Marketing Engagement Index

Team: Pierre-Louis Bourlon and Noe Mikati
Faculty Advisor: Prof. Rama Ramakrishnan
CTI Advisors: Ashok Mehta, Manish Nigam

Project Context and Overview

Problem Statement

CTI Marketing wants to develop a data-driven tool to quantify Financial Advisors' degree of cross-channel marketing engagement. This could help in providing qualified marketing leads to the Sales team. Utilize analytics to gain insights on engagement, and determine:

- **WHO:** are the most engaged advisors
- **WHY:** engagement improves or decreases
- **WHAT:** marketing actions lift engagement
- **WHEN:** to flag marketing quality lead for Sales

Dataset Description

FA Attributes

Email engagement

Website engagement

Events Engagement

Target

Predict Advisors prone to engage in meaningful conversations with CTI's Sales in near future

Project Objectives

Develop analytical approach to calculate a Marketing Engagement Index (MEI) for every financial advisor:

- Problem framing with Marketing/ Sales leaders
- Extracting useful data from raw data
- Predictive Modeling
- First version of the MEI

- Use MEI to get insights on advisors' engagement journey
- Provide guidance on marketing actions to lift advisors to the next engagement stage
- Provide Sales with frequently updated quality leads

Project Timeline

March-April: Build data retrieval pipeline, Get familiar with CTI's data sources and datalake

May: Identify key interactions and performance indicators, write first version of feature engineering code

June-July: Build and evaluate first version of the model on a subset of advisors, build the scoring pipeline and evaluate score

July-August: Generalize model to all advisors, retrain model with only significant predictors, identify insights from score

Methods

A credit score-like approach

We want to build an **interpretable** and **additive** model that scores marketing interactions. One method to do so would be setting scores that replicate business knowledge on engagement contribution. We instead want to **use machine learning to learn the points from the data.** This context is analog to credit-scoring, where models are required to have those same properties.

[Aggregated interactions features at advisor/observation window level] → **Weight of evidence Transformation** → [Transformed interactions WoEs as features] → **Classifier** → β → **MEI**

Weight of Evidence Transformation

For **additiveness** and **interpretability** characteristics, we transform features using the Weight of Evidence transformation. We bin each variable and transform it based on the following formula for feature *i* and bin *k*:

$$WOE_{i,k} = ln\left(\frac{DistrGood_{i,k}}{DistrBad_{i,k}}\right) * 100$$

Classification and Scoring

Training first model with all marketing features

We built the whole data pipeline for model training, predictive performance evaluation, and score computation from model outputs. We performed exploratory analysis on the score and built verification protocols.

Identify statistically significant drivers

Machine Learning and statistical methods to select top variables with highest predictive value. Simpler models generalize better on unseen data and are more interpretable.

Retrain final model with retained predictors

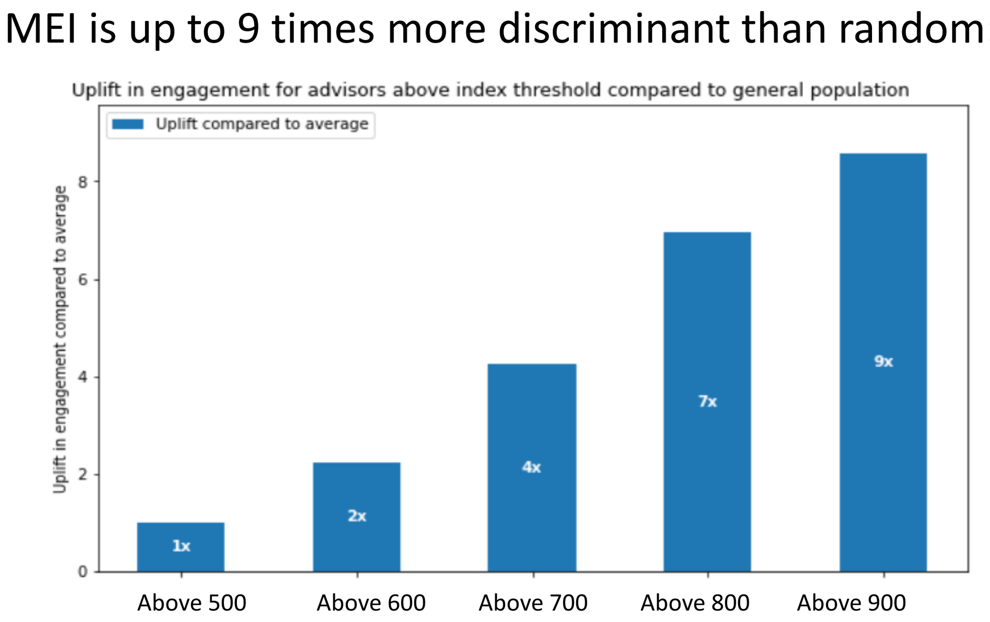
Predict outcome for every advisor on a monthly basis

We train a Weighted Classifier on the transformed features. We predict meaningful interactions that represent high engagement. Finally, from the weights of evidence and the Logistic Regression, we derive the number of points associated with variable i and bin k:

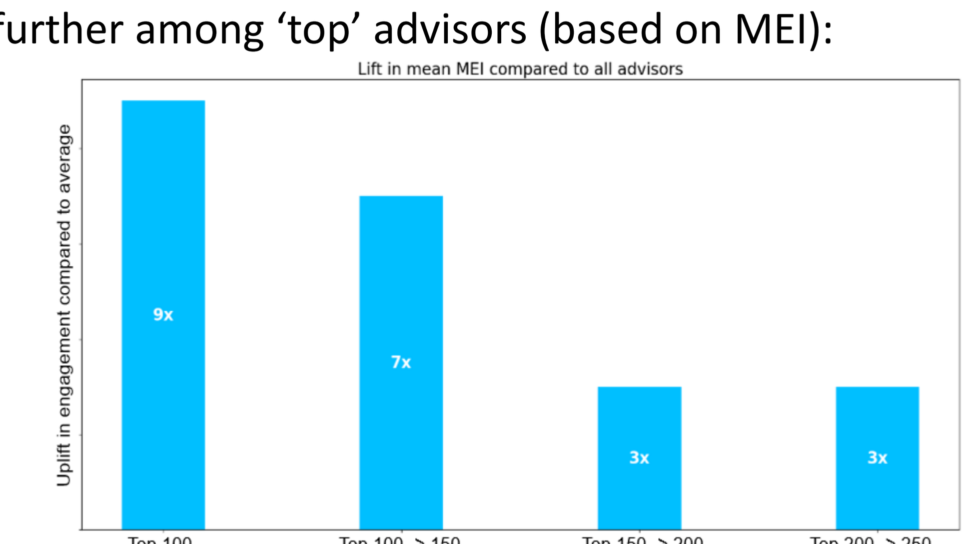
$$Points_{i,k} = (\beta_i * WOE_{i,k} + \frac{\alpha}{n}) * Factor + \frac{Offset}{n}$$

Evaluation of Modeling

Discriminative power of the MEI:
For each threshold in the index, we take out-of-sample advisors whose index is above the threshold. We compute the ratio between the proportion of engaged advisors, and the general population, showing that the MEI is up to 9 times more discriminant than random:



The model differentiates propensities to convert even further among 'top' advisors (based on MEI):

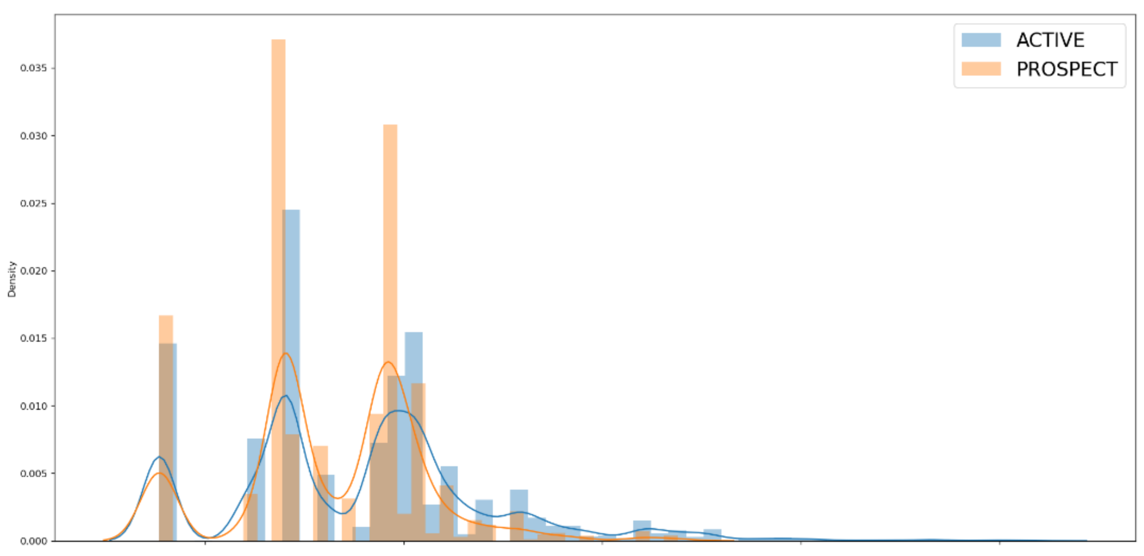


Conclusions and recommendations

Scoring overview, Engagement Insights and Conclusions

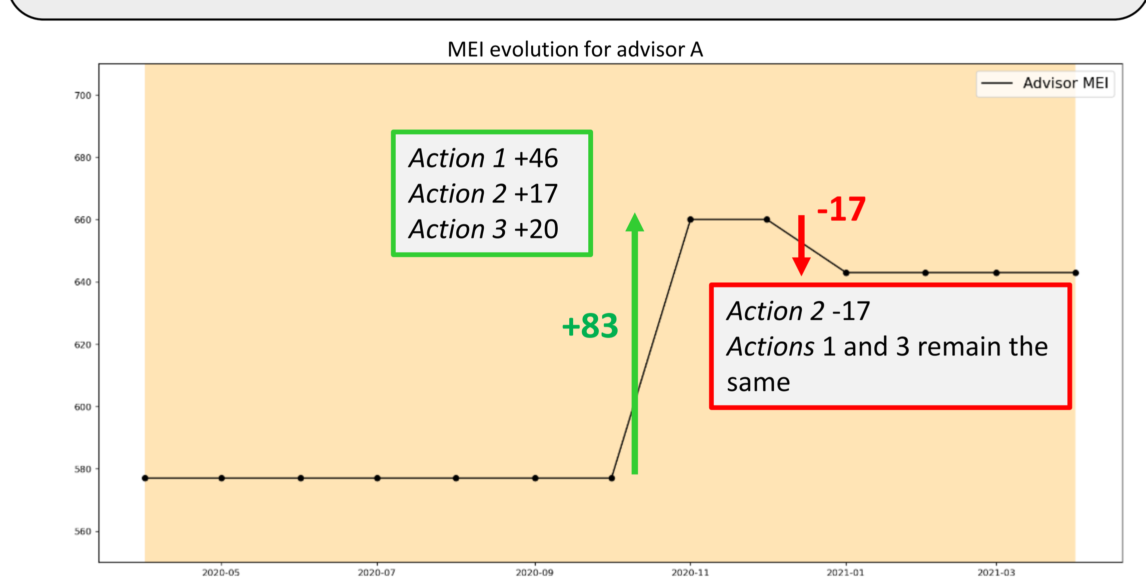
Let's look at a few properties of the scores. Those MEIs were all computed on our holdout subset of advisors, for monthly observations windows from April 2020 to April 2021.

Score distributions: The global distribution is skewed towards 500-600 score range. Let's focus on the top 5% of advisors in terms of MEI:



- Active advisors are more engaged than prospects
- 3 engagement levels seem to appear from data

MEI tracking at advisor level: Ability to evaluate every FA's score at every point in time, insights into engagement patterns and journeys:



- Interpretation of *fine-grain* marketing behaviors across channels : business-actionable
- Insights from absolute and relative score variations

Identify Quality Leads: Ability to flag identify high-propensity-to-convert leads for Sales in a regular, repeatable and understandable fashion.

Next steps

Model is currently being put into production by CTI Data Engineering team

Strategic meeting with Sales and Marketing Leaders to discuss how to integrate model to operations

Our contribution

Built a **robust** and **production-ready** model pipeline, from feature engineering to scoring output

Demonstrated **confidence** in model performance and scoring behavior, through an extensive **verification procedure** and statistical analysis. Provided insights on who are engaged advisors and what are the most effective drivers of engagement

Built clear visuals to **communicate results** and **embark** Sales and Marketing audience on model characteristics and range of applications

Provide **recommendations** on future data **enhancements**, that can get seamlessly folded into the modeling process

Handed out model that **can be "cloned" and adapted** to address other key opportunities in the marketing-sales ecosystem at CTI