

CLIENT'S NAME

Hawthorne, California 12345 • 123.456.7890 • www.infotech@gmail.com

QUALIFICATIONS PROFILE

Highly analytical and process-oriented senior data scientist, leveraging advanced expertise and a Ph. Ph.D.-level understanding of data modeling frameworks. Expert in creating complex models to drive insights-driven solutions and software, adeptly tackling real-world business challenges. Proficient in managing diverse data models and metrics to ensure integrity, accessibility, and usability across the organization. Key leadership experiences include guiding teams in developing end-to-end solutions that integrate cutting-edge data models with scalable software applications. Known for establishing rapport across all organizational levels and backgrounds in fast-paced environments.

AREAS OF EXPERTISE

| | |
|---|---|
| <u>Data Manipulation and Analysis</u> | Data Mining Data Wrangling Data Modelling Data Visualization Exploratory Data Analysis (EDA) Feature Engineering Business Intelligence (BI) Data Science Model Building and Troubleshooting |
| <u>Machine Learning</u> | Probabilistic Programming with Bayesian Inference Dimensionality Reduction Support Vector Machine (SVM) LightGBM Decision Tree Random Forest K-nearest Neighbors Algorithm (KNN) Anomaly Detection Neural Networks Clustering – PCA, K-Means Naive Bayes Natural Language Processing (NLP) Linear/Polynomial Regression Multivariate Logistic Regression Time Series Analysis Algorithm Supervised Unsupervised Generative Models Ensemble Learning Hyperparameter Tuning Bias-Variance Tradeoff |
| <u>Statistical/Mathematical Methodologies</u> | Dispersion Skewness Correlation Variance Significance Bayesian Inference Monte Carlo Simulation Optimization Multivariable Calculus Relational Algebra Advanced Linear Algebra Abstract/Modern Algebra Computational Algebra |
| <u>Business Acumen</u> | Project Oversight and Leadership Team Supervision and Training Risk Assessment and Mitigation Continuous Process Improvement System Implementation Technical Communication |
| <u>Languages</u> | English Gujarati Hindi Marathi |

TECHNICAL ACUMEN

Big Data/Cloud: AWS (EC2, S3, Dynamo DB, Redshift) | Azure (Kubernetes Service, DevOps, Databricks)
Snowflake | MLFlow | HPC Cluster | Grid Computing | Docker | VMWare

Management/Collaboration Frameworks: GIT | Confluence | JIRA | ADO Board | Monday.com

Languages: Python | SQL | C++ | Fortran | Perl | R | Linux | DAX | Shell | MATLAB | HTML

Libraries: TensorFlow | PyTorch | LightGBM | FastAPI | Pandas | Numpy | SciPy | Sci-Kit Learn | Matplotlib
Seaborn | Plotly | ggplot | Theano | Jupyter

Tools: Pycharm | VScode Powershell | Jupyter Notebooks | Tableau | Microsoft Power BI | Postman | Flask
Microsoft Office Suite (Word, Excel, and PowerPoint)

Databases: SQL Server | MySQL | PostgreSQL | Dynamo DB | SQLite

PROFESSIONAL EXPERIENCE

COMPANY NAME | CITY, STATE

Senior Data Scientist, Pricing: Dec 2021–Present

- Assume responsibility for the development of advanced pricing models for sports and live entertainment events utilizing deep learning, generative, and probabilistic programming techniques, including Bayesian inference
- Expertly manage the Price Model Team, as well as train junior data scientists, interns, and analysts, and create key performance indicators (KPIs) to monitor and enhance model/product performance
- Generate complex SQL queries to extract, transform, and load (ETL) data, for conducting data analysis and designing model features
- Execute application programming interface (API) to auto-seed and broadcast tickets for live events in real-time, encompassing development, deployment, maintenance, and enhancement of the end-to-end broadcasting solution

Key Highlights

- ✓ Succeeded at creating and implementing pricing models for 120 leading sports franchise teams across all US major leagues (MLB, NBA, NFL, NHL, and MLS), as well as for NCAA college sports (football, basketball, and soccer) teams, diverse touring events (Monster Jam, Supercross, Disney on Ice, etc.) and music concerts; with each model generating revenue from \$100K to up to \$1M annually
- ✓ Engineered a broadcasting algorithm for ticket selection, leveraging inventory data and sales trajectory
- ✓ Drove the automation of the entire process of building price models from the ground up, significantly reducing the time required from over 2 days to less than 2 hours, utilizing Azure framework and integrating MLflow

COMPANY NAME | CITY, STATE

Associate Data Scientist, Pricing: Oct 2020–Nov 2021

- Directed the end-to-end discount engine, leading a team of professionals while establishing strategic relationships with the Product Development Team and the User Analytics Team utilizing Git, Jira, and Confluence
- Demonstrated proficiency in handling queries against Amazon Redshift involving nested arrays and utilizing clickstream analytics to track, assess, and improve user experience
- Expertly managed end-to-end code deployment process within the Git framework, including regular maintenance and enhancement
- Facilitated the development of proof of concept (POC) by providing necessary data and insights into its financial impacts

CLIENT'S NAME

Hawthorne, California 12345 ▪ 123.456.7890 ▪ www.infotech@gmail.com

Key Highlights

- ✓ Led development and implementation of an analytics-driven pricing and discount solution across 70 distribution centers in the US, along with optimization, automation, and scaling of its end-to-end process utilizing the AWS framework
- ✓ Formulated pricing-discount implementation logic while optimizing its effectiveness across multiple dimensions
- ✓ Successfully created, monitored, and improved the discount framework's overall performance using Tableau and Power BI

COMPANY NAME | CITY, STATE

Data Scientist, Performance-HSE and IT Department: Sep 2018–Sep 2020

- Held accountability for developing analytical tools, encompassing machine learning algorithms, process automation, and BI reporting using Power BI
- Championed in utilizing various techniques, such as feature engineering, model selection, data training, optimization, and visualization for both structured and unstructured data
- Collaborated with the Executive Committee as a developer of complex data models and visualization utilizing advanced Power Query and DAX for reporting,
- Trained team members to utilize analytical tools, including machine learning algorithms for generating oil-rig performance rankings, advanced DAX queries, and web scraping with Power BI framework
- Generated reports and data models for various functions, including HSE, Drilling Performance, Operations Management, Integration, and Human Resource teams

Key Highlights

- ✓ Designed and introduced an innovative and effective data model employing unsupervised clustering algorithms, which is now implemented company-wide for benchmarking performance KPIs across a fleet of 47 offshore oil-rigs spanning multiple regions and continents
- ✓ Conceived and established the concept of using a recommender system, while guiding a team of analysts through the implementation of NLP techniques for analyzing daily rig activity logs
- ✓ Constructed a temperature-based equipment failure model utilizing a supervised learning algorithm, achieving an accuracy of 80%
- ✓ Steered key efforts to create an optimized safety incidence model using a supervised learning algorithm, attaining an accuracy of 85% in predicting safety-related incidents across 42 offshore rigs, which resulted in significant downtime reduction for the rigs
- ✓ Optimized BI reporting by automating the post-processing of real-time data using Databricks within the Azure framework
- ✓ Facilitated data governance initiative by developing foundational concepts using descriptive data analytics
- ✓ Implemented machine learning models within Azure for scheduled execution

COMPANY NAME | CITY, STATE

Research Fellow/Scientist: Oct 2014–Aug 2018

- Supervised a team of researchers in the investigation of the biological response of cancer cells to heavy-ion beams

Key Highlights

- ✓ Carried out design, planning, scripting, and testing of a new treatment planning software, calibrating its Monte Carlo-based dose engine using simulations
- ✓ Conceptualized and established analysis techniques to boost dose calculation accuracy by up to 80% using Monte Carlo simulations combined with machine learning techniques, which generated and analyzed over 10 GB of data with the support of high-performance computing (HPC) clusters
- ✓ Contributed to the investigation to translate the use of the space dosimeter Timepix, which was created in collaboration with CERN, into a clinically relevant detection system

EDUCATION

Doctor of Philosophy (PhD) in Nuclear Physics ▪ *University Name | City, State*

PROFESSIONAL DEVELOPMENT

Machine Learning, Jun 2017 | *Coursera*
Advanced Power BI and Power Query, Jan 2019 | *Microsoft*
Tableau, Dec 2020 | *Udemy*

PUBLICATIONS AND ACOLADES

Springer Book Publication as a part of Springer Thesis Award in Physics, 2015
24 Journal Publications with 4 First Author Publications in Reputed Journals
Notre Dame Physics Research and Dissertation Award, 2015
The Cornelius P. and Cynthia Browne Fellowship, 2013
Sigma XI Grant Award, 2009–2011
SET (India) qualified, 2006