ARUP ARCALGUD

Software Engineer

Phone: +1-847-962-0938 **Email:** arup.arcalgud@gmail.com **Project Portfolio:** https://aruparc.github.io

WORK EXPERIENCE

Software Engineer **May 2022-Present**
Microsoft – Redmond, WA

Developing and scaling Microsoft Copilot for Security, the generative AI assistant for operations in security and IT.

- Iterative feature development, prompt engineering, and platform regression testing pertaining to core Copilot for Security functionality - Microsoft Sentinel, Microsoft Defender Threat Intelligence, and Microsoft Intune (C#, .NET, Microsoft Azure)
- Designing and implementing a prompt evaluation/feature regression testing pipeline for the Copilot for Security platform, allowing for continuous monitoring, triaging, and fixing of numerous feature regressions ahead of EAP and GA releases
- Independent end-to-end ownership and implementation of third-party integration features for Microsoft Copilot for Security, allowing customers to integrate their licenses for external cybersecurity platforms such as IBM X-Force, PulseDive, CheckPhish

Developing an ML-aided multivariate anomaly detection and mitigation platform for cloud infrastructure.

- Testing and evaluating multivariate anomaly detection ML models hosted by Azure Cognitive Services using Python, TensorFlow
- Designing and implementing Azure Logic Apps for incident ticket creation, enrichment, and triaging using Azure Logic Apps to trigger on anomalies detected by the model

Software Engineer **July 2020-May 2022**
Northrop Grumman Mission Systems – Woodland Hills, CA

Developing and integrating machine learning models for multivariate anomaly detection.

- Developing and deploying Tensorflow neural network models for prediction on time-series data as part of an ensemble learning model
- Driving the development and integration of a one-class support vector machine model within an ensemble model pipeline for boosted model performance via binary classification
- Exploring data pre-processing techniques such as PCA, temporal chunking, and time series decomposition for eliminating noise in time series data

Data pipelining using Microsoft Azure tools and Apache NiFi for time-series geolocation data.

- Used Azure ML and Azure Stream Analytics to build models and gain data insights from multiple streams of time-series data
- Used Apache NiFi and Azure EventHubs to feed multi-stream time-series data from Azure Blob Storage to trained AutoML model for live object classification

Designed and implemented a ground-up database-agnostic pipeline for multi-stream data ingestion.

- Used SQLAlchemy with Python for batch ETL data operations on various SQL databases (SQLite, Postgres, Oracle)
- Designed and developed a database agnostic, intuitive web dashboard with Python Flask, JavaScript, CSS, using SQLAlchemy to run dynamic user queries on the database through a RESTful API
- Generated data visualizations based on user-selected features from geolocation data with Python's Bokeh library

Graduate AI Research Student **May 2019-May 2020**
Georgia Institute of Technology – Atlanta, GA

Worked in a project group under Professor Ashok Goel on designing a knowledge-based natural language AI agent with the focus on generating feedback for startup founders regarding their business model.

- Used the Python natural language toolkit (NLTK) library to parse user inputs and referenced the ConceptNet API, allowing our AI agent to form accurate questions and suggestions regarding the user-input phrases.
- Performed literature reviews and competitive market analysis to assess user pains and needs for similar platforms and services
- Planned roadmap, knowledge-base architecture, and algorithmic comparisons for recommender systems

Full-stack Developer **Feb-Aug 2018**
Allianz Life – Minneapolis, MN

Planned for transitions to newer web frameworks with focus on scalability and support with enterprise architecture teams.

- Planned, designed, and delivered a prototype full-stack web application within 2 months, using MongoDB, Node.js, Embedded JavaScript, and Express.js to showcase scalable web technologies
- Used continuous integration via Docker containers and Jenkins to reduce deployment time

- Front-end web development as part of a cross-functional Agile feature team, focused on the Allianz Life customer portal with AB testing and experimentation

Software Development Intern **Jun-Aug 2016**

Zephyr – Newark, CA

Used Agile methodology to develop and test Linux and Microsoft server-side scripts for facilitating customer database backup and restore. Collaborated across departments (Product Development, QA, Client Services) to gather customer requirements and deliver effective solutions

- Debugged and solved 5-10 issues a week while testing RESTful API calls to test robustness and functionality of Zephyr Cloud 4.8 web app
- Provided 2 new features for the Zephyr Cloud 4.8 product release (Aug 2016) – database backup and database restore on Windows, Linux servers using JDBC API
- Delivered customer-centered features to help facilitate ease of use – extended database backup compatibility to numerous database vendors (MySQL, PostgreSQL, Oracle, NoSQL)

EDUCATION

Georgia Institute of Technology – M.S. Computer Science **Aug 2018-May 2020**

- Specialization in Interactive Intelligence, 3.77 GPA
- Core Courses: Machine Learning, Knowledge-based AI, Social Computing

University of Wisconsin – Madison – B.S. Computer Science **Aug 2013-Aug 2017**

PROJECTS

Mock Pinterest: Image Recommender System

Python, Keras, Python Image Library

Exploring and testing pre-trained Keras image classification models for building an image recommender system with cosine similarity.

- Keras image models explored include VGG16, VGG19, NASNetMobile, and InceptionV3, among others.

Toxic Comments Classification

Python, Natural Language Toolkit, Scikit-Learn

Over the span of 4 months, worked with a fully-remote team to explore various data pre-processing methods, unsupervised and supervised learning models for multi-class classification of toxic texts.

- Specifically built and analyzed K-means, multinomial naïve-bayes, linear SVC, and logistic regression with TF-IDF vectorization and truncated singular value decomposition

CAPABILITIES SUMMARY

Data Science:

- Python, TensorFlow, Keras, Pandas, Numpy, Matplotlib, SciKit-Learn, Natural Language Processing, Recommender Systems, Neural Networks, K-Nearest Neighbors, Clustering, Feature Selection, Dimensionality Reduction, Apache NiFi, Azure AutoML, Azure ML Studio

Web and Mobile:

- Python Flask, SQL, HTML5, CSS, Bootstrap, RESTful APIs, Google Firebase, Slack, Atlassian JIRA, GitHub Tools, Docker, NodeJS, Embedded JS, C# .NET

Social:

- Strong and clear verbal and written communication, strong positive cross-functional collaboration between key technology and business-focused teams, efficient task organization skills, coherent presentation skills