Exploring the Trade-Off: Centralized vs Embedded Models for ML Scientists

Matteo Ruffini
ruffinim@amazon.de
Amazon Music ML
Berlin, Germany

Joos-Hendrik Boese
jooshenb@amazon.de
Amazon Music ML
Berlin, Germany

DESCRIPTION OF THE TALK

Abstract

In the rapidly evolving landscape of machine learning (ML) management, organizations face a critical decision when it comes to structuring their ML teams. This talk aims to compare and analyze the merits and drawbacks of two prominent models: the centralized ML science team model the and embedded ML scientists model. In the centralized model, a Core ML team owns a central ML component, consulting the teams that own customer-facing features on its usage. In the embedded model, scientists are directly assigned to a feature-owning team, and own all the ML-related aspects of that team. We compare advantages and disadvantages of the two approaches, present case studies and discuss potential hybrid solutions that can combine the advantages of both approaches.

Potential Discussion Points

- Definition and Key Characteristics: Understand the fundamental differences between the centralized and embedded models, including team structure, reporting lines, and ownership of ML products.
- Deep Dive: the Centralized Model. Explore the benefits, such as enabling long-term research, fostering a strong team spirit, and developing scalable platform solutions. Discuss the challenges, including slow adaptability to changing business needs.
  - Business Case study: Amazon Music Core ML team. Describe high level team structure and the goal — consisting in delivering a centralized experimentation platform for ML. Describe benefits, success stories and challenges.
- Deep Dive: the Embedded Model. Examine the advantages, such as quick fulfillment of diverse business needs and the opportunity for scientists to develop valuable product and customer obsession skills. Discuss the limitations, such as potential limited long-term research opportunities and difficulties in scaling solutions horizontally.
  - Business Case study: Skyscanner Data Science team. Describe high level team structure as a pure embedded organization, success stories and issues.
- Hybrid Approaches: Explore the possibility of combining elements from both models to strike a balance - including having both a core team with additional embedded ML scientists working directly with partner teams. Discuss risks — including an increased risk of churn for scientists not working in a central team — and potential mechanisms for seamless transition between centralized and embedded roles, mitigating frustration and friction among scientists.

Relevance to Workshop

As machine learning applications become increasingly prevalent in various industries, effective management of ML teams, products, and projects becomes crucial for achieving business impact. Understanding the advantages and disadvantages of different organizational models will equip participants with valuable insights and methodologies to optimize their ML management practices.

By delving into the complexities of ML management structures, this talk aims to provide attendees with a comprehensive understanding of the trade-offs involved in adopting a centralized or embedded model for ML scientists. The insights shared and discussions held will contribute to the development of management methods and best practices in the field, empowering participants to make informed decisions and drive successful ML initiatives within their organizations.

INFORMATION ABOUT THE PRESENTER

A short bio of the main presenter

Matteo Ruffini is an Applied Science Manager at Amazon Music, where he leads a team of Machine Learning scientists working on Music Personalisation. He holds a PhD in Artificial Intelligence from Barcelona Tech University (UPC) and his publication record includes papers presented at NeurIPS, The Web Conference, ML Journal and ECML. Before joining Amazon, Matteo held the position of Data Science Manager at Skyscanner, where he focused on personalization strategies and spearheaded experimentation protocols.

A brief company or project portrait

Amazon Music is a music streaming service offered by Amazon. It provides users with access to a catalog of 100 million songs and thousands of expert-programmed playlists and stations, albums, and podcasts. With a focus on personalization, Amazon Music utilizes machine learning and data-driven algorithms to tailor recommendations and create customized audio experiences for its customers.