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

   Ans: Based on the coefficient values from the below screen shot,the top three variables are
   i) LeadOrigin
   ii) LastNotableActivity
   iii) CurrentOccupation

| | coef | std err | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| const | -0.8512 | 0.086 | -9.953 | 0.000 | -1.019 | -0.684 |
| Do Not Email | -1.0852 | 0.193 | -5.629 | 0.000 | -1.463 | -0.707 |
| Total Time Spent on Website | 1.1199 | 0.041 | 27.139 | 0.000 | 1.039 | 1.201 |
| LeadOrigin_Landing Page Submission | -0.2882 | 0.091 | -3.176 | 0.001 | -0.466 | -0.110 |
| LeadOrigin_Lead Add Form | 3.4075 | 0.213 | 15.974 | 0.000 | 2.989 | 3.826 |
| LeadSource_Olark Chat | 1.1437 | 0.123 | 9.287 | 0.000 | 0.902 | 1.385 |
| LeadSource_Welingak Website | 2.1860 | 0.746 | 2.932 | 0.003 | 0.725 | 3.647 |
| LastActivity_Converted to Lead | -1.1773 | 0.213 | -5.525 | 0.000 | -1.595 | -0.760 |
| LastActivity_Email Bounced | -1.1929 | 0.375 | -3.179 | 0.001 | -1.928 | -0.458 |
| LastActivity_Not Sure | -1.5439 | 0.453 | -3.406 | 0.001 | -2.432 | -0.655 |
| LastActivity_Olark Chat Conversation | -1.4204 | 0.167 | -8.497 | 0.000 | -1.748 | -1.093 |
| CurrentOccupation_No Information | -1.1827 | 0.090 | -13.212 | 0.000 | -1.358 | -1.007 |
| CurrentOccupation_Working Professional | 2.6548 | 0.203 | 13.051 | 0.000 | 2.256 | 3.053 |
| LastNotableActivity_Had a Phone Conversation | 3.2750 | 1.150 | 2.848 | 0.004 | 1.021 | 5.529 |
| LastNotableActivity_SMS Sent | 1.4162 | 0.082 | 17.294 | 0.000 | 1.256 | 1.577 |
| LastNotableActivity_Unreachable | 1.6289 | 0.552 | 2.952 | 0.003 | 0.547 | 2.710 |

2. 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:Top 3 categorical/dummy variables are below
   i) LeadOrigin_Lead Add Form
   ii) LastNotableActivity_Had a Phone Conversation
   iii) CurrentOccupation_working Professional

| | coef | std err | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| const | -0.8512 | 0.086 | -9.953 | 0.000 | -1.019 | -0.684 |
| Do Not Email | -1.0852 | 0.193 | -5.629 | 0.000 | -1.463 | -0.707 |
| Total Time Spent on Website | 1.1199 | 0.041 | 27.139 | 0.000 | 1.039 | 1.201 |
| LeadOrigin_Landing Page Submission | -0.2882 | 0.091 | -3.176 | 0.001 | -0.466 | -0.110 |
| LeadOrigin_Lead Add Form | 3.4075 | 0.213 | 15.974 | 0.000 | 2.989 | 3.826 |
| LeadSource_Olark Chat | 1.1437 | 0.123 | 9.287 | 0.000 | 0.902 | 1.385 |
| LeadSource_Welingak Website | 2.1860 | 0.746 | 2.932 | 0.003 | 0.725 | 3.647 |
| LastActivity_Converted to Lead | -1.1773 | 0.213 | -5.525 | 0.000 | -1.595 | -0.760 |
| LastActivity_Email Bounced | -1.1929 | 0.375 | -3.179 | 0.001 | -1.928 | -0.458 |
| LastActivity_Not Sure | -1.5439 | 0.453 | -3.406 | 0.001 | -2.432 | -0.655 |
| LastActivity_Olark Chat Conversation | -1.4204 | 0.167 | -8.497 | 0.000 | -1.748 | -1.093 |
| CurrentOccupation_No Information | -1.1827 | 0.090 | -13.212 | 0.000 | -1.358 | -1.007 |
| CurrentOccupation_Working Professional | 2.6548 | 0.203 | 13.051 | 0.000 | 2.256 | 3.053 |
| LastNotableActivity_Had a Phone Conversation | 3.2750 | 1.150 | 2.848 | 0.004 | 1.021 | 5.529 |
| LastNotableActivity_SMS Sent | 1.4162 | 0.082 | 17.294 | 0.000 | 1.256 | 1.577 |
| LastNotableActivity_Unreachable | 1.6289 | 0.552 | 2.952 | 0.003 | 0.547 | 2.710 |

3. 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: The leads whose score is 100,the conversion chance is high for those leads.

| | LeadID | Converted | Conv_Prob | final_predicted | lead_score |
|---|---|---|---|---|---|
| 1678 | 1678 | 1 | 0.996471 | 1 | 100 |
| 238 | 238 | 1 | 0.997326 | 1 | 100 |
| 1088 | 1088 | 1 | 0.996471 | 1 | 100 |
| 1507 | 1507 | 1 | 0.996471 | 1 | 100 |
| 587 | 587 | 1 | 0.996471 | 1 | 100 |
| 1773 | 1773 | 1 | 0.995736 | 1 | 100 |
| 513 | 513 | 1 | 0.996471 | 1 | 100 |
| 90 | 90 | 1 | 0.996471 | 1 | 100 |
| 1323 | 1323 | 1 | 0.997874 | 1 | 100 |
| 1004 | 1004 | 1 | 0.996471 | 1 | 100 |
| 170 | 170 | 1 | 0.996471 | 1 | 100 |
| 1359 | 1359 | 1 | 0.997736 | 1 | 100 |
| 946 | 946 | 1 | 0.995146 | 1 | 100 |
| 2294 | 2294 | 1 | 0.996471 | 1 | 100 |
| 1367 | 1367 | 1 | 0.997279 | 1 | 100 |
| 1424 | 1424 | 1 | 0.996471 | 1 | 100 |
| 1378 | 1378 | 1 | 0.998179 | 1 | 100 |

4. 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: Sales team should call the people based on their probability of conversion; they can focus on the probability around 70% to 100% and avoid calling all the potential leads. We can fine tune our model and change the rules in the model so that it predicts students as Leads if their probability of conversion is more than 70%.