**Rahul Kotagiri**

**krahultej@gmail.com**

**Phone: +1 419 378 (0118)**

**Chatham, NJ**

Over 14+ years of experience in the areas of Digital Analytics, Digital Marketing, Market Segmentation, Channel Mix, Personalization, Web journey, new business, and customer lifecycle/engagement analysis. ***Core areas*** like Computer Vision, GenAI, LLM, Machine Learning, Data Mining with large datasets, Data Acquisition, Data Validation, Predictive Modeling, and Data Visualization. Worked extensively on a wide range of projects including NLP, Image Processing, Deep Learning, Statistical analysis, Machine Learning, Optimization, and Simulations. I am proficient in languages like Python, SAS, R, SQL, Hive, PySpark, etc.

**TECHNICAL SKILLS**

**Languages:** PySpark, Python, Golan, SAS, Java, C++, R, SQL, HIVE, PIG,, Linux, H2o, Angular JS, C#, .Net, C++, TensorFlow, Caffe, Pytorch , Keras, Scala, Shell

**Machine Learning** XGB, GBM, Ensemble models, K-NN, Decision Tree, Random Forest, K-Means clustering, Sequential Rules, SVM, Regression models,

**ML Techniques** Regularization, Hyper Parameter Tuning, Feature Engineering, Dimensionality Reduction.

**Deep Learning (DL)** GenAI, LLM Models, OpenCV, Neural Networks, Deep Neural Networks (CNN & RNN) and LSTM,

**Cloud & Frameworks** AWS, Azure, Google Cloud, TensorFlow, Kera’s, and Pytorch, Databricks.

**Data Visualizations** Tableau, MicroStrategy, R, Python

**Big Data**  Hive, PIG, Hadoop, MongoDB, MLFlow, Airflow, PySpark, Jupyter, Kubernetes, NoSQL databases, Kubeflow.

**AI/ML Ops** Docker, Git, CI/CD pipelines (Jenkins / GHA), Terraform, R-Shiny, HTML, JS

**Data Science Skills:** Predictive Modeling, Machine Learning, Image Processing, Analytics, Statistics, Data Sciences, Data Mining, Text Mining, NLTK, Exploratory Data Analysis, Forecasting, Optimization, Simulations, Mathematics, AIOps, MLops.

**Data Skills: Structured**/Unstructured Data, Images, Videos, Graph, Big Data Technology, Data Mining & Visualizations

**Management Skills:** P&L owner, Geo Program lead, Executive and client presentation, pre-sales and sales, planning, delegation, training, mentoring, project management, six sigma, etc

**PROFESSIONAL EXPERIENCE:**

**Verizon USA Lead Data Scientist - (Architect) December 2022 – Present**

This platform is to develop machine learning models and develop a Subject Matter Expertise and central AI/ML team for the Verizon business and product teams.

## Responsibilities:

* Use advanced statistical and machine learning algorithms to enhance the prediction accuracy of churn and disconnect behavior for Consumer and Business Segments.
* Build new Machine Learning (ML) models to predict Bad-Debt, fraud losses as well as fraud call volumes in support of Bill2Cash and Finance Operations strategy using API design and RESTful services alongside including source control (Git), CI/CD pipelines, testing, and documentation.
* Identify areas to optimize bad debt and recoveries by using customer behavioral data science as well as deeper understanding of existing processes.
* Guide and support decisions on Value Based Marketing & Digital Transformation by optimizing promotion and pricing strategy.
* Hands on active participation in analytical projects delivering improvements in predicting desired metrics to enable robust financial planning & decision guidance.
* Create POC models which help drive value for users, advertisers, and our company, focusing on using and adapting LLMs models for both text and multimodal challenges. Design proof of concepts and new features based on new LLM capabilities.
* Prepare presentation materials and formal documents for use with senior management to promote findings and drive science-based decision support recommendations.
* Develop and deploy both traditional statistical models and machine learning techniques and algorithms: Regression, clustering, neural networks, Anomaly Detection, Random Forest etc. to improve prediction and forecasts of desired metrics.

**American Express USA**

**Staff Data Scientist - (Architect) Aug 2020 – Nov 2022**

This Center of Excellence (Coe) team integrates with cross-functional teams to enable core products and initiatives to be rolled out in part of the digital initiatives to boost overall web journey analysis, user engagement, decision making, increase NAA, BBA and new product launches. Executive reporting, developing machine learning models, and leading an analytics team called Digital Experimentation and Insights (DE&I) for the digital web journey lifecycle.

## Responsibilities:

* Develop and maintain **Machine Learning** models and build financial models to prevent fraud transactions and have a smooth purchase cycle. ***(Patent Application*** *in* ***Process)***
* Build and maintain applications with requirements in building scalable Machine Learning powered applications using **PySpark, Cloud and Big data (AWS & Azure)** ecosystems and using data tools like **Adobe Analytics, Dynatrace, Content square, Mix panel, Adobe Audience manager.**
* Strong proficiency in Python and modern AI/LLM development. Experience with OpenAI's GPT models and API integration along with expertise in prompt engineering and LLM optimization
* Experience with Git version control and collaborative development
* Develop production-grade applications with unstructured data like Text and images. Build a **NLP workflow** that leverages Machine Learning to power its decisions based on text using Databricks.
* Employ **data analysis** using Machine Learning, predictive analytics, data mining, pattern recognition, data modeling, and statistical methodology to solve large scale business optimization problems and identify meaning derived from large data sets.
* Designed and revamped web **personalization (UX/UI)** experience for Web/App platforms from Home page to form pages (112+ enhancements with low to high KPI deliverables) through A/B testing, MVT, Market Basket Analysis, Association Rules/C-filters, Hybrid Models, Gap/ Competitor Analysis, etc. Notably, an overall **increase of 7% in US consumer cards NAA** (new accounts acquired) **and 13% increase in BBA** (business bill accounts)**.**
* Ensure data practices like data ethics, data compliances like **GDPR and CCPA,** etc. are followed and in line with the company data sharing principles in the application of innovative solutions and new.
* Build and perform **prototype analysis** and **build pipelines** iteratively to provide insights at scale.
* Develop web-based KPIs and metrics, advocating for changes where needed for both product development and business/sales activity.

**Verizon USA -Dallas Sr. Data Scientist June 2018 – July 2020**

This platform is to develop machine learning models and develop an integrated analytics platform called VOC (Voice of Customer) for the digital web purchase lifecycle to achieve the 30% digital channel mix.

## Responsibilities:

* Develop and maintain **Computer Vision**, **Machine Learning** and **fraud detection** models to ensure smooth purchase flow in digital sales. Maintaining Hadoop ecosystem pipelines and running automation scripts for Hive, Hbase using Scala and PySpark in AWS.
* Build pipelines for automated transformation of data into data marts in order to support reporting, predictive analytics, and targeting.
* Responsible for designing and building new Big Data systems for turning data into actionable insighs using python, java.
* Work on actionable insights through a wide variety of Big Data and Analytics tools to achieve results (i.e., Python, Hadoop, HIVE, Scala, Impala, and others).
* Analyze user purchase flow and optimize purchase cycle flow and streamline the sales purchase flow and increase channel mix based on the distribution patterns. The insights are put in a dashboard for the executive teams to monitor in tableau and PowerBI for business and marketing decision-making.
* **Achieved** reduction in recurring chats and calls up to **11% and 7%** respectively by filtering redundant inquiries and implementing analytics workflow for Natural language processing on the customer chats/calls and design pipeline to output chat insights from large volumes of chats data (Voice of Customer) using techniques like RNN, bag of words, topic modelling etc.

**L’Oréal USA Clark NJ**

**Data Scientist (Research Programmer) February 2017 – April-2018**

**Gloria L’Oréal: Intelligent Machine Learning Engine for Labs – Digital Incubator**

This platform is to develop Computer Vision and Machine learning models into an integrated product platform for all the scientists and end users that can use it for product development in the L'Oréal digital landscape.

## Responsibilities:

* Develop Image processing structures using OpenCV in Python and fed into a Convolution Neural Network and machine learning models to predict the Expert(dermatologist) scores of the facial and hair features before and after L’Oréal Cosmetic application using AWS Sage maker, Databricks.
* Created a web platform called SPRINT- simulation for hair color and formula detection using RGB/LAB values to detect closest formula using clustering and simulations for research purposes.
* Created an Acne detection algorithm and released it in app store to detect if the facial image has acne or not using OpenCV and CNN model using video cache using C++ and java.
* SWOT analysis of the different developed models and techniques of machine learning models using automated bash daily scripts.
* Build Ontologies, knowledge graphs, and NL/chat query parsers Experience with RNN / LSTM / Encoder-Decoder deep learning architectures.
* Analyze statistical inferences and predict customer and product behavioral dynamics.
* Develop Machine Learning models, validate and push to production. Prediction model-based projects over different expanses of projects to meet the needs of marketing & scientific teams in L’Oréal USA using REST API, AWS lambda, API gateway, Jenkins, AngularJS for webapp and with git versioning.
* Involves core research and developing scalable data science products for making key decisions. Following data science CRISP-DM protocols for project executions.
* Apply statistical and econometric models on datasets to measure results and outcomes, identify causal impact, attribution and predict future performance of users or products in an agile development

**Bowling Green State University- Decision optimization and Simulation Lab OH**

**Operations Research– Graduate Assistant January 2016 – August 2016**

* Optimization real-life **Case Study**, for Sports Analytics: [***https://goo.gl/GukH3x***](https://goo.gl/GukH3x)
* Negative Outcomes Risk Prediction Model: Analyzed Medicare resource utilization groups (RUGs) and Managed Care insurance claims data from a healthcare provider and predicted residents with negative margins using Regression and CART.
* Performed clustering analysis on historical patient-level data to classify them into payment (total expense per stay) groups and identified parameters expenditures and provided recommendations to drive reimbursements.

**Vizury – A Big Data & Data Science Ad-tech Firm Bangalore India**

**(Clients: Fintech in Asia, USA and Australia) July 2014 – July 2015**

**Data Scientist (Geo Program Manager - SEA/ANZ/USA)**

**Intel ibid: Programmatic Bidding Algorithm – Data Sciences:** Intel ibid is a Real Time Bidding (RTB) Engine developed by Data Science team of Vizury. This Engine predicts the worth of an Ad Impression in Real Time & sends bid accordingly. It was built to predict Click-through-Rate (CTR) & conversion rate (CVR) and is state of art Realtime advertising technology.

## Responsibilities:

* Product owner for the site performance KPI and tag integration for the campaign deployment. Monitor web KPI in Omniture and google analytics for clients in retargeting segments and click tracking.
* Develop and improve the performance of the prediction models using different regularization, hyperparameter tuning to minimize the cost function, using C++, Golan, and Python environments for scaling the product.
* Product enhancements with improved prediction metrics of the CTR models using feature engineering, outlier detection, Class Imbalance problems, and missing value detection.
* Leverage analytics to optimize real-time bidding strategies on major Ad-Exchanges e.g. Google, Facebook, for ad campaigns in order to maximize revenues and profit margins for the company.
* Implemented Logistic Regression, Linear Regression, Bagging, Boosting, Decision Trees, Clustering Algorithms, Support Vector Machines (SVM), optimization, and Stochastic Processes to predict CTR and CVR models (Click through rates and Conversion rates) and to increase overall sales
* Analyze customer user data using big data tools Hive, and Pig to understand click through patterns & buying behavior that helped increasing customer engagement & conversions for E-Commerce Clients.
* Work closely with product engineering team to propose, validate and iteratively build data-driven product features like recommendation system, Rules, Product proximity match etc.

**iQuanti, Inc -Digital Consulting & Analytics to leading clients**

**(Discover, American Express, Allstate) Data Analyst(web) India**

**Transaction Lifecycle Segmentation Analysis for Client: American Express**  **June 2012 – June 2014**

The project objective was to provide deep insights into web Transactions and inventory data that will empower clients to make better decisions in each of the stages across the lead to order lifecycle. Digital mix across SEO, Paid, social media, organic and Affiliate marketing.

## Responsibilities:

* Analyze the customer behavioral patterns using classification and segmentation techniques using Omniture and google analytics web data.
* Optimization Recommendations: Proactively identify and quantify business opportunities using web analytics data. Measure the success of various campaigns as well as A/B tests occurring on our sites.
* Calculate Customer Lifetime value (CLTV) and churn/attrition rates for different demographic data.
* Analyze data and identify financial trends in customer behavior, acquisition, retention, and customer movement through the website.
* Predicted the google search volume for keywords (keyword tool) based on several factors using linear regression
* Interacted with client-side Business Analysts and Technical Leads for requirement analysis and to define Business and Functional Specifications for a leading Finance Company
* Analyze structured data & identify top profile customers using Correlation & Regression Techniques.

**EDUCATION**

* Master of Science in **Analytics (Data Science)**, Bowling Green State University **(BGSU)**, Ohio,**2016**
* Big Data Analytics & Advanced Analytics degree, **Carnegie Mellon University,** PA, (INSOFE), **2015**
* Materials Engineering, **National institute of technology - Warangal (NITW) ,2013**

**CERTIFICATION**

* Google Analytics Certified **Google Inc, 2013-2014**
* Foundations for Data Science **Stanford University 2016**
* Big Data, Strategic Decisions, Analysis to Action  **Stanford University 2016**

**ACHIEVEMENTS:**

* Achieved phenomenal increase in segment based new accounts acquired and net billed business acquired using customer journey analysis and product enhancements across web and app.
* Patent application process in progress for identifying unique fraud and anomaly detection architecture at Amex.
* Instrument in cost reduction models for Stores and launching breakthrough product initiatives at Verizon with significant revenue impacts and reduction in customer attrition.
* Gloria platform at Loreal overall improved time and cost reduction in survey and research hours upto $1.2 million yearly.
* Increased Vizury global Revenue from $6 to $7.2 million MoM increase by enhanced model deployment.
* Recognized for exceptional performance in JFM-2015 Quarter in Vizury.
* Active contributor to Kaggle and GitHub community coding and open-source development.