesser than 0 and less than 6.
4. Split the **pickup_datetime** column into date and time.
    a. From date, we get the day of the week and set 1 if its a weekend and 0 if its a weekday
    b. From time, we try to figure out the night shift(10pm - 6am) and set that time to 1 and others as 0

Gradient Boosting vs Adaboost

- Both AdaBoost and Gradient Boosting build weak learners in a sequential fashion. Originally, AdaBoost was designed in such a way that at every step the sample distribution was adapted to put more weight on misclassified samples and less weight on correctly classified samples. The final prediction is a weighted average of all the weak learners, where more weight is placed on stronger learners.
- **Gradient Descent is a generalization of Adaboost where the objective function is now not constrained to the exponential loss** and where the weak learner are learned in a greedy fashion.
- In Complex problems, stumps are not enough, when **strong base learners are used, Adaboost falls behind and never reaches the performance of Gradient Boosting**
- **AdaBoost can be sensitive to noisy data and outliers.**
- Adaboost re-weights the training data at each iteration over these weights while gradient boosting simply does a regression over the negative gradients.

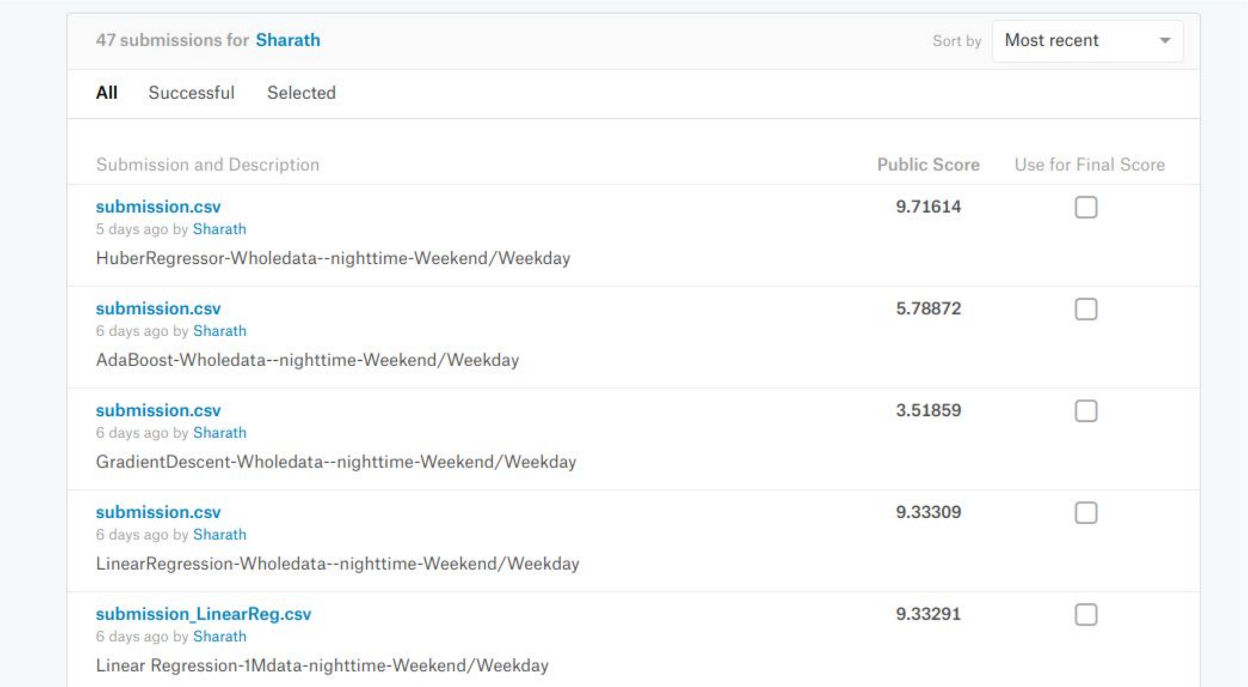
Ensemble methods vs Linear Regression

- Ensemble methods are meta-algorithms that combine several machine learning techniques into one predictive model in order to **decrease variance**(bagging), **bias** (boosting), or **improve predictions** (stacking).
- ensemble methods, such as boosting and blending, work by taking the outputs from individual models, together with the training data, as inputs to a bigger model
- Linear regression models are also very much affected by outliers.

Scores obtained (Tried these 3 models for multiple feature sets but found these feature sets to be the best among them):

3.51859-- For Gradient Boosting Regressor
5.78872-- For AdaBoost Regressor
9.33309-- For Linear Regression

Screenshot of the submissions:



47 submissions for Sharath

Sort by Most recent

All Successful Selected

Submission and Description	Public Score	Use for Final Score
submission.csv 5 days ago by Sharath HuberRegressor-Wholedata--nighttime-Weekend/Weekday	9.71614	□
submission.csv 6 days ago by Sharath AdaBoost-Wholedata--nighttime-Weekend/Weekday	5.78872	□
submission.csv 6 days ago by Sharath GradientDescent-Wholedata--nighttime-Weekend/Weekday	3.51859	□
submission.csv 6 days ago by Sharath LinearRegression-Wholedata--nighttime-Weekend/Weekday	9.33309	□
submission_LinearReg.csv 6 days ago by Sharath Linear Regression-1Mdata-nighttime-Weekend/Weekday	9.33291	□