**Amit Kheterpal** akheter1@asu.edu Portfolio
Tempe, Arizona (open to relocation) (623) 209-4223 LinkedIn

## EDUCATION

**W. P. Carey School of Business at Arizona State University** August 2022 – May 2023
*Master of Science in Business Analytics (MSBA)*, GPA: **4.0/4.0** *Tempe, AZ*

- *Relevant Coursework:* Database Management, Supply Chain Analytics, Data-Driven Quality Management, Machine Learning, Data Mining, Natural Language Processing, Deep Learning, Analytical Decision Modeling, Marketing Analytics.

**Guru Gobind Singh Indraprastha University** August 2014 – May 2018
*Bachelor of Technology, Electrical and Electronics*, First Division with **Distinction** *Delhi, India*

## PROFESSIONAL EXPERIENCE

**Data Science Intern** December 2022 – May 2023
*Avnet Inc.* *Phoenix, Arizona*

- Created a Python-driven supply chain Demand forecasting system, merging advanced time-series models (**Prophet, Auto ARIMA, Exponential Smoothing**) for **30 %** lower forecast errors enabling precise prediction for the upcoming 12 months.
- Devised interactive Python dashboards using the **Streamlit** Library, amplifying inventory management by effectively conveying performance metrics and time-series model insights, significantly ameliorating operational efficiency.

**Statistical Analyst** November 2020 – July 2022
*ComScore Technologies Ltd.* *Pune, India*

- Led **SAS** to **Python** migration with **Spark** integration for weighting and sample selection processes. Automated the process via **Jenkins**, saving **$40,000** annually and substantial efficiency boost by minimizing reliance on SAS.
- Resolved panel health issues, including data inconsistencies, missing values and outliers utilizing **SAS**, **Python**, and **PostgreSQL** by analyzing and problem-solving, reducing data anomalies by **50%** and improving overall data accuracy and reliability.
- Tracked and analyzed market-specific demographic trends and market penetration in the US and EMEA regions, leveraging **Excel** and **Tableau** to create insightful visualizations that contributed to a **10%** increase in data-driven decision accuracy.
- Spearheaded the automation and implementation of **Python-based QA** processes for sample selection and weight generation, resulting in a notable **25%** increase in daily production efficiency.
- Boosted query performance by **30%** through enhanced **SQL** scripts migration from **Greenplum** to **Snowflake** data warehouse.

**Big Data Engineer** July 2018 – November 2020
*Infosys Ltd.* *Pune, India*

- Reduced data storage costs by **25%** for a **Fortune 500 Fintech** client through effective project management and technical expertise, leading the migration of **100 petabytes** of large on-premises datasets to AWS cloud servers.
- Implemented an **AWS**-based data solution, utilizing **Airflow** for workflow orchestration. Streamlined data ingestion from **S3**, processing with **Glue** and **PySpark**, and storage of outputs in **Redshift**, achieving a significant **40%** increase in processing speed.
- Executed event-driven data processing using **AWS Lambda** triggers, leading to an impressive 25% decrease in manual labor. Additionally, optimized data processing performance by 10% through **AWS Glue** job run-time reduction via Spark code refactoring.
- Collaborated closely with cross-functional teams and external partners to align analytical projects with business strategies, efficiently gathering and documenting requirements that resulted in a **20%** reduction in project turnaround time.
- Curated structured and unstructured data such as **JSON** and **Parquet** files, and seamlessly integrated it into the Redshift data lake.

## PROJECT EXPERIENCE

- **Predicting Hotel Booking Cancellations:** Designed a potent ensemble-based classification system using **LR, KNN, DT, RF**, and **ANN** models for a **97%** accuracy. Empowered proactive customer churn prediction and targeted retention strategies.
- **Real-Time Parking Space Detection:** Developed and fine-tuned a custom **YOLOv5** model for parking space detection, leveraging transfer learning techniques to achieve high accuracy and robust performance.
- **Sentiment Analysis and Topic Modeling:** Employed **NLP** modules such as Vader Lexicon and Latent Dirichlet Allocation to perform sentiment analysis and topic modeling on smart watch brand reviews from Reddit API.
- **Lean Six Sigma Project:** Applied Lean Six Sigma to reduce SLA time by 15% and save $1.2M. Utilized **Tableau** for impactful data visualizations, revealing bottlenecks and driving actionable insights.

## SKILLS & ACTIVITIES

- **Languages:** Python (Pandas, NumPy, Scikit-Learn, Plotly, NLTK, Spacy, TensorFlow, Keras), C++, SQL, HIVE QL
- **Analytical Tools:** AWS, SAS, Microsoft SSIS, SSRS, Power BI, Tableau, Advanced Excel, SPSS, Sqoop, Shell Scripting, ETL
- **Technical Skills:** Statistical Analytics, Big data, Machine Learning, Deep Learning, NLP, Data Visualization, Decision Modelling.
- **Statistical Skills:** Hypothesis testing, A/B Testing, Time Series Forecasting, Regression Analysis, Design of Experiment
- **Key Skills:** Storytelling, Analytical Reasoning, Problem Solving, Project Management.
- **Certifications:** Global Agile Developer, Hacker Rank SQL Advance, NLP with Python Bootcamp (Udemy.com)
- Lifetime Member of **Beta Gamma Sigma**, The International Business Honor Society