Chamila C. Dharmawardhana, PhD

ccdxz8@mail.umkc.edu +1(509)201-7496

LinkedIn Google scholar GitHub

Authorized to work in USA (GreenCard holder)

I am a dedicated **data scientist** and a **computational physicist** with 14+ years of diverse experience in numerical simulation, scripting, data analysis and on HPC and other UNIX based systems. Excellent team player with aptitude for planning and conducting interdisciplinary projects.

Skills & experience

Languages:	Python, SQL, C++, Java, Fortran, BASH, Mathematica
Frameworks:	Pandas, numpy, scipy, SciKit-learn, TensorFlow, Keras, beautifulsoup, GitHub
DataBases:	Amazon Redshift, Athena, S3
HPCC:	Supercomputers (NERSC, XSEDE, ALCF). PBS, LSF and SLURM batch queuing systems
Management:	Effectively designed interdisciplinary research projects with diverse research collaborations

Projects

Smart campaign scheduler via classification models

- Building appropriate read patterns and ETL for feature engineering and analysis to generate appropriate training dataset. Data cleaning handling Nulls. of numerical and categorical features
- Assessing statistical significance of features and EDA to understand business value and predictability of target
- Training and hyper-parameter tuning XGBoost (**Sci-Kit learn**)
- Deploying model(s), handling errors, optimizing memory consumption and code refactoring to maximize readability

Facial emotion recognition via deep learning (GitHub)

- A facial expression recognition model to serve as a personal journal of emotion
- Base CNN model developed with FER2013 Kaggle dataset (**Tensorflow, Keras**)
- Images scraped from Google images (**Selenium, bs4**) were used for transfer learning
- **Django app** build to service the CNN model and provide a journal app

Pneumonia prediction from X-ray images with CNN techniques (GitHub)

- CNN model developed with Kaggle dataset (**Tensorflow, Keras**)
- Hyper parameter tuning to increase accuracy.
- Transfer learning from VGG19 model to increase accuracy

Customer churn prediction of TelCo data via supervised learning (GitHub)

- CNN model developed with Kaggle dataset (**SciKit-Learn, StatsModels**)
- Customer churn prediction with **logistic regression, Decision Tree, Random Forest, SVC, verity of boosting** algorithms.
- Hyper parameter tuning and network optimization toincrease model accuracy and recall

Employment

Associate Data Scientist, Promoboxx.com 2020 – present

- Developed a new A/B testing framework for clients to test marketing campaigns
- Developed recommendation system of digital campaigns for retailers for effective marketing
- Natural language processing (NLP) tooling to identify consumer sentiments for digital marketing campaigns

Senior Research Associate, Illinois Institute of Technology 2019 - 2020

- Development FFT based algorithms for protein docking (**numpy, pandas**) to evaluate binding affinity of protein complexes
- Automated structure retrieval using **protein data bank API** and building scripts to streamline docking simulations of large databases

Post Doctoral Fellow, Georgetown University (2 Publications) 2017 - 2019

- Development of accurate and efficient computational methods for protein simulations
- Building python scripts to streamline FF parameterization and simulation
- Developed and implemented potential function for the SSMP model

Post Doctoral Fellow, University Colorado Boulder (3 Publications) 2016 - 2017

- Understanding corrosion of high temperature super-alloys in 4D: experimental/theoretical collaboration
- Developed python tools for force fields building, data handling and visualization of results
- Automated calculations and visualizations via python and bash scripting

Education

2020 **Certificate, Data Science, Flatiron School of Computing**
2015 **PhD, Physics, University of Missouri - Kansas City**
2008 **MS, Physics, Central Michigan University**
2005 **BS, Physics/Mathematics, University of Colombo, Sri Lanka**