Nice questions!

You can consider looking into a few state-of-the-art models from literature with similar purposes in order to get inspired by the methods they propose. Also pay attention to the imbalanced nature of your dataset for which literature will help as well.

A good data analysis is suggested before implementing the methods you list in order to understand the patterns in the data. Also your evaluation of the important features as you suggest can be compared against those insights from your data analysis so you see if they are inline.

Is there any connection between the clustering and the regression models you list? would they help one another in any way?

Your regression models you list are not as specific yet. What are you going to estimate exactly? You mention random forest and neural networks, you will need to justify the use of them. Also consider a simple linear regression for example to use as a baseline and compare against as part of your evaluation.

Consider different versions of your methods (different parameter settings, number of layers, neurons, activation functions etc.) in order to understand the impact on the performance as part of your evaluation.

Enjoy your work!