Sumit Gupta

Bloomington, IN 47404

Email: sumit743@gmail.com
Mobile: (812) 391-4082
Website: http://sumitg.com

Summary: 4+ years of combined research and development experience in Machine Learning – including 2+ years of research in Deep Learning (for Computer Vision and NLP).

EDUCATION

Indiana University Bloomington *Bloomington, Indiana*
Master of Science in Computer Science (GPA: 3.82) *May 2016*
Thesis: Convolutional Neural Networks for Infrared, Fine-Grained, and Egocentric Scene Classification

Dhirubhai Ambani Institute of Information and Communication Technology *Gandhinagar, India*
Bachelor of Technology in Information and Communication Technology *May 2010*

TECHNICAL SKILLS

Languages & Technologies: Python, C++, Java, R, MATLAB, Octave, C, SQL, Linux, Git, LaTeX.
Libraries: Tensorflow, scikit-learn, Keras, Caffe, OpenCV, Torch, Theano, NumPy, Pandas.

WORK EXPERIENCE

Software Developer *(Machine Learning / Deep Learning)* *Aug. 2016 – Present*
Bloom Insurance *Bloomington, Indiana*

- Developed a library to perform Optical Character Recognition (OCR) on photos of non-flat labels from prescription drug bottles and tubes to recognize various information like drug name, type, dosage, and contents.

Research Intern *(Deep Learning)* *May 2015 – July 2015*
IU Computer Vision Lab *Bloomington, Indiana*

- Developed & Applied machine learning techniques to images in Infrared spectrum for human recognition.
- Trained Deep Learning models on first-person images with text to automatically generate image captions.

Associate Instructor *(Computer Vision / Artificial Intelligence)* *Aug. 2014 – May 2016*
Indiana University Bloomington *Bloomington, Indiana*

- Taught graduate level courses in Computer Vision and Artificial Intelligence.

Software Engineer *(Machine Learning / Data)* *Aug. 2010 – Aug. 2014*
Hewlett-Packard *Bangalore, India*

- Developed machine learning models for predicting departing customers by processing Churn rate data.
- Designed & developed distributed data processing system capable of handling several terabytes per day.
- Implemented ML and ETL algorithms to generate insights by processing high volume data.
- Technologies: Machine Learning, Data Processing/Analysis, Python, Java, SQL.

PROJECTS (see all: **http://sumitg.com/projects**)

Kaggle Microsoft Malware Classification Challenge *(Machine Learning)* *Apr. 2015*

- Designed and trained a classification model on 500 gigabytes of malware source code using Extreme Gradient Boosting & Random forest. Extracted features based on byte 4-grams frequency and instruction count.
- Languages & Tools: C++, Python, Caffe, scikit-learn.

Kaggle Right Whale Recognition *(Deep Learning, Image Classification)* *Nov. 2015*

- Developed a models to classify individual whales using C++, Python and deep learning techniques.
- Extracted features from a Convolutional Neural Network (CNN) and trained an SVM to identify individual whales.
- Fine tuned pre-trained CNN models to Right Whale data and combined different techniques to improve accuracy.

First-person Scene Classification *(Deep Learning, Image Classification)* *Feb. 2016*

- Trained a multi-label deep learning system to classify images from wearable camera into several categories based on Location, Activities and Objects (like indoor, outdoor, restaurant, eating, driving).
- Languages & Tools: C++, Python, Caffe, scikit-learn.

RECOGNITION

- Recognized by National Innovation Foundation (Department of Science and Technology, Govt. of India) for creativity and innovativeness.