April 2023

TECHNIQUES TO FACILITATE LEAD PRIORITIZATION FOR ONLINE LEAD GENERATION SYSTEMS

Chris Shatara

Aseem Banshidhar Asthana

Rui Brettes

Anbarasi Chinnaiah

TECHNIQUES TO FACILITATE LEAD PRIORITIZATION FOR ONLINE LEAD GENERATION SYSTEMS

AUTHORS:
Chris Shatara
Aseem Banshidhar Asthana
Rui Brettes
Anbarasi Chinnaiah

ABSTRACT

Lead generation is a very important aspect of business development and revenue generation for any company. For a higher number of deals to be closed, a lot of leads need to be generated. With a high volume of leads, it is impossible for a sales team to review each lead. Teams using a best-guess or first-come/first-served basis might not identify the best leads available. Similarly, using prior experience has the potential to cause important leads to be ignored. The number of features that can result in a favorable outcome for a given lead is large, complicated, and not deterministic, which, thus, introduces a problem that cannot be solved basic computational logic. Presented herein are techniques through which a number can be assigned to a lead using a machine learning (ML) algorithm that outputs a number that is directly proportional to probability of the lead being converted to a favorable outcome (e.g., sales, etc.). Such prioritization of leads can help a sales team converge on important leads first and not miss out on an opportunity to convert the leads. Further, by setting the ML algorithm in a continuous learning mode, the ML model will always be up-to date and can learn newer features. Feedback mechanisms can also be utilized (e.g., using a correlation of features with outputs) to further improve lead quality, thereby improving lead convergence.

DETAILED DESCRIPTION

Lead generation is a very important aspect of business development and revenue generation for any company. Most companies have systems in place that can generate leads (either manually or in an automated fashion) that are then passed on to sales teams. A sales team can engage with a potential customer and may eventually close a sales deal, or the deal may not go though and be dropped. Closure of any lead depends on various factors that cannot be described by simple rules. Generally, the closure rate in a given industry can be anywhere between 10% to 30%. Improving the closure rate even by 5% is incentive to improve lead conversion. Additionally, for a higher number of deals to be closed a lot of leads need to be generated. However, this presents a challenge. Most companies have a fixed-size team that handles leads. When a team is overwhelmed with thousands of leads, it is impossible for the team to go over each lead. Currently, leads are handled by sales teams using a best-guess or first-come/first-served basis. However, such approaches have potential of ignoring some very important leads.

This proposal provides for the ability to solve this problem by prioritizing generated leads based on a probabilistic outcome for each lead. In particular, each lead can be assigned a number between 1 and 100 that reflects the probability of each particular lead being closed, with a value of '1' indicating a very poor probability and a value of '100' indicating a very high probability. Thus, techniques of this proposal can help a sales team to focus on important (highly probable closing deals) deals and bucket lower ranked deals for further lookup at a later time. Accordingly, techniques as provided by this proposal can help improve closure rates for sales teams. Additionally, in some instances, feedback can be provided indicating "what works" such that the quality of leads can be further improved.

It is possible for some sales team to generate thousands of leads on a monthly basis. In such cases, it can be very difficult for a sales team to act on each lead without applying some type of prioritization scheme to leads. Further, it is imperative to send out only very high-quality leads to a sales team. Thus, every lead needs to have a score that identifies the likelihood of the lead getting converted with only the very high-quality leads being sent to the sales team and lower score leads being funneled back into the system for nurturing customers. Nurturing outcomes can result in removing leads or resending leads at a later time with higher scores.

Techniques of this proposal involve a three-step process for determining features necessary for lead generation. The first step consists of identifying a feature set that might be useful in determining a lead score. Using the feature set, a score can be generated that is based on an arbitrary estimation. For the second step, feedback data is collected from the leads in order to identify which leads were actually converted and identify key features responsible in lead conversion. Machine learning (ML) algorithms can be used to determine this relationship, as discussed in further detail herein. For the third step, an ML model is built using the features and the trained model can be deployed to automate lead scoring and generation.

Consider various techniques that can be utilized to determine features and static scoring of leads. An engagement score can be calculated that helps determine a lead's online behavior and can indicate how interested the lead (potential customer) is based on activity data, such as form submission(s), responsiveness to email, etc. In some instances, engagement scoring can be reviewed with the sales team on a periodic basis (e.g., weekly) to obtain feedback on scores. Figure 1, below, illustrates example interest indicators that can be used to calculate an engagement score.

Group	Behavior	Frequency (times)	Engagement Score			Max. point for group
			Within 7 days	Within 14 days	Within 30 days	
Hand Raiser	Contact Sales	At least 1 time	75	60	30	75
	3rd Party Sales Appointment		75	60	30	
Events	Event Attend / Engage	At least 1 time	75	60	30	75
	Webinar Attend / Engage		75	60	30	
	Event Reg		15	11	4	
	Webinar Reg		15	11	3	
Email	Email Click	At least 1 time	15	10	5	15
Content	Gated Asset Form Submit	At least 1 time	70	60	30	75
	Content Syndication: Engaged Content		60	30	11	
Online Behavior	Buy Flow Revisit, Funnel Drop Off	At least 1 time	75	60	30	
	Power User /Power Company		50	25	10	

Engagement Score	Interest Level
More than 74	1
51 - 74	2
26 - 50	3
Less than 26	4

Figure 1: Engagement Score Example

Next, a profile score can be calculated that is a measure of how well a lead fits into a target audience of the sales team based on information such as job title of a lead, phone number of a lead, availability of a lead, email address domain, and/or the like. Such (demographic) data can help to determine the prospect of a lead and its fit. Figure 2, below, illustrates example profile information that can be used to calculate a profile score for a potential lead, in which a potential lead is assigned a fit level of A, B, C, or D based on the information/profile score.

Profile	Field Values	Profile Score
Job Title	CEO, President, Management / Executive, C-level executive	40
	Director	20
	Is not blank	10
Phone Number	Is not blank	40
Email Address	Email address other than personal	20
	Personal email	10
	test email	0
Known Account	Commercial or Small	40
	Enterprise or Service Provider	40

Profile Score	Fit Level
More than 75	A
51 - 75	B
26 - 50	C
Less than 26	D

Figure 2: Profile Score Example

Next, the engagement score and profile score can be combined to generate a static lead score, as shown below in Figure 3. The lead score can be represented as a 4x4 matrix [D1 .... A1, ... D4 ... A4], where only the top row A1, B1, C1, D1 scores are to be used as leads.

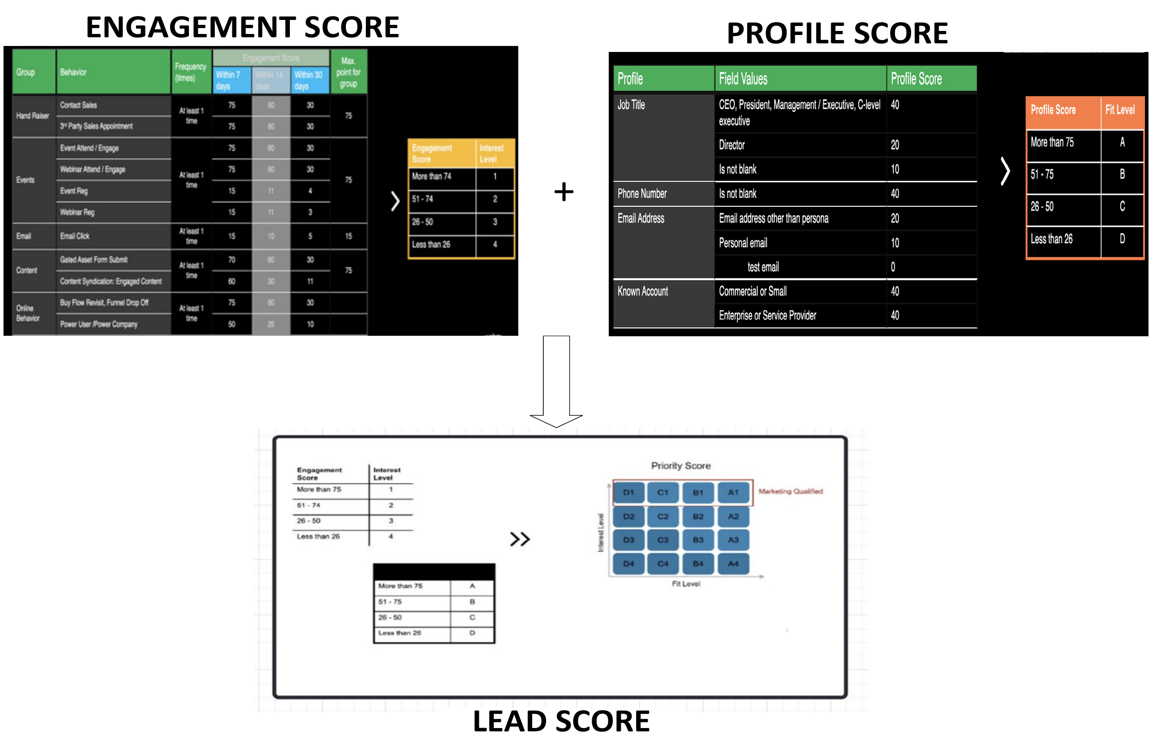

Figure 3: Lead Score Generation

As a part of the first phase process for generating lead scoring, the stating model as described above is used, however, there are some drawbacks of using the static model, such as:

- Arbitrary Score values are chosen for Profile and Engagement;
- Considerations for other features are ignored; and
- Not all features will be contributing to the result.

In order to address such drawbacks, feedback data is collected from the leads to determine which leads were actually converted and to further identify the key features responsible for such lead conversion. Machine learning operations can be utilized for this purpose. Several models can be selected, and the best performing model can be used to determine feature relevance. It should be noted that deep learning models may not be used for feature determination.

Using the lead feedback data, important features for converted leads can be identified. Figure 4, below, illustrates example features that may be identified for converted leads utilizing a Random Forest machine learning model.

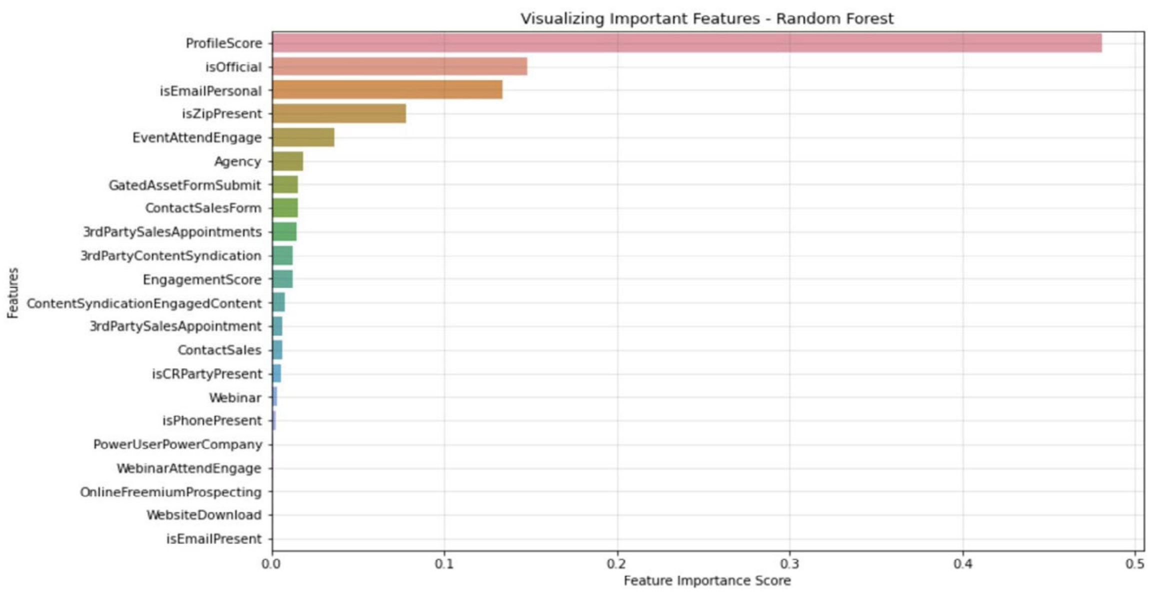

Figure 4: Feature Identification using Machine Learning

Once key features are identified, an ML model can be built that can automate lead scoring. Figure 5, below, illustrates an architecture diagram of an ML model-based lead scoring module that may be utilized to generate lead scores in accordance with the techniques of this proposal.

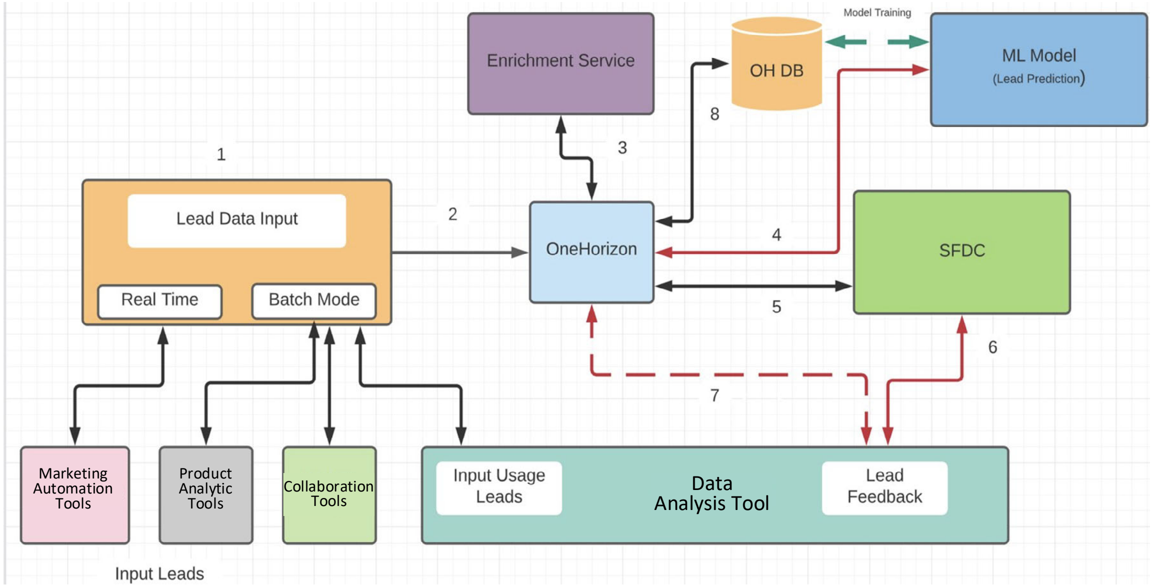

Figure 5: Example Lead Scoring Module Architecture

The ML model-based lead scoring service as proposed herein can be used to extend the static score by adding artificial intelligence (AI) capabilities in order to generate a dynamic score. The ML model can be continuously trained using feedback from the latest data. High score leads generated based on this model that fail to get converted can be utilized to identify new features that might have been missed out in the current model. The model will then be re-trained using the new feature set.

In one example, enrichment data can be an important source for identifying new features that can enhance lead conversion and result in better scoring. Figure 6, below, illustrates example details associated with calculating lead scores using enrichment scores such that engagement scores, profile scores, and enrichment scores can be combined to generate lead scores. Lower lead scores can be funneled back into the system for nurturing customers. Nurturing outcomes can either remove leads or resend leads with higher scores.

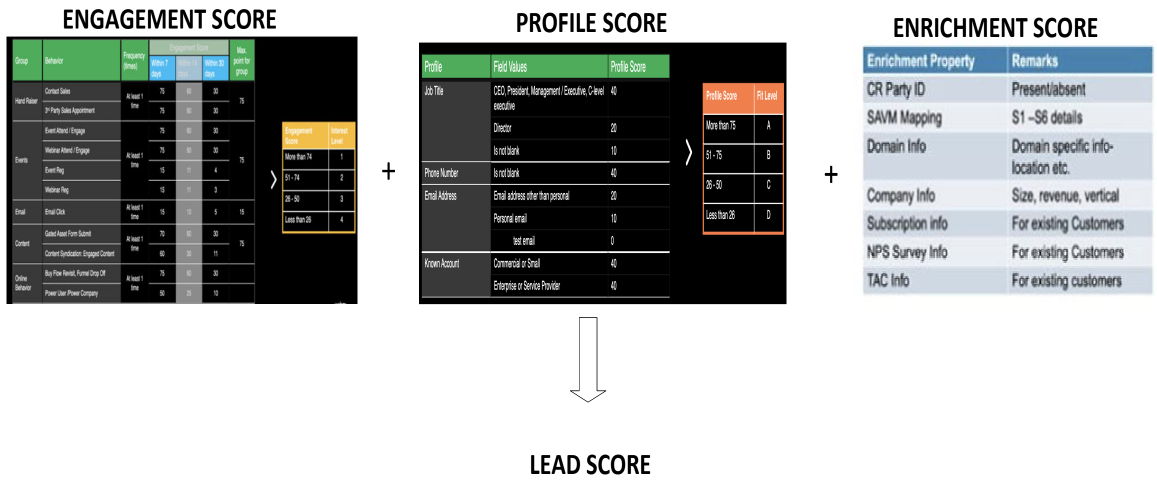

Figure 6: Lead Score Generation Using Enrichment Scores

Accordingly, techniques of this proposal can be used to provide accurate lead data that can result in lowering lead volume while leading to higher lead conversion. In one real-world implementation of the techniques proposed herein, lead volume has been decreased by 70% while doubling lead conversion efficiency, which has resulted in significant revenue generation. Further, by continuously monitoring feedback, additional improvements in lead quality and conversion can be achieved utilizing the techniques proposed herein.