




The Producers Chair

How to Effectively Combine Generative
AI and Predictive AI

April 24, 2024



Luke Shulman

Director of Ecosystem Integration @ 

- 15 Years experience in data analytics and data science at multiple firms
- Contributor to DataRobot MLFlow flavor, Altair, Zebras and other open source data science libraries

Luke.Shulman@datarobot.com

617-826-9820

What we will do today?

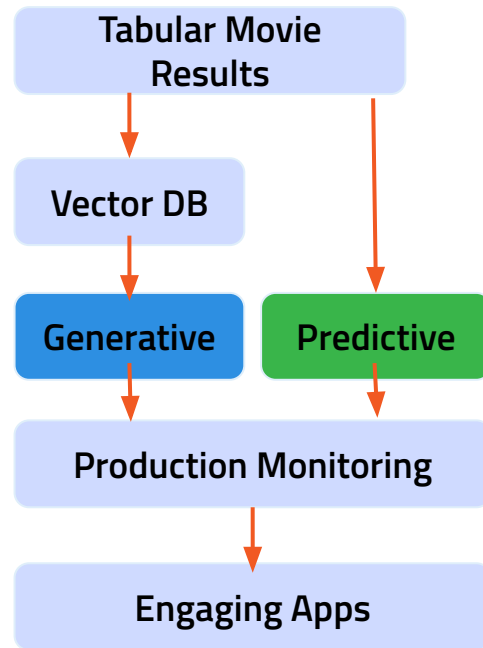
The Data

- Dataset of movies and their past performance from OpenML ([43603](#), [43113](#))
- Transformations for regression and classification of movie performance
- Vector Database of movies: Title, Actors, Director, Genre, and Plot Summary

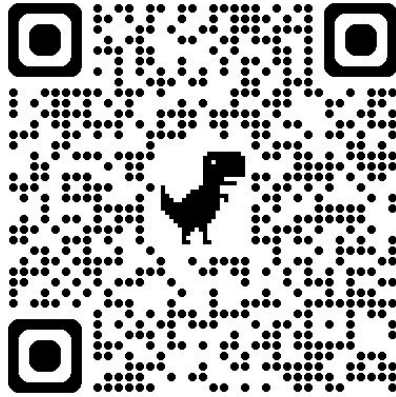
Modeling

- Classify whether the movie will earn above average
- Regression target how much this movie might earn
- RAG LLM to help generate ideas for the films.

Use the DataRobot Free Trial. All examples will run outside DataRobot too.



Materials for Today



<https://github.com/kindofluke/producers-chair-odsc>

- Notebooks
 - Data
 - More Directions
- (Use Python3.10 or better)

Why use Predictive AI and Generative AI?

Generative AI

- + Imaginative
- + Capable of parsing meaning from boundless amounts of text and knowledge
- + Excellent summarization and language marshalling
- Shaky relationship with truth
- Incredibly sensitive to order and nuance
- Prone to simple mistakes

Predictive AI

- + Reproducible
- + Evaluated via metrics
- + Fast
- Uninterpretable numeric values
- Needs discipline in development
- Can often fail to converge for some data sets.