

TAMKEEN, MOHAMMAD SAMAN

keensam04@gmail.com | [+91-8971779822](tel:+91-8971779822) | [linkedin.com/in/samantamkeen](https://www.linkedin.com/in/samantamkeen) | samantamkeen.netlify.app

8+ years experience in JVM languages and distributed big data technologies. Hands on with software architecture for developing scalable enterprise software and solving problems in ranking and recommendation space. Pursuing MS in ML & AI from LJMU. Currently building reporting capabilities for small businesses & self-employed.

Education

MS IN MACHINE LEARNING & AI | 2023 (PURSUING) | LIVERPOOL JOHN MOORES UNIVERSITY

- Distance learning program via UpGrad of 6 months duration ending in May, 2023

PG DIPLOMA IN MACHINE LEARNING & AI | 2019 | IIITB

- Distance learning program via UpGrad of 11 months duration ending in September, 2019

M.TECH. | 2016 | BITS. PILANI

- Major: Software Engineering
- Work integrated M.Tech. course of 2 year duration along with working with SAP
- Submitted a dissertation of "Evaluation of timeseries data management in HANA vs distributed big data store"

B.TECH. | 2014 | CHANDIGARH UNIVERSITY

- Major: Electronics and Communication Engineering
- Related coursework:
 - 6 week summer internship at IIT-Bombay where I worked in a project under the National Mission on Education through ICT (MHRD NME-ICT)
 - 6 week training on Embedded systems and wireless technology (microcontrollers & sensors)

Certification

RECOMMENDER SYSTEMS SPECIALIZATION | 2020 | COURSERA

- The specialization is offered by the University of Minnesota by the same professors who run movielens project.

Presentation

Presented "Akash Bazar" – an appstore for Aakash tablets at IIT-Bombay in 2013 (<https://youtu.be/WD2EOsiQDGY>)

Key Skills

- Languages known: Python, PySpark, Java, Scala, Javascript, SQL, RAML
- Components worked with: AWS (Redshift, DynamoDB, S3, etc.), Kafka, Cassandra, Spark, Databricks, SJS, SAP HANA, MySQL, Aerospike, Redis, Cloud Foundry, Openstack swift
- Frameworks used: Pandas, Numpy, Scipy, Matplotlib, Seaborn, Scikit-Learn, Spring, Jersey JAX-RS, Swagger, Node.js, ANTLR, SAP UI5, Cordova
- Micro service based application development and operation on cloud
- Application development using distributed system & CQRS

Work Experience

SENIOR SOFTWARE ENGINEER | INTUIT | MAY 2021 ONWARDS

- Worked on realtime ingestion pipeline to bring in data in the reporting data mart
- Built from scratch a parity tool to detect data issues between any 2 given data sources
- Phased region-wise deployment of realtime reporting capabilities to all customers
- Currently working on context-based report recommendation

SENIOR DATA SCIENTIST | GO-MMT | FEBRUARY 2020 - APRIL 2021

- Developed L2R system using XGBoost algorithm to enhance user's experience while searching for a hotel
- Worked on a Multi Arm Bandit based approach to auto-select the best performing model parameters dynamically

SENIOR SOFTWARE ENGINEER | GO-MMT | JUNE 2018 - JANUARY 2020

- Worked on ranking algorithms (Collaborative filtering based model / Word2Vec based model) to improve the conversion rate of hotels
- Implemented A/B framework to test and compare the various ranking algorithms
- Analyzed users segments based on the device used to be leveraged in multiple systems
- Business anomaly detection tool

DEVELOPER | SAP LABS INDIA PVT. LTD. | APRIL 2016 - MAY 2018

- Worked on the core services of the IoT Application Enablement Platform – a PaaS offering
- Built the big-data enabled timeseries store using different distributed stores
- Constructed ingestion pipeline for timeseries data streaming into the system from the devices
- Migrated live timeseries data for a customer from non-bigdata store to big-data store

SCHOLAR | SAP LABS INDIA PVT. LTD. | AUGUST 2014 - MARCH 2016

- Pursued M.Tech. along with working for different teams at SAP (Work Integrated Learning Program)
- Developed UI for managing work packages in integration flows for HANA Cloud Integration
- Developed UI and back-end services for an internal job-portal

Projects

REPORTING AS A CAPABILITY (MAY 2021 ONWARDS)

Part of a team which is responsible for building reporting as a foundational capability to be used by different applications

LEARNING TO RANK (FEBRUARY 2020 - APRIL 2021)

Developed a machine learning model to rank hotels for new users based on search context

- Using spark (python and scala api) to build the end-to-end training pipeline
- Using XGBoost model to train the L2R model

BUSINESS ANOMALY DETECTION TOOL (JUNE 2019 - JANUARY 2020)

A tool to detect business anomalies by monitoring and predicting transaction metrics

- Deployed timeseries forecasting algos (Holt-Winters algorithm, STL decomposition, fbprophet) to predict key metrics
- Built aggregation capabilities to detect anomalies across different dimensions

RESTAURANT SEARCH CHATBOT (APRIL 2019)

Build a restaurant search chatbot deployed on slack on top of Zomato api's. The bot can search for restaurant taking location, cuisine and budget as input from the user

- The chatbot was implemented as an assignment for the PGD ML & AI course
- The chatbot was built using RASA framework

REINFORCEMENT LEARNING BASED SYSTEM FOR ASSISTING CAB DRIVERS (AUGUST 2019)

Build an RL-based algorithm to help cab drivers maximise their profits by improving decision-making process on the field

- The system was implemented as an assignment for the PGD ML & AI course
- Implemented the Deep Q-learning (DQN) algorithm to implement the Markov Decision Process

EXPERIMENTATION PLATFORM TO RUN A/B TESTS (JUNE 2018 - DECEMBER 2018)

A platform based on facebook planout to define, run and track server-side experiments

- Implemented drools powered rule based engine to run experiments on targeted subset of users
- Implemented JAVA sdk for planout script management

DATA INGESTION PIPELINE MONITORING APPLICATION (JANUARY 2018 - MAY 2018)

Build an application to monitor the various components of data ingestion pipeline

- The application could be used to debug issues in the pipeline (kafka/sjs/spark cluster)
- The team won product innovation award for the project

DATA INGESTION PIPELINE FOR TIMESERIES DATA (JANUARY 2017 - MAY 2018)

An ingestion pipeline written in spark to read device data from kafka topics and persist in big-data store

- Created stateful streams to read from Kafka and write to big-data store
- Implemented notification mechanism for dynamic update of spark states (cache)
- Enabled checkpointing and spark relevant configuration for fault tolerance and high availability of the streaming job
- Wrote custom monitors to monitor the status of the ingested message on Kafka
- Implemented metering of ingested data to enable pricing

THING MODEL SERVICES FOR IOT APPLICATION ENABLEMENT (APRIL 2015 - DECEMBER 2016)

A set of REST APIs to enable modelling of real-world things into the system (as digital twin)

- Defined REST APIs using RAML and wrote implementation of the API using Jersey JAX-RS framework.
- Helped in development of DAO layer to talk to persistence using JDBC template and DBCP2 connection pools
- Wrote data definition files to create persistent artifacts (tables) and created views to handle complex transactions

BIG-DATA MIGRATION (MARCH 2017 - OCTOBER 2017)

A worker application written in Java to migrate data from non big-data store to big-data store

- Developed and deployed the application as worker node and enabled auto-scaling based on number of instances
- Planned the 2-phased strategy to execute migration on customer's LIVE landscape to minimize downtime
- Ran the migration end-to-end on customer's landscape to move data for 300+ tenants

SCHOLAR PORTAL (MARCH 2015 - JULY 2015)

An internal job portal where managers would post job openings and scholars could apply for the same

- Created the UI view for scholars using SAP UI5
- Implemented REST endpoints to capture changes to user profile in XSJS (application layer of HANA)
- Designed and created tables in HANA to save the updated profiles