Vibhas Kumar Vats

vkvatsdss@gmail.com | vkvats.github.io

EDUCATION

Indiana University Bloomington, IN
Ph.D. in Computer Science, with a minor in Computer Science *2021 – 2026 (expected)*
Indiana University Bloomington, IN
Master of Science in Data Science (with thesis) | GPA: 3.97/4.0 *2019 - 2021*
Master's thesis: Response-Based Knowledge Distillation

National Institute of Technology (NIT) Patna Bihar, India
Bachelor of Technology in Electrical Engineering | CGPA: 8.77/10.0 *2011 - 2015*

RESEARCH EXPERIENCE

Research Assistant, Indiana University Computer Vision lab

Deep Learning (DL) - Case Based Reasoning (CBR) survey paper: *June 2020 – Present*
- Collaborated with DL-CBR research group in writing two sections of the survey paper- retrieval and DL models with memory.

Exploring Accuracy and Explainability of DL-CBR Hybrid Architectures *Jan 2022 – Present*
- Exploring hybrid system leveraging knowledge-engineered and network learned feature in concert.
- Analyzing different DL architectures to study the impact of learned features in the DL-CBR hybrid system.

Multi-View Stereo (MVS) Method for High-resolution Depth-map Prediction and Point Cloud Generation *Aug 2021 – Present*
- Improving the application of MVS algorithm for high-resolution depth map prediction
- Applying MVS algorithm at multiple stages of DL architecture to improve the accuracy and completeness of 3D point cloud

Roof-area Segmentation and Orientation Detection *Aug 2021 – Present*
- Detection of roof orientation and plane area on 3D point clouds data by applying RANSAC algorithm
- Improving roof segmentation and orientation detection using satellite images

Master's Thesis on "Response-based Knowledge Distillation" *Aug 2020 – May 2021*
- Analyze the knowledge distillation process and proposed the soft-label hypothesis to explain the behavior of distillation process.
- Proposed special consideration for pre-training teacher models for retaining similarity information in soft-labels for better knowledge distillation.
- A condensed version of the thesis was published at AAAI-2022 workshop

External Research Fellow, NIT Patna (Electrical Engineering Lab). *July 2018 – June 2019*
Project title: "Sustainable Smart Grid Framework for Energy Management System Incorporating Available Renewable Resources." funded by the Science & Engineering Research Board, Government of India.
- Successfully conceptualized and implemented a model to mitigate the Communication-link failure in smart meter-based load forecasting system using various classification methods.
- Programmed and implemented an electrical load forecasting system with one year of data using polynomial regression model.
- Presented paper at 4th IEEE International Conference on Computing Communication and Automation (2018).

PUBLICATIONS

- **V.K.Vats** and David Crandall, "Controlling the Quality of Distillation in Response-Based Network Compression", Association for the Advancement of Artificial Intelligence (AAAI – 22) workshop.
- **V.K.Vats**, M. De, S. Rai and S. De, "Mitigating Effect of Communication Link Failure in Smart Meter based Load Forecasting", Springer, 4th International Conference on Nanoelectronics, Circuits & Communication Systems 2018.
- **V.K.Vats**, M. De, S. Rai and D. Bharti, "Very Short-Term, Short-Term and Mid-Term Load Forecasting for Residential Academic Institute: A Case Study", IEEE Xplore, 4th IEEE International Conference on Computing Communication and Automation 2018.

WORK EXPERIENCE

Course Development and Instructor – Computer Vision Paper Discussion Section, Indiana University *Jan 2021 – May 2021*
- Designed and was the instructor of paper discussion section, first of its kind at IU, in computer vision course
- Lead the discussion of DL based seminal papers covering development of CNNs, Transformers and MLP based models
- Received Associate Instructor of the year award, 2021-2022 for my distinguished contribution in this course

Associate Instructor, Indiana University
- Computer Vision *Spring 2021*
- Elements of Artificial Intelligence *Fall 2020, Spring 2021*
- Held weekly office hours on Zoom to work one-on-one with undergraduate, graduate students and working professionals.
- Manage and help more than 100 graduate students with their doubts on Piazza/InScribe

Senior Manager, Tata Motors Ltd. Pantnagar. *Aug 2015 – Aug 2017*
- Analyzed and optimized the maintenance frequency of Generator yard equipment using past breakdown and maintenance data.
- Overhauled and systematized power transformer oil filtration frequency of 36 units of transformer, individually, by predicting oil characteristics threshold value against number of operation hours of each transformer using eight years of past records.

SKILLS & CERTIFICATIONS

Languages & Tools: Python(advanced), SQL, R, Rstudio, PostgreSQL, C – (intermediate)
Frameworks & Libraries: Pytorch(intermediate), TensorFlow(advanced), Keras, NumPy, Pandas, NLTK, Pillow, Scikit-learn.
Certifications: "Machine Learning" and "Deep Learning" specialization by Prof. Andrew Ng on Coursera,

- "TensorFlow in practice" by Laurence Moroney from Google.
- "Statistical Learning" by Prof. Hastie and Prof. Tibshirani on Stanford online

COURSE PROJECTS

Zero-shot learning (ZSL) and outlier detection using C2C-Siamese network (C2C-SN) *Jan 2020 – May 2020*

- Used Siamese network to learn the similarities and differences between two classes with 13 hand-engineered features for 10 classes.
- Successfully detected outlier classes with accuracy of about **72%** and confidence threshold value of **0.999** on MNIST dataset.
- Assessed the potential of doing ZSL by feeding the features learned from C2C-SN and its semantic embedding into Fully connected layer and doing nearest neighbor search. The set up did not prove to perform up to its theoretical understanding.
- Tech Stack: Python, PyTorch, TensorFlow, NumPy, Pandas, Siamese networks. Model accuracy: **72%**

Exploratory analysis of Pima-Indian (women) diabetes dataset *March 2020 – May 2020*

- Analysis of eight medical parameters to explore and establish the cause of rise in diabetic patients in the Pima Indian community.
- Used Generalized Linear Model to train and then find the efficacy of model (78%) to predict diabetic women in Pima Indian community.
- Tech Stack: R, RStudio, Ggplot2, Tidyverse. Model accuracy: **78%**

Optical Music Recognition (OMR) and text annotation *Feb 2020 – March 2020*

- Implemented 2D convolution from scratch with separable kernels on scanned (greyscale images) musical notes.
- Detected the musical notes by using hamming distance and also through template matching score using edge maps (Sobel operator).
- Used Hough transformation to detect the staff lines and draw bounding box around symbols and tagged text with it.
- Tech stack: Python, NumPy, Pillow.

Study on sentiment Analysis methods *Dec 2020*

- Implemented Naïve Bayes, Logistic regression, Recurrent Neural Networks (RNN) from scratch.
- Fine-tuned Google's BERT sequence classification model on sentiment140 dataset.
- Tech Stack: TensorFlow, Huggingface, Python, NumPy. Model accuracy: **86%**

Transfer Learning: Neural style transfer (TensorFlow), Car detection with YOLO (TensorFlow), Face recognition (TensorFlow).

AWARDS & HONORS

- Associate Instructor of the year award- 2021-2022, Indiana University Bloomington
- Best Graduate National Institute of Technology Patna, batch of 2015.
- Honored by Tata Sustainability group for distinctive work in CSR under Pro-engage program in 2016-17.
- Awarded under eNoble program for valuable contribution as CSR volunteer in 2016-17.
- Team member, winner of national level Golden Peacock Environment Management Award 2016 in the manufacturing sector.
- Coordinator, CSR cluster of Graduate Engineer Trainee club of Tata Motors, Pantnagar.