

How to create Preprocessed data for ML Works -

Machines don't understand free text, image or video data as is, this information has to be converted to 1s and 0s, this is the same with ML Works.

For our libraries to work seamlessly for the trial/community version please ensure that you upload preprocessed data (transformed and/or encoded version of the data that can be easily parsed by the machine). The datatypes in the input data should be numeric or int i.e. all the string fields have been resolved to binary where as numeric or continuous fields can be as is.

A few suggestions from the team:

Make sure the uploaded data frame or dataset has gone through the following checks for optimum result:

1. Data Quality Assessment (Missing value imputation, reduced duplicity, etc.)
2. Feature Aggregation (if needed)
3. Dimensionality Reduction (if needed)
4. Feature Encoding (if needed)
5. Feature Selection