# Prateek K. Agnihotri

C-66, Aravali House, IIT Delhi
New Delhi, India - 110016
Phone: +91 8826382362
Date of Birth: 27/09/1997
Email: prateekkumaragnihotri@gmail.com
LinkedIn: https://www.linkedin.com/in/prateek-kumar-agnihotri-18b498157/

## Education

B.Tech IIT Delhi
Biochemical Engineering & Biotechnology (**8.09/10**)

## Courseworks

Machine Learning | Probability
Deep Learning | Linear Algebra
Statistics | Stochastic Processes
Calculus | Differential Equations

## Languages/Softwares

Python | C++ | Matlab

## Python Libraries

Tensorflow | Keras | Pytorch
OpenCV | Numpy | Pandas
Sklearn | Matplotlib | Imgaug
NLTK | Flask | BeautifulSoup

## Exposure

Neural Networks | CNN | RNN
LSTM | GAN | Auto-encoders
Transformers | Adversarial attacks & defenses | SVM
Bayesian Models | Decision Tree | Random Forest | KNN
XGBoost | LGBM | Catboost
PCA | t-SNE | UMAP | FLDA

## Volunteer Experience

iGEM, IIT Delhi
- Genetically modified bacteria
- Competed at MIT, USA

Group Representative, DBEB
- Organizes seminars
- Maintains lab budgets

NCC, IIT Delhi
- Participated in regular drills and special parades

## Achievements

- **Kaggle TensorFlow 2.0 Question Answering, 2020 - Bronze Medal (International Rank 121)**
  - Employed hard negative sampling to increase the difficulty of the candidate-level training
  - Fine-tuned Bert Q & A model on natural questions dataset and optimized thresholds
- **International Data Analytics Olympiad, 2020 - Ranked 45th among 2756 from 83 countries**
  - Clustered on temporal features to identify phase and applied linear regression techniques within each phase to identify phase-specific features for each satellite (Track 1)
  - Trained a LightGBM model to predict the position of space objects using simulation data
- **Goldman Sachs Quantify 2019 Lining Up Logs -** Ranked **1st** in campus (Machine Learning)
- **Flipkart's 2019 AI ML Challenge Stage 1 -** Ranked 73rd among 6737 Teams Pan India
- **International Genetically Engineered Machine Competition, 2018 – Bronze Medal**
- **Qualifying Examinations -** JEE Advanced | Ranked 3402nd among 1.3 million students

## Internships and Research Projects

- **MyWays |** Machine Learning Research and Development Trainee [Dec'19 – Jan'20]
  - Built a transformer model based **Open-domain Chatbot** for career consults and queries
  - BERT based NN model to **detect personality types in subjective personality test**
- **Domain Generalization and Domain Adaptation |** Prof. A P Prathosh [May'19 – Nov'19]
  - Used GANs, auto-encoders and adversarial training strategies for SOTA tumor classification models so that they can either adapt biases or learn invariant features across data domains
  - Used **CycleGAN** and **DefenseGAN** to get inter domain transformation models
  - Generated classifier's **adversary images** and re-trained it on them making it robust
  - Used **multitask learning** and **gradient reversal layer** to minimize tumor classification loss and maximize domain classification loss to extract domain invariant representations
- **Lensless On-chip Microscope |** Prof. RaviKrishnan Elangovan [May'18 - Sep'18 ]
  - Built a setup using CCD arrays and raspberry pi to sample light scattered by bacterial cells
  - Numerically reconstructed the bacterial image from the sampled light
  - The reconstructed image was then fed to a CNN to classify the bacterial type
  - The prototype was proposed for **Design innovation Summer Award 2018** for its state-of-the-art technology and ability to drastically reduce the cost of microscopes

## Academic Projects

- **Advanced Machine Learning Course |** Prof. A P Prathosh [Jan'19 - Apr'19 ]
  - **Adversarial FGSM and Carlini-Wagner Attacks** to fool CNN based digit classifier
  - Lipschitz constraint and Distillation Network as **defenses against adversarial attacks**
  - **WGAN and β-VAE** to generate image samples and compressed data representations
  - CNN to extract features and a LSTM on top of it to identify relevant frames in MOOCs
- **Brain Tumor Segmentation in MRI Scans |** Prof. A P Prathosh [Jan'19 - Feb'19 ]
  - Used histogram matching to match different sourced test-train image data statistics
  - Performed Bayesian hyperparameter optimization over Unet with skip connections
- **Explainable CNN based Cancer Detecting Classifier |** Prof. Anup Singh [Jan'20 - ongoing]
  - Trained Nasnet model on Pcam dataset with on-fly data augmentation
  - Making outcomes of the CNN based Nasnet classifier interpretable and explainable by disentangling Nasnet's representations into explanatory decision trees

## Personal Projects

- **Kaggle – Data Science Bowl 2019 – Silver Medal on Public LeaderBoard (78$^{th}$ Rank)**
- **Denovo Drug Design for Coronavirus using LSTM decoder and ChEMBL25 dataset**