LIVE: Interview Questions

[Productionization , Deployment,
Retraining , Testing , ...

→ Very scenario based

→ focus on your projects

→ engineering - hacks

→ simpler Questions

→ end - end systems

→ design mixed up.

[DISCUSSION ORIENTED]

① GBDT on tabular data

— how would you productionize?

— retrain in real-time

— test cases for automated testing

— 10 TB of train data

— nightly evaluation $\longrightarrow$ GBDT vs logistic Regr.

— Mobile/Web/smartWatch

*** — Improve models over time

— Testing for model-bias

— Incorporate features from your 'team'

— A/B testing infrastructure

— Retrain scheduling

— System-design block diagram.

— AWS/Heroku/ on-site

— Data pipeline : prod servers to ML

— Cost vs Speed trade off

② CNN model

→ all the above

— non GPU system @ run time

③ RecSys for music

— recency & frequency