

# ANIMESH SENGUPTA

DATA SCIENTIST | Amherst, MA

 [Animeshsg](#) |  [asg-animesh](#) | [asg.animesh@gmail.com](mailto:asg.animesh@gmail.com) | +1 908-219-2250

## PROFESSIONAL SUMMARY

A polyglot engineer working in the intersection of machine learning and cloud systems. A recent graduate in data analytics with 3 years of experience in building and automating cloud applications. A curious data science researcher exploring possibilities to leverage data for society. Also highly passionate about exploring the power of neural networks and machine learning to identify hidden patterns.

## WORK EXPERIENCE

### Machine Learning Intern: Operations

06/2023 - 08/2023

*ReSurfX Inc - A BioTech outcomes intelligence company* — Cambridge, MA

- Implemented large file upload (100s GB) workflow from client to S3 and EC2 using Kafka and SQS with AWS boto3.
- Researched Model interpretability of CNN layers using captum to understand Layer contributions
- VitalDB Wave signals processing in AWS Spark EMR and built transformer model to predict patient mortality rate.

### Sr Software Developer: Cloud Migration Automation

12/2021 - 07/2022

*Societe Generale GSC* — Bangalore, India

- Spearheaded and Co-built a utility software - Infra Autoprovisioner API that provides Terraform and Kubernetes blueprints as a PR. 95% Uptime achieved.
- Packaged and parameterized re-usable infra components. Reduced Infra provisioning TAT by 70%.
- Built and automated the DevOps pipeline for utility streams and enabled 15+ cloud migrations.

### Software Developer: Cloud Engineer

12/2021 - 07/2023

*Societe Generale GSC* — Bangalore, India

- Managed and Spearheaded the setup and deployment of private cloud OpenStack over multiple modes and regions.
- Created a 10+ self-service XaaS Flask API using Openstack native APIs.
- Created an Angular Webapp and pySpark Data pipeline to process and display incident data for reporting. Achieved 90% Uptime.

## PROJECTS

**IaC Autoprovisioner API for Azure and Openstack cloud:** *FastAPI, Terraform, Kubernetes, Helm Charts, Ansible*

- Built 15+ self-service APIs to provide cloud templates which enable cloud migration
- Created 8 Architectures worth of parameterized blueprints of terraform, ansible and helm charts for deployment in K8s of Azure and Openstack clouds.

**Populist rhetoric flow in Media:** *BERT, Huggingsface, PyTorch, LLM*

- Used transfer learning to train BERT on print media data to perform topic modeling and semantic emotion analysis
- Network Analyser of topic/rhetoric flow across media organizations

**Multi Output+Class Transfer learning on Hate speech:** *PyTorch, RNN, Transformer, tensorboard*

- Fine-tuning DistilBert towards multi-output and multi-class classification on hate speech dataset
- Achieved 70+ accuracy in 6/8 output classes with 85% accuracy in the hate speech class. Identified annotator bias towards classes.

**DeepSplines Based Neural Network Modelling:** *PyTorch, WandB, Splines, R*

- Comparative analysis of de facto Activation functions against higher order splines for a time-series dataset. Observed comparative behavior of splines with de-facto Activation functions..
- Improved efficiency of neural nets with deep splines by 20% in compared to sigmoidal activation functions.

## TECHNICAL PROFICIENCY

**Languages:** Python, R, Java | **ML Frameworks:** PyTorch, Keras, Tensorflow, MLFlow, Captum, WandB | **Data Pipeline:** pySpark, Serverless EMR, Hadoop | **Database:** SQL, MongoDB, PgSQL | **API:** FastAPI, Flask | **Cloud:** AWS OpenStack | **IaC:** Kubernetes, Terraform, Helm Charts, Ansible, Jenkins | **Visualization:** Tableau, Grafana | **Misc:** Airflow, Boto3 | **OS:** Linux, Unix

## SKILLSETS

Agile Scrum Development | Experience in RNN, Transformers | Hyperparameter tuning | Containerization | CRUD API development | CI/CD Automation | Deep Learning | EMR and Big data pipeline | OOP

## EDUCATION

### MS in Data Analytics

08/2022 - 12/2023

GPA: 3.67/4.0

**University of Massachusetts, Amherst, MA**

**Concentration:** Computational Social Sciences

**Relevant Coursework:** Advanced Machine Learning, Regression Modelling, Fair Algorithms, Semi Parametric Regression

### B.Tech In Computer Science and Engineering

08/2015 - 05/2019

GPA: 9.14/10

**VIT University, India**

Ranked 17/500 in my course.

Best research publication award 2018.

## RESEARCH PUBLICATIONS

- [IoT-based healthcare monitoring system for war soldiers using machine learning](#) - Elsevier Procedia Computer Science, 2018. 64. Citation
- [Edge Information Based Image Fusion Metrics Using Fractional Order Differentiation and Sigmoidal Functions](#) - IEEE Access, 2020. 29. Citation
- [Facial Identification using Haar Cascading with BRISK](#) - Conference ic-ETITE, IEEE 2020

## CERTIFICATIONS

Tensorflow Developer Certificate, Udemy, 2023  
Intro to SQL certification, Databricks 2023

## LEADERSHIP ROLES & RESPONSIBILITIES

- SG GSC Cloud professional services expert
- Course Assistant for Machine Learning for Social Sciences
- Committee Head of Funding for Android Club
- Research Assistant for Prof Wayne Xu's Social Media Lab.
- Upgrad's University Mentor Evangelist