**Sanjay M Tamrakar**

**• (832) 359 6523 • tamrakar\_Sanjay@hotmail.com**   **•Green Card Holder**

Experienced Data Scientist with a demonstrated history of working in the research industry. Highly skilled in machine learning, deep learning and statistical models including but not limited to survival analysis, mixed-effect models, time-series, general linear models, neural-networks (CNN, RNN, ANN, GNN, LSTM), natural language processing and text mining.

**EDUCATION**

**MS Statistics Miami University, Ohio Aug 2019**

***GPA: 3.75/4.0***

**BS Mathematics University of Texas at Tyler, Texas May 2015**

(Minor: Business)

**KEY SKILLS**

* Knowledge in **MS Office suite** (MS-Word, Access, Excel, PowerPoint, Outlook)
* Advanced Knowledge of Programming in **R**, **SAS, Python, SQL, Tableau, MATLAB** and **JMP**
* Building a pipeline from data collection, extraction and verification with building various machine learning, deep learning models and then presenting meaningful results on Dashboard
* Experienced with version control tools like **GitHub**, **Gitlab**, **TortoiseSVN**
* Experienced working on **High Performance Computing** along with cloud-computing platform like **Amazon AWS**

**RELEVANT EXPERIENCE**

* ***Data Scientist II Battelle Memorial Institute July 2018 – current***
* Worked with the Health Analytics team in supporting government and private industry clients in the healthcare, defense, environmental and energy sectors
* **Communicated** efficiently with internal and external clients in describing technical results in a clear and succinct manner in the form of presentations and technical reports
* Worked in a fast-paced environment with multiple projects and met deadlines
* Supported business development efforts by contributing to technical sections of the proposals
* Worked **independently** as well as in **teams**, including **leading** and **mentoring** junior staffs in the projects and overall career development
* Automated multiple tasks to reduce time in multiple projects by efficient coding strategies
* Built a pipeline in assessing the human-pattern recognition using time series clustering (**Permutation Distribution Clustering**, **Dynamic Time Wrapping**) using Python
* Implemented and tested methodologies to solve machine learning/deep learning models for the classification problem related to REBS (Resource Effective Bio-identification System) in R and Python
* Tested various deep-learning methods and architecture search algorithms like **DARTS** (Differential Architecture Search) with the Neuro-Engineering Team using Python
* Implemented various anomaly detection algorithms like **OneCLassSVM**, **IsolationForest** , **Autoencoders** in Python to detect anomalies in the data
* Performed meta-analysis, survival analysis, cross-design and various mixed models in drug study as well as time to death analysis using SAS and R
* Contributed to delivering the dataset, codebook, data dictionary as well as written well-versed code in the survey analysis data
* Performed **CART** analysis, disparity analysis, descriptive statistics reported in the Wayfinder platform
* ***Graduate Assistant/ Projects Miami University Aug 2016 – July 2018***
* Held office hours and helped students understand more about Applied Statistics and Business Statistics
* ***Client-based Project*** – (i) Difference in Heart Rate Variability for Sherpa and College student,

(ii) Organizational sponsorship in contexts of poverty: exploring the determinants of venture growth

* ***R-Shiny Project*** –(i) Model building with regression analysis on New York taxicab data,(ii) Behavior of various commodities during Great Recession
* ***SOA Case Study Challenge 2017*** – Akua Island Coastal Development and Zone Allocation
* ***DataFest 2017*** – Expedia website data to find meaningful conclusions from large data
* ***Kaggle Project*** – (i)Predicted sales prices for houses and practiced feature engineering, RFs, and gradient boosting, (ii) Logistic regression to predict who survived in the Titanic data
* ***Undergraduate Project*** – Statistical Models for Discrete Data of DNA samples
* ***Grader and Tutor University of Texas at Tyler Aug 2013– Aug 2016***
* Taught advanced calculus classes and lab sessions, graded assignments and tutored students
* Held office hours and proctored tests

**PUBlICATIONS**

* Niemuth NA, Fallacara D, Triplett CA, **Tamrakar SM**, Rajbhandari A, Florence C, Ward L, Griffiths A, Carrion R Jr, Goez-Gazi Y, Alfson KJ, Staples HM, Brasel T, Comer JE, Massey S, Smith J, Kocsis A, Lowry J, Johnston SC, Nalca A, Goff AJ, Shurtleff AC, Pitt ML, Trefry J, Fay MP. **Natural history of disease in cynomolgus monkeys exposed to Ebola virus Kikwit strain demonstrates the reliability of this non-human primate model for Ebola virus disease**. PLoS One. 2021 Jul 2.
* Colachis SC 4th, Dunlap CF, Annetta NV, **Tamrakar SM**, Bockbrader MA, Friedenberg DA. **Long-term intracortical microelectrode array performance in a human: a 5 year retrospective analysis**. J Neural Eng. 2021 Aug 23;18(4). doi: 10.1088/1741-2552/ac1add. PMID: 34352736.