SAURAV SINGLA

+91-995-879-3952 | sauravsingla08@gmail.com | linkedin.com/in/saurav-singla-5b412320/ | Gurgaon, India

Senior Data Scientist | Author | Instructor | Mentor

PROFILE SUMMARY

- Accomplished and high-performing Analytical professional with 15+ years of deep expertise in the application of analytics, business intelligence, machine learning, deep learning, natural language processing, and statistics in multiple industries and 4 years of consulting experience and 5 years of managing a team in the data science field. Creative problem solver with a unique mix of technical, business, and research proficiency that lends itself to developing key strategies and solutions with significant impact on revenue growth, profitability, cost base efficiency, and ROI. Working experience in machine learning, statistics, natural language processing, and deep learning with extensive use of Python, R, SQL & Tableau. Demonstrated excellence in teaching the learners to grasp data science knowledge more engagingly by providing courses on Udemy marketplace with more than 21K learners enrolled. Adept at mentoring data practitioners, enterprises, and startups to build high-performing teams, transform their business, and grow the industry. Author of "Machine Learning for Finance" for Asia's largest publisher called "BPB publications."

TECHNICAL SKILLS

- **Data Science skills** - Statistics, Machine Learning, Artificial Intelligence, Deep Learning, Ensemble Algorithm, Boosting, Bagging, AdaBoost, Gradient Boosting Machines - GBM, Random Forest, XGBoost, Light GBM, Neural Network, Deep Neural Network - DNN, Convolutional Neural Network - CNN, RNN, LSTM, GRU, Support Vector Machine - SVM, k-Nearest Neighbor - KNN, Self-Organizing Map, Decision Trees Algorithm, Classification and Regression Tree - CART, C4.5, C5.0, Chi-squared Automatic Interaction Detection - CHAID, Bayesian Algorithm, Naive Bayes, Gaussian Naive Bayes, Multinomial Naive Bayes, Clustering, k-Means, Hierarchical Clustering, Apriori algorithm, Eclat algorithm, Regression Analysis, Linear Regression, Logistic Regression, Categorical Regression Analysis, Dimensional Reduction Algorithm, Principal Component Analysis - PCA, Linear Discriminant Analysis – LDA, Factor Analysis, Recursive Partitioning, Generalized Linear Model - GLM, DBScan, Genetics Algorithm, Random Grid Search, Hyperparameter optimization
- **Python Libraries**- Pandas, matplotlib, scikit-learn, scipy, Numpy, TensorFlow, Keras, PyTorch, fast.ai
- **Tools & languages** – Python, R, SQL, NLP, h2o, Minitab, Palisade, @Risk 7.5
- **Visualization**- Tableau, ggplot in R, Altair, seaborn, mpl_toolkits, plotly, bokeh
- **Simulation**- Simul8, Insight Maker, Agent-based Modelling, System Dynamics, Discrete Event Simulation, Any Logic
- **Statistical Analysis**- Prescriptive Modelling, Predictive Modelling, Cognitive Modelling, Classification, Linear programming, Optimization Algorithms, Gaussian processes, Markov Models, A/B Testing, Multivariate Testing, parametric modeling, non-parametric modeling, Quantitative and Qualitative Research, Econometric Applications, Time Series forecasting, ARIMA, Prophet Modelling, Hypothesis testing, Chi-Square, t-test, ANOVA, Probabilistic Models, optimization techniques

WORK EXPERIENCE

Senior Data Scientist (Full-Time) at JBSA, London, UK | Oct 2017 - Sep 2020

- Directed and coordinated the capture, management, analysis, and insight of data as a key commercial asset
- Achieved driving a project to completion and collaborated with other team members to ensure work is done correctly and quickly
- Facilitated a culture with an emphasis on exploration, intelligence, and capabilities instead of tools
- Coached a team of 9 members that focuses on solving real problems with scalable, fit for purpose applications of AI and advanced analytics capabilities
- Planned, developed, and applied cutting-edge machine learning systems and statistical modeling to extract insight from vast amounts of structured and unstructured data
- Addressed a broad range of internal and external data sources and alternative datasets
- Guided in gathering and analyzing information, formulating, and testing hypotheses, and developing and communicating recommendations

Significant Accomplishments:

- Implemented ML solution to the customer retargeting platform of the client resulted in 100% improvement in campaign performance within 6 months of implementation and 2 x improvements in CTR performance and campaign cost reduced up to 20% because of automation
- Executed ML solution to the digital channel platform of the client for brands resulted in the enabled discovery of customer attributes through the mining of unstructured data as against no information before. This resulted in intelligent marketing campaigns with higher ROI compared to baseline. 10% increase in customer engagement from digital interactions in 2 months of testing with different marketing campaigns
- Implemented ML solution for credit scoring platform resulted to lift of 60-70% within first three deciles for 4 customer segments in out of time validation and removed the subjectivity and improved TAT for a loan application to within few minutes thereby improving overall customer experience
- Implemented optimization solution helped the bank with daily forecasts for deposits and withdrawals to more efficiently distribute its money across ATMs, RCMs and branches. Forecasts were accurate to the tune of +/-15% on average. It reduced its transportation cost and saved interest on borrowed money by optimizing cash deposits to branches, ATM's and RCM's once the forecast for deposits and withdrawals was complete
- Implemented Propensity algorithm for the leading private life insurance firm resulted to tailored marketing campaigns across modes of marketing and incremental Revenue of USD 100,000 in 3 months and lower cost of Marketing Campaigns
- Implemented for Marketing Spend Optimization for pharmaceutical enterprise in which modeling has given result decrease investment & hold revenue steady in which marketing mix can be optimized to generate previous year revenue while decreasing investment by -44%. Hold investment steady & increase revenue in which using the last year's marketing budget, revenue can be increased by 30 million euros in this year. Maximize profitable investment in which using maximized investment scenario, an additional 45 million euros in revenue can be captured with an additional investment of 15.4 million euros
- Implemented discrete event simulation models that describe complex medical processes to understand and evaluate the strategic positions of treatments and demonstrate the value of interventions beyond efficacy, to elucidate the economic and operational consequences associated with the given healthcare intervention

Tools & Technologies: Machine Learning, Deep Learning, Python, Tableau, H2O, XGBoost, LightGBM, Catboost, Stack Ensemble, NLP, Glove, Word2Vec, Named Entity Recognition, Simul8, Anylogic, Minitab, Palisade, Discrete Simulation, Agent-Based Modelling Simulation

Consultant Data Scientist (Full-Time) at Valiance Solutions, Noida, India | Nov 2016 - Aug 2017
- Directed in Developing algorithms to efficiently go through large datasets
- Recommended code for Econometric, Statistical, Machine Learning, Deep Learning models
- Developed meaningful data visualizations from the analytics to enable effective communication
- Directed in research to develop prototypes and proof of concepts
- Enlisted end-to-end implementation of client needs for econometric analyses to generate forecasts, optimizations, and simulators
- Associated with the client to comprehend, design, and execute data science-driven solutions
- Led a team of 4 members to deliver reliable analytical services and actionable insights to clients by leveraging proper tracking technologies, data mining techniques, and visualization tools

Significant Accomplishments:
- Built an ML-powered predictive customer churn system helped telecom client reduce churn from 40% to 25%
- USD 1.5 million per annum gross cost savings to the client from fraud mitigation via launching ML-powered Fraud Detection Engine and almost 40% additional cost savings as the reliance on third-party collection agencies decreased significantly
- Risk Detection platform lead the client by reducing credit default from 6% to 2%, and with credit default coming under control, the client was able to increase loan disbursements by 10% to USD 55 million per year

Tools & Technologies: Machine Learning, Deep Learning, Natural Language Processing, Python, Tableau, SQL, NLP, GBM

Data Scientist/Data Scientist Manager (Full-Time) at Osram, A Siemens Company, Gurgaon, India | Nov 2008 - Sep 2014/Dec 2014 - Oct 2016
- Developed Post Campaign Analysis, rate/price sensitivity, cross-sell and up-sell identification, Customer Segmentation, Customer Churn, Customer Profiling, Propensity Modelling, Multi-Channel Attribute Models Customer lifetime value, and Customer Engagement modeling
- Collaborated the building and managing the team roadmap and backlog, and delivered actionable data insights to sales, marketing, logistics, supply chain, finance, and senior leadership
- Managed a cross-functional team of 15, ranging from entry-level to analysts, and closely collaborated with business development, data analysis, operations, and marketing teams
- Developed strategic roadmap for analytic capabilities to drive business and incremental revenue
- Validated project conclusions and present results to stakeholders using advanced data visualization techniques
- Enlisted code to develop automated processes include activities such as predictive modeling via supervised or unsupervised machine learning and statistical techniques
- Analyzed the huge amount of transactional and behavioral data, and implemented probabilistic, machine learning techniques to build models to improve targeting and revenue
- Identified and integrated new datasets that can be leveraged through product capabilities. Worked closely with the engineering team to strategies and executed the development of data products
- Executed analytical experiments to help solve various problems, making a true impact across various departments
- Devised and utilized algorithms and models to mine data and error analysis to improve models, and cleaned and validated data for uniformity and accuracy
- Analyzed data for trends and patterns and interpreted data with a clear objective in mind
- Communicated analytics solutions to stakeholders and implement improvements as needed to operational systems

Significant Accomplishments:
- Improved data mining processes, resulted in a 20% decrease in time needed to infer insights from customer data used to develop marketing strategies
- Implemented ML-enabled Marketing Mix resulted in improved ROI on the marketing campaign by 6%
- Implemented Hidden Markov Model for incremental measurement and multi-channel marketing attribution modeling to help drive more efficient and effective channel optimization which resulted in improved customer conversion by 11%
- Created visual dashboards which helped in creating better marketing strategies, leading to 14% growth in customer acquisition
- Implemented innovative use of internal and external datasets and analysis strategies led the company to make marketing adjustments to increase efficiency by 12%
- Identified strong promotion opportunities using machine learning which led to increasing sales of long-tail products by over 30%, and forecasting and visual dashboard to analyze the effectiveness of promotions
- Executed ML-enabled Customer Feedback & Sentiment Analysis System using NLP led to 85% reduction in efforts by and a 10x improvement in turnaround time of the customer service team and assisted in improved customer retention by 6%
- Improved demand forecasting led to a 15% enhancement in fill rate, 11% of reduction in backorders of retail partners, and up to 85% boost in the forecasting accuracy for new products
- Applied Intelligent Risk Assessment solution bring about a 15% improvement in credit limit extension for channel partners and reduced delay in payment by a whopping 10%
- Identified best product mix on transaction data and implemented Market Basket Analysis for identifying best product mix using Apriori, Eclat algorithm, FP Growth, Association Rule led to 8% improvement in sales of new and distorted products
- Increased Customer LTV through the development of customer contact strategies informed by rigorous testing and customer analytics in areas of retention, frequency, and spend and by building Frequency/Recency Matrix, Gamma-Gamma Model, BGF Models
- Utilized Cohort Analysis, Survival Analysis, Cox Proportional Hazard Model, Kaplan Meier, Pareto, Stobachoff index, Regression modeling to develop benchmarks for Customer Retention, Customer Loyalty, Customer Profitability
- Implemented customer segmentation to restructure the delivery service using PCA, K-Means clustering, Gaussian Mixture Model clustering
- Optimized marketing campaign channel-wise using Greedy Algorithm, IBM Decision Optimization CPLEX Modelling and predicted the likelihood of customers buying a recommended product channel-wise
- Designed Customer Propensity score using Naïve Bayes classifier for recommending the best channel for contacting each prospective customer which helped in improving customer acquisition
- A/B testing on price increase effect on projected revenue and the conversion rate
- Applied random forest, support vector machine, gradient boosting to train the transactional dataset and help identify the likelihood of late delivery, returns

Tools & Technologies: Machine Learning, Deep Learning, Natural Language Processing (NLP), Python, R, Tableau, SQL, Minitab, C5.0, Random Forest, SVM, Neural Network, NLP, GBM, Linear & Logistic Regression, Decision Tree, CART, CHAID, K-means, Gaussian Mixture Model, KNN, Naïve Bayes, PCA, Apriori, ECLAT algorithm, FP Growth, Association Rule, Cohort Analysis, Survival Analysis, Cox Proportional Hazard Model, Kaplan Meier, Pareto, Stobachoff Index, optimization, objective function, constraint, linear programming, Greedy, CPLEX

Senior Data Analyst (Full-Time) at Harvey Norman, Melbourne, Australia | Aug 2003 - Jun 2007
- Built custom segmentation to target and deliver a deeper understanding of customers
- Performed analytics across internal and external data sets to understand the wider utility
- Provided analytical support to the business on ad-hoc analytical requests, enabling the business to investigate problems, identify new opportunities, guide planning processes, and influence decision making
- Produced regular reports to ensure the business is well informed and able to construct intelligent decisions on results of all direct and digital campaigns
- Initiated data checks and investigation of data anomalies
- Ensured accuracy and consistency in both data manipulations and conclusions
- Created presentations through findings for internal & external stakeholders
- Successfully developed new statistical predictive models to forecast demand, including pricing models
- Efficiently developed metrics, KPIs, analytics, and dashboards to provide marketing measurement and drive marketing performance through deep business insight
- Optimized statistical efficiency & data quality via developing and implementing data collection systems and other strategies
  Tools & Technologies: Excel, Minitab, SQL, ARIMA, Exponential Smoothing methods, Holt Winter, Moving Average Methods, Frequency/Recency Matrix, Gamma-Gamma Model, BGF Models

BOOK, COURSE, ARTICLES, PUBLICATIONS (With Weblinks)
- **Book**
  o Machine Learning for Finance BPB Publications
- **Course**
  o Data Analysis for Business and Finance Udemy
  o Complete Outlier Detection Algorithms A-Z: In Data Science Udemy
- **Article**
  o How important is Explainable AI? Towards Data Science
  o Big Data in Financial Services Towards Data Science
  o The Importance of Hypothesis Testing Hackernoon
  o Insiders' Guide to Generative and Discriminative Machine Learning Models KDnuggets
  o Batch Normalization in Deep Neural Networks Kdnuggets
  o Machine Learning to Predict Credit Risk in Lending Industry AI Times Journal
  o Why is the Central Limit Theorem Important to Data Scientists? Towards Data Science
  o For what reason Probability Important to Machine Learning? Data Science Central
  o Introduction to Dropout to regularize Deep Neural Network Data Science Central
- **Academic Research Paper**
  o Demand and Capacity Modelling in Healthcare using Discrete Event Simulation
  o Comparative Study of the Deep Learning Neural Networks on the basis of the Human Activity Recognition
  o Comparative Analysis of Transformer based Pre-Trained NLP Models
  o Multi-Class Sentiment Classification using Machine Learning and Deep Learning Techniques
  o Integer Optimization for Dream 11 Cricket Team Selection

EDUCATION
- University of Westminster, Master's Degree with Distinction in Business Intelligence & Analytics (2018)
- Institute of Management Technology (IMT), Executive Program with Distinction in Marketing (2016)
- Kurukshetra University, Bachelor's Degree in Commerce (1999)