Aneri Rana

arana@umass.edu | (718)808-5622| linkedin.com/in/aneri-rana-09bb87137 | github.com/anerirana

EDUCATION

University of Massachusetts Amherst, USA Exp. Graduation **May 2024**
Expected to graduate with a Master of Science in Computer Science (*Data Science Concentration)*

Pune University, India Graduated **May 2018**
Secured a Bachelor of Engineering in Computer Science with *First Class with Distinction*

TECHNICAL SKILLS

- Data Science Skills: Machine Learning, Natural Language Processing, Deep Learning, Data Visualisation.
- Programming Languages: Python, C++, Java, JavaScript, R, MATLAB, BigQuery, SQL, HTML, Swift.
- Tools and Technologies: PyTorch, TensorFlow, HuggingFace, Apache Spark, Data Studio, Google Cloud Platform, Xcode, Android Studio, Spring Boot Framework.

PROFESSIONAL EXPERIENCE

HSBC Technologies, Pune, India

Senior Software Engineer | *Data Scientist* **Jan 2021 – Aug 2022**

- Spearheaded pilot projects for innovative Data Science use cases in the financial sector with a focus on payments processing, such as fraud pattern detection and cash flow forecasting.
- Developed a **multi-task deep learning** model for automatic repair of transaction fields that pruned the manual processing of more than 200k failed transactions/month in the UK and is planned to further extend to 34 countries. Achieved **0.93 and 0.96** macro average precision and recall of the auto repair model respectively.
- Coordinated a team of 5 Data Scientists, Data Engineers and Business Analysts to develop the engineering solution for **long-term deployment and monitoring** of the Auto Repair ML model for multiple countries.
- Trained a **BiLSTM** model for classification of addresses in payment messages with an F1-Score of 0.94. The predictions are assigned a confidence score and used to generate Data Studio reports for audit in HSBC.
- Constructed a GCP solution for batch training and prediction of Machine Learning models in production using **Airflow** on **Google Kubernetes Engine**.
- Built a distributed cloud computing platform for parallel processing of Machine Learning models on Apache **Hadoop YARN** Clusters using **Apache Spark**.
- Orchestrated the construction of end-to-end API infrastructure on Google Kubernetes Engine for **serving Machine Learning models** in production. Enabled the first two Machine Learning models built on the platform, to go live via the API platform.

Senior Software Engineer | *Full Stack Mobile Developer* **March 2020 – Dec 2020**

- Devised HSBC Kinetic iOS app using Swift; Designed the Open Banking, Cards, Transaction, and Login session management modules of the application (HSBC Kinetic webpage: https://www.business.hsbc.uk/en-gb/everyday-banking/business-accounts/kinetic).
- Designated as iOS Champion for performing code reviews, conversing with business stakeholders to understand requirements and propose feasible solutions, creating reusable components, and supporting DevOps tools.

Software Engineer | *Full Stack Mobile Developer* **Aug 2018 – Feb 2020**

- Implemented an encryption micro-service in Java utilizing Spring Boot framework.
- Enabled smooth integration by resolving issues in the transaction categorization data pipeline connecting mainframe copybooks, Dataflow, BigQuery, Kubernetes micro-services, and iOS application.

PROJECTS, PATENT, AND PAPER

Mars Spectrometry | Detect Evidence for Past Habitability

ML Challenge, NASA, 2022

- Ensembled various combinations of **metric learning** algorithms like LMNN, NCA, ITML and classifiers such as **Extra Trees** and **LightGBM** to predict composition of chemical compounds in geological samples from Mars.
- Evolved gas analysis mass spectrometry data from Mars exploration missions by curiosity rover was given by NASA to predict chemical compositions which indicate past livable conditions on Mars.
- Attained **18th position globally** among 713 participants with a log loss of 0.14 on the test data set.

System and Method for Classification of Hate Speech

- Engineered and patented an innovative system for the classification of audio/video content on social media as hate speech. Patent pending, **Indian Patent Application 201921052170.**
- Collected and annotated 1k video dataset from YouTube and Twitter to train a **multimodal deep learning** model that combines emotional features of speech with semantic features of the text.
- Overcame the challenge of a limited dataset by leveraging **transfer learning** from pre-trained text and emotion models.
- Extracted text features from the last layer of a **BERT** model fine-tuned using PyTorch on a hate speech Twitter dataset.
- Extracted emotion features from a **CNN** model pre-trained on an existing audio dataset to predict emotion attributes, i.e., valence, arousal, and dominance.
- Authored a paper on the same, currently under review by the ACM Transactions on Intelligent Systems and Technology.

Easylend | Mobile Application

Open Banking Hackathon, HSBC, 2018

- Built a peer-to-peer lending platform, that connects loan seekers with less fortunate backgrounds to the willing investors, when it is difficult to obtain loans from banks through conventional means.
- Coordinated a team of 4 to develop the android application in Java.
- Qualified as national winners among 43 teams in India at the HSBC global hackathon that aims to promote innovation in Open Banking.

Document Classification and Information Extraction

Code Grind Hackathon, HSBC, 2019

- Scanned reimbursement documents for Fusion platform using GCP cloud vision API and classified them using **SVM classifier** into categories like food, travel & hotel.
- **Extracted information** such as date and cost using regular expression for auto-filling forms to save employee time.
- Secured second place in CodeGrind, HSBC's largest hackathon.

COURSES AND CERTIFICATIONS

- Machine Learning, by Stanford University, Coursera in Nov 2019.
- Google Cloud Certified "Associate Cloud Engineer" in Aug 2019.
- Developing applications with Google Cloud Platform specialization, by Google, Coursera in Mar 2019.

AWARDS AND ACCOMPLISHMENTS

- Recipient of the "all stars" award for teamwork and well-rounded performance throughout 2019, HSBC.
- Received the "Kinetic Star" recognition for identifying and solving critical issues in Kinetic app, HSBC, 2018.
- Awarded 1st in Hidden leaf Inter-college competition, organized as a treasure hunt using the knowledge of the Linux file system, Techtonic, 2015.
- Placed 2nd in JAM, Tesla, 2016, an inter-college competition to identify and solve bugs within a minute in each of the 30 java code samples.

VOLUNTEERING

- Programmed new features, fixed bugs when identified, and resolved issues raised by the users for **Keras**, **Snorkel**, and Avalanche Open Source Projects, May 2021 – March 2022.
- Managed the website and developed new features (in python) as the Super Volunteer and Web maintainer for Women in Machine Learning (WiML) which aims to enhance the experience and impact of women in ML, Oct 2021 – Aug 2022.
- Facilitated a one-hour breakout session on "Leveraging Open-Source Tools for NLP" at the **International Conference for Machine Learning** (ICML, a top-tier ML conference) with research associates from the Georgia Institute of Technology and the University of Toronto in July 2021.
- Selected as an **emcee** to host talk and keynote sessions for an entire day at the Open Data Science Conference in Sept 2021. Introduced speakers, handled live broadcasting, managed issues during the session, raised questions on behalf of attendees, and concluded the session.