

Xingfan Xia

8502 134th Ct NE, Redmond, WA 98052
xingfanxia@gmail.com • +1 (507) 403-1689 • <https://xiax.dev>

EDUCATION

Carleton College, Northfield, Minnesota

- B.A. in Computer Science

Graduated Jun 2018

WORK EXPERIENCE

Amazon Web Service – Athena, Redmond, WA

- Software Engineer Apr 2022 – Present
 - **Building the next-gen storage infrastructure for large scale distributed systems with billions of rows and terabytes of data, improving performance of AIML and other big data analytics applications.**
 - Tech Stack: Java, Python, Spark, Iceberg, AWS Services(Athena, S3, CodePipeline, CodeBuild, Cloudwatch, etc.)
 - Implemented custom table redirection handling mechanism for Iceberg tables in internal Athena fork of Trino query engine and enabled exclusive CTAS(Create Table As Select) feature for Iceberg Tables. Pushed the release of Iceberg Table GA Preview for Athena query engine at RE:INVENT 2022 and millions of CTAS queries being executed since the GA announcement.
 - Overhauled the existing Hive JSON Serde library from an outdated open-source fork. Improved the query execution performance by up to 30% and mitigated type casting compatibility issues.
 - Improved the Athena Engine release pipeline with automation and Slack notifications, which saved 20% of engineering hours dedicated to maintaining and monitoring the deployment.
 - Implemented a virtualized table connector on top of the Iceberg table connector so users can leverage the performance of the Iceberg table engine to access other table formats, such as Governed Table with scalability issues. This improved the query execution speed to these table formats by approximately 40%, avoided congestion, and is projected to save 4 million in hardware costs.
 - Worked on a cross-team endeavor to overhaul the query routing logic from Athena engine to multiple external services, reduced excessive retry behaviors and reduced service congestion by up to 32%.

Apple, Cupertino, CA

- Machine Learning Engineer Aug 2020 – Apr 2022
 - **Improving users' experience with proactive AI assistant on iOS, macOS, watchOS, tvOS.**
 - Tech Stack: Python, Objective-C, Sklearn Kit, Pytorch, Airflow, Hive, Presto, Spark
 - Implemented an on-device user profiling framework which consists of several models to predict users' overall interest and preferences as a upstream service which provides signal to downstream consumers like Apple Map, Apple News, etc.
 - Built listener services to collect feedback and loggings to better track user engagement rate of suggestions and prediction provided.
 - Improved the on-device location model using more diverse data sources with better filtering and data sanitation, increased user engagement rate by 9%.
 - Overhauled the topic prediction model with more up to date data source and sanitized features, increased user engagement rate by 11%.
 - Developed an automatic machine learning pipeline which retrains models with data source updates, launches experiments A/B testing the models, and deploys if a specified improvement is observed, saving engineering hours across the team.
 - Worked on action prediction service which predicts user intention at a particular time and location by aggregating various signals from multiple models input which is used in Siri Suggestions and Siri Shortcuts.

Airbnb, San Francisco, CA

- Software Engineer Jan 2019 – Aug 2020
 - **Fighting financial fraud with machine learning models and rule based engine.**
 - Tech Stack: Python, Java, Scala, Sklearn Kit, Airflow, Hive, Presto, Spark
 - Developed a machine learning model on millions rows of data to tackle down the fake review problem at Airbnb with Sklearn and Pandas, cutting the volume of fake reviews to 1% of the peak volume and prevented 2 million loss in coupon and credits.
 - Trained a machine learning model on millions rows of data to produce generic risk score with Sklearn and Pandas, evaluating the overall risk of users on Airbnb especially on Airbnb hosts, similar to Alipay's Zhima/Seasame score.
 - Collaborated with team to launch the delayed payout policy which delays the payout of risky Airbnb hosts, prevented 60% of fraudulent payouts being made and hence in an estimated 10 mmillion in annual savings.
 - Built a java service which provides downstream service with risk signals easily from a single source of truth without having to dive into hundreds of services and APIs using Redis, Hive, and Apache Kafka.
 - Designed and Implemented a scoring java service which produces risk score of certain event and scenario supporting 6 downstream consumers using Redis, Hive, and Apache Kafka.

Lino Network, Cupertino, CA

■ Software Engineer

Mar 2018 – Oct 2018

- **Building a decentrailized video and live streaming platform.**
- Tech Stack: Python, Golang, Javascript, VueJS, Splunk, ELK, AWS(S3, Kinesis, Lambda, Route53, Cloudfront, CodePipeline, CodeBuild, CodeDeploy, Elastic Beanstalk)
- Performed a major upgrade of DLive, largest live streaming DApp on Blockchain; implemented a twitch-like chatting experience with Javascript, VueJS, GraphQL, and Web Socket. Dramatically increased user retention rate by 150% and average daily session length by 350 %.
- Architected and developed a video platform that includes ingest, transcoding, and distribution of video files in Golang with AWS S3, SQS, and Cloudfront from scratch. Serving terabytes of video content in different qualities globally.
- Built GPU-accelerated transcoding cluster for the video platform in Golang with FFmpeg and Nvidia CUDA. Increased transcoding efficiency by 15 times and expanded the supported video formats and codecs.
- Improved users' experience uploading video by deploying video ingest service in 7 major regions of the world to reduce uploading wait time and the possibility of transfer errors with AWS API Gateway and AWS Lambda.
- Secured content on the DLive platform by implementing signed-cookie protection on AWS Cloudfront in Python to defend third-party malicious attack.
- Designed and implemented data pipeline at Lino from scratch to monitor user behaviors, user QoE metrics, and streaming service robustness with AWS Kinesis, AWS Lambda, ELK Stack, and Splunk. Performed real-time analytics of gigabytes of logs from users and aggregated into comprehensive reports and dashboards that provides both technical and business insides.

TECHNICAL SKILLS

- Programming Languages: Proficiency in Python, Java, Golang, Javascript, Scala, Objective-C
- Experienced AWS Services: Athena, S3, SQS, EC2, Lambda, API Gateway, Auto Scaling Group, Cloudfront, Route 53, CodePipeline, CodeBuild, CodeDeploy, Kinesis
- Machine Learning & Data Engineering: Scikit Learn, OpenCV, Pytorch, Matplotlib, Scipy, Numpy, Pandas, ELK, Apache Iceberg, Apache Spark, Apache Kafka, Airflow