AD Click Prediction Pipeline
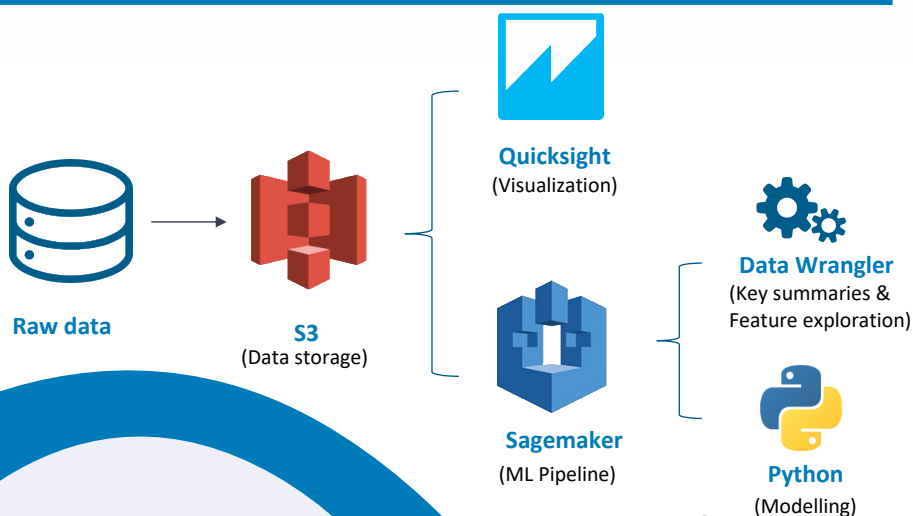
AWS S3, Sagemaker & Quicksight

We aim to provide **a robust data exploration and ML pipeline** that can be readily utilized by users **across the business spectrum** for their predictive needs.

Why care about CTR?

- Click-through rate (CTR) is used for evaluating ad performance and adjudging its outreach.

- Click prediction systems are essential and widely used for sponsored search, consumer econometrics and ROI management.

How do we utilize AWS services to predict CTR?

Raw data

S3
(Data storage)

Quicksight
(Visualization)

Sagemaker
(ML Pipeline)

Data Wrangler
(Key summaries & Feature exploration)

Python
(Modelling)

Source: **Kaggle**

Original source: **Avazu**

Size: **6 GB**

Records: **41 million**

- Training: 10 days of click-through data

- Testing: 1 day of ads data to predict clicks

Features: **21**

Why AWS?

Amazon S3 (Simple Storage Service)

Object storage is built to store and retrieve **any amount of data from anywhere**.

Amazon Quicksight (Visualization & Dashboards)

Cloud-scale business intelligence (BI) service that can be used to deliver easy-to-understand insights through **dynamic and interactive** dashboards.

Total users clicking on the ads out of total targeted users

Comparison of app categories getting maximum clicks

Visualization depicting the time of day fetching maximum clicks

Key Outcomes

- Predicts Ad clicks with **56%** precision
- Overall prediction accuracy of **83%**

Amazon Sagemaker

Helps data scientists and developers to prepare, build, train, and deploy **high-quality machine learning** models quickly by bringing together **a broad set of capabilities purpose-built for machine learning**.

Business Value

Dynamic & User friendly

- Easy to use
- Interactive dashboard for quick insights
- A broad set of capabilities purpose-built for ML

Secure, collaborative & cost-efficient

- Secure data storage
- Allows collaboration
- Conduct the entire project in the AWS ecosystem

Grow business revenue

- Better target the potential customers
- Help ad channels achieve monetization
- Improve ROI management

Aashish Gulabani
gulab001
@umn.edu

Jiarui Hu
hu000476
@umn.edu

Shashwat Khare
khare04
@umn.edu

Akshita Srivastava
sriva232
@umn.edu

Runan Cao
cao00235
@umn.edu

Contact us