SHIVIKA K BISEN

248-229-6074 • shivikakbisen@gmail.com • LinkedIn • GitHub • Portfolio • San Jose, CA

PROFESSIONAL EXPERIENCE

Sr Data Scientist – **Machine learning**, PAXAFE Inc May 2021 – Present

San Jose, California

- Architected **end-to-end machine learning** and **statistical** model, to **predict trend** real-time for **Johnson & Johnson.** Coordinated with business and dev team to deploy models at global scale, utilizing Python, MySQL, Docker, AWS S3 and Redshift, and Flask APIs.
- Created interactive global-scale **dashboards** and **data visualizations** using Superset, DynamoDB, SQL, and Tableau to monitor insights and optimize business metrics, enhancing product performance by 25% annually.
- Performed **A/B testing** to evaluate performance of machine learning models and updating the model in event of data-drift.
- Implemented **data models** on large scale dataset of 10 TB+. Automated machine learning model retraining pipelines from AWS S3 to Redshift, employing agile methodologies.

Software Engineer – **AI/ML,** iCatalysts Jul 2020 – Apr 2021

San Francisco, California

- Implemented **GPT-3**-based text classification in a proof of concept, contributing to deep learning algorithm development.
- Developed D3.js code for California Energy Commission's app, focusing on **interactive data visualization**, analytics.
- Enhanced product user experience by 19 % by testing new features, identifying bugs, and debugging using Python and Git.

Data Scientist – **Microsoft project**, University of Michigan – Ann Arbor May 2020 – Apr 2021

Ann Arbor, Michigan

- Led end-to-end development of United Nations search engine on **unstructured** and **structured** reports, achieving a 0.87 Mean Average Precision and refined **ranking** and **relevance** (outperforming Google for UN-related queries) (Paper link) .
- Deployed BM-25 ranking on Flask after data scraping, PDF text extraction, and **BERT**-based text summarization.

Machine Learning Scientist – **Google.org Project**, University of Michigan – Ann Arbor Jan 2020 – Nov 2020

Ann Arbor, Michigan

- Designed an NLP/NLU **classifier** using logistic regression to predict **customer behavior** for UpTogether platform aimed to optimize **advertising** and **audience targeting**. Achieved an F1 score of 82.
- Collaborated with Prof. David Jurgen to develop **advanced feature extraction** algorithms using GloVe, LIWC, Stanford CoreNLP sentiment analysis and **RoBERTa** Transformers (TensorFlow/PyTorch).

EDUCATION

University of Michigan – **Ann Arbor** **Ann Arbor, Michigan**

Master of Science in Data Science, Computer Science & Engineering Jan 2019 – Dec 2020

VIT University **Vellore, India**

Bachelor of Technology Biomedical Engineering Jul 2009 – Aug 2013

RELEVANT PROJECTS

E-commerce product recommendation based on transfer learning | Python, AWS Sagemaker, Tensorflow, Keras

- Led the research and development of algorithm for image similarity-based recommendation on large scale ecommerce dataset link.

Advanced data visualization on complex large dataset on US population | Python, Javascript (D3.js), Big data

- Implemented data processing for geospatial and temporal data for visualizing large scale CDC dataset to identify trends and get correlation analysis of multiple variables link.

Sentiment analysis on autonomous cars | Python, Scikit learn, Big data, Data mining, Information technology

- Extracted insights from large scale unstructured dataset using doc2vec, advanced topic modeling NLP techniques GitHub.

TECHNICAL SKILLS

Technical Languages: Python, SQL, JavaScript, R, HTML5, CSS3

Tools and Frameworks: AWS (S3, Redshift, Sagemaker), Docker, Git/Bitbucket, Pandas, NumPy, SciPy, Scikit-learn, TensorFlow, Keras, Pytorch, Matplotlib, NetworkX, genism, React, Flask, Django, Node.js, D3.js, Hadoop/Spark, Jira, Excel, Jupyter, MySQL, PostgreSQL

PUBLICATION & AWARDS

- Presented at ML Conference 2022, sharing insights on successfully deploying machine learning solutions to production.
- Secured top rank 2/70 in a Kaggle competition at UMich, focusing on ML search algorithms for COVID queries; featured on Medium for GPT-3 vs BERT | Stanford coreNLP| GloVe embedding
- Published "An Improved Segmentation Technique for Breast Infiltration/Tumor Detection from Mammograms" in International Journal of Engineering and Technology, 2013, Vol. 5, Issue 3, pp. 2565-2574.
- Awarded UMich Scholarship as Data Science Coordinator for 100+ LSAMP members, providing workshops on machine learning and deep learning