Day - 5

06 December 2024 10:10

1. Event Handler in Dashboard.
2. Integration is with Dashboard.
3. Prediction in splunk.
4. Basic of MLTK Toolkit.
5. Command - eventcount
6. Knowledge object - field extraction.
7. DBConnect.
8. Data ingestion linux/Windows Server.

(1) Eventcount:- | eventcount summarize = true/false index=*

(2) Field Extraction:- (1) Regular Expression.
(2) Delimiter type.

(1) Regular Expression:- Pattern we will pull the data.

(2) Delimiter:- Split the event on the basis of the certain symbols like space, comma, special character etc.

(3) Predict Command:- Future value, historical Data point.

LL - Local level
LLP - Local Level Trend
LLP - Seasonal Local Level
LLB - Bivariate local level
BiLL - Bivariate local Level

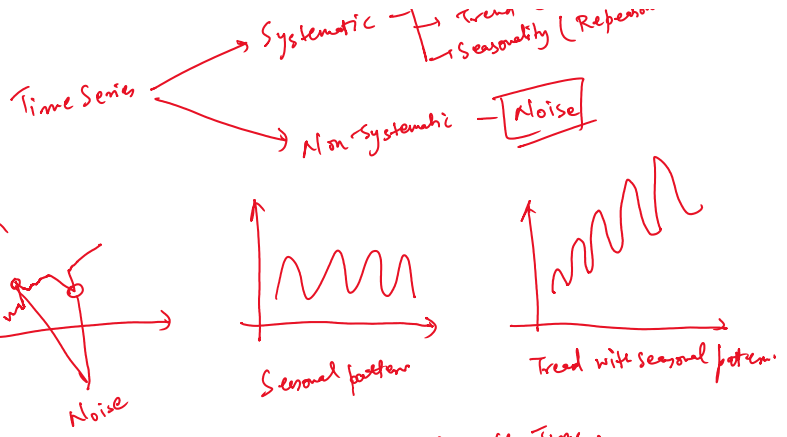
lower 95 - 95% confidence
upper 95 - 95% confidence

95
100
105

holdback = <n> = Exclude last n data point from Prediction.
future_timespan = <n> = how far into the future to predict (No. of Time Unit)

(1) LL (Local Level) Algo.:- Value = Noise.
↓ ↓
True value. fluctuation/random Variation.

Time Series → Systematic → level (Avg. value of series)
Trend (Increase/decrease value in the series)
Seasonality (Repeation short-term cycle in the series)



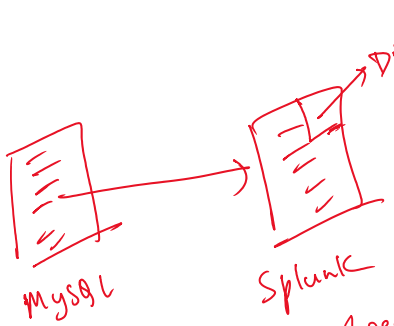
Univariate:- Only Variable is varying over Time.

ex → Collect temp. data of the room every second.

MultiVariate Time Series:- Multiple Variable are Varying over time.

Ex → tri-axial acceleration.
(x, y, z)

DB Connect:-

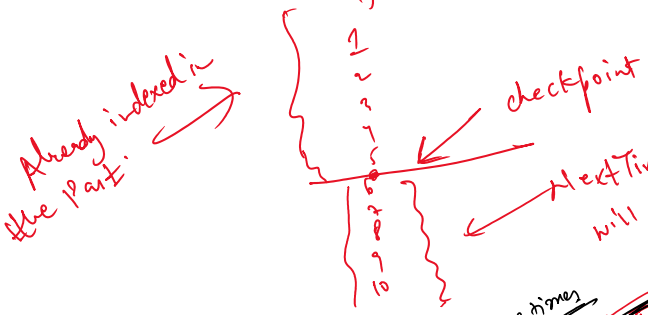


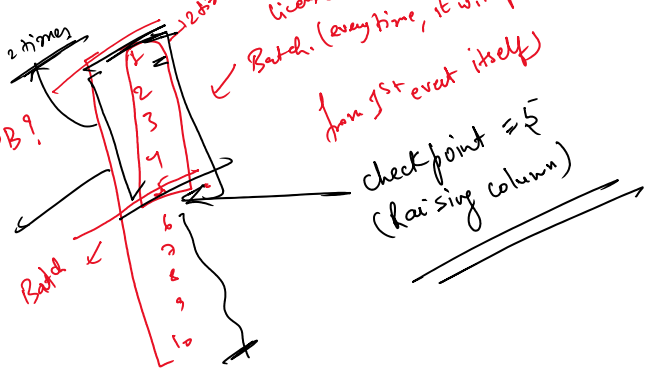
1 event

API → Query & fetch data :- Interval Basis

① Batch Process → Pull all the day from starting everytime. Duplication will occur.

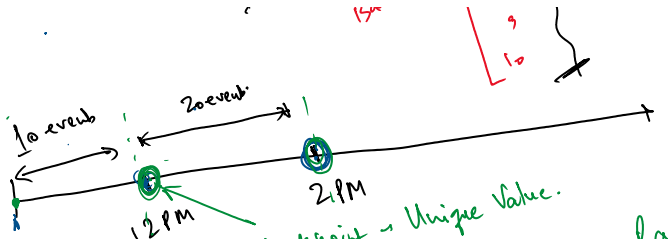
② Rising Column → Checkpoint. Next time when the query will run, it will not fetch the already index data.





① JRE

② Drivers → What type of DB?



Checkpoint → Unique Value.

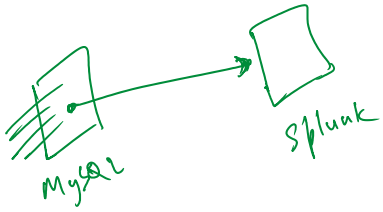
20 event will be pulled → Batch Process.

Raising column.

↓

Avoid Duplicate data.

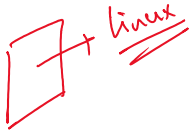
via setting the checkpoint to define the last ingested data.



② Batch. → Data in one go.

② Raising → specific interval / Repetitive

① App-Add-on Linux Server



Splunk

linux

job.resultCount == 0 → done? done?

Machine learning:- Dataset → Train your model

TestDataset → Analysis the o/p

① Supervised →

② Unsupervised →

depend one variable ∝ size of flat

② Unsupervised -

① Liner Regression - depend one variable ∝ size of flat

② Multi Regression - depends on multiple Variable ∝ size, location, (High, low), Corner flat, (Brand New/Resale)