MRUDUL BANKA

682.551.9129 | mrudulbanka@gmail.com | Portfolio: http://www.mrudulbanka.ml
San Francisco - Dallas: Open to Relocation | LinkedIn: https://tinyurl.com/yyzk7a2v

EDUCATION

- University of Texas at Arlington, Arlington, Texas   May 2019
  Master of Science in Industrial Engineering, GPA: 3.8 (Awarded for academic excellence by University)
  Courses: Statistics, Regression Analysis, Data Mining, Operations Research, Design and Analysis of Algorithms, Decision Analysis, Simulation, Engineering Economics, Experimental Design, Big Data Analytics, Production-Inventory control.
- **MOOCs:** Machine learning & Deep learning - Andrew Ng & Fast.ai, Social Network Analysis- Lada Adamic & Michael Kearns
- SVKM's NMIMS University, India - **Bachelor of Technology in Mechanical Engineering**   Aug 2015
  Ranked in list of top 12 solvers globally conducted engineering competitions by Ennomotive.com.

WORK EXPERIENCE

Big Data Analytics Co-Op, Apilation.ai, Mill Valley, CA - Dallas, TX - Remote   Aug 2018 - May 2019
- **Circuit Billing Reconciliation Engines:** Designed, built and tested *NoSQL aggregation pipelines* of tariffs for a top 5 telecom company in jSonar, working in an agile environment which would help achieve additional revenue of upto $6 million and save manual audit by 6000 hours. Created project plan and *visualizations* to provide *statistical insight* of results.
- **TV Program Recommendation Engine**: Built a *collaborative filtering* model in python to recommend programs a customer would watch based on consumer behaviour history and *micro-segmentation*. Used IMDB genres to enrich data.
- **Set Top Box Failure Prediction:** Created *xgboost* prediction model pipeline to *predict failure* based on *6 KPI's* for a top 5 telecom company using R and python. *Productionized* model scoring refreshes every 15 minutes on streaming data.
- **Data Profiling:** Explored data, *developed hypotheses* for digital advertising value generation for a large industry.

Junior Data Analyst, Shree Jeenmata Dyeing & Printing Mills Private Limited, Mumbai, India   Dec 2016-July 2017
- **Forecasting:** Implemented *time series* modelling to understand *seasonal* flow of demand and requirement of raw material. Formed *forecasting patterns* to predict inventory levels.
- **Process Improvement:** *Optimized* resource allocation in production system by applying *LEAN SIX SIGMA*, KAIZEN and 5S, obtaining a *11% reduction* in monthly resource cost.

Assistant Engineer, Rean Watertech Private Limited, Kolkata, India   Dec 2015-Nov 2016
- **Business Development:** Part of a four-member team successful in bagging 3 rural water supply projects of *$20 million* - assessed tender price bids, conducted *procurement risk analysis*, trend analysis, *supply chain* communication.
- **Engineering:** Designed sections of conveyor and bulk material handling equipment.

SKILLS

Statistical Modeling and Analysis: Python, R, SAS, SQL, MATLAB, Advanced Excel, Tableau, Power BI.
Machine learning libraries: Weka, Scikit-learn, H20, TensorFlow, Keras, Pandas, NumPy, SciPy, Matplotlib, PyTorch, OpenCV
Programming Languages and Version Control: Java, C, Embedded C, C++, JavaScript, Git, Flask – Heroku - model deployment
Database Management: MongoDB-NoSQL, MySQL, Hadoop – MapReduce, Spark, ML, Hive, Impala, Scoop, Flume

PROJECTS

- **The New York Times Comments Analysis**: To predict which comments provide more insights of customer opinion using natural language processing, logistic regression, multilayer Perceptrons, clustering in pyspark and topic modeling.
- **Credit card Fraud Prediction:** Trained and test machine learning (ML) models to predict based on customer details. Obtained F1 score in a range of 0.90 – 0.93 by using Random Forest and other machine learning models in spark
- **A/B testing:** Design and analysis of a two-factor experiment to test coffee pH level from 3 roasts and 2 coffee machines.
- **Cryptocurrency Prediction**: Trained model with multiple activation functions using Keras - validation loss of 0.29 – 0.37.
- **Simulation**: Developed model of delivery and installation of solar panels for Sunrun Inc. using Witness Horizon. Optimized KPI's to improve process lead time from a multiple-day to one-day installation system and overall customer experience.
- **Image Classification and Object Detection**: Built models with help of TensorFlow, OpenCV – cats vs dogs, food, mobiles.
- **Audio Classification**: Trained Keras convolutions neural network to classify urban street sounds.
- **Twitter streaming analysis:** Built models to detect hate speech in tweets, sentiment analysis and data scrapping.
- **House price prediction:** Used advanced regression, stacking multiple algorithms to obtain RMSE of 0.93 – Kaggle Top 10%.
- **Model Deployment:** Deployed model to predict spam text messages https://murmuring-sierra-45580.herokuapp.com/
- **3D Printer:** Designed and fabricated 3d printer of build volume 200 mm^3 to produce prototypes in PLA and ABS.

SUMMARY

- Self-driven approach to problem-solving in complex systems with the ability to develop and implement quantitative models.
- Capable of working with multiple programming languages, unstructured large data, and data-driven products.
- Adept at communication and can seamlessly integrate into a multidisciplinary team environment or in a collaborative role.