Ashish Sharma

+1-214-986-0577 | hello@sharma-ashish.com | sharma-ashish.com | LinkedIn | GitHub

EDUCATION

The University of Texas at Dallas **May 2020**
Master of Science, Data Science GPA 4.0
Jaypee University of Information Technology, India **June 2014**
Bachelor of Technology, Electronics and Communication Engineering GPA 3.88

SKILLS AND COMPETENCIES

Analytic Tools and Languages: Python, PySpark, R, Advanced Microsoft Excel, GitHub, JIRA

Database and Visualization: SQL, Oracle DB, Power BI, Tableau, Matplotlib

Data Science and Analytics: A/B Testing, Machine Learning, Deep Learning, Microsoft Azure, Spark, Hadoop, pandas, Scikit-learn, Regression, Random Forest, XG Boost, LGBM, Clustering, K-means, Deep Neural Networks, TensorFlow, Embedding, keras, Azure Databricks, Natural Language Processing (NLP), Statistical Modeling

PROFESSIONAL EXPERIENCE

Data Scientist – Cotiviti, Dallas, TX **June 2020 – Present**

- **Predictive Modeling**: Improved the payment accuracy solution, expediting the claim review process by 25%, automating 150k claims per month that resulted in $10.5 million annual savings
- **Machine Learning:** Created and deployed ML models using LGBM algorithm and a deep learning model utilizing the deep cross-network architecture on a cloud platform using 100 million claims for with over 1300+ features
- **BI Dashboards**: Derived the optimal threshold for binary classification models by doing a cost-benefit analysis and created an interactive Tableau dashboard to track key performance indicators
- **A/B Testing:** Led A/B testing to evaluate the model performance in production, comparing automation vs. manual results to maintain >95% precision for each model

Data Scientist Co-op – Cotiviti, Atlanta, GA **August 2019 – April 2020**

- **Distributed Computing:** Created data pipelines in Azure Data Factory using PySpark, wrangling 12 million records in Spark to identify linkages between providers and payers
- **Business Intelligence:** Developed dynamic Power BI dashboards that fed into a downstream application that helped the payers understand the overall health of their provider network based on 14 different metrics

Data Science Intern – Divergence.ai, Dallas, TX **June 2019 – August 2019**

- **NLP**: Created a POC for a chatbot with language understanding capability using Microsoft bot framework for a food packaging client, and built NLP models for topic modeling using Gensim to identify frequently asked questions

Data Analyst - Cognizant Technology Solutions, India **January 2015 – February 2018**

- **Recommendations**: Increased customer conversion rate by 17% by providing insights to the marketing team
- **Data Wrangling**: Extracted historical auto insurance data of 3.8 million customers using SQL, transformed and cleansed it using Python to conform to business requirements
- **Customer Segmentation**: Created a logistic regression model using 30 features to predict the purchase likelihood of a customer with 86% accuracy, created meaningful customer segments to come up with a targeting strategy

HACKATHONS AND ACADEMIC PROJECTS

Sign Language Interpreter using CNN - UNT 2019 Hackathon **– Winner [GitHub | YouTube]**

- Implemented a sign language interpreter to help people with hearing impairment with their daily communication needs using Deep Learning
- Built a CNN model with 95% accuracy on 105k greyscale images of 44 gestures in American Sign Language in Python

Los Angeles Restaurant Data Analysis - Informs Analytics Challenge **– Winner [GitHub]**

- Built an XG Boost classifier to predict the health grade of the restaurant in the 88 cities of the Log Angeles County using only its name, address, and zip code with a precision of 73%