Chitra Lekha Sura

chitra22lekhas@gmail.com | +1 (425)-375-1786 | Redmond, Washington | linkedin | portfolio

EDUCATION

Stevens Institute of Technology	Master's of Science (MS)	Computer Science	Sep 2022– Dec 2023
Osmania University	Bachelor's(BE)	Computer Science	July 2018– May 2022

EXPERIENCE

Intone Networks| Data Analyst — March 2024 – present

- Deployed a **Tableau dashboard** integrating **data models** including **star schemas** enabling efficient **data querying.**
- Designed **SQL**- data models for optimized data source management, resulting in **12% improvement** in data-driven decisions.
- Automated **data aggregation** and analysis using **PostgresQL** and **Python**, streamlining cross-departmental data evaluation.
- Developed **Machine learning** models and **time series forecasting** to forecast customer growth rate.

Verzeo EdTech Pvt. Ltd. | Machine Learning Intern — March 2020 – April 2020

- **Data Cleaning** and genre analysis using **Python** and **Pandas** to identify high-rated gaming categories on the App Store.
- **Visualized** price-rating trends with **Matplotlib** and **Seaborn**, highlighting **correlations** between price and ratings.
- Applied **one-hot encoding** for genre **classification** and used **regression** models to **predict** user rating trends.
- Utilized **decision trees** and **random forests** to predict app success based on genre, price, and user ratings.

Association of Computing Machinery (ACM), MVSR | Joint-Secretary > Vice-chair — September 2020 – May 2022

- Represented ACM at several events, conferences showcasing leadership in organizing and promoting technical sessions.
- Conducted over 20 IT/Analytics workshops and seminars on campus.
- Led workshops on machine learning and **data analysis** fundamentals, and topics like **regression models, decision trees,** and **clustering** techniques**.**

END NOW Foundation | Cyber Rakshak - cybercrime prevention ambassador — February 2020 – June 2022

- Conducted awareness sessions on cybercrime for students to educate on online safety and **digital responsibility.**
- Promoted **cybersecurity** best practices to prevent crimes by using techniques like **reverse image searching** and completed **Cisco Cybersecurity Essentials** and **CyberOps Associate**.

PROJECTS

Data Analysis of COVID-19(2020&2021) world dataset:

- Conducted **data exploration** on **MySQL** correcting over 500 data inconsistencies, revealing a 2.5% global infection rate, and highlighting a 50% reduction in new cases post-vaccination.
- Developed interactive **Tableau** dashboards and **calculated fields** to visualize trends in infection rates, vaccination coverage.
- Utilized **Python** for **correlation analysis** revealing significant correlations(0.75 between mobility and cases)
- Analyzed dataset using **R (knitr, ggplot2, tidyr, dplyr)** to find trends and determine cases based on demography and time.

Image Captioning

- Integrated **CNNs** and **RNNs** in **deep** learning for **image captioning**, achieving 92% accuracy in generating contextually-rich descriptions. Utilized an **Inception V3 module**, outperforming **ResNet-50** by 4% in **feature extraction.**
- Optimized **feature extraction** to reduce image disturbances during **CNN encoding**, resulting in a 3x reduction in model training time. Leveraged **TensorFlow, Pandas,** and other advanced libraries for implementation.

Data Analysis of "Cyclistic - bike share"

- Executed **analysis** on **normalized, standardized and preprocessed** Cyclistic bike-share app's data set of 5 million records, utilizing **Python libraries (Pandas, NumPy)** to eliminate 99% of **outliers** and ensure data integrity.
- Performed **geospatial analysis** by mapping 100,000+ latitude and longitude points using **Folium**, on real-time maps to visualize ride patterns.
- Generated visualizations, histograms and **time series graphs(SARIMA model),** to forecast trends and insights that drove data-informed decisions

Predictive Analysis of Job Change Intentions

- Executed strategic data analysis on 10,000+ candidates' intentions on their job, utilizing machine learning models like **K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest,** and **Gradient Boosting**.
- Performed a comparative study among the model performances, analyzing metrics like **precision, recall accuracy, f1-score**, with **Random Forest** achieving an accuracy of 92%.
- Applied **Gradient Boosting** improving prediction precision by 15%, optimizing post-training candidate retention strategies.

SKILLS AND TOOLS

- Programming Languages: Python, Java, SQL, R
- Machine Learning libraries: TensorFlow, Keras, PyTorch, Scikit-Learn, pandas, PySpark, NLTK, XGBoost
- Databases: MySQL, PostgreSQL, NoSQL, Tableau, PowerBI, T-SQL, Hadoop, Excel, PowerQuery
- Certifications: Data Analytics & visualisation job simulation, Google Data Analytics, Cisco Cybersecurity essentials.