

Capstone Project – Lead Scoring Case Study



Q1) Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

Ans.1) According to my model, the 3 variables contributing the most toward the probability of converting a lead are –

- a) Lead Origin_Lead Add Form (coeff 4.04)
- b) Last Activity_Had a Phone Conversation (coeff 2.89)
- c) Last Notable Activity_Unreachable (coeff 2.46) (this does not make sense in business perspective)
- d) Lead Source_Welingak Website (coeff 2.05)

Q2) What are the top 3 categorical/dummy variables in the model which should be focused the most on to increase the probability of lead conversion?

Ans.2) According to my model, the 3 categorical variables contributing the most toward the probability of converting a lead are –

- a) Lead Origin_Lead Add Form (coeff 4.04)
- b) Last Activity_Had a Phone Conversation (coeff 2.89)
- c) Last Notable Activity_Unreachable (coeff 2.46) (this does not make sense in business perspective)
- d) Lead Source_Welingak Website (coeff 2.05)

Q3) X Education has a period of 2 months every year during which they hire some interns. The sales team 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 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.3) In technical terms, we need to find the best combination of specificity and sensitivity from our logistic regression model, that will help in converting maximum leads. Let us see the table below –

	prob	accuracy	sensi	speci
0.0	0.0	0.481731	1.000000	0.000000
0.1	0.1	0.572517	0.986040	0.188149
0.2	0.2	0.689980	0.949744	0.448529
0.3	0.3	0.765523	0.898557	0.641869
0.4	0.4	0.787267	0.805491	0.770329
0.5	0.5	0.791078	0.739879	0.838668
0.6	0.6	0.773593	0.668683	0.871107
0.7	0.7	0.740417	0.550954	0.916522
0.8	0.8	0.707241	0.442066	0.953720
0.9	0.9	0.659493	0.312238	0.982266

We can see that a sensitivity (lead conversion) of 94.9% is achieved for a lead score of 20, whereas it has a specificity (relevant calls) of 44.85%. Similarly, for a lead score of 30, we have 89.8% sensitivity and 64.18% specificity.

Hence, a middle value, say **lead score of 25 will be ideal strategy** for the sales team to expect maximum conversions, according to our logistic regression model.

Q4) 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.4) In this scenario, we need to maximize the specificity of the model, without compromising the sensitivity much. From the table above in Ans.3, we can see that –

A lead score of 70 gives us a specificity of 91% with a decent conversion rate of 55%, whereas a lead score of 80 gives us a specificity of 95% with a decent conversion rate of 44%. Hence, **a lead score of 75** will be ideal for the team to avoid making unnecessary calls.

Disclaimer: All content and material on the upGrad Campus website is copyrighted material, either belonging to upGrad Campus or its bona fide contributors and is purely for the dissemination of education. You are permitted to access, print, and download extracts from this site purely for your own education only and on the following basis:

- *You can download this document from the website for self-use only.*
- *Any copies of this document, in part or full, saved to disk or to any other storage medium, may only be used for subsequent, self-viewing purposes or to print an individual extract or copy for non-commercial personal use only.*
- *Any further dissemination, distribution, reproduction or copying of the content of the document herein or the uploading thereof on other websites, or use of the content for any other commercial/unauthorized purposes in any way which could infringe the intellectual property rights of upGrad Campus or its contributors, is strictly prohibited.*
- *No graphics, images, or photographs from any accompanying text in this document will be used separately for unauthorized purposes.*
- *No material in this document will be modified, adapted, or altered in any way.*
- *No part of this document or upGrad Campus content may be reproduced or stored on any other website or included in any public or private electronic retrieval system or service without prior written permission from upGrad Campus.*
- *Any rights not expressly granted in these terms are reserved.*