

Shardul Suryakant Rane

shardul.rane0913@gmail.com | +1(716)602-3269 | <https://github.com/shardul0913> | <https://www.linkedin.com/in/shardul-rane-06142b115>

EDUCATION

- **Master of Science in Data Science** **GPA: 3.56/4.0**
University at Buffalo, State University of New York
Aug 2019 - Feb 2021
Relevant Coursework: Machine Learning, Statistical Data Mining, Deep Learning, Data Intensive Computing, Probability, Data Modelling
- **Bachelor of Engineering in Computer Engineering** **GPA: 7.52/10.0**
University of Mumbai
Jun 2013 - May 2017
Relevant Coursework: Artificial Intelligence, Software Engineering, Data Structures, Database Management, OS, Analysis of Algorithms

TECHNICAL SKILLS

- **Skills:** Predictive Models, Hypothesis Testing, EDA, Time-Series Models, Anomaly Detection, Topic Models, Marketing Analytics
- **Programming and Technologies:** Python, R, SQL, NoSQL, MATLAB, Java, C, JavaScript, Spark, Hadoop, Docker, RESTful API, Google Cloud Console, AWS Lambda, Sagemaker, IBM SPSS, Qlik, Tableau, GitHub, Adobe Analytics, Google Analytics, Excel
- **Libraries:** Scikit-Learn, PyTorch, Keras, TensorFlow, Scipy, Pandas, Numpy, Matplotlib, dplyr, DAAG, FB prophet, R-Shiny

PROFESSIONAL EXPERIENCE

Artificial Intelligence Intern, Hewlett Packard (HP) -Boise, USA **Jun 2020 – Present**

Developed Forecasting Tool for Printer Toner Usage at 93% Accuracy

- Built Time series model for 5 years monthly usage prediction in 10 markets to help product managers with printer feature selection in new product design. **Tools :** Logarithmic Regression, R-Shiny, AWS Redshift, MS SQL, Qlikview
- Tweaked accuracy of 82% by 11% with Stepwise Selection and Regularization for better model interpretation with PMs

Analyzed Impact of Pandemic on Product Usage and Share in 10 Markets Worldwide

- Performed Hypothesis Testing on Time-Series dataset with Granger Causality, Dickey-Fuller test to understand impact of COVID-19 on product sales and page shares in Americas, Europe and Asia

Traced Patterns in Counterfeit Product Reviews with 85% accuracy with NLP

- Designed Bi-gram Topic Model and Sentiment Engine for Clone Product Usage. Intimated PMs by highlighting the topic wise keywords for positive/negative reviews. **Methods:** Sklearn, NLTK, Non-Negative Factorization, SVD, TF-IDF, Watson sentiment

Data Analyst, Performics, Publicis Groupe -Mumbai, India **Dec 2017 – Jun 2019**

Implemented User Behavior and Intent Prediction Model with 84% accuracy on Web Traffic data

- Developed predictive model to track user purchase intent based on search term & demographics with Data studio dashboard. Improved Display Targeting revenue by 12% for 3 Clients. **Methods:** GCP, Python, KMeans, PCA, Random Forest Classifier

Designed Website Performance Reporting BI Framework for 10+ Ecommerce, Travel and Finance Firms

- Analyzed Clickstream data to build Business Intelligence dashboard for sales, product performance, A/B tests using Python, JavaScript in Google Analytics. Helped improve user conversions by ~16% for Thomascook, Bloomberg and Reliance Insurance.

Established Custom Ecommerce Revenue Attribution Reports To Reduce Task Time by 79%

- Updated Traditional First touch attribution model based on user session duration and conversion rules with an automated reporting dashboard in Google Analytics. **Tools :** Python, Excel, Google Analytics, BigQuery, GCP

ACADEMIC PROJECTS

- **Classification Model Project with ‘Dublin Business School’ students to Predict Game Outcome** [\[Link\]](#)
 - Implemented Ensemble Learning with 78% Recall accuracy using XGBoost, Random Forest Classifier and SVM on multiclass dataset with to predict soccer game result. Implemented Grid Search for Parameter Tuning for the models.
- **Anomaly Detection in Machine Sensors Using Supervised Machine Learning** [\[Link\]](#)
 - Traced machine failures with 89% accuracy for varying threshold with ensemble of Auto-encoder LSTM Neural Network and ARIMA model on ‘Pressure sensor’ data. Improved accuracy by 15% by Stabilized Time-series by Time Differencing.
- **Apache Spark Multi-label Natural Language Classifier to track Movie Genres** [\[Link\]](#)
 - Overcame the limitation of Multi-label classifier in Spark by implementing multiple Logistic Regressors for each class
 - Improved accuracy of 71% by 13% with feature engineering by validating TFIDF, Word2Vec in Spark Mllib
- **Fraud Detection Web-tool for Employees Travel Expenses Auditing** [\[Link\]](#)
 - Tracked frauds with SVM & predicted expenses with L2 Regression at 96% accuracy. Tackled data imbalance with SMOTE
- **Convolutional Neural Network model with Transfer Learning to Identify Stars with Exoplanets** [\[Link\]](#)
 - Classified the exoplanet stars with 97% accuracy using CNNs for Identifying light flux Spatial Pattern
 - Implemented Transfer Learning between VGG16 and traditional CNN in Keras with Flask API for model support

ACHIEVEMENTS & LEADERSHIP

- Achieved AI Explorer badge as AI intern at Hewlett Packard & participated in InternStellar Awards for print usage prediction tool
- As a Member of Data Analytics seminar group at Publicis, Trained employees on Advanced Analytics and Statistics
- **Certified:** Coursera Neural Networks, AWS Machine Learning, Apache PySpark, Tableau Certification, IBM SPSS