

Marketing Colloquia

2020 Spring – Schedule

Dokyun (DK) Lee

Assistant Professor of Business Analytics
Carnegie Mellon University, Tepper School of Business

Tuesday, February 25, 2020

741 JMHH
12:00 PM to 1:20 PM

FOCUSED CONCEPT MINER (FCM): INTERPRETABLE DEEP LEARNING FOR TEXT EXPLORATION

ABSTRACT: We introduce the Focused Concept Miner (FCM), an interpretable deep learning text mining algorithm to (1) automatically extract interpretable high-level concepts from text data, (2) focus the mined concepts to explain user-specified business outcomes, such as conversion (linked to read-reviews) or crowdfunding success (linked to project descriptions), and (3) quantify the correlational relative importance of each concept for business outcomes against one another and to other explanatory variables. Compared to 4 interpretable and 4 prediction-focused baselines that partially achieve FCM's goals, FCM attains higher interpretability, as measured by a variety of metrics (e.g., automated, human-judged), while achieving competitive predictive performance even when compared to prediction-focused blackbox algorithms.

The relative importance of discovered concepts provides managers and researchers with easy ways to gauge potential impact and to augment hypotheses development. We present FCM as a complimentary technique to explore and understand unstructured textual data before applying standard causal inference techniques.

Applications can be found in any setting with text and structured data tied to a business outcome. We evaluate FCM's performance on 3 datasets in e-commerce, crowdfunding, and 20-NewsGroup. Plus, 2 experiments investigate the accuracy-interpretability relationship to provide empirical observations for interpretable machine learning literature along with the impact of focusing variables on extracted concepts. The paper concludes with ideas for future development, potential applications, and managerial implications.