somnathr.com
linkedin.com/in/somnathrakshit

SOMNATH RAKSHIT

(512) 798-3552
somnath52@gmail.com
github.com/somnathrakshit

EMPLOYMENT

Senior Data Scientist, Walmart Inc. **Dallas, TX, July 2021 – Ongoing**

- Built linearly additive time series models to forecast sales for different granularities, time periods, and channels.
- Collaborated with other data scientists to explore the data, experimented with different models, and evaluated them using appropriate metrics.
- Reduced model inference time from 20 seconds to 0.1 seconds using vectorization.
- Architected, implemented, and deployed 100+ models to production using FastAPI and Docker through a REST API.
- Developed data quality checks, and CI/CD pipelines for training and inference services.

Research Assistant, Univ. of Warsaw **Warsaw, Poland, Jan 2019 – Aug 2019**

- Ranked genetic biomarkers for breast cancer subtypes (Mentors: Dr. I. Saha and Dr. D. Plewczyński).
- Developed a novel ranking algorithm using multi-objective genetic algorithms.
- Demonstrated the relation between up/down regulation of genes and survival probability of a population.

Software Engineer – ML, Cyware Labs **Bangalore, India, Jul 2018 – Nov 2018**

- Built a Convolutional Neural Network to cluster similar articles resulting in 2x no. of articles selected.
- Determined trending keywords using Named Entity Recognition from news articles.
- Deployed models in AWS EC2 using Celery, Gunicorn, and Django.

EDUCATION

The University of Texas at Austin **Austin, TX, Aug 2019 - May 2021**

- **M.S. in Information Studies, GPA: 3.89**
- **Graduate Coursework:** Mathematical Statistics for Applications, Linear Models, Data Mining, Introduction to Machine Learning, Database Management
- **Teaching Assistant:** User Generated Content Analytics (Fall 2019), Data Science Principles (Spring 2020 & 2021), Advanced Predictive Modeling (Fall 2020).

Jalpaiguri Govt Engineering College **WB, India, Aug 2014 - May 2018**

- **B. Tech in Computer Science and Engineering, GPA: 8.68**
- Artificial intelligence, Data Structures, Calculus, Discrete Mathematics, Probability and Statistics, Design and Analysis of Algorithms, Object Oriented Programming

LANGUAGES AND TECHNOLOGIES

- Python, SAS, Java
- Scikit-learn, Numpy, Prophet, Pandas, Matplotlib, Spacy, PyTorch, Tensorflow, Keras, Git
- FastAPI, Docker, Kubernetes, Apache Spark, Linux, AWS, Azure, GCP

PROJECTS

Academic Projects

- **A GPU-accelerated MRI Sequence Simulation for Differentiable Optimization and Learning:** Wrote an open source GPU accelerated simulator in PyTorch for the Extended Phase Graph Algorithm for certain sets of sequences. After successfully parallelizing the simulator, we increased its speed by almost 100,000 times. This work was accepted as an abstract in ISMRM virtual conference. Guide: Dr. Jon Tamir
- **Credit Risk Model and Scorecard:** Built a credit risk model using logistic regression after preprocessing and feature engineering in Python to predict the probabilities of loan default and assigned credit scores to potential borrowers.
- **Deep Stock Predictions:** Built a multi-stack Bidirectional LSTM model strengthened by the addition of Attention units. Proposed a data driven approach for optimal selection of window length and multi-step prediction length, and added analyst calls as technical indicators to the model. Guide: Dr. Edison Thomaz

Personal Projects

- **geograpy3:** geograpy3 is a Python library that is used to extract place names from a URL or text, and add context to those names, e.g., distinguishing between a country, region or city. It can be installed from PyPi using pip and has over 100,000 downloads.