SHRITEJ SHRIKANT CHAVAN

Richardson, TX, 75080 | 945-247-4245 | sxc220078@utdallas.edu | LinkedIn |GitHub

EDUCATION

The University of Texas at Dallas, Richardson, Texas — **Expected – Dec 2023**
Master of Science, Business Analytics Data Science Cohort – Minor, Applied Machine Learning — **GPA 4.00**
Awards - Dean's Excellence Cohort Scholarship and Graduate Teaching Assistant

Indian Institute of Technology Madras, Chennai, India — **Aug 2014 – Sep 2019**
Dual Degree - Bachelor of Technology and Master of Technology, Biological Engineering — **GPA 3.54**

PROFESSIONAL EXPERIENCE

Data Science Intern – T-Mobile USA Inc, TX, United States — **May 2023 – present**

- Deployed ML evaluation tool for T-Mobile's internal data hackathon to score and rank participants ML models using Azure Databricks, AutoML, MS Power Apps and Azure Data Lake by collaborating with cross-functional team of engineers

Data Scientist - Target Corporation India, Bangalore, India — **Jan 2020 – Jul 2022**

Marketing – Bid Price Optimization

- Led project on optimizing opening bid price to promote specific online searched products and increased revenue by $5M
- Conceptualized methodology consisting of **Monte Carlo** simulation on Bid Price distribution and **SARIMAX** forecasts of metrics like clicks, click rate and cost-per-clicks of sponsored product ads with **MAPE less than 25**%
- Evaluated Incremental Sales using campaign pre- and post-analysis and statistical techniques such as **A/B testing**, **KL Divergence,** and Mann-Whitney U Test and **HiveSQL** for data wrangling across all the Loyalty Marketing campaigns in 2019 of **$160M**
- Employed **Bayesian Optimization** technique for Hyper-parameter Tuning and improved model's runtime by **50%**

Merchandising – Store Product Placement Optimization

- Developed quasi-experimental framework consisting of **Difference-in-Difference** and **Synthetic Control** methods for causal inference to evaluate product positioning strategy and product placement and improve ROI by 14%
- Built SKU-level Forecasting engine based on **Interrupted Time Series** analysis by using ProphetFB for time series on top of **Classification and Regression Trees (CART), XGBoost** and **Bayesian Networks** to group similar items and got **MAPE less than 0.3**
- Redesigned Hive-based data pipeline by **PySpark** framework to improve scalability and reduced execution time by 35%

Merchandising – Store Remodel Assessment

- Executed **Principal Component Analysis** to reduce dimensions and implemented **Hungarian Optimization** to optimize store layout and **Elastic Net** regression to estimate its financial impact which lifted our sales by **10%**

Machine Learning Intern - Microspin, Chennai, India — **Jun 2019 – Jul 2019**

Dynamic systems simulation and control implementing recurrent neural network-based architecture.

- Designed consecutive **LSTM**-based architecture that forecast future rpm to forestall any motor failures and got **RMSE of 2%**

ACADEMIC PROJECTS

Market Mix Modelling using Genetic Algorithms and Multivariate Regression, HPE-UTA 2023 Analytics Student Competition (link)

- Secured **1st** place among 52 teams in HPE-UTA 2023 Analytics Competition and presented at UTA's 7th Annual Analytics Symposium
- Formulated **non-linear optimization** problem of marketing campaign's budget allocation using XGB Regressor to predict campaign return like click-through rate and conversion rate and optimize the spending using robust approach of **Genetic Algorithms**

Modeling of human gait through Reinforcement Learning, master's Thesis (link)

- Analyzed **neuronal mechanism** of gait by modeling postural sway during upright human posture as a 1-D inverted pendulum
- Deployed Deep Deterministic Policy Gradient, **Deep Reinforcement Learning** algorithm for continuous action (torque) space and performed **Hyper-parameter Tuning** for 120 epochs to limit 95% of sway frequencies in the range of -1 to 1 and sway to 0.5 cm

A Panel Data Study of the Determinants of Life Expectancy in Sub-Saharan Africa (SSA), Econometrics Research Project (link)

- Collated data from **World Bank** and performed **EDA** on panel data of 19 SSA countries from 2005 to 2019 to analyze the impact of country's spend on Healthcare and Education on life-expectancy using **Fixed Effects** and **Random Effects** with R-squared > 0.9

Accented Text-To-Speech Synthesis: Natural Language Processing Project (link)

- Built a Text-to-Speech system capable of accented conversion using L2Arctic dataset (27 hours recorded speech)
- Designed architecture incorporating **Tacotron2** with **Variational Autoencoder**, implementing 2 variants based on labels inclusion
- Achieved Mel Cepstral Distortion (MCD) of 7.1 and Word Error Rate (WER) of 0.25 for No-label model, and 6.98 and 0.24 for label model respectively, outperforming **GMVAE**, GST and GT models to produce state-of-the-art quality speech

TECHNICAL SKILLS

Programming: Python (TensorFlow, Keras, Pytorch, Streamlit, NLTK, Matplotlib, Scikit-Learn), R, Hive SQL, PySpark, MongoDB, C++
Software: AWS, Snowflake, Apache Hadoop, Kubernetes, Databricks, Airflow, GitHub, Tableau, STATA, SAS, Excel, JIRA, Linux