MRUDUL BANKA 682.551.9129 | mrudulbanka@gmail.com | Portfolio: https://www.bankamrudul.ml
Plano, Tx: Open to Relocation | LinkedIn: https://tinyurl.com/yyzk7a2v

SKILLS

- **Modeling**: Linear, Non-Linear, Integer, dynamic - Programming, Convex Optimization, Game theory, Fuzzy logic, Decision trees, Markov Chains, Queuing Theory, Regressions, NLP, Timeseries, Neural Networks, Classification, and heuristics.
- **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.
- **Programming Languages and Version Control**: Java, C, Embedded C, C++, JavaScript, Git, Flask, Heroku.
- **Database Management**: MongoDB-NoSQL, MySQL, Hadoop – MapReduce, Spark, ML, Hive, Impala, Scoop, Flume,
- **Cloud Services**: Azure Data Factory, Azure Databricks, Presto.

WORK EXPERIENCE

- ***Data Scientist,*** *Apilation, Inc*, Plano, TX Aug 2019 - Present
  - **360 Omnichannel Churn Prediction:** Performed diagnosis of **cancellation and deactivation data** to create **customer segments**. Used multiple sources of data to **create an interactions 360 data lake to create profiling** of customers in conjunction with diagnosis. Created interpretable **xgboost, anomaly detection and a black box deep learning model** to find if a customer **churns and its reason using H2O.ai, python.**
  - **Home services customer churn**: Created multiple machine learning models to understand which sales agreements are at a risk of churn in the **next 2 months**. Trained 4 models, 2 each for residential and commercial customers with the best model having 356 features and an area under curve of the **precision recall curve (auc_pr) of about 0.86 and live production results of 0.62 average auc_pr**.
  - **Home services orders and technician capacity forecasting:** Performed forecasting of number of work orders and number of technicians required to fulfil the forecasted work orders, using neural networks and stacked ensemble for multiple branches and business locations for year 2021.
  - **PepsiCo Consultation**: Part of team helping Pepsi revolutionize their existing data processes and architecture. Performed analysis to **understand bottlenecks** in their processes, **found root cause of issues** and pain points across the organisation's architecture. Helped design the new architecture and flow of data by mapping the processes from source to consumer. Developed **solutions** which could improve and accelerate the entire modelling workflow
  - **PepsiCo analytics Development**: Designed and executed ingestion scenarios for data ingestion from 50 data sources. Performed end to end testing of dev, QA and prod environments which included testing of Azure Data Factory, Databricks, h2o driverless ai, presto engine, and React server. Executed strategy to orchestrate ad hoc machine learning models. Acted as a facilitator to simplify handoff's between various teams.
  - **Call fraud detection:** Modeled call risk severity to predict whether a call being made is fraudulent before the call is accepted with **severe risk prediction accuracy of 0.77** with Call data records and additional feature like blacklist flag, robocalls flag, etc.
  - **Performance analysis for customer service agents**: Created a Performance and **sentiment analysis** system to understand the performance of the branch, branch managers, and agents. Used open-source natural language processing libraries to train sentiment analysis and **topic modeling** models from verbatim field of **survey data**.
  - **IoT projects**: Performed in depth analysis and created dashboards for multiple **IoT devices** like energy drink coolers, shipping containers, routers, residential gateways, DSLAM, and financial point of sales devices. For some devices created machine learning models to **predict device failure**.
  - **Conducted training sessions for data professionals, interns to train** on Apilation's tools. Participated and presented in client demos and technology validation sessions.

- ***Big Data Analyst,*** *Apilation, Inc*, Plano, TX Aug 2018 - May 2019
  - **Circuit Billing Reconciliation:** Designed, built, and tested ***NoSQL aggregation pipelines*** of tariffs for a top 5 telecom firm in an agile environment to save manual audit by 6000 hours. performed ***root cause analysis*** *and created dashboards to show the results.*
  - **TV Program Recommendation Engine**: Built a ***collaborative filtering*** model in python to recommend programs 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 in 15 minutes on streaming data.

- ***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*** production resource allocation with application of – ***FMEA, LEAN SIX SIGMA*** & Kaizen, obtaining a monthly ***11% cost reduction***.

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

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 and optimization, Engineering Economics, **Design of Experiments, Big Data Analytics**, Production-Inventory control.

- SVKM's NMIMS University, India - **Bachelor of Technology in Mechanical Engineering** Aug 2015
  - Ranked in the list of top 12 solvers for globally conducted engineering competitions by Ennomotive.com.

PROJECTS

- **Model Deployment:** Deployed a model to predict spam text messages - https://murmuring-sierra-45580.herokuapp.com
- **Venture Capital Investment Analysis:** Forecasted traffic on Jet Rail for 7 months to help Unicorn Ventures make an investment decision using time series Analysis –submission root mean squared error (RMSE) = 216, Top 20%.
- **Jigsaw Unintended Bias in Toxicity:** Used word embeddings with LSTM, GRU and word2vec with Kaggle score of 0.93. At the time of submission was ranked in the top 10%.
- **AI Chatbot:** Built a chatbot in pyspark which understands the sentiment of the conversation and answers to move the conversation towards a positive sentiment.
- **A/B testing:** Design and analysis of a two-factor experiment to test coffee pH level from 3 roasts and 2 coffee machines. Performed the experiment, collected the data, and analysed the data in SAS.
- **House price prediction:** Used advanced regression stacking many algorithms to obtain RMSE 0.93, Kaggle Top 10%
- **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.
- **Hand Posture Recognition**: Trained and tested machine learning (ML) models to recognize 5 different hand postures from a dataset of 12 users using MATLAB and Python. Obtained best model accuracy of 0.975 by using bagging trees classifier.
- **3D Printer:** Designed and fabricated 3d printer of build volume 200 mm^3 to produce prototypes in PLA and ABS.