Some thoughts on the report:

Classification: Spam

Remove link🡪 spam emails almost always have links, should it not be good to keep a binary attribute contains links yes/no?

Regression: Automobile data set

For 4 cars there is no price. These should be excluded from the regression model.

Should we exclude nominal losses as there a high number of missing values. Check how many cars have missing values if we exclude the once without a price.

We have a couple of categorical attributes; we will have to change this in multiple binary attributes.

I also checked the scatter plots attribute vs price (automobile data), here we can see that some attributes have a clear linear relation, some are more parabolic shaped, and probably should be transformed into a quadratic attribute, or we should include quadratic and linear.