

# JAGAN SHANMUGAM

Röntgenstraße 9, 10587 Berlin.

+49 15166256754

✉ [jaganshanmugam@outlook.com](mailto:jaganshanmugam@outlook.com)

in [linkedin.com/in/jaganshanmugam](https://www.linkedin.com/in/jaganshanmugam)

github.com/js

## Experience

---

### MBiton GmbH

Jun 2020 – Present

*Data Scientist (Senior → Expert)*

*Berlin, Germany*

#### Internal Knowledge Assistant | *Langchain, LLaMaIndex, PGVector, Azure OpenAI, RAG*

- Devised the business case for the Knowledge Assistant, enabling easier and more efficient access to useful information instead of wasted effort on searching & processing information from silo-ed and sometimes outdated sources.
- Designed the architecture for the Interface and Backend considering data security and governance.
- Implemented Data Ingestion pipelines to vectorize multimodal data from Confluence, JIRA, Gitlab, and internal tools.
- Deployed a FastAPI backend to serve requests from multiple frontends including Rocket.Chat and a web app.
- Extended RAG to support Agentic workflow and also implemented continuous adversarial testing, request/response guardrails and evaluated the solution on the usefulness of answers and relevancy of references.

#### ExploreMe - HeyMercedes Voice Assistant | *BERT, MLFlow, Databricks, PromptFlow, Azure CognitiveSearch*

- Devised and implemented a BERT-based binary classifier to identify Car Info Requests from all utterances.
- Trained & Evaluated the classifier based on the dataset curated from real driver utterances for six languages.
- Integrated the classifier in the Mercedes Voice Assistant Platform via Kubernetes to serve requests from drivers on-road.
- Having solved the needle-in-the-haystack problem with domain-specific binary classifier, assisted in the development of an ensemble-based domain classifier and User-Manual Augmented Generation module.
- With data-driven feature development and prioritization, identified 6 features that led to an increase in feature coverage by 51%, Analyzed PROD utterances and benchmarked NLU providers, detected feature scope gaps, and derived improvements for existing voice assistant domains.

#### Telelog - Driver & Vehicle Data Analytics | *Tableau, Databricks, PySpark*

- Analysed millions of trips from the Mercedes fleet for a variety of features from the infotainment system.
- Closely collaborated with the business teams to drive data-driven product & feature decisions.

### Scientific Games, India Pvt Limited

Oct 2014 – Aug 2017

*Software Engineer - Business Intelligence*

*Bengaluru, India*

- Integration of Business Intelligence (BI) suite with a customer data hub - Developed SSIS-SQL packages for integrating customer data hub (Ontario Lottery Gaming Corporation, Toronto) with Business Intelligence suite.
- Designed and developed enhancements for player analytics in SQL server and supported User Acceptance Testing on-site.

## Education

---

### University of Stuttgart

Grade: 1.5, Oct 2017 – Jan 2020

*Master of Science in Information Technology - Focus on Machine/Deep Learning*

*Stuttgart, Baden-Württemberg*

#### Master Thesis & Research Internship | *Bosch Center for Artificial Intelligence, Renningen* | Apr 2019 – Dec 2019

- **Latent space Bayesian Optimization with Transfer Learning** - Developed a method for warm starting Bayesian Optimization (BO) of the target task with previous optimization data from related tasks for finding optimal parameter settings in industrial processes.

#### Student Research Assistant | *Machine Learning and Robotics Lab, Stuttgart* | Dec 2018 – Mar 2019

- Collected motion capture data using ROS (Robot Operating System) and developed test environment for Hierarchical Inverse Reinforcement Learning algorithms.

#### Work Student | *Daimler AG, Sindelfingen* | Mar 2018 – Jan 2019

- Validated ADAS attributes in NDS (Navigation Data Standard) databases to detect discrepancies.

### College of Engineering, Guindy

Grade: 85%, Aug 2010 – May 2014

*Bachelor of Engineering in Electronics and Communication*

*Chennai, India*

## Technical Skills

---

**Languages/Data Science Tools:** Python, SQL, Scala, PyTorch, Pandas, Transformers (HuggingFace), FastAPI, Streamlit  
**Cloud/Technologies/Databases:** Databricks, Azure, K8s, Docker, Postgres, MS SQL Server

## Leadership / Awards

---

- Mentored interns & junior data scientists and supervised bachelor and master thesis students in MBiton.
- Acted as an interim Product Owner (PO) and ran bi-weekly sprint planning and review meetings.
- Recipient of Deutschland Stipendium - winter semester 2019/20, for excellent academic performance in Masters.
- Project Head of Robotics club of College of Engineering, Guindy (CEG) Chennai, India.