

EDUCATION

New York, NY	New York University	Sep 2019 – May 2021
MS in Data Science at Center for Data Science . CGPA: 3.45 / 4.0		
Coursework: Probability and Statistics; Optimization and Computational Linear Algebra; Machine Learning; Deep Learning; Big Data		
Vellore, India	Vellore Institute of Technology	Jul 2015 – May 2019
B. Tech. in Information Technology. CGPA: 9.15 / 10.0		

EXPERIENCE

[Noodle.ai](#)

Data Science Intern (Spring 2019)	Jan 2019 – May 2019
<ul style="list-style-type: none">Designed the pipeline that was internally used by the data science team for pre-processing and structuring different time-series datasets for problems like forecasting and anomaly detection.Developed a modular, reusable, and comprehensively documented python framework (called “sequel”) for creating and training sequence-to-sequence models using Keras.Used “sequel” to benchmark different deep learning architectures (like vanilla RNNs, sequence-to-sequence models) for multi-variate time series forecasting (on internal datasets of the company).	
Data Science Intern (Summer 2018)	May 2018 – Jun 2018
<ul style="list-style-type: none">Performed exploratory data analysis on a steel manufacturing plant’s data, with the aim to understand and model the steel manufacturing process.Augmented the existing system (which was used to predict the time required by an order of steel coils to get ready) with a conditional probabilistic model that improved the prediction metrics by ~20%.	

PROJECTS

“Mydia” – An opensource framework for working with videos in python	View on GitHub
<ul style="list-style-type: none">Developed and currently maintaining a small, simple, and efficient python library for reading videos as NumPy tensors. The library provides some useful features such as custom frame selection in videos, frame resizing, grayscale conversion, pixel normalization and much more.The library follows the paradigm “simple things should be simple, and complex things should be possible”. Exhaustively documented the code and provided numerous examples to get started with.Set up the CI pipeline for efficient testing and distribution of the library (via python package managers like pip and conda).	
Human activity recognition using 3D convolutional neural networks	View on GitHub
<ul style="list-style-type: none">Made use of temporal (3-dimensional) CNNs for video recognition for identifying 6 different human actions (like running, jogging, walking, clapping, etc).Tested different model architectures and tuned various hyper-parameters to get the best performance out of the model, and achieved an overall accuracy of about 65% on the testing data.	
Dog breed classifier using deep learning	View on GitHub
<ul style="list-style-type: none">Built an algorithm to identify the canine breed, given an image of a dog. For fun, if given an image of a human, the algorithm identifies the resembling dog breed.Used transfer learning and compared different pre-trained CNN architectures (like ResNets, Inception, VGG). Got an accuracy of about 80% in classifying dogs into their breeds (~120 breeds in total).	

TECHNICAL SKILLS

Programming Languages	Python (proficient), C++ (familiar), Java (familiar)
Tools and Frameworks	
• Python Data Science Ecosystem	Scikit-learn, NumPy, Pandas
• Deep Learning	TensorFlow and Keras
• Image and Video Processing	OpenCV, FFmpeg
• Database Management Systems	SQL