

Comparing Model Development Practices in B2B vs B2C Machine Learning Teams

SALONI POTDAR*, IBM Watson, USA

NAVNEET RAO*, Thumbtack, USA

1 ABSTRACT

Machine learning (ML) model development practices can greatly vary based on factors like the maturity of the ML practice within an organization or the scale at which the company operates in terms of the number of end consumers. Business to Business (B2B) ML teams either build models or build platforms for building models for tens, hundreds, or even thousands of enterprise customers. Customers then deploy the models for the end users of their product. B2B ML teams optimize their model development processes to satisfy most if not all of the needs of their enterprise customers based on feedback from their end users. Business to Consumer (B2C) ML teams build models that get deployed directly for their end users. Their model development processes aim to directly optimize for the satisfaction of their end users.

Due to the nature of the various operating constraints in B2B vs B2C ML teams, the model development practices and the decision frameworks that evolve in these teams can greatly vary. For example, in an ML team evolving a consumer marketplace, a new classification model might demonstrate during offline analysis an improvement in a metric like precision. The model might then get deployed and A/B tested as an experiment. Post experiment analysis might reveal a significant lift in conversion resulting in the new model getting deployed to production. Now compare this to the case of demonstrating an offline improvement in a metric like precision, for an ML team evolving a B2B AI platform. The new model has to be tested on not one but many heterogeneous customer datasets. You might see offline improvements on a subset of data and a degradation in performance across others. Since enterprise customers generally have to choose to make available new changes to their end users, there may not be an easy way to get early feedback on the impact of this new model via an A/B test. Determining the viability of a launch is likely down to the decision framework within the team. From opportunity sizing to model development, from offline analysis to model deployment, the model development practices and the decision making criteria around it can vary in B2B vs B2C ML teams. In this talk, the authors will present comparative insights across model development processes across these two types of ML teams based on their experience refining ML practices across different ML teams.

2 POTENTIAL DISCUSSION POINTS

- Planning & Prioritization: external customer requests vs internal opportunity sizing
- Impact Criteria: model metrics, business metrics
- Model versioning, deployment & testing
- Model scaling challenges
- Feedback Loops: direct vs indirect
- Data privacy challenges

*Both authors contributed equally

Authors' addresses: Saloni Potdar, IBM Watson, USA, potdars@us.ibm.com; Navneet Rao, Thumbtack, USA, navneet@thumbtack.com.

AUTHORS

Saloni Potdar is an Engineering Manager and Senior Technical Staff Member at IBM Watson where she works on Natural Language Processing and Machine Learning. Her team designs and develops algorithms for IBM's conversational AI product - Watson Assistant. She got her Masters degree at the Language Technologies Institute at Carnegie Mellon University in 2014. Her work spans intent classification, few-shot classification, transfer learning, multilingual models, named entity recognition and spell-correction in the product which are supported in multiple languages. The algorithms are designed to be custom-trained for customers globally, deployed at scale with hundreds of thousands of models in production and serves more than 1.9% of the world's population every month. She is a Master Inventor with over 25 filed patents, has over 400 citations with publications at top ML/NLP conferences.

About IBM Watson. IBM Watson creates the AI technology that can power chatbots across the world in multiple domains and languages. Watson's AI algorithms are at the core of creating AI models that can understand and converse with humans. Watson powers the most widely used chatbots in customer care experiences in the world: from scheduling COVID-19 vaccine appointments at CVS to car financing options at GM Financial.

Navneet Rao is a senior engineering manager at Thumbtack, a late-stage Sequoia backed consumer facing marketplace for local services. His team is responsible for matching customers on Thumbtack to professionals & local businesses in their neighborhood using data science & machine learning. They work on challenges in areas like search, ranking & marketplace efficiency. Before Thumbtack, he was an engineering lead for the conversational AI team at Watson Assistant, IBM's largest AI product. He focused on machine learning challenges like language understanding, intent & entity detection, transfer learning & machine bias detection. He also holds several patents in this space. He is an alumni of the Language Technologies Institute (LTI), at the School of Computer Science at Carnegie Mellon University.

About Thumbtack. Thumbtack (www.thumbtack.com) is a local services marketplace where customers find and hire skilled professionals. Our app intelligently matches customers to electricians, landscapers, photographers and more with the right expertise, availability, and pricing. Headquartered in San Francisco, Thumbtack has raised more than 400 million USD from Baillie Gifford, Capital G, Javelin Venture Partners, Sequoia Capital, and Tiger Global Management among others. Available in all 3,143 U.S. counties, more than 4M customers have used Thumbtack in the last 12 months to find and hire professionals and local businesses.