

VARGHESE PAUL

Data Science Practitioner

Portfolio: <http://www.vpaul.in>



Contact : +91 9677258592

NLP experience : Sentiment Analysis, **Semantic** Similarity, **Embedding** Representations - **Word2Vec**, Sentence-**BERT**, Fast Text, Spam Classification, Log Analytics, **Bixby** User Utterance Analytics.

ML experience : Linear **Regression**, SVM, **Decision Trees**, Random Forest, **XGBoost**, Naive Bayes, **Neural Networks**.

Cloud knowledge : Data Science knowledge on Azure (**Microsoft Certified: Azure Data Scientist Associate - Certification**) and basic practitioner knowledge on AWS.

Languages/Libraries/Skills : Python, SQL, MongoDB, SpaCy, TensorFlow, Keras, PyTorch, Fast Text, GenSim nltk, scikit-learn, NumPy, Pandas, Jupyter Notebook, VS Code, Exploratory Data Analysis, Data Visualization and Tool Development for POC.

Recognitions : Received **Samsung Excellence Award** and **Spot Award** from **Samsung Research India Bangalore** for contribution in Bixby and On-Device Bixby Productization.

Education:

- **MBA in Business Analytics**
Bits Pilani || 2021 to 2023
- **B.E Computer Science Engineering**
SSN College of Engineering / Anna University || 2013 to 2016
- **Diploma in Computer Engineering**
Panimalar Polytechnic College
2010 to 2013

Summary

Data Science Practitioner with **6 years** experience in Software engineering and **NLP modeling** for **Samsung Bixby** - NLU and NLG, On Device AI solutions, **Cloud based Chatbot Solution**. Ability to **Build and Train Models** to production

Roles and Responsibilities

Lead Engineer

Samsung Research Bangalore || Mar 2021 to Present

- Developed NLP Models for Bixby Chat for Entity Recognition and Capsule (Algorithmic Handler) Classification.
- Developed NLG (Response) Model which will generate responses based on Semantic Similarity with Root Sentences using S-BERT.
- Performed Descriptive Analytics and Visualization on Market Utterances.
- Performed ETL on High Confidence Utterances for Bixby Training Data.
- Release Management and Deployment of Production Data in AWS.
- Bixby Log Analytics using Entity Recognition using LSTM and Fast Text.

Senior Software Engineer

Samsung Research Bangalore || Jul 2019 to Mar 2021

- Developed Entity Recognizer using fasttext library, for dynamic validation of on-device models.
- Performed model optimization and compression for edge deployment of on-device AI.
- Performed model CI-CD using circle CI in staging environment.
- Developed tool for model CI-CD monitoring in production environment using NodeJS.
- Model drift analysis using NLG diff algorithms based on fasttext embeddin.
- Data drift management on utterances (chatbot sentences) database.

Senior Software Engineer

Tata Elxsi Bangalore || Aug 2016 to Jul 2019

- Developed a module for Calculating the thumbnail loading time and video playback time using Python and Image Processing libraries.
- Designing MySQL schema and writing MySQL queries and integrating data with Grafana Dashboard.
- Developed a Web based Tool to perform Acceptance testing of all the components before integrating components in a headend/server environment.

Certifications: Coursera - DeepLearning.AI,
Udemy - NLP Courses, Microsoft Certified: Azure Data Scientist Associate.