



Viraj Bagal
Machine Learning Research Fellow

PROFILE

Address
64, Safalya Bungalow, Dnyaneshwar Nagar
Ekvira Chowk, Pipeline Road
Ahmednagar, Maharashtra - 414003

Phone Number
7219652039

Email Address
virajbagal12@gmail.com

Twitter
Viraj Bagal

ACHIEVEMENTS

KVPY Scholar
Indian Institute of Science (IISc), Bangalore

KVPY program aims to identify and support talented and motivated students in research. I secured **All India Rank 69** in their written exam+interview. My research will be supported by them until 06/2021.

National Top 1% in National Physics Graduate Examination (NGPE) 2019
Indian Association of Physics Teachers (IAPT))

Selected in Mathematics Madhava Competition 2018
Homi Bhabha Centre for Science Education, TIFR

Exam included topics like Calculus, Algebra, Combinatorics. On account of selection, I attended the prestigious Madhava Camp in Indian Statistical Institute, Bangalore, India

Silver Medal (Top 2%) in PANDA Competition on Kaggle
Kaggle

The system achieved 0.928 Cohen Kappa score and secured 16th (top 2%) position on the leaderboard among 1010 participants.

2 x Kaggle Expert
Kaggle

Only 8% of total Kaggle competitors are at this or above this rank.

SKILLS

Python Expert	Pytorch Expert
TensorFlow Intermediate	OpenCV Expert
Scikit Learn Expert	

LANGUAGES

English Professional Working Proficiency	Hindi Professional Working Proficiency
Marathi Native	

ABOUT MYSELF

I am currently a final year BS-MS student pursuing major in Physics and minor in Mathematics. I am interested in Computer Vision (CV), Natural Language Processing (NLP), generative networks and robustness. I am experienced in multimodal data representation learning (CV + NLP), generative networks and interpretability.

EDUCATION

Indian Institute of Science Education and Research (IISER), Pune (July 2016 - April 2021)
Integrated BS-MS Physics and Mathematics GPA - 9.3

- Courses in Physics**
 - Classical & Quantum Physics, Statistical Physics, Condensed Matter Physics, Quantum Field Theory, Atomic & Molecular Physics, Particle Physics, Optics
- Courses in Mathematics**
 - Linear Algebra, Single & Multivariable Calculus, Probability, Statistics, Set theory

WORK EXPERIENCE

CCNSB Lab, International Institute of Information Technology (IIIT), Hyderabad (May 2020 - April 2021)
Research Fellow

- Generated ligands conditioned on multiple physicochemical properties as well as scaffolds using custom small Generative Pre-Training (GPT) and SMILES representation of molecules.
- Our model achieved new state-of-the-art results in terms of validity, uniqueness, and novelty of generated molecules on the Guacamol dataset and competitive results to graph-based approaches on the Moses dataset.
- Our model can generate molecules having particular values of certain properties like QED score, logP, TPSA, SAS, Molecular weight by conditional training. Obtained mean absolute difference (MAD) scores lie within 10% of the desired value.
- Moreover, our model can generate molecules of certain scaffolds as well.
- Interpretability of the generative process addressed by saliency maps.[Link to paper](#)

CVIT Lab, International Institute of Information Technology (IIIT), Hyderabad (May 2020 - April 2021)
Research Fellow

- Built interpretable medical visual question answering system on radiology images
- Implemented self-supervised training with pretext tasks such as Image-Text Matching and Masked Visual- Language Modelling
- Achieved new state-of-the-art (SOTA) performance on ImageClef 2019 and VQA-RAD datasets
- My single model outperforms the ensemble of previous SOTA models
- Interpretability addressed via Attention maps
- Paper accepted at ISBI 2021.**[Link to the paper](#)

Indian Institute of Science Education and Research (IISER), Pune (December 2019 - May 2020)
Research Student

- Analysed Drell Yan process using nanoAODs
- Analysed different particle collections using distribution plots
- Created dataset from the images of Calotowers within a cutoff dR of electrons from collision data
- Implemented CNNs for the identification of fake electrons in LHC
- Analysed results using probability histograms and ROC curves
- Achieved 81% accuracy in identifying fake electrons

PERSONAL PROJECTS

Mixed Sample Data Augmentations (MSDAs) (May 2020 - January 2021)
<https://github.com/VirajBagal/FMix-Paper-Implementation>

- Reproduced, Ethan Harris et al. FMix: Enhancing Mixed Sample Data Augmentation paper.
- Compared the performance of FMix, Cutmix , Mixup and Baseline on the Fashion MNIST dataset.

Efficient Resizing & Highly Imbalanced Multilabel Classification of ChestX-rays (June 2020 - January 2021)
<https://github.com/VirajBagal/ChestXRay14-Reimplementation>

- Reproduced, Ekagra et al.Jointly Learning Convolutional Representations to Compress Radiological Images and Classify Thoracic Diseases in the Compressed Domain. ICVGIP 2018.
- Going one step further, tried FMix to improve the performance.
- Added Grad-CAM in the pipeline. Model not only predicts but even highlights the decisive region.

Kaggle: Prostate cANcer graDe Assessment (PANDA) (July 2020 - January 2021)

- Built robust ensemble of models for predicting Gleason scores and ISUP grades for WSIs of prostate biopsy.
- The system achieved 0.928 Cohen Kappa score and secured 16th (top 2%) position on the leaderboard among 1010 participants.