|  |
| --- |
| **SAI SUDEEP REDDY ANNA** |
| Technical skills**Languages & Tools**Python, R, Tableau, SAS, Spark, Power BI, Google Analytics, AWS, HTML**Databases**Snowflake, MySQL, MongoDB, Cassandra, PostgreSQL, Redshift, DynamoDB**Learning**Data Science Stack – Pandas, NumPy, Scikit-Learn, SciPy, Matplotlib, FlaskOpenCV, TensorFlow, XGBoost,**Data Modelling**ER-Modelling, Logical Modelling**Version Control**GitHub, BitbucketEducation**Master of Science** June 2020**Data Analytics & Engineering** **3.6/4**Northeastern University, Boston**Bachelor of Engineering May** 2017**Computer Science Engineering 7.2/10**Satyabhama University, Chennai. | EXPERIENCE**Data Analyst** Jan 2020 – July 2020**Ninja Shark, Boston*** Developed dashboards on **Tableau** to share insights and recommendations on customer characteristics and customer behavior
* Created **data pipelines** and optimized the workflow for Form Recognition, Segmentation and **OCR** to extract data.
* Used **SQL** store the entire extracted data and built **computer vision models** to detect potholes and anomalies in steel manufacturing in real-time.

**Junior Business Data Analyst** July 2017 – Dec 2018**Kiara, Hyderabad, India*** Successfully **preprocessed large amounts of data using Python** and interpreted data to draw insights into customer behavioral pattern.
* Providing technical expertise on data storage structures, data mining, and data cleansing and used **ETL** process to standardize data.
* Utilized **R**, Tableau and **SQL** to perform predictive modeling and Exploratory Analysis on more than 100,000 customer sales databases.
* Proposed solutions which **increased the conversion rate from 1.9% to 10.35%** using **Multivariate Time-series Analysis** and **A/B testing.**

**Data Analyst Research Intern** May 2016 – Dec 2016**Indian institute of Management (IIM), Lucknow*** Worked with Data Engineers and Data analysts to web scrape data, clean and filter data.
* Used **SQL and Excel** to ensure data quality and integrity. Identified and eliminated duplicates and inaccurate data records.
* Created a database from scratch and Implemented **indexes for efficient storage using SQL.**
* **Analyzed and Processed data using SAS, Tableau and R.**
* Developed data model, pipelines and reported **dashboards** using **Quick Sight** as per requirements and used **JIRA** agile communication.
 |
| WEB <https://www.linkedin.com/in/sudeep-reddy-855535122/>Contact (857)-763-7475 saisudeepreddyanna@gmail.comStrengths**Collaborator****Delivering Results****Self-Motivated and Curious****Problem-Solving** | Academic Projects**Red Crow Start-up data** **Python | Tableau | SAS | Machine learning | R | June 2020**Created a dataset by **Web scraping.** Detected and Evaluated highest costing details of the invested start-ups. visualized and predicted the Highest paying insurance using **Tableau, SAS and R**. Used Predictive models like **Regression, Decision Tree** to identify low and high-risk factors**Uber-Lyft Fair Prediction****R | Knime | Hyperparameter Tuning | SQL | Regression | Tableau |Dec 2019**Performed **Data cleaning, merging** of data sets visualization on various variables. Implemented methods like **regularization** and **regression** to identifythe important factor variables that effect the fair of the cab. Built Predictive models and found the **top 5 influencing factors** that affect the **price** and used **XGBOOST** and obtained and **accuracy** of 93%**Airbnb Dataset-New York City****Python | Tableau| Google Cloud | Data Visualization | Data Mining | Regression | Sept 2019**Performed **Exploratory Data analysis** to find the density of different kind of Airbnb and **feature engineering** for amenities and predicted price of Airbnb in New York city with final model **Random Forests** with an **accuracy** of 96.94% and **RMSE of 0.65.** |