

Predictive Analysis of Lead Conversion Using CRM Data in the EdTech Industry: A Study at Intellipaat

Akshatha N¹, Dr. Kali Charan Sabat²

¹ AKSHATHA N Student, Department of Management Studies (MBA), Dayananda Sagar College of Engineering, Bengaluru, India

² Dr. Kali Charan Sabat Professor (Guide), Department of Management Studies (MBA), Dayananda Sagar College of Engineering, Bengaluru, India

Abstract - Customer Relationship Management (CRM) systems are central to how Education Technology (EdTech) companies manage enquiries, yet the data captured in such systems is not always analyzed for its predictive value. This study examines lead conversion at Intellipaat, a Bengaluru-based EdTech company, by combining CRM-related organizational context observed during an internship with a structured questionnaire administered to 103 individuals who had enquired about Intellipaat's certification programmers, of whom 88 confirmed an actual enquiry. Frequency and percentage analysis, Cronbach's Alpha, Chi-Square tests, Pearson/Spearman correlation and binary logistic regression were used to examine counsellor-interaction quality, lead source, learner satisfaction/trust, and counselling and decision-factor ratings in relation to conversion. Lead source was significantly associated with enquiry outcome ($\chi^2 = 25.287$, $df = 12$, $p = 0.014$), while overall satisfaction was strongly correlated with trust ($r = 0.848$) and willingness to enroll again ($r = 0.822$). Counsellor satisfaction alone was not significant ($\chi^2 = 6.745$, $df = 8$, $p = 0.564$), and the logistic regression model was not significant ($p = 0.672$). The findings indicate that structured, CRM-recorded behavioral data offers a more promising foundation for predictive lead scoring than survey-based perception composites alone.

Key Words: Customer Relationship Management, Lead Conversion, Predictive Analytics, EdTech, Lead Scoring, Logistic Regression.

1. INTRODUCTION

The Indian Education Technology (EdTech) industry has expanded rapidly with rising internet penetration, smartphone affordability and demand for career-oriented upskilling. As course offerings become increasingly similar across providers, effective lead management and counselling have become important differentiators. Customer Relationship Management (CRM) systems allow EdTech companies to record prospective learners' enquiries, interactions, follow-ups and outcomes. Predictive analytics can extend this function by estimating which leads are more likely to convert, enabling counselling resources to be prioritized.

This study focuses on Intellipaat, a Bengaluru-based EdTech company offering certification programmers in data science, cloud computing, artificial intelligence and related domains. During an internship with the Business Development team, it was observed that CRM data is collected extensively but is not systematically analyzed to understand what drives conversion. The study therefore combines organizational CRM context with learner-reported enquiry experience to examine

factors associated with conversion and assess the feasibility of predictive analytics using the available information.

2. LITERATURE REVIEW AND RESEARCH GAP

Prior research shows increasing use of predictive lead scoring in sales and marketing. Wu, Andreev and Benyoucef [1] reviewed 44 studies and found a shift from experience-based qualification toward predictive, machine-learning-based lead scoring. Yim et al. [2] demonstrated that online professional-education enrolment data can support high-accuracy conversion models, while Ulhe and Mohite [3] reported the potential of AI-powered lead scoring in B2C EdTech. Education-analytics reviews by Almalawi et al. [4], Aldisa et al. [5] and Sghir et al. [6] similarly show strong use of predictive models, although much of this work concerns academic outcomes rather than commercial conversion. Alfredo et al. [7] emphasize human-centered use of AI, supporting the view that predictive tools should assist rather than replace counsellor judgement.

CRM theory treats CRM as a strategic, customer-centered capability involving the identification, attraction, retention and development of customers [8–10]. The present study also draws on the Technology Acceptance Model, Theory of Planned Behavior and Relationship Marketing to interpret counselling, decision and relationship factors. The research gap is threefold: limited evidence on predictive lead scoring in Indian certification-focused EdTech; separation of organizational CRM data and learner perceptions in prior research; and absence of a prior study focused specifically on Intellipaat.

3. METHODOLOGY

The study used a descriptive-cum-analytical, cross-sectional design. Primary data were collected through a structured Google Form questionnaire using convenience sampling. A total of 103 valid responses were received, of which 88 respondents confirmed an actual enquiry with Intellipaat. Demographic and course-search questions used all 103 responses, whereas counselling, decision-factor and enquiry-outcome analyses used the 88 confirmed enquirers.

The questionnaire contained 22 closed-ended items across six sections covering demographic profile, learning preferences and lead source, counselling experience, decision factors, satisfaction/trust/future-enrolment intention, and suggestions. Counselling and overall-opinion items used five-point scales; decision factors used a four-point importance scale. Data were cleaned and coded in Microsoft Excel and analyzed in IBM SPSS using frequency and percentage analysis, means and standard deviations, Chi-Square tests, Pearson and Spearman correlations, and binary logistic regression.

Reliability was assessed using Cronbach's Alpha. The counselling-experience scale showed high internal consistency ($\alpha = 0.966$), as did the overall-opinion scale ($\alpha = 0.940$). The decision-factor scale returned $\alpha = -0.040$, indicating that the eight factors should not be treated as a single underlying construct. The study is limited by the modest convenience sample, self-reported responses, and lack of direct access to Intellipaas's raw CRM export data for confidentiality reasons.

Table 1: Hypotheses Tested in the Study

Code	Null hypothesis
H01	No significant association between counsellor-interaction quality and enquiry outcome.
H03	No significant association between lead source and enquiry outcome.
H04	No significant relationship between overall satisfaction and trust / willingness to recommend.
H05	Counselling-experience and decision-factor ratings do not significantly predict lead conversion.

4. RESULTS AND ANALYSIS

The respondent profile indicates that the largest age groups were 36–40 years (22.3%) and 20–25 years (21.4%), while more than two-thirds were aged 26 or above. Male respondents represented 62.1%, female respondents 34.0%, and other respondents 3.9%. Occupational categories were unemployed (29.1%), working professional (28.2%), self-employed (26.2%) and student (16.5%). Artificial Intelligence was the most sought course category (17.5%). Email (20.4%), Google Search (16.5%) and Instagram/Facebook (16.5%) were the leading sources of first awareness.

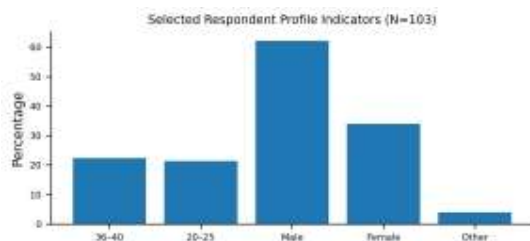


Fig. 1: Selected Respondent Profile Indicators (N=103)

Among the 88 confirmed enquirers, 83.0% were contacted by a counsellor, while 17.0% received no recorded counsellor contact. Of those contacted, 69.3% were satisfied or very satisfied, 17.1% were dissatisfied and 13.6% were neutral. Mean ratings across the five counselling-experience statements ranged from 3.30 to 3.41 on a five-point scale; understanding learning and career goals was highest ($M = 3.41$) and clear explanation of course details was lowest ($M = 3.30$).

Course curriculum had the highest mean decision-factor importance ($M = 3.09$), followed by placement assistance and course fee ($M = 3.00$ each). Reviews and ratings (21.6%),

course fee (20.5%) and placement assistance (19.3%) were the most frequently selected single decisive factors.

Table 2: Mean Importance of Decision Factors (N=88)

Decision factor	Mean	SD
Course curriculum	3.09	0.85
Placement assistance	3.00	1.03
Course fee	3.00	0.96
Trainer's expertise	2.97	0.94
Reviews and ratings	2.97	0.96
EMI / payment options	2.93	0.99
Brand reputation	2.86	1.06
Flexible learning schedule	2.86	0.94

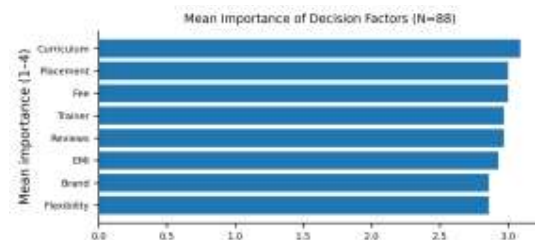


Fig. 2: Mean Importance of Decision Factors (N=88)

For current enquiry status, 35.2% had enrolled, 35.2% were still comparing platforms, 14.8% were awaiting more information and 14.8% had decided not to enroll. Mean overall satisfaction was 3.34, trust in Intellipaas was 3.40, and willingness to enrol again was 3.34.

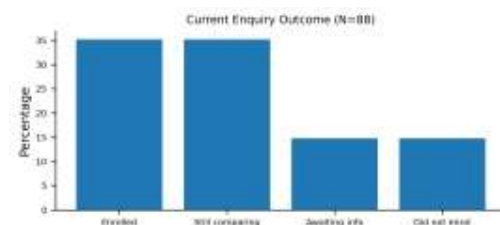


Fig. 3: Current Enquiry Outcome (N=88)

Table 3: Reliability of Multi-Item Scales (N=88)

Scale	Items	Cronbach's Alpha
Counselling experience	5	0.966
Overall opinion – satisfaction, trust, intention	3	0.940
Decision-factor importance	8	-0.040

5. HYPOTHESIS TESTING AND DISCUSSION

The Chi-Square test of counsellor-satisfaction category against enquiry outcome was not significant ($\chi^2 = 6.745$, $df = 8$, $p = 0.564$); H01 was not rejected. In contrast, lead source was

significantly associated with enquiry outcome ($\chi^2 = 25.287$, $df = 12$, $p = 0.014$); H03 was rejected. Email and Google Search leads showed comparatively strong enrolment patterns, whereas LinkedIn-sourced leads showed a higher proportion of non-enrolment in the observed sample. Because some contingency-table cells fell below the recommended minimum expected frequency, the Chi-Square results should be treated as broadly indicative rather than precise probability estimates.

Overall satisfaction was strongly and positively correlated with trust (Pearson $r = 0.848$; Spearman $\rho = 0.787$; $p < 0.001$) and willingness to enroll again (Pearson $r = 0.822$; Spearman $\rho = 0.774$; $p < 0.001$). H04 was therefore rejected. A binary logistic regression combining counselling-experience and decision-factor composites was not significant overall (LR $\chi^2 = 0.795$, $df = 2$, $p = 0.672$; McFadden's pseudo $R^2 = 0.007$). Neither counselling experience (OR = 1.182, $p = 0.411$) nor decision factors (OR = 1.293, $p = 0.702$) was an individually significant predictor; H05 was not rejected.

Table 4: Summary of Hypothesis-Testing Results (N=88)

Hy pothes is	Test	Statistic	p- value	Deci sion
1	H0 Chi-Square	$\chi^2 = 6.745$, $df = 8$	0.564	Not rejected
3	H0 Chi-Square	$\chi^2 = 25.287$, $df = 12$	0.014	Rejected
4	H0 Pearson correlation	$r = 0.848$	<0.001	Rejected
4	H0 Pearson correlation	$r = 0.822$	<0.001	Rejected
5	H0 Binary logistic regression	LR $\chi^2 = 0.795$, $df = 2$	0.672	Not rejected

The findings present a mixed but coherent picture. The significant association between lead source and outcome is consistent with prior lead-scoring literature [1–3]. The absence of a significant relationship between counsellor satisfaction and immediate outcome does not mean counselling is unimportant, because satisfaction is strongly related to trust and future-enrolment intention. The non-significant regression is also consistent with the reliability result: the decision-factor items are conceptually distinct and therefore unsuitable as a single composite predictor. The evidence supports a progression toward predictive analytics based on structured behavioral CRM variables rather than perception composites alone.

6. PRACTICAL AND INDUSTRY IMPLICATIONS

The 17.0% no-contact rate should be addressed through CRM contact alerts that flag an enquiry without a counsellor response within a defined service window and escalate unactioned leads. Lead qualification can use source-level performance, with

higher-converting sources receiving timely follow-up and lower-converting sources receiving closer qualification. Counselling should be personalized to learner needs rather than relying on a single generic script.

A structured follow-up sequence should target the 35.2% of enquirers still comparing platforms, supported by verified learner reviews and outcome stories because reviews and ratings were the most frequently selected decisive factor. Fee discussions should proactively address EMI or instalment options and placement assistance. Intellipaat should also record structured behavioral CRM fields such as response time, follow-up count and counsellor assigned, and monitor contact rate and outcome-by-source continuously to create a stronger basis for future predictive modelling.

7. CONCLUSION

This study examined lead conversion at Intellipaat by combining CRM-related organizational context with a structured questionnaire administered to 103 respondents, of whom 88 confirmed an actual enquiry. Lead source was significantly associated with enquiry outcome, and overall satisfaction was strongly linked to trust and future-enrolment intention. However, counsellor satisfaction alone and the survey-based logistic regression model did not significantly predict conversion. The 17.0% no-contact finding represents an immediately actionable CRM gap. Overall, predictive analytics is a realistic goal for Intellipaat, but a reliable lead-scoring model will require a larger evidence base and more structured CRM-recorded behavioral data.

ACKNOWLEDGEMENT

The author gratefully acknowledges the guidance of Dr. Kali Charan Sabat, Professor, Department of Management Studies, Dayananda Sagar College of Engineering, Bengaluru, and the support provided during the internship at Intellipaat Software Solutions Private Limited.

REFERENCES

- Wu, M., Andreev, P., & Benyoucef, M. (2024). The state of lead scoring models and their impact on sales performance. *Information Technology and Management*, 25, 69–98. <https://doi.org/10.1007/s10799-023-00388-w>
- Yim, W. Y., Khaw, K. W., Lim, S. T., & Chew, X. (2024). Enhancing conversions and lead scoring in online professional education. *International Journal of Management, Finance and Accounting*, 5(1), 15–63. <https://doi.org/10.33093/ijomfa.2024.5.1.2>
- Ulhe, M. S., & Mohite, S. S. (2025). Automated customer segmentation: AI-powered lead scoring for EdTech. *International Journal of Advanced Research*, 13(5), 204–211. <https://doi.org/10.21474/IJAR01/21055>
- Almalawi, A., Soh, B., Li, A., & Samra, H. (2024). Predictive models for educational purposes: A systematic review. *Big Data and Cognitive Computing*, 8(12), Article 187. <https://doi.org/10.3390/bdcc8120187>
- Aldisa, R. T., Rochim, A. F., & Triayudi, A. (2026). Student achievement prediction models: A PRISMA-based systematic literature review. *Journal of Information Systems and Informatics*, 8(2).

6. Sghir, N., Adadi, A., & Lahmer, M. (2023). Recent advances in predictive learning analytics: A decade systematic review (2012–2022). *Education and Information Technologies*, 28, 8299–8333. <https://doi.org/10.1007/s10639-022-11536-0>
7. Alfredo, R., Echeverría, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martínez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence*, 6, Article 100215. <https://doi.org/10.1016/j.caeai.2024.100215>
8. Buttle, F., & Maklan, S. (2019). *Customer relationship management: Concepts and technologies* (4th ed.). Routledge.
9. Ngai, E. W. T. (2005). Customer relationship management research (1992–2002): An academic literature review and classification. *Marketing Intelligence & Planning*, 23(6), 582–605. <https://doi.org/10.1108/02634500510624147>
10. Payne, A., & Frow, P. (2005). A strategic framework for CRM. *Journal of Marketing*, 69(4), 167–176. <https://doi.org/10.1509/jmkg.2005.69.4.167>
11. IBEF. (2025, January 20). India's edtech market likely to reach Rs. 2,50,850 crore (US\$ 29 billion) by 2030. India Brand Equity Foundation. <https://www.ibef.org/news/india-s-edtech-market-likely-to-reach-rs-2-50-850-crore-us-29-billion>