

# Shivika K Bisen

248-229-6074 | <https://sbisen.github.io/> | <https://www.linkedin.com/in/shivika-k-bisen/> | <https://github.com/sbisen>  
|[sbisen@umich.edu](mailto:sbisen@umich.edu) | Troy, MI, 48084

## EDUCATION

### University of Michigan Ann Arbor

Ann Arbor, MI

**Master of Science in Data Science**, GPA: 3.69/4.00

January 2019 – December 2020

*Coursework:* Applied Machine Learning, Regression Analysis, Database Systems, Advanced Data Mining, Information Visualization, Models of Operation Research, Time Series Analysis, Agile Software Development, Information Retrieval.

### VIT University

Vellore, India

**Bachelor of Technology Biomedical Engineering**, GPA: 3.98/4.00

July 2009 – August 2013

## SKILLS

Python, JavaScript, R, PHP, HTML/CSS, SQL, Pandas, Numpy, Sklearn, TensorFlow, Keras, Pytorch, NLTK, CoreNLP, spaCy, Matplotlib, seaborn, NetworkX, gensim, Altair, Plotly, Pydeck, Django, Flask, Elasticsearch, SQLite, D3.js, MySQL, Neo4j, AWS Sagemaker, AWS, Spark, Hadoop, GCP, Jupyter, Mallet, MATLAB, Tableau, ERP, CRM, Jira, Airtable, Microsoft Excel, Docker, Git

## WORK EXPERIENCE

### PAXAFE

Indianapolis, IN (Remote)

#### Data Scientist

May 2021 – Present

- Leading machine learning model development for advanced temperature prediction from real-time IoT data. Developed **regression** models and tested it against the baseline model of time series forecasting.
- Responsible for building web applications and data analysis dashboard for two multinational companies (clients).
- Worked on geospatial data analysis and advanced 3D visualization.

### iCatalysts,

San Francisco, CA (Remote)

#### Data Science Developer

July 2020 – April 2021

- **Accomplishment:** Implemented GPT-3 (NLP deep learning) and developed interactive web app- network analyzer for California Energy Commission.
- Implemented **natural language processing, classification**, sentiment analysis into the product codebase.
- Worked with Neo4j, Node2Vec and developed D3.js **interactive visualization** for Empower Innovation website.
- Worked in prototyping, debugging, testing, maintaining application in scalable software development process (ETL).

### University of Michigan

Ann Arbor, MI (Remote)

#### Data Science Researcher (Data Scientist, Search Engine)

May 2020 – April 2021

- **Accomplishment:** Deployed Chetah **Search Engine** for **United Nations** reports with MAP- 0.78 (better than the Google for UN docs queries)
- **Semantic Search: Team lead** for **end to end** data science process. Web crawled and extracted unstructured text from PDF using Apache Tika, Haystack, Elastic search and ETL. Performed **information extraction** on the data.
- Experimented with **chatbot** development and analyzed the data with Hadoop, SQL and Spark.
- Deployed BM-25 index/ranking systems (information retrieval code) and recommendations systems algorithms with deep learning algorithm BERT. Compared results with baseline model- word2vec cosine similarity, BM-25L.

- **Accomplishment: Predictive ML model** with an 82 F1 score, to predict for the types of social interactions on social networking site- Uptoegether.
- Mentored by Prof. David Jurgen and Prof T. Dillahun. Lead the team and developed ML supervised models, semi supervised and unsupervised.
- Performed **Feature engineering** using LIWC, Stanford CoreNLP and GloVe word embedding, Word2Vec and LIME. Implemented deep learning architectures/compiler like RoBERTa.
- Performed **topic modeling** on a big data set using Mallet. Also, experimented with **clustering** and **semi-supervised** models like Doc2Vec using the gensim library.

### RELEVANT PROJECTS

- Search Engine for pdf documents based on Word2vec, BM25 and BM 25L Article. [Video](#)
- Sentiment analysis on autonomous cars (NLP- supervised, semi supervised, clustering, topic modeling). [Github](#)
- Deep Learning based recommendation system with AWS Sagemaker, VGG-16 and ResNet. [Github](#)
- Advanced data visualization on complex dataset CDC [Demo](#)

### PUBLICATION AND CONFERENCES

- **An Improved Segmentation Technique Based on Delaunay Triangulations for Breast Infiltration/Tumor Detection from Mammograms** [International Journal of Engineering and Technology](#).
- Publication featured on **Medium**: [Topic Modeling](#) | [D3.js interactive line charts](#) | [Stanford coreNLP \(python\)](#).
- MIDAS Annual DPG Conference 2021, UMich 2021 | WiDS Conference 2021, Stanford University 2021

### ACHIEVEMENTS AND LEADERSHIP

- **Top 2nd rank in Kaggle competition- UM 2020** (Search engine algorithm for Covid-19 queries)
- **College of Engineering and CS Graduate Scholarship, UMich 2019.**
- **Graduated with Honour, Magna Cum Laude.**
- **Awarded MI- LSAMP Scholarship for exceptional STEM achievement, UMich 2020.** Lead for ML-workshops at Michigan Institute for Data Science (MIDAS) for Universities like MSU, Wayne State, WMU and UMich.