

Lead Scoring

Contexte

Une société d'éducation nommée **Yotta Executive Education (EXED)** vend des cours en ligne aux professionnels du secteur. Chaque jour, de nombreux professionnels intéressés par les cours atterrissent sur leur site web et recherchent des cours.

L'entreprise commercialise ses cours sur plusieurs sites web et moteurs de recherche comme Google. Une fois que ces personnes arrivent sur le site web, elles peuvent parcourir les cours, remplir un formulaire pour le cours ou regarder des vidéos. Lorsque ces personnes remplissent un formulaire en renseignant leur adresse électronique ou leur numéro de téléphone, elles sont classées comme des prospects. En outre, l'entreprise obtient également des *Leads* grâce à des références antérieures. Une fois ces leads acquis, les employés de l'équipe de vente commencent à passer des appels, à rédiger des courriers électroniques, etc. Grâce à ce processus, certains leads sont convertis alors que la plupart ne le sont pas. Pour l'EXED, le taux de conversion typique de leads est d'environ 30 %.

Aujourd'hui, bien que Yotta EXED reçoive beaucoup de leads entrants, son taux de conversion est très faible. Par exemple, s'il acquiert 100 prospects en une journée, seule une trentaine d'entre eux sont convertis. Pour rendre ce processus plus efficace, l'entreprise souhaite identifier les leads à fort potentiel, également appelées *Hot Leads*. Si elle parvient à identifier cet ensemble de leads, le taux de conversion des leads devrait augmenter car l'équipe de vente se concentrera désormais davantage sur la communication avec les *hot leads* plutôt que d'appeler tout le monde.

De nombreux leads sont générées au cours de la phase initiale (en haut), mais seuls quelques-uns d'entre eux apparaissent comme des clients qui finissent par acheter. Au stade intermédiaire, vous devez bien entretenir les prospects potentiels (c'est-à-dire les informer sur le produit, communiquer en permanence, etc.).

Yotta EXED veut sélectionner les prospects les plus prometteurs, c'est-à-dire les prospects qui ont le plus de chances de se transformer en clients qui payent. L'entreprise vous demande de construire un modèle dans lequel vous devez attribuer un score à chacun des leads de telle sorte que les prospects ayant un score plus élevé aient une plus grande chance de conversion et que les prospects ayant un score plus faible aient une chance de conversion plus faible. Le PDG, en particulier, a donné une approximation du taux de conversion cible des prospects à environ 80 %.

Dictionnaire de données

Description de la table : `data.csv`

Variables	Définition
ID_PROSPECT	L'identifiant prospect
ORIGINE_LEAD	L'identifiant d'origine avec lequel le prospect a été identifié comme étant un lead
SOURCE_LEAD	Source du lead
NIVEAU_LEAD	Niveau attribué à chaque prospect en fonction de son profil.
QUALITE_LEAD	Indique la qualité du lead basé sur l'intuition de l'employé et les données qui ont été affectés au lead.
CONTACT_PAR_MAIL	Variable booléenne sélectionnée par le prospect dans laquelle il choisit s'il veut ou non recevoir un email sur le cours
CONTACT_PAR_TELEPHONE	Variable booléenne sélectionnée par le prospect dans laquelle il choisit s'il veut ou non recevoir un appel sur le cours
STATUT_ACTUEL	Indique si le prospect est un étudiant, un travailleur indépendant ou un salarié.
CONVERTI	Variable cible. Indique si le prospect a été converti ou non.
NB_VISITES	Le nombre total de visites effectuées par le prospect sur le site web.
DUREE_SUR_SITEWEB	Le temps total passé par le prospect sur le site web.
NB_PAGES_VUES_PAR_VISITE	Nombre moyen de pages du site Internet consultées lors des visites.
DERNIERE_ACTIVITE	Dernière activité effectuée par le prospect. Inclut l'ouverture des courriels, la conversation Chat, etc.
DERNIERE_ACTIVITE_NOTABLE	
PAYS	Pays du prospect
VILLE	Ville du prospect
SPECIALISATION	Domaine de l'industrie dans lequel le prospect travaillait auparavant. Comprend le niveau "Select", ce qui signifie que le prospect n'a pas sélectionné cette option en remplissant le formulaire.
TAGS	Étiquettes attribuées aux prospects indiquant l'état actuel du prospect.

Variables	Définition
INDEX_ACTIVITE	Un indice attribué à chaque prospect en fonction de son activité.
INDEX_PROFIL	Un indice attribué à chaque prospect en fonction de son profil.
SCORE_ACTIVITE	Un score attribué à chaque prospect en fonction de son activité.
SCORE_PROFIL	Un score attribué à chaque prospect en fonction de son profil.
ANNONCE_VUE	Indique si le prospect a vu l'annonce dans l'un des articles énumérés.
MAGAZINE	Variable booléenne
ARTICLE_JOURNAL	Variable booléenne
FORUM	Variable booléenne
JOURNAUX	Variable booléenne
PUB_DIGITALE	Variable booléenne
RECOMMANDATION	Indique si le prospect est venu par le biais de recommandations.
Comment avez-vous entendu parler de nous ?	La source par laquelle le prospect a entendu parler de nous.
Souhaitez-tu recevoir plus d'infos sur notre cours ?	Variable booléenne
Souhaitez-tu recevoir des mises à jour sur nos programmes ?	Variable booléenne
Souhaitez-tu recevoir des mises à jour par message privé ?	Variable booléenne
Souhaitez-tu payer par chèque ?	Variable booléenne
Souhaitez-tu recevoir une copie de notre livre blanc ?	Variable booléenne