Pradeep Singh

www.pdeep.xyz | pdeepsingh094@gmail.com | +1-619-538-4719

Education

- **San Diego State University** — San Diego, CA
  Master of Science in Computational Data Science; GPA: 3.70/4.0 — *Aug. 2017 – May. 2020*
  - **Thesis**: Neural Mechanism for Target (Object) Tracking in Visual System
  - **Publication**: Models for Propagating Facilitation in Visual System. Accepted in ICIV, 2019
- **University of Mumbai** — Mumbai, India
  Bachelor of Engineering in Electronics and Telecommunication; GPA: 3.4/4.0 — *Aug. 2011 – July. 2015*

Technical Skills

Languages/Libraries: Python (Numpy, Scipy, Pandas, Matplotlib), C/C++, MATLAB, PySpark, SQL
Machine Learning: TensorFlow, Keras, Scikit-learn, MLlib, XGBoost
AWS: Storage (S3), Computing (EC2, EMR), Amazon SageMaker
Tools: Git, Docker, Gitlab, Flask

Experience

- **Dassault Systemes | Machine Learning Research Intern | [Paper] [Poster]** — June - Dec 2019
  - **(Graph) Deep Learning**: Researched & designed novel deep learning (CNNs, GCNs) models for super-resolving CFD simulations both on structured and unstructured grids.
  - **Framework**: Developed SRCFD, a generalized and platform agnostic framework (in TensorFlow) for super-resolving coarse simulations into fine simulations.
  - **Python Package**: Developed a python package to automate, generate, extract, process and convert (unstructured and structured) mesh data (simulation) into graph data and vice-versa.
  - Delivered a **dataset** of low and high resolution simulations. Built custom docker images to containerize ML models.
- **San Diego State University | Graduate Research Assistant | [Thesis] [Code]** — 2017 - 2020
  - **Research**: Researched mechanisms (response facilitation, selective attention) for target tracking in Visual System
  - **Modeling**: Built computational models of neurons and (networks) astrocyte in Matlab.
  - **Simulation**: Simulated models of facilitation, calcium waves, calcium pumps, in biological cells.
  - **Analysis**: Analyzed gigabytes of data in Matlab. Carried out comprehensive parametric study of our models.
  - **Results**: Poster accepted in ICIV, 2019. One journal paper (in-progress)
- **HERE Technologies | Data Analyst** — 2016 - 2017
  - Worked in data-processing team, building data pipelines for cleaning and processing data (mostly images) using python data science stack. Our team role was to ensure quality of data coming in for various tasks.
- **Raman Research Institute | Research Software Intern** — 2015 - 2016
  - Worked on different testing methodologies for 8 Tile Digital Receiver System and proof of concept project on development of signal processing algorithms using OpenCL for FPGA based architecture.

Projects

- **Visual Recognition using CNNs | [Report] [Code]**
  - **Classification**: Built Image classification system using Convolutional neural networks in Tensorflow. Desgined architectures like VGG, ResNet. Achieved accuracy of 95% with VGG model.
  - Build and deployed Flask web app to serve model in real time. Improved the performance by using techniques like data augmentation, transfer learning and batch normalization.
  - **Object Detection**: Built Object detection model YOLO in TensorFlow for detecting objects in Images.
  - **Few-shot Learning**: Implemented SOTA Few-shot learning models like, Siamese neural network, Matching Networks and Prototypical Networks in TensorFlow.
- **Increasing the Resolution of Images | [Code]**
  - Implemented SOTA Image super-resolution research papers – SRCNN, FSRCNN, ESPCN, SRGAN, EDSR and WDSR in TensorFlow. Explored approaches like adversial training, sub-pixel convolution.

- **Autoencoders | [Report] [Code]**
  - Implemented different forms of Autoencoders: Sparse, Denoise, Contractive and Variational Autoencoders.
- **Neural Machine Translation | [Report] [Code]**
  - Implemented three different Encoder-Decoder models: 1.) Seq2Seq with no attention and 2.) Seq2Seq with attention mechanism using RNNs (LSTM/ GRU) and 3.) Transformers for translating English sentence into French. Explored various forms of attention mechanism.
- **Churn Prediction | [Code]**
  - Built different ML models: SVM, Tree based models, logistic regression for predicting customer churn rate using PySpark and Scikit-learn. Dataset: Telco Customer churn dataset.
- **Hyperparameter Optimization using Bayesian Learning | [Report] [Code]**
  - Implemented an hyperparameter optimization algorithm using bayesian methods that finds better hyperparameters for machine learning models in less number of steps as compared to random/grid search.

Open Source

- **DeepClean**: Python package to clean and pre-process text and image data.
- **ML Notes**: Quick notes on Machine Learning. Read on the fly.
- **Data Science 101**: Notes and tutorials on how to use python, pandas, seaborn, matplotlib, scipy for data science.

Leadership & Achievements

- **Scholarship**: Tuition scholarship for the academic year 2017, 2018 and 2019.
- **ResearchX**: Founder & Author of a blog on Research opportunites in India. 1 Million+ monthly views by April 2017.
- **KC Xplore**: Founded and lead e-Newspaper of my undergraduate college [Video].
- **IEEE Club**: Vice-Chairperson of IEEE student chapter for the year 2013-14. Organized technical event, conferences.