

After reviewing all the 43 rules of Machine Learning, we are going to discuss about few rules that are relevant to our project. As per rule 3, Once we have collected the data and we know what output we want to achieve, it is best to use machine learning as it is easier for updating and maintaining. When rule 8 is taken into consideration, A system must be updated and monitored every day for a better product quality. In our project, we always try to use the latest and updated models to achieve significant results. According to rule 9, before exporting any model to serving, it is very important to detect problems with it and it is important to conduct sanity checks before exporting the model and if we find any issue, it is advisable not to export it. As a team we always try and follow this method and conduct checks. We reviewed rule 19 and it says that it is simpler to use many simple features rather than few complex ones. We tried to use groups of many simple features that apply to parts of data. Rule 22 suggests that we clean up the unnecessary features that we are not using in our infrastructure because having unused features will delay usage of more significant features. As a team, whenever we find ourselves not using a feature or if it is not working in combination with others, we try and remove it. When rule 25 is considered, if you find a feature that improves log loss but decreases the system's performance, you need to look for another feature. Depending what our final goal is, we will choose either performance or predictive power. As per rule 33, when we produce a model on a day, we need to start testing the data on the model which comes from the next day of production. This might not give as good and exact results as the data tested on the same day the model is produced, but the results will not be worse either. The results will still be positive and in the expected range which is good.