Yash Katariya

yash.katariya10@gmail.com | yashk2810.github.io/ | github.com/yashk2810 | (919)455-8284

EDUCATION

NORTH CAROLINA STATE UNIVERSITY August 2017 - Present
MS Computer Science GPA: N/A
Coursework: Foundations of Data Science, Automated Learning and Data Analysis, Algorithms

PUNE INSTITUTE OF COMPUTER TECHNOLOGY 2013 - 2017
BEng Computer Engineering GPA: 3.48 / 4
Coursework: Business Intelligence, Data Mining, High Performance Computing, Database Management

EXPERIENCE

E2OPEN | DATA SCIENCE INTERN Pune, India | Dec 2016 – March 2017

- Analyzed time series data to forecast retail sales with multiple store multiple item combinations.
- Achieved an accuracy of 89.47% using an ensemble of machine learning models for predicting the optimal price to maximize sales.

SKILLS

LANGUAGES	Python, C++, R, Java, HTML, CSS, Javascript
DATABASES	MySQL, MongoDB, SQlite3
FRAMEWORKS / TOOLS	Keras, Pytorch, Django, scikit-learn, Pandas, Jupyter Notebook, Express.js

PROJECTS

IMAGE CAPTIONING | KERAS, TENSORFLOW, NUMPY, PANDAS, MATPLOTLIB, TQDM
Generates automatic captions for images. Trained the model on the Flickr8K dataset. Extracted the features of images using pretrained InceptionV3 model and used beam search to predict the captions. Achieved a loss value of 1.5987 after training for 50 epochs.

SEMANTIC IMAGE SEGMENTATION | KERAS, TENSORFLOW, NUMPY, MATPLOTLIB
Implementation of the 100 Layer Tiramisu paper on the camvid dataset. Trained the model for approximately 500 epochs and achieved an accuracy of 89% by segmenting the image into 32 segments classes.

DEEP CONVOLUTIONAL GAN | KERAS, NUMPY, MATPLOTLIB, TQDM
Implemented the DCGAN paper and trained the generator and the discriminator on the MNIST dataset. Trained the generator and the discriminator for 50 epochs obtaining very good results.

NEURAL STYLE TRANSFER | KERAS, MATPLOTLIB, NUMPY
Used VGG-16 model trained on imagenet and calculated the content loss and the style loss to create artistic style photos. Wrote a blog post explaining the technique used to achieve style transfer.

PREDICTING VISITOR ATTENDANCE IN A PARK | SCIKIT-LEARN, PANDAS, MATPLOTLIB
Built an ensemble of XGBoost and Gradient Boosting model to predict the number of visitor attendance count of the National Park. Cleaned and analyzed the data and engineered new features using the insights obtained by analyzing the data. Achieved an error rate(RMSE) of 95.59.

AWARDS

RANK 16 / 1494
Analytics Vidhya Ultimate Student Hunt Machine Learning Competition.

RANK 7 / 926
MiniHack: Machine Learning Competition held by Analytics Vidhya