**LEAD SCORING ASSIGNMENT SUBJECTIVE QUESTIONS**

1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

Ans. The top three categorical/dummy variables in the final model are ‘Lead Origin\_Lead Add Form’, ‘Lead Profile\_Lateral Student’, ‘Last Activity\_Other Activity’ with respect to the absolute value of their coefficient factors.

1. What are the top 3 categorical/dummy variables in the model which should be focused the most on in order to increase the probability of lead conversion?

Ans. The top three categorical/dummy variables in the final model are ‘Lead Origin\_Lead Add Form’, ‘Lead Profile\_Lateral Student’, ‘Last Activity\_Other Activity’ with respect to the absolute value of their coefficient factors.

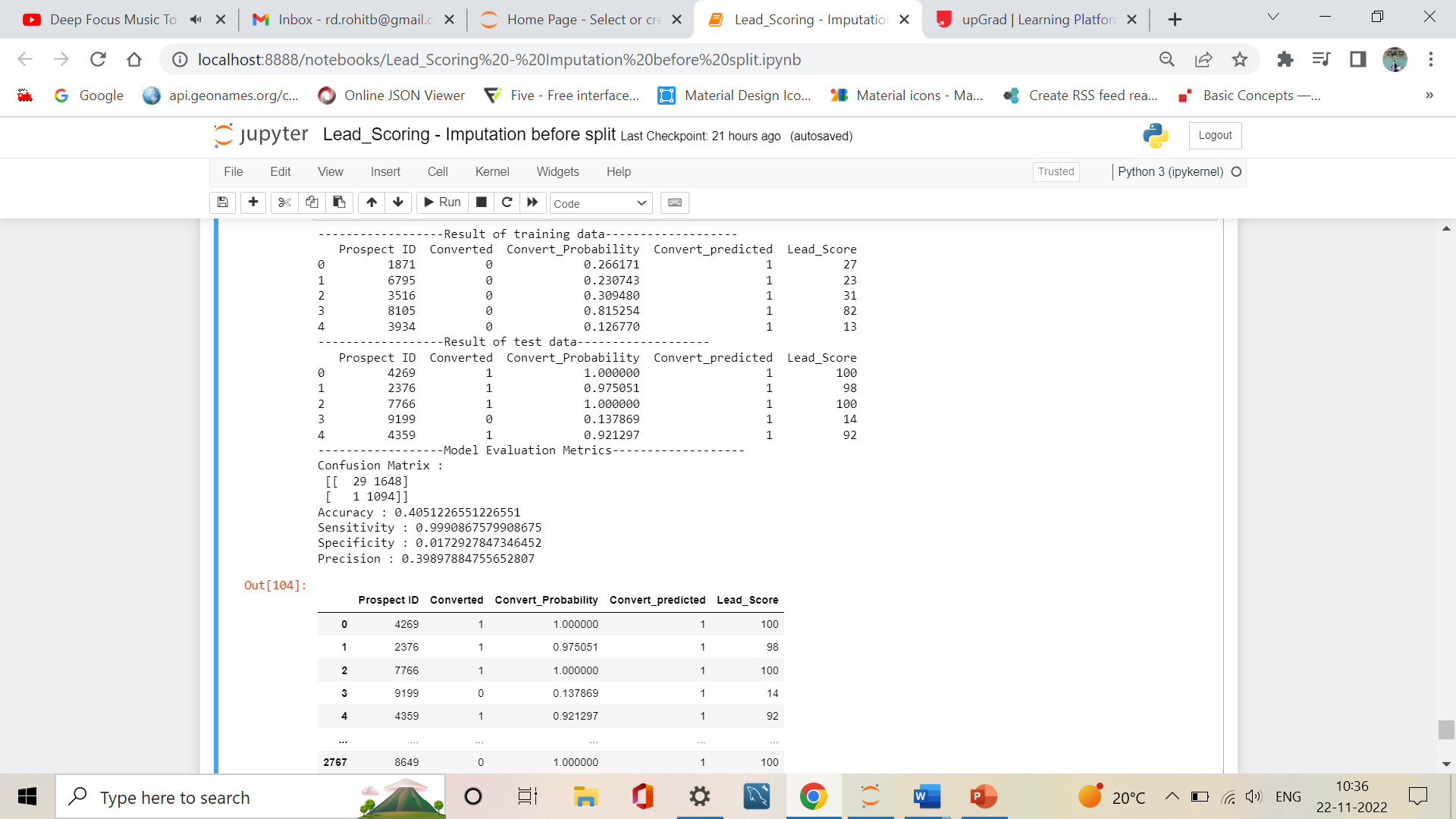
1. X Education has a period of 2 months every year during which they hire some interns. The sales team, in particular, has around 10 interns allotted to them. So during this phase, they wish to make the lead conversion more aggressive. So they want almost all of the potential leads (i.e. the customers who have been predicted as 1 by the model) to be converted and hence, want to make phone calls to as much of such people as possible. Suggest a good strategy they should employ at this stage.

Ans. # Here in this scenerio cutoff of 0.1 should be used as We have built a reusable model which will predict Convert value and Lead Score given training, test data and a cut-off.

#Different cutoffs can be used depending on the use-cases (for eg. when high sensitivity is required, when model have optimum precision score etc.)

build\_model\_cutoff(X\_train[col], y\_train, X\_test[col], y\_test, cutoff=0.1)

A snippet of when it was run on the jupyter notebook:



1. Similarly, at times, the company reaches its target for a quarter before the deadline. During this time, the company wants the sales team to focus on some new work as well. So during this time, the company’s aim is to not make phone calls unless it’s extremely necessary, i.e. they want to minimize the rate of useless phone calls. Suggest a strategy they should employ at this stage.

Ans. # In such scenerios the cutoff should be kept 0.9, # Here in this scenerio cutoff of 0.1 should be used as We have built a reusable model which will predict Convert value and Lead Score given training, test data and a cut-off.

#Different cutoffs can be used depending on the use-cases (for eg. when high sensitivity is required, when model have optimum precision score etc.)

build\_model\_cutoff(X\_train[col], y\_train, X\_test[col], y\_test, cutoff=0.9)

A snippet of when it was run on the jupyter notebook:

