

SIDDHARTH CHAUHAN



DATA SCIENTIST | MACHINE LEARNING ENGINEER



PERSONAL DETAILS



Mobile: +91 6301765124



Hyderabad, India



siddharthc96@gmail.com



<https://siddharthchauhan.github.io/>



<https://www.linkedin.com/in/siddharth1992/>



<https://github.com/siddharthchauhan>



<https://www.kaggle.com/siddharthpchauhan>



<https://medium.com/@siddharthc96>



PERSONAL PROFILE

Data Scientist with 3+ years of broad-based experience in building data-intensive applications, overcoming complex architectural and scalability issues in diverse industries. Proficient in predictive modeling, data processing and data mining algorithms as well as scripting languages like Python. Capable of creating, developing, testing and deploying highly adaptive diverse services to translate business and functional qualifications into substantial deliverables.



WORK EXPERIENCE

DATA SCIENTIST



MAXISIT INC. | FEB 2018 - PRESENT

Clinical Data Anomaly Detection System

- Developed an end-to-end ML pipeline from data preprocessing of clinical data to deployment using unsupervised machine learning algorithms like Clustering-Based Local Outlier, K Nearest Neighbors, Principal Component Analysis and Isolation Forest. (Pandas, Pycaret)

Exploratory Data Analysis Application

- Developed an Exploratory Data Analysis (EDA) tool which helped clinical researcher and Statisticians in exploring data and reducing time on performing data analysis of raw data coming from CROs. (Pandas, Matplotlib, Plotly)

Recommendation System of Clinical Reports

- Developed a recommendation engine using content-based filtering technique which recommends clinical and operational data reports depending on similar reports created in application by user. (Pandas, Sklearn)

Conversational AI Using Amazon Lex

- Developed a chatbot using Amazon-Lex and AWS lambda converting natural language to SQL which will be used by statistician to find out meaningful insight from clinical data. (Amazon Lex, AWS Lambda)

Predictive Forecasting of Subject Data

- Built Predictive forecasting system using Statistical model and Deep Neural Network which will be used by CRO in forecasting enrolment of subjects at site for given clinical study making sure they are satisfying their goal. (Pandas, Scipy, Tensorflow)



AREAS OF EXPERTISE

- Business Understanding
- Data collection
- Data Preparation
- Exploratory Data Analysis
- Data Modelling
- Model Evaluation
- Model Deployment



SOFT SKILLS

- Excellent Communicator
- The ability to analyse complex technical information
- Detail oriented
- Excellent problem solver
- Good at working with large teams as well as individually



CERTIFICATIONS

- IBM Data Science Specialization
- IBM Applied Data Science
- Create Interactive Dashboards with Streamlit and Python
- Rasa Developer Certification



TOOLS & TECHNOLOGIES



EDUCATION HISTORY

GEORGE WASHINGTON UNIVERSITY

Masters in Computer Science, 2016

NIRMA UNIVERSITY

Bachelors in Information Technology, 2013