SANKARSHAN ACHARYA

sankarshansacharya@gmail.com

26238 Wyndemere Court
Escondido, CA 92026

(760) 443-0901 *(Cell)*

HIGHLIGHTS

- Strong analytical and problem-solving skills. Able to lead projects with rock-solid commitment to customer satisfaction
- Experience using Python, JavaScript and Tableau for visualization and reporting
- Significant experience with relational and non-relational databases. Significant experience with MapReduce, Spark and Hadoop for the analysis and storage of large datasets
- Solid background in statistics and modeling methodology. Able to work with complex data and drive discovery. Knowledge of supervised and unsupervised learning
- Good programming knowledge including Python, R, Java and MATLAB

EDUCATION

- University of California, San Diego
  - Masters' in Data Science and Engineering
    - Capstone project involved predicting real estate rental rate using Airbnb data. Model yield: R^2 of 85% and MAE of 50. Commercial potential being explored
  - B.S. Applied Mathematics
  - B.S. Chemical Engineering
- Coursera
  - Deeplearning.ai
  - UCSD Java Programming: Big Data
  - Johns Hopkins University: Data Science
- Udemy: Excel Analytics

SKILLS

- Python: pandas, seaborn, geopandas, matplotlib, numpy, scikit-learn, statsmodels, nltk, TensorFlow, Keras, PyTorch
- R: dplyr, diffEq, ReacTran, rsconnent, quanteda, RWeka, tm
- PostgreSQL, Oracle, Map/Reduce, Hadoop, Spark, Hadoop, AWS
- Java, Ruby on RailsBridge, Ember, Shell scripting, Node-RED
- Microsoft Office: Word, Excel PowerPoint
- Visualization with Tableau, Power BI
- Data Files: JSON, XML, CSV, Graph Data
- Knowledge of ETL, data cleaning
- Proficient in Software Testing Life Cycle (STLC) methodology
- Statistical modeling, Deep learning, Ensemble Methods, Time series analysis, NLP

EXPERIENCE

Transcore, Data Analyst *June 2019-August 2019*

- Worked on Tableau and SSRS based reporting
  - SSRS based reporting used procedures written in SQL Server and RDLs (Report Definition Language) to take the tables created by the procedures and format them in a way the customers want

- Tableau based reporting is similar in that it uses the same procedures (and maybe tables) and uses them to create visualizations which are put together in a way as to present a Gestalt (whole picture) for the report

Initiative towards venture funding *February 2019– present*

- Product to provide insights into the financial health of the companies for investors
  - Financial data (e.g. income statements, balance sheets, and cash flow statements) for publicly traded companies is graphically displayed to enable decisions
  - Prototyping reports using R, Power BI, Apache Tomcat web server to create a PoC.
  - Browser-based app to receive requests from the user for the reports is under construction. Apache Tomcat webserver is being used for the purpose

Contract Programmer and Modeler *April 2015 – present*

- Assisting clients with modeling and programming
  - Ormat tech is a renewable energy company. Using a combination of ARIMA and regression, predicted abnormal operating conditions. Great initial results. Work in progress
  - Tru Halo is an online market place for health, wellness and beauty products. Working on a product recommendation engine for use by customers
  - Mercury Health aims to predict falls in patient populations. Used a combination of convolutional neural network and feedforward neural networks to predict likelihood of fall. Great initial results. Work in progress
- Training professionals in mathematics (statistics, calculus), programming and engineering

Verisk, Data Science Intern *June 2017 – September 2017*

A diverse analytics company focused on risk across many domains

- Developed statistics-based analytics to predict length of stay in a healthcare facility. Enabled at least 50% improvement in the accuracy of predictions
  - Employed linear regression and Gradient boosting techniques. Developed the modeling approach to include many sub-models in lieu of a single global model
- Developed initial models for assessing energy disruption risk
  - Conducted a survey of the NLP methods for detecting fake news stories related to energy disruption. Created an initial model using PyTorch with an accuracy of over 90%. Model needs further refinement

Betasoft, Development Trainee *April 2015 – Sept 2015*

Software development organization

- Enabled criteria-based testing of created models for public use on a robust platform
  - Created Test Cases and designed test plans for smoke, sanity, functional, system, integration and unit testing. Tested Android and IOS mobile devices and applications with Selendroid, Appium, and Robotium in conjunction with Selenium

Alphaszenszor, Contract Modeler *August 2014 – Feb 2015*

Sensor chip design and production start up

- Worked on the temperature control system for the sensor chip
  - Derived the model based on principles of electrical circuits and heat transfer. Solved the derived equations using Mathematica Dsolve and Manipulate
  - The evolved analysis represented significant improvement over published literature and was used in seeking series B funding for the company

ACTIVITIES

- Member, Tau Beta Pi, National Engineering Honors Society January 2014 – present