Pradeep Singh

https://pdeep.xyz | pdeepsingh094@gmail.com | +1 619 538 4719

Last Updated on 9th February 2020

EDUCATION

SAN DIEGO STATE UNIVERSITY
M.S. in Computational Data Science
Exp. May 2020 | San Diego, CA
GPA: 3.7/ 4.0

UNIVERSITY OF MUMBAI
B.S. in Electrical Engg.
Aug 2011 - June 2015 | India

COURSEWORK

MACHINE LEARNING
Deep Learning
Machine Learning
Numerical Optimization
Parallel Computing

DATA SCIENCE
Statistical Inference
Big Data: Tools and Methods
Techniques of Data Science
Foundations of Data Science

SKILLS

PROGRAMMING
Python • Matlab • C/C++ (basic)

MACHINE LEARNING
TensorFlow • Keras • Scikit-learn

DATA & VISUALIZATION
SQL • Pandas • Matplotlib
Numpy • Seaborn • Scipy

PLATFORMS
macOS • Linux • Git • Docker

GITHUB PROJECTS

Deep Clean : Python package to clean multimodal data (image and text).
Quick ML : Read ML notes on the fly.
Data Science 101 : How-to-Guide for data science tools and packages.

ACHIEVEMENTS

2017-19 Tuition Scholarship
2015 Research Fellowship at RRI
2014 U.G. Research Fellowship
2014 Editor-in-chief at KC-Xplore
2013 Vice-chairperson - IEEE Branch

EXPERIENCE

DASSAULT SYSTÈMES | Deep Learning Research Intern | Project Report
June - Dec 2019 | Boston, MA

- Owned, researched and proposed a novel end-to-end deep learning pipeline for accelerating CFD simulations using super resolution techniques.
- Developed SRCFD, a generalized and platform agnostic framework, based on CNN and GCN, that can super resolve coarse-to-fine simulations.
- Extended SRCFD to unstructured meshes using graph convolution network.
- Developed different architectures in SRCFD using Tensorflow.
- Built and generated in-house dataset for Dassault Systemes.
- Built custom docker images and dockerfiles to dockerize applications and tools.
- Laid foundation for long-term and short-term future research in ML for CFD.

RESEARCH

COMPUTATIONAL SCIENCE RESEARCH CENTER | Research Assistant
Aug 2017 - May 2019 | San Diego, CA

- MS Thesis: Models for propagating Facilitation in the Insect Visual System.
- Built mathematical models of network of neurons and astrocyte cells in Matlab.
- Studied facilitation mechanism, calcium dynamics in biological cells and analyzed gigabytes of data in Matlab.
- Carried out comprehensive parametric study of our models and thus characterizing facilitation in visual system.
- Published our research in an international conference and journal (in-progress).

PROJECTS

IMAGE RECOGNITION USING CNN | Computer Vision (code | report)

- Built 5 different Convolutional Neural Networks using Keras and TensorFlow to classify 70,000 fashion images into 10 labels.

BAYESIAN OPTIMIZATION | Numerical Optimization (code | report)

- Built an hyperparameter optimization algorithm that finds better hyperparameters for machine learning models in smaller number of steps than random or grid search.

PARALLELIZED DEEP LEARNING | Parallel Computing (code | report)

- Implemented sequential and parallel neural network model using data based parallelism in Python using MPI and achieved 50% improvement in training time with parallel model.

CHURN PREDICTION | Data Science (code)

- Built different ML models: SVM, Tree based models, logistic regression for predicting customer churn rate using PySpark and Scikit-learn.

NEURAL MACHINE TRANSLATION | NLP/ Deep Learning (code | report)

- Developed an end-to-end machine translation pipeline using recurrent neural network based models: simple RNN, RNN with Embedding, Bidirectional RNN, Encoder-Decoder RNN & achieved accuracy of 98%.