WEEK 1 – RESEARCH

Find the advantages and disadvantages of the methods to be used for prediction.

|  |  |
| --- | --- |
| Name of method | Attributes Suitable |
| Gaussian Processes | Numeric attributes, Unary attributes, Empty nominal attributes, Binary attributes, Missing values, Nominal attributes |
| Linear Regression | Numeric attributes, Nominal attributes, Binary attributes, Date attributes, Missing values, Empty nominal attributes, Unary attributes |
| Multilayer Perception | Empty nominal attributes, Numeric attributes, Binary attributes, Missing values, Date attributes, Unary attributes, Nominal attributes |
| SMO-reg | Binary attributes, Nominal attributes, Empty nominal attributes, Unary attributes, Missing values, Numeric attributes |
| Kernel Regression | Numeric attributes, Nominal attributes, Binary attributes, Date attributes, Missing values, Empty nominal attributes, Unary attributes |

How many previous values can be used for the prediction.

WEKA does not provide any options for selecting the number of values to be used for prediction within a column.

So an alternative which can be used is to ask the user to enter which dates, he/she wants to be used for prediction. Then once the particular dates are obtained, bring or select only those from the database and feed it to WEKA for prediction.

Is there any provision for using future values of a column for prediction of another column.

Yes WEKA allows this using the concept of overlay data. Overlay data basically help to select a column for which we have a future value and using that value, allows to predict another value which may be required.

But if are using Overlay data and don’t have a future value for some period, then the required prediction for that period which also not be provided.