dheeraj.chand@gmail.com
+1 2025507110
GitHub: github.com/dheerajchand

Dheeraj Chand

Austin, TX (30.2672°N, 97.7431°W)

PROFESSIONAL SUMMARY

Senior Data Analysis & Analytics Professional with 21 years of expertise in statistical analysis, machine learning, and big data processing. Expert in advanced statistical modeling, predictive analytics, and data visualization with experience serving major brands, organizations, and political candidates. Deep specialization in fraud detection, entity resolution, and pattern analysis across multi-terabyte datasets.

KEY ACHIEVEMENTS AND IMPACT

Advanced Analytics & Machine Learning

- Developed parametric spatial clustering algorithms for voter analysis, resulting in 88% improved targeting efficacy
- Developed meta-analytical techniques to resolve ambiguous dimensions, resulting in discovery of 170% more viable targets
- Created fraud detection systems for campaign finance data analysis across multi-terabyte datasets
- Modernized legacy ETL processes by implementing dbt and PySpark workflows, reducing processing time by 57%

Big Data & Data Engineering

- Built multi-tenant data warehouse and data lake using Snowflake, dbt, and AWS for longitudinal analysis
- Transformed small data team into big data engineering team using Hadoop Clusters and Hive on AWS
- Developed advanced data pipelines for machine learning applications enhancing consumer segmentation
- Introduced version control and Agile methodologies, improving project delivery timelines by 40%

Statistical Analysis & Research

- Built the first collaborative and multi-actor contributed poll of polls used by the Democratic Party
- Developed RACSO platform for pollsters to fully administer research, analyzing bids from 1,200 vendors
- Pioneered the integration of advanced mapping techniques into standard reports, including choropleths and hexagonal grid maps
- Led multi-million dollar research projects involving sensitive consumer data with privacy compliance

For a more detailed, full description of my experience, please visit my LinkedIn and Personal Site.

CORE COMPETENCIES

Statistical Analysis & Machine Learning: *Advanced Statistical Modeling* (Regression, Clustering, Segmentation, Machine Learning); *Predictive Analytics* (Time Series Analysis, Forecasting, Risk Modeling); *Data Mining* (Pattern Recognition, Anomaly Detection, Entity Resolution); *Machine Learning* (SciKit-Learn, TensorFlow, PySpark, Spark MLlib)

Big Data & Data Engineering: *Big Data Processing* (Apache Spark, PySpark, Hadoop, Snowflake, dbt); *Data Warehousing* (Multi-tenant Architecture, ETL/ELT Pipelines); *Cloud Platforms* (AWS (EC2, RDS, S3), Google Cloud Platform, Microsoft Azure); *Data Pipeline Optimization* (Performance Tuning, Scalability, Monitoring)

Data Visualization & Reporting: *Data Visualization* (Tableau, PowerBI, Seaborn, Matplotlib, d3.js); *Geospatial Analysis* (ArcGIS, Quantum GIS, GRASS, OSGeo, PostGIS); *Interactive Dashboards* (Real-time Analytics, Custom Reporting); *Business Intelligence* (KPI Development, Performance Metrics)

PROFESSIONAL EXPERIENCE

Siege Analytics, Washington, DC | PARTNER 2005 – Present

Advanced Data Analysis and Statistical Modeling

- Developed parametric spatial clustering algorithms for voter analysis, resulting in 88% improved targeting efficacy
- Developed meta-analytical techniques to resolve ambiguous dimensions, resulting in discovery of 170% more viable targets
- Created fraud detection systems for campaign finance data analysis across multi-terabyte datasets
- Built multi-tenant data infrastructure supporting concurrent access from diverse client organizations

Helm/Murmuration, Washington, DC | DATA PRODUCTS MANAGER June 2021 – May 2023

Big Data Analytics and Machine Learning

- Built multi-tenant data warehouse and data lake using Snowflake, dbt, and AWS for longitudinal analysis across attitudinal, behavioral, demographic, economic and geographical dimensions
- Modernized legacy ETL processes by implementing dbt and PySpark workflows, reducing processing time by 57%
- Developed advanced data pipelines for machine learning applications that enhanced consumer segmentation and predictive modeling capabilities
- Managed teams of seven to eleven engineers, designers, analysts, and external stakeholders using Agile methodologies and modern DevOps practices

GSD&M;, Austin, TX | ANALYTICS SUPERVISOR November 2019 – June 2020

Data Science and Business Intelligence

- Transformed the small data team into a big data engineering team, going from working on small datasets on laptops to using Hadoop Clusters and Hive on AWS
- Implemented spatial analysis and consumer segmentation methodologies that revealed new insights about existing customers
- Introduced version control and Agile methodologies to the data team, improving project delivery timelines by 40%
- Managed three analysts, mentoring them in advanced market research techniques and data analysis

Mautinoa Technologies, Washington, DC | SOFTWARE ENGINEER August 2016 – February 2018

Data Science and Econometric Modeling

- Developed SimCrisis, a GeoDjango web application using Python, PostgreSQL/PostGIS, and NetLogo for multi-agent modeling and econometric simulations of crisis economies
- Built modular application using Python, Django, and GRASS accepting rules extensions for ethnic strife, different crisis types, supply failures, and disaster scenarios
- Conceived and built application using Python, Pandas, and Jupyter to predict how crisis economies respond to different humanitarian interventions
- Applied agent-based modeling, statistical analysis and machine learning systems for humanitarian impact assessment