# Pranav Pankaj Jaipurkar

8125 48th Avenue, #622-A2 Parkside Apartments, College Park, Maryland, 20740
(202)-415-1749, pjaipurk@umd.edu, https://www.linkedin.com/in/pranav-pankaj-jaipurkar

## EDUCATION

**University of Maryland, College Park, MD** — **GPA – 3.883/4.0**
Master of Engineering, Electrical and Computer Engineering — **Expected-May 2020**

**College of Engineering Pune, India**
Bachelor of Technology, Electrical Engineering — **May 2018**

## SKILLS

**Programming Languages**-C, C++;
**Scripting Languages**-MATLAB, Python;
**Database languages**-MySQL, MongoDB;
**Statistical Computing Languages**-RStudio;
**Desktop Skills**-MS Office (Word, Excel, PowerPoint);
**Other Software**-Tableau, Arduino, Proteus;
**Python Libraries**: Numpy, Tensorflow, Keras, Pandas, Pytorch, sklearn, scikit-learn, matplotlib
**Skilled Topics**-Machine Learning, Deep Learning, AI, Data Analytics, Advanced Statistics, Big Data.

## CERTIFICATIONS

**DataCamp-**
1) Introduction to Python — **Credential ID - #9708249**
2) Introduction to R — **Credential ID - #9708764**
3) Introduction to SQL for Data Science — **Credential ID - #9752628**

**Quality Enhancement in Engineering Education (QEEE)-**
1) Object Oriented Programming — **May 2016**

## RELEVANT COURSEWORK

**ENPM600 –** Probability and Stochastic Processes for Engineers (**GPA – 4.0/4.0)**;
**ENPM808W** – Data Science (**GPA – 4.0/4.0)**; **ENPM809N -** Data Mining **(GPA – 4.0/4.0)**;
**ENPM809K -** Fundamentals of Artificial Intelligence and Deep Learning (**GPA – 4.0/4.0)**;
**ENPM611** - Software Engineering; **ENPM613** - Software Design and Implementation

## PROFESSIONAL WORK EXPERIENCE

**CODEWIZARDS TECHNOLOGIES PVT. LTD** — **Nagpur, India**

**Big Data Intern** — **Dec 2017–Jan 2018**
- Mastered the basic concept and need for advanced methods for analysis of Big Data.
- Skilled the software RStudio for analysis of Big Data.
- Performed a data analysis for a commercial hospital and inferences were drawn.

**Data Analyst Intern** — **May–Jul 2017**
- Learnt the basic concept and need for Data Analytics.
- Skilled the use of Python and RStudio for Data Analysis.
- Produced an introductory data analysis report on the data given about sales of different motor vehicles in different countries and hence inferences were drawn.

**Database Administrator** — **Dec 2016–Jan 2017**
- Tutored the responsibilities of a database administrator.
- Performed installation, configuration and wrangling of datasets from different sources using MySQL.

**GRAYCELLS WEB DESIGN AND DEVELOPMENT** — **Nagpur, India**

**Web Developer** — **May–Jul 2016**
- Trained with the use of Python software for designing web pages.
- Designed the following web pages-www.altisales.com; www.andersonlongetivityclinic.com; www.kaliperchoice.com

## ACADEMIC PROJECTS AND ASSIGNMENTS

**University of Maryland** — **College Park, MD**

- **ENPM 809K- Fundamentals of AI and Deep Learning** — **Jan – May 2019**

**1) Logistic Regression Model for Image Classification Problem-**
- Modeled a Logistic Regression Image Classifier to predict whether a given image is a cat or not.
- Training accuracy of 90% and testing accuracy of 65% was obtained.

**2) Digit Classier Model-**
- Designed a logistic regression model for classifying digits using kNN function for the MNIST dataset.

- Adopted the function GridSearchCV() to find the optimum value of the hyperparameter k for which cost function was least was chosen.
- Testing accuracy of 75% was obtained for the optimum value of the hyperparameter.

**3) Two Layer Neural Network for Image Classification Problem-**

- Designed a Two Layer Neural Network to predict whether a given image is a cat or not.
- Training accuracy of 95% and testing accuracy of 74% was obtained.

**4) Shallow Neural Network for Image Classification Problem-**

- Structured a Shallow Neural Network to predict whether a given image is a cat or not.
- Training accuracy of 97% and testing accuracy of 76% was obtained.

**5) Feedforward Neural Network for Image Classification Problem-**

- Programmed a Feedforward Neural Network to predict whether a given image is a cat or not.
- Training accuracy of 98% and testing accuracy of 79% was obtained.

**6) Deep Neural Network for Image Classification Problem -**

- Programmed a Deep Neural Network with 3 hidden layers to predict whether a given image is a cat or not. Activation function for layers 1 and 2 was 'Relu' and for layer 3 it was 'Sigmoid'.
- Training accuracy of 98% and testing accuracy of 80% was obtained.

**7) Residual Network for Image Classification Problem-**

- Prepared Identity block and convolutional block of the Residual Network with 3 retrieve filters and 'Relu' activation.
- Drafted residual Network model with 50 layers for the image classification to be performed on the 'Signs' dataset.
- Training accuracy of 89%, testing accuracy of 48% and loss of 0.33 was obtained.

**8) Neural Style Transfer using Deep Learning-**

- Generated an image of the Louvre museum in Paris (content image) mixed with a painting by Claude Monet 9styling image) using Convolutional Neural Network.
- Optimized the cost function of content matrix and style matrix using Adam optimizer to update the pixel values rather than the parameter values.

**9) Recurrent Neural Network-**

- Built a single RNN forward cell and iterated it to perform forward propagation of the Neural Network which gave predicted output, shape of the predicted output matrix and caches which stored values for backward propagation.
- Studied the concept of Long Short Term Memory (LSTM), different gates such as forgot gate, update gate and output gate present in it and implemented forward propagation of the Neural Network.

**10) Autonomous Driving Car Detection using Deep Learning-**

- Performed car detection using a popular object detection algorithm: YOLO (you only look once).
- Implemented Non-Max Suppression (NMS) and Intersection Over Union (IOU) concept to accurately define the box boundaries.

**11) Neural Machine Translation using Deep Learning-**

- Translated human readable dates into machine readable dates (YYYY/DD/MM) using attention mechanism and two LSTM blocks.

**12) Face Detection with Keras-**

- Modelled a Neural Network with Keras to detect whether given face is a happy face or not.
- Obtained a test accuracy of about 88%.

**13) Neural Networks with Tensorflow-**

- Implemented a Neural Network using Tensorflow on the 'Signs' dataset.
- Training accuracy of approximately 99%, testing accuracy of approximately 71% and cost after epoch 100 of around 1 was obtained.

**14) Football Problem using Deep Learning-**

- Proposed some recommendations as to where should a goal keeper hit ball so that the chances of scoring a goal are increased.
- L2 regularization and dropout were the two methods used to avoid overfitting in which 3 layered NN with dropout had a testing accuracy of 95% as against to 3 layered NN with regularization which had a testing accuracy of 93% and 3 layered NN without regularization that had a testing accuracy of only 91.5%.

**15) Optimization Methods-**

- Implemented and compared the output of 4 different optimization methods on Neural Networks namely- Batch Gradient Descent (GD), Stochastic Gradient Descent (SGC), Mini Batch Gradient Descent and Adam in terms of accuracy, cost function, speed and momentum/oscillations made to reach the convergence.
- Adam increased the testing accuracy from 79% obtained in Mini Batch Gradient Descent to about 94% with less oscillation and more speed.

**16) Operation on Word Vectors using Deep Learning-**

- Defined and implemented cosine similarity function as a measure of degree of similarity between words.
- Mapped words to their respective word vectors and performed word analogy task successfully to predict the answer of the respective analogy. (example: boy -> actor then girl -> ?)
- Debiased word vectors with gender words 'woman' and 'man' to find gender specific words.
- Performed equalization algorithm for gender specific words which made the equalized words to be at the same dstance from reference point.

- **ENPM 809N- Data Mining** **Aug-Dec 2018**

**1) Linear Regression Model-**

- Prepared a linear regression model to predict the output (Response) from the 18 dependent variables.
- Suggested the most non-collinear dependent variables that contributed to the output (Response) using multi-collinearity test and vif function present in the car library in RStudio.
- Programmed the linear regression model again with the relevant dependent variables and used the model to predict the output of testing dataset.
- Improved the accuracy from 69% to 78% with above method.

**2) Cluster Analysis using k-means and k-medoids-**

- Conducted Cluster Analysis on UCI Machine Learning Wine Dataset using k-medoids technique to divide the dataset into different number of clusters.
- Performed Silhouette Width analysis and hence optimum number of clusters for that given dataset were found to be 3.

**3) Principal Component Analysis-**

- Performed principal component analysis on a dataset having one output variable and 18 dependent input variables to reduce the redundant variables in the dataset using princomp() and prccomp() functions in RStudio.
- Reduced the number of independent variables from 18 to 4 using the functions.

**4) Association rules-**

- Derived association rules for a car manufacturing dataset using 'arules' package and apriori function in RStudio.
- Discarded the redundant rules and the remaining rules were sorted according to the descending value of lift.

- **ENPM808W- Data Science** **Aug-Dec 2018**

**1) World University Ranking Dataset-**

- Conducted initial exploratory Data Analysis on 'world university ranking' dataset found on Kaggle website.
- Applied machine learning algorithms to build linear regression, logistic regression, SVM and decision tree models to predict whether any particular university is a top ranked university (world rank below 200 or US national rank below 40) or not.
- Drafted two types of models namely research-based models and employment-based models with the dependent variables research, citation, patents and teaching, international, income, number of students, student-staff ratio respectively.
- Modelled a new feature by adding weights derived from Artificial Neural Network algorithm to the existing dependent variables and predicted the top rank only from that feature which increased the predictive accuracy of the model**.**

**2) New York Times Dataset-**

- Performed initial exploratory Data Analysis on 'New York Times' dataset and made visualization plots.
- Studied trend and behavior analysis for the number of clicks and impressions between age groups and also between male and female.

**3) IMDB Movies Dataset-**

- Performed initial exploratory Data Analysis on 'IMDB Movies' dataset having data for all the movies between 2006-2016 and made visualization plots.

**4) Housing Price Prediction Models-**

- Programmed Linear Regression models and predicted the housing prices across different states and counties in the US in 2013 from the training data available from the year 2007.
- Training accuracy of about 86% and testing accuracy of about 77% was obtained.

**5) Classifier for Quiz bowl Questions-**

- Prepared Logistic Regression, SVM and Decision Tree models to predict the output of Quiz-Bowl questions.
- Modelled a new feature that matched the guess produced with the words in the questions revealed so far.
- SVM model had a testing accuracy of about 81% as against testing accuracy of 69% Logistic regression and 61% in Decision trees.

## College of Engineering **Pune, India**

### 1) Forecasting Output Power of Solar PV System **Aug 2017- Apr 2018**

- Modelled a data logger (hardware) for collecting data (panel temperature, ambient temperature, humidity, solar irradiation, current (I), voltage (V), power rating (P)) across the solar panels.
- Coded Arduino ATMega32 in the data logger to collect the data after every 1 hour during the day.
- Recorded data was in an excel sheet format and stored in an SD card which could be removed from the micro-controller.
- Used MATLAB software for the analysis of recorded data from Artificial Neural Network and Wavelet Decomposition toolbox available.
- Testing accuracy of 76% was obtained in forecasting clouds/clear sky/rain.