Rochish G S K

Data Scientist

hyderabad, India
+918106833102 · raptortreats@gmail.com

GitHub, Linked in, kaggle

Date of birth

23 Apr 1997

Skills

- JavaScript
- Git
- Predictive Modeling
- Machine Learning
- Python
- Deep Learning
- Tensorflow
- pytorch
- Microsoft Power BI
- Tableau
- Numpy
- Pandas
- Jupyter Notebook
- Seaborn
- Matplotlib

Profile

Accomplished Data Scientist with a passion for delivering valuable data through analytical functions and data retrieval methods. Committed to helping companies advance by helping them to develop strategic plans based on predictive modeling and findings. Bringing forth a proven track record of analyzing complex data sets and serving as a strong advisor.

Education

Bachelor of Technology, MLRIT, Hyderabad

June 2014 — May 2018

Post Graduation, IIIT-B, Bengaluru

September 2022 — September 2023

Employment History

Data Scientist, PharmCADD, Hyderabad

September 2021 — Present

- Developed Machine learning models for Herg predictions.
- Developed Deep learning model using LSTM-GRU for RNA small molecules predictions
- Performed various data visualizations for understanding the data of RNA small molecules.
- Developed consensus model using XG boost, Random forest and DNN for better model and predictions for Triple Re-uptake data.
- Generated visualizations for codon frequency in RNA small molecules.

Machine Learning Expert, Sai Life Sciences Ltd, Hyderabad

April 2021 — September 2021

- Performed Data analytics and data visualization for few drug data.
- Developed various machine learning and deep learning models using Python RDkit and jupyter notebook on various data to predict half-life, solubility, Caco2, bioavailability, clearance hepoticyte, Ld50 Zhu, Cyp2d6 and etc.
- Performed similarity maps on low and high half life molecular data and predicted the accuracy of the data as 88%.
- Developed Deep learning models for drug discovery.
- Worked on attention transformer for molecular prediction.

Hobbies

singing, playing guitar, trekking, Badminton

Languages

English

Hindi

Telugu

Machine Learning Intern, Sai Life Sciences Ltd, Hyderabad

December 2020 — March 2021

- Developed a machine learning model using python, RDkit and jupyter notebook for blood brain barrier permeability.
- Performing data analytics for adverse event analysis

Projects

RNA small molecule prediction

- A deep learning model for predictions of RNA small molecules.
- This is an LSTM – GRU model which is a sequence to sequence model .
- Later this deep learning model was used for all the RNA sequences .

Triple Re-uptake models

- A machine learning model which predicts if the molecule is active or inactive
- This is a classification model used for predicting active or inactive.
- Triple re-uptake has three sets of data SERT, NET and DAT.
- Used XG boost classifier model for these predictions.
- Later these models were used for predicting of active and inactive molecules.

Descriptor Visualizations

- Created various types of plots for distribution of molecular descriptors.
- Used matplotlib and seaborn libraries for these visualizations.
- Used Scatterplot, Barplot, Boxplot, violinplot, pie chart etc.
- Written an own snippet for creating scatterplots for distribution of one target to all.

HERG toxicity model

- A machine learning model which predicts if the molecule is herg toxic or not.
- This is a classification model used for predicting toxic or intoxic.
- Used Deep Neural Networks for these predictions.
- Later these models were used for predicting of toxic and intoxic molecules.
- Performed similarity maps where it predicts the toxicity of every part of a molecule.

Attention Transformer

- Designed a single neural network architecture that performs competitively across a range of molecule property prediction.
- Developed this deep learning model for prediction of properties of molecules
- This deep learning model was further used for the machine learning models developed.

Deep Docking

- A deep learning model for augmentation of structure based drug discovery.
- Developed this model to predict the docking score for molecules.
- Libraries used are tensorflow, keras, rdkit and sci-kit learn.

Half- life Prediction

- Performed data analytics and visualization for human and rat half-life data.

- Predicted machine learning models and performed similarity maps for low and high half-life data.

Blood Brain Barrier Permeability

- Developed a machine learning model to predict Blood Brain Barrier Permeability using various classification algorithms and observed best model by Extreme Gradient Boosting (XG boost) Algorithm.
- Performed data pre-processing on the dataset gathered from various publications.
- Analyzed data using rdkit, numpy, pandas and seaborn

Chemistry Python Libraries

- Understanding python packages for Chemistry.
- Machine Learning and Deep Learning (rdkit, deepchem, and reinvent) Libraries.

Solubility Prediction tool

- supported in development of a machine learning models to predict and assess compound solubility and decision making using multiple regression and classification algorithms and further validated using externaldataset
- Performed data pre-processing on the dataset gathered from various publications.
- Analysed data using rdkit, numpy, pandas and seaborn.

Dual Steganography

- Encrypted a message into an image file and embedded it into a videofile.
- Measured Peak Signal to Noise Ratio between original host and generated stenographic files.

Design of Reconfigurable Antenna for Cognitive Radio Advanced Design System

- Designed frequency-reconfigurable antenna system for cognitive radio to overcome spectrum utilization problems.
- Fabricated the resulted Antenna using Antenna Prototyping Machine

Courses

Data Science Specialization, IIT Roorke

September 2021 — September 2022

Data Science and Machine Learning, Udemy

November 2020 — December 2020

python, Solo learn

January 2020 — March 2020

Complete SQL Bootcamp,, Udemy

June 2020 — July 2020