**Approach: Jobathon (Sep-2021)**

**By Shubhadip Biswas**

Features of Date has been created like 'Weekend', 'quarter’, 'month', 'year', ‘day’ , ‘weekday’ etc.

Trend of Sales with different features has been observed:

Chart

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with low confidence

Graphical user interface

Description automatically generated

**Four models have been developed after parameter tuning. Weighted average models and individual models have been checked for MSLE.**

Lags of sales, moving average of lags, lags difference, and dy (lags difference)/dt have been calculated and used in the model.

For Final model following variables were considered:

Discount,Holiday,weekday,Weekend,quarter,weekofmonth,Sales\_lag\_1,Sales\_lag\_2,Sales\_lag\_3,Sales\_lag\_4,Sales\_lag\_5,Sales\_lag\_6,Sales\_lag\_7,Sales\_lag\_8,Sales\_lag\_9,Sales\_lag\_10,Sales\_lag\_11,Sales\_lag\_12,Sales\_lag\_13,Sales\_lag\_14,Sales\_lag\_15,Sales\_lag\_16,Sales\_lag\_17,Sales\_lag\_18,Sales\_lag\_19,Sales\_lag\_20,Sales\_lag\_21,Sales\_lag\_22,Sales\_lag\_23,Sales\_lag\_24,Sales\_lag\_25,Sales\_lag\_26,Sales\_lag\_27,Sales\_lag\_28,Sales\_lag\_29,Sales\_lag\_30,Sales\_lag\_31,Sales\_lag\_32,Sales\_lag\_33,Sales\_lag\_34,Sales\_lag\_35,Sales\_lag\_36,Sales\_lag\_37,Sales\_lag\_38,Sales\_lag\_39,Sales\_lag\_40,Sales\_lag\_41,Sales\_lag\_42,Sales\_lag\_43,Sales\_lag\_44,Sales\_lag\_45,Sales\_lag\_46,Sales\_lag\_47,Sales\_lag\_48,Sales\_lag\_49,Sales\_lag\_50,Sales\_lag\_51,Sales\_lag\_52,Sales\_lag\_53,Sales\_lag\_54,Sales\_lag\_55,Sales\_lag\_56,Sales\_lag\_57,Sales\_lag\_58,Sales\_lag\_59,Sales\_lag\_60,Sales\_lag\_61,Sales\_rolling\_15,Sales\_rolling\_10,Sales\_rolling\_9,Sales\_rolling\_8,Sales\_rolling\_7,Sales\_rolling\_6,Sales\_rolling\_5,Sales\_rolling\_4,Sales\_rolling\_3,Sales\_rolling\_2,Sales\_Diff\_15,Sales\_Diff\_10,Sales\_Diff\_9,Sales\_Diff\_8,Sales\_Diff\_7,Sales\_Diff\_6,Sales\_Diff\_5,Sales\_Diff\_4,Sales\_Diff\_3,Sales\_Diff\_2,Sales\_Difft\_15,Sales\_Difft\_10,Sales\_Difft\_9,Sales\_Difft\_8,Sales\_Difft\_7,Sales\_Difft\_6,Sales\_Difft\_5,Sales\_Difft\_4,Sales\_Difft\_3

Model has been developed for each store. To calculate lags of future data, step wise model has been implemented.

The MSLE of training, and test has been evaluated to reach to the best fit model.

Finally LightGBM model has been considered for final solution with following parameter:

colsample\_bytree= 0.66, learning\_rate= 0.045, max\_depth= 10, max\_features= 93, min\_impurity\_decrease= 0, min\_impurity\_split= 1e-07, min\_samples\_leaf= 1, min\_data\_in\_leaf=2,verbose=-1,

min\_samples\_split= 95, n\_estimators= 150, num\_leaves= 15, reg\_alpha= 1, reg\_lambda= 1,

subsample= 0.4,random\_state=1,metric='mae',max\_iter=4.