Wild Wild Tests

Keynote

Jacopo Tagliabue[1,2]

[1]*Bauplan Labs, New York, USA*
[2]*New York University, New York, USA*

Abstract

Recommender systems are typically evaluated through performance metrics computed over held-out data points. However, real-world behavior is more nuanced: we explore the shortcomings of a simplistic, one-metric-fits-all approach and fight the false testing dichotomy quantitative-and-scalable vs. qualitative-and-manual. We introduce RecList to address the engineering challenge of a rounded evaluation of recommender systems, and discuss the motivation and opportunities behind two recent data challenges, EvalRS 2022 and 2023.

Speaker biography

Educated in several acronyms across the globe (UNISR, SFI, MIT), Jacopo was co-founder and CTO of Tooso, a NLP startup in San Francisco acquired by TSX:CVO. He led Coveo's AI roadmap from scale-up to IPO, and built out Coveo Labs, an applied R&D practice rooted in collaboration (with Stanford, Netflix, Microsoft, NVIDIA) open source and open science. Co-creator of the RecList library for trustworthy evaluation of recommender systems, his work appeared in venues such as NAACL, RecSys, WWW, SIGIR. While building his new company, Bauplan, he is teaching ML Systems at NYU, which is mostly notable because it is the only job he ever had that his parents understand.

CEUR Workshop Proceedings (CEUR-WS.org)