

Data Mining Assignment 1 [Date: 25 Jan21]

Marks-15

Transform crop production dataset into following format

State Name	District Name	Crop_Year1				Crop_Year2				...
			Crop	Area	Production		Crop	Area	Production	
S	A									
		Kharif				Kharif				
		Rabbi				Rabbi				
		Wholyear				Wholyear				

- Slice the transform dataset in two different dataframe
 - Df1= data[:50%]
 - Df2=data[51%:]
- Find the count of NaN values attribute(feature) wise in both dataframes
- Use 'ffill' to replace NaN values in Df1
- Use 'bfill' to replace NaN values in Df2
- Find the mean value of every attribute and replace with NaN in df1
- Find the median value of every attribute and replace with NaN in df2
- Find correlation in df1, df2 and visualize it
- Convert df1 into Min Max normalized format.
- Convert df2 into standardization format.