SHIVANG ARYA

arya.shivang@gmail.com

github.com/NotShivang | shivangarya.com | linkedin.com/in/shivangarya

EDUCATION

The University of Texas at Austin	Master of Science in **Information Technology and Management**	May 2020
	Overall GPA: 3.88/4	
Narsee Monjee Institute of Management Studies, Mumbai	Master of Business Administration in **Technology Management** & Bachelor of Technology in **Computer Engineering**	May 2019

EXPERIENCE

Overhaul Group – *Data Integrations Team;* Austin, Texas   Sep 2020 – Present

- Managerial/Leadership Milestones
  - Working with international clients to design and implement data integration and reporting solutions
  - Actively involved in interview processes to expand the team
  - Managed a team of interns responsible for creating a data lake that ingests over 1 million events per day
  - Made key process improvements in requirement gathering and prioritization that reduced the average integration time > 50%

- Technical Milestones
  - Used prefect.io for data orchestration and deployment of ETL / ELT pipelines
  - Built an MVP for an intelligence project to process, query and visualize over 500 million rows of data
  - Worked with a large variety of IoT devices and API's – REST, SOAP and GraphQL
  - Interacted with MongoDB, DynamoDB and PostgreSQL databases for either data loading or data reporting requirements
  - Introduced the use of Python for ETL scripts reducing the code size of an integration 1000 times and making it 3 times faster
  - Assisted in the development of an Infrastructure as Code solution using AWS CloudFormation for easy and efficient Command Line Interface deployment of Python code
  - Technical documentation – responsible for documenting the codebase

Claira Technologies – *Data Science Intern, Natural Language Processing;* Remote   Jul 2020 – Aug 2020

- Designed and developed an algorithm that takes in the user's skills and creates competency statements from them
- Algorithm is based on Google's BERT (Bidirectional Encoder Representations from Transformers) language model
- Deployed this on Sagemaker and exposed an endpoint to be accessed via Lambda and API Gateway

Dell Technologies – *Research Scientist Intern, Capstone Project;* Austin, Texas   Jan 2020 – May 2020

- Implemented an imputation model for time series data based on deep convolutional generative adversarial networks
- Achieved 98% cosine similarity on average when comparing synthetic data from the algorithm with real values
- Created a Docker application to deploy this solution

HDFC Capital Advisors – *Management Trainee, Investments;* New Delhi, India   May 2018 – Sep 2018

- Calculated an average rate of return of about 16%, with the help of market analysis and financial modeling, on private equity and debt investments in real estate projects worth over $25M

Bharti Airtel Limited – *Intern, Information Technology Department;* Gurgaon, India   Jun 2016 – Jul 2016

- Developed an automated system for formatting and uploading data dumps, increasing efficiency of the process by 67%

PROJECTS

- **Business Data Science – FIFA Manager:** Finding the optimal replacement for a football (soccer) club's outgoing player using similarity scores and recommender systems. Moreover, suggesting the best playing 11 given a budget and a formation.
- **User Generated Content Analytics – YouTube Clickbait:** Topic modeling and sentiment analysis on over 2500 YouTube thumbnails and titles, obtained by web scraping (Selenium) and labeled by Google Cloud Vision API calls, to predict if the video will go viral. Found that a thumbnail depicting travel and having a negative title increases engagement by 20%.
- **Data Management – PointsGiving:** A webapp designed to implement a monthly points-based reward system along with admin functionalities such as dashboards to generate reports that summarize all the activity and giving out bonus rewards.

- **Personal Website – shivangarya.com:** My personal website hosted on GitHub pages and made using a combination of Jekyll, HTML, CSS, and JavaScript. Incorporates Continuous Integration, Continuous Delivery and Continuous Deployment.
- **Marketing Analytics:** Finding the customer lifetime value, time series forecast for sales using Winters' method and an optimum marketing budget allocation strategy for Olist e-commerce website.
- **Big Data and Distributed Programming:** Music recommendation system using Spark based on a user's song history and cosine similarity with other users; Movie recommendation system using item based and user based collaborative filtering.
- **Emerging Technologies (IoT):** Programmed a raspberry pi to receive step count from a microbit and weather data from an android application. This data is used to tune a machine learning model to obtain daily goals which are displayed on the application
- **Advanced Programming and App Development – Event Management System:** Android application and a website to create, join, modify, and delete events with different admin and user functionalities.

ADDITIONAL INFORMATION

Publication: Does 'Femvertising' matter? A study in the Indian context **–** *International Journal of Research and Analytical Reviews*

REFERENCES

Dr. Jason Vertrees
CTO, Stronghold Investment Management
jason.vertrees@strongholdim.com

Dr. Xinchi He
Senior Staff Engineer, Autonomy
hexinchi@gmail.com