

# Shivani Dandir

(520)-270-3879 | shivanidandir17@gmail.com | [LinkedIn](#) | GHC 2022

## SUMMARY

Data Science professional with 2+ years of experience in designing and modeling Machine Learning Algorithms. Demonstrated a history of independent problem-solving, data-driven thinking, and making critical decisions during challenges. Excellent collaborative skills and ability to clearly articulate perspective with a desire to learn.

## EDUCATION

### Master of Science in Management Information Systems (MIS)

Eller College of Management, University of Arizona, Tucson, AZ, USA

Dec 2022

GPA 3.7

Coursework – Machine Learning, Data Mining, Business Intelligence, Data Warehouse, Big Data, Statistics

### Bachelor of Engineering in Information Technology (IT) MITCOE, Pune, India

GPA 3.75

May 2019

## SKILLS

- **Frameworks and Libraries:** Pandas, NumPy, Scikit-Learn, TensorFlow, Keras, NLTK, NLP, Spark SQL, PySpark
- **Models:** CNN, LSTM, ARIMA, Logistic regression, KNN, Naïve Bayes, SVM, Decision Tree, Supervised & Unsupervised
- **Languages and Databases:** Python, SQL, Java, R, Spark, Scala, Hive, NoSQL, Oracle, Hadoop, MySQL
- **Tools and Services:** Tableau, Git, PowerBI, RedShift, AWS EC2, AWS S3, AWS DynamoDB, Glue, Lambda, Airflow
- **Data Warehousing:** Performance tuning, ETL, Data modeling, Data pipeline, Data standards and policies, Reporting and analysis, Data Integration, Data Analysis, Validation, Optimization, Technical documents, Design documentation
- **Professional Skills:** Excellent communication, Problem-solving, Analytical, Critical thinking, Creative approach

## WORK EXPERIENCE

### Data Scientist | University of Arizona, Tucson, USA

Aug 2021– Present

- Created 20 reporting dashboards in Tableau for administrative data to provide insights to university business teams including the senior leadership, contributed a unique perspective and creative approach to solving problems
- Designed & Implemented Usage & Performance dashboard from scratch resulting in 50% relocation of cluster resources
- Developed a self-serving ETL pipeline - transforms data, cleans it, and feeds it to the dashboard after data analysis
- Built a reputation for delivering ahead of schedule and for effectively integrating new technologies

### Business Intelligence Engineer Intern | Amazon, Seattle, USA

May 2022 – Aug 2022

- Created 5 enhanced metrics to evaluate customer engagement for the Alexa app using SQL queries & Tableau
- Translated business needs into data, created weekly/monthly business reviews for Alexa App using ETL in Redshift
- Optimized existing SQL codes using data aggregation & temp tables which saved 70% of resources per execution
- Worked with product & engineering team to build efficient & scalable data pipelines on a large dataset (500TB+)
- Analyzed critical data problems to optimize reports and installed data quality validations to report data accurately
- Dealt with a degree of ambiguity and developed a quick proof of concepts; designed 15 interactive Tableau dashboards

### Business Consultant | Samsung, Tucson, USA

Jan 2022 – May 2022

- Researched key areas to find business solutions, used a hypothesis-driven approach after problem breakdown
- Collaborated with stakeholders & executives to frame a story around business cases to present it to C-suite
- Designed a product recommendation to increase 20% revenue by customer & industry research on focused group

### Data Engineer | Deloitte, Mumbai, India

Dec 2019 – Jul 2021

- Spearheaded 2 modules, data collection & correspondence, independently in the highest revenue-generating public sector agile project, collaborated with 150+ practitioners, generated \$200k+ in revenue, and managed 3 interns
- Analyzed key issues provided data fixes in SQL, impacted 1M+ cases in production; worked in a cross-functional team
- Reduced team efforts by up to 20 hours/week by developing JIRA dashboards to enable time and usage tracking
- Devised Batch Report Automation Macro using Excel & MySQL for performance optimization, saving 120+ hours/week
- Continued to learn and develop machine learning and data science skills through the Deloitte Machine learning team

## TECHNICAL PROJECTS

### Data Scientist | Converting audio to text, Research project

Python, Machine Learning

- Evaluated machine learning models (DeepSpeech, coqui) to transcribe audio of children with Down Syndrome
- Scripted extraction and splitting code in Python and Linux to feed it to the deep learning models for transformation
- Processed 100GB+ real-time audio data in python for noise cancellation, & documented the results for further research

### Data Scientist | Predicting Customer Spending Scores

R, Machine Learning

- Modeled 5 Machine Learning algorithms – Logistic Regression, KNN, Naïve Bayes, Decision Trees, Neural Networks
- Chose the neural networks as the confusion matrix resulted in the lowest false-negative rate on 100,000+ rows of data
- Extracted and cleaned the data using R, represented confusion matrix, and designed correlation plot.

### Data Scientist | Stock Market Prediction

Python, Time Series Forecasting

- Developed Deep Learning & Time series model on 2 years of data ARIMA & LSTM to predict prices 10 mins in future
- Evaluated performance using mean squared error; visualized through time-series decomposition

### Data Scientist | Identifying racist/sexist Tweets

Python, NLP, ETL, Deep Learning

- Cleaned data by parsing & tokenizing 50,000+ tweets, visualized in WordCloud, used natural language processing
- Built bag of words & TF-IDF models with Naïve Bayes algorithm to determine the nature of tweets

### Data Architect | Enterprise Database Management

Oracle | JAVA | AWS | MS Visio

- Created a Data Model and architecture for a food delivery app, after gathering requirements, and mapping relations
- Devised data manipulation language (DML) in SQL for retrieving data on the front end using Java, in AWS S3
- Designed relational database, ER diagram, normalized the database, performed ETL, created procedures in PLSQL