

# Pradeep Singh

pradeepsinngithub.github.io | pdeepssingh094@gmail.com | 619.538.4719

## EDUCATION

### SAN DIEGO STATE UNIVERSITY

#### M.S. IN DATA SCIENCE

Expected Dec 2019 | San Diego, CA

GPA: 3.6/ 4.0

### UNIVERSITY OF MUMBAI

#### B.S. IN ELECTRICAL ENGG.

Aug 2011 - June 2015 | India

## LINKS

Github:// [pradeepsinngithub](#)

Webpage:// [pradeepsinngithub.github.io](#)

LinkedIn:// [pradeepsinngithub](#)

Twitter:// [@pradeep\\_sinngithub](#)

## COURSEWORK

### MACHINE LEARNING

Deep Learning (NLP)

Machine Learning

Numerical Optimization

Parallel Computing

Computer Vision

### DATA SCIENCE

Statistical Inference

Bayesian Statistics

Techniques of Data Science

Foundations of Data Science

## SKILLS

### PROGRAMMING

Python • C • C++ • Matlab • R

### MACHINE LEARNING

TensorFlow • Keras • Numpy

NLTK • SKLearn • Gensim

### DATA & VISUALIZATION

SQL • Pandas • R Studio • Shiny

StatsModels • Scrapy • Tableau

Matplotlib • Plotly • Seaborn

### PLATFORMS

Google Cloud • AWS • macOS

Linux • Windows • GitHub

## ACHIEVEMENTS

2018 Tuition Wavier + Stipend

2017 Tuition Wavier + Stipend

2015 Research Fellowship at RRI

2014 Undergrad. Research Fellowship

2014 Editor-in-chief at KC-Xplore

2013 Vice-chairperson - IEEE Student Branch

## RESEARCH

### COMPUTATIONAL SCIENCE RESEARCH CENTER | RESEARCH ASSISTANT

Aug 2017 – Till date | San Diego, CA

- Working on principal based approaches to target detection in visual neurons in dragonflies using neural computation.
- Working on modeling and characterizing a mechanism (response facilitation) by which dragonflies detect and react to small moving targets by increasing sensitivity to small objects that move along continuous paths.
- Parallelizing models using GPUs and analyzing gigabytes of data in Matlab.
- Comprehensive parametric study of our model outlining how and why it behaves with different values of parameters and thus characterizing the facilitation mechanism.

## PROJECTS

### NEURAL MACHINE TRANSLATION | NLP/ DEEP LEARNING (CODE | REPORT)

- Built 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%.

### IMAGE RECOGNITION USING CNN | COMPUTER VISION (CODE | REPORT)

- Built and trained 5 different Convolutional Neural Networks using Keras and TensorFlow to classify 70,000 fashion images into 10 labels.
- Achieved accuracy of 95% with VGG model + batch normalization.
- Used data augmentation & feature engineering techniques to add more images to the dataset, which increased the accuracy by 20%.

### BAYESIAN OPTIMIZATION | NUMERICAL OPTIMIZATION (CODE | REPORT)

- Build an hyperparameter optimization algorithm using bayesian methods that finds better hyperparameters for machine learning models in smaller number of steps than random or grid search.
- Comprehensive analysis and study of different hyperparameter tuning algorithms in machine learning.

### PARALLELIZED DEEP LEARNING | PARALLEL COMPUTING (CODE | REPORT)

- Implemented sequential and parallel neural network model using data based parallelism in Python using MPI.
- Trained both models on CPU and GPU and achieved 50% speed up in training time with parallel model.

### PREDICTIVE MODELING AND ANALYSIS | DATA SCIENCE [CODE]

- Cleaning and analyzing Countries of the World data set using Pandas.
- Built a predictive model using regression models, random forest model to predict the GDP of countries using SKLearn.
- Visualized the results using Matplotlib and Seaborn.

### SENTIMENT ANALYSIS | NLP/ DEEP LEARNING [CODE]

- Built an end-to-end sentiment classification system using 3 different classifiers: Recurrent neural network, Naive bayes and Gradient boosting.
- Preprocessed the data using beautiful soup, regular expressions and nltk, which increased performance by 10%.