**Vaibhav Kumar Katturu**

New York City, NY 11213| [vaibhav.k.katturu@pace.edu](mailto:vaibhav.k.katturu@pace.edu) | Phone: 813-585-4565 |[LinkedIn](http://www.linkedin.com/in/vaibhav-kumar-559787185)| [GitHub](https://github.com/vaiv26/learner)

**EDUCATION**

**Pace University, Seidenberg School of Computer Science and Information Systems New York, NY**

Master of Science (MS) in Computer Science | **Concentration:** Artificial Intelligence | **GPA:** 3.9/4.0 June 2021

**Amity University Kolkata Kolkata, India**

Bachelor of Technology in Computer Science and Engineering | **GPA:** 3.5/4.0 May 2019

**RELEVANT COURSEWORK**

Data Mining with Python | Algorithms and Computational Theory | Artificial Intelligence (AI) | Pattern Recognition | DBMS | Deep Learning | Probability and Statistics | Distributed Systems | Machine Learning |Cloud Computing

**TECHNICAL SKILLS**

**Programming Languages:** C++, Java, Pytorch, TensorFlow, Python, Scikit-Learn, HTML, CSS, JavaScript, skills with MATLAB

**Operating Systems:** Windows, iOS, Unix/Linux Virtual Machine environment: Ubuntu and Kali.

**Database Management:** Relational database handling using MySQL, Distributed data processing, SQLite, Microsoft Access

**Software:** Unity Game Engine, Microsoft Office (Word, PowerPoint, Excel), REACT, AJAX

**Cloud Computing Platform:** AWS, Google Cloud Platform (Google Collab)

**Machine Learning Tools:** Decision Trees, Linear Regressions, Logistic Regression, Support Vector Machines, Supervised Learning, data driven statistical modeling, graphical models

**ACADEMIC PROJECTS**

**A New Ensemble method to combine multiple Classifiers**  **September 2019 – December 2019**

* As a member in a team of two, demonstrated problem solving skills to developed a new Ensemble method to combine multiple CNN classifiers that process images using Computer in Pytorch framework. Using modified Borda count method the outputs of multiple classifiers were assigned with points which helped to determine the disease class. With the new Ensemble method 90% accuracy was achieved,
* In this project, a total data set of 512 plant images were used where we had to extract ,transform data, perform feature extraction and analysis to clean the large data set. I Built 20 machine learning models using supervised learning with CNN, that could process data, match images of plants with disease with the images in the database by development of algorithms where after designing series of theory and practice in design of experiments and results, familiar enough with techniques and limitations of observational studies.

**Created a library system in MySQL Server**

* Developed a library system using data driven statistical modelling in SQL management studio to replace manual checking of books in the library.
* Extract transform and cleaned multi-TB dataset in SQL Management Studio using query language MySQL and gave output with a statistical analysis of the results.

**Classifying One- or five-star yelp reviews based on text content using Natural language Processing December 2019**

* Classified one star to five-star reviews using Data Science tools like Natural language in scripting language Python on the yelper reviews data given by users to businesses an open source machine learning project.
* In google cloud platform google Collab used seaborn to exploit and visualize the data, used scikit-learn, count vectorization for splitting the data into training and testing data.
* Used Naive Bayes’ modelling for prediction and evaluations of data.

**EXPERIENCE**

**Webtek Labs Pvt.Ltd** **Kolkata, India**

Intern**June 2018 – July 2018**

* Collaborated with a team of 10 to design and develop an android application using SQLite, Java and XML in Android Studio which processes data and stores information of all the items in a pharmacy to replace paper-based checking system.
* Created the relational database with all the editing functionality to improve the experience of the user.

**AWARDS**

**Pace University, Machine Intelligence Day,** Presentation Award 1st Place Winner **December 2019**

Presentation [MID2019](http://csis.pace.edu/~scha/MID2019/program.html?) where as a participant my responsibility was to present and communicate the project idea and my beliefs clearly and compelling in both verbal and written form, discuss the architectural concepts and the algorithms used in our research, with my technical team member demonstrate presentation skills and critical thinking skills by defending our research thesis. Also, standing at the podium, I discussed my journey where I had to formulate, test, and discard or revise many hypotheses to reach the final goal.

**PUBLICATIONS**

P. Thind, V. Katturu, S. Cha. “A New Ensemble Method for Convolution Neural Networks and its Application to Plant Disease Detection.” MLDM conference, Forthcoming.