Siddha Ganju

(412) 789 6469 • sganju1@cs.cmu.edu • https://www.linkedin.com/in/sidgan

Education

Carnegie Mellon University (CMU), School of Computer Science — **Pittsburgh, PA**
Master of Computational Data Science — *December 2016 (Expected)*
Major: Analytics

National Institute of Technology (NITH) — **Hamirpur, India**
Bachelor of Technology, GPA – 8.08 on 10 — *May 2015*
Major: Computer Science and Engineering

Areas of Interest

Multi-Modal Deep Learning, Machine Learning, Natural Language Processing, Computer Vision, Data Science

Experience

Carnegie Mellon University — **Pittsburgh, PA**
Research Assistant, Mentors: Olga Russakovsky, Abhinav Gupta — *May 2016 - Present*
Ongoing research on test time supervision offered by visual question answering as opposed to standard computer vision paradigms like image segmentation and object detection.

CERN — **Geneva, Switzerland**
Openlab Intern, Mentors: Valentin Kuznetsov, Tony Wildish, Manuel Martin Marquez, Antonio Romero Marin — *June 2015 - August 2015*
Used Apache Spark to streamline different predictive prototypes by gathering information from CMS, ran predictive models and proposed datasets which will become popular over time. Evaluated quality of individual models, performed component analysis and selected best predictive model for new set of data.

Publication

Ganju, Siddha *et al.* (2015). Evaluation of Apache Spark as an Analytics framework for CERN's Big Data Analytics.
DOI: 10.5281/zenodo.3186

Skills

Languages	Lua, Python, JAVA, C, R, Octave, C++, LaTeX
Libraries	Torch, TensorFlow, Caffe, Theano, scikit-learn, pandas, NLTK, Weka
Frameworks	Spark, Hadoop, Amazon Web Services, .NET, Azure

Select Projects

Open Advancement of Question Answering Consortium — **CMU**
Capstone Project, Mentors: Eric Nyberg, Matthias Grabmiar — *Jan 2016-Present*
Pursuing active research on Multi-modal QA systems based on an ensemble of Deep Learning and Rule-based systems.

Automated Pipeline for Machine Learning Problems — **NITH**
Summer Internship, Mentor: Anirudh Koul, Data Scientist, Microsoft — *May 2014-August 2014*
Created a Python command line toolkit using scikit, numpy, pandas and matplotlib libraries to solve machine learning problems automatically. Imputation and hyper parameteric optimization placed my model among the top 10% of the Titanic kaggle.com challenge (Rank 198/2035 in July 2014). Experimented with large data sets and deployed on Hadoop cluster over AWS. Presented at **Grace Hopper 2015**

Achievements

2016 Speaker, 'Atom Smashing using Machine Learning at CERN', **Strata+Hadoop World**, San Jose, USA
2015 Winner, Best Innovative Outreach, CERN WebFest, Geneva, Switzerland
2014 Winner, Grace Hopper Conference Hackathon, Bangalore, India
2014 Finalist, New York University International Hackathon, Abu Dhabi
2013 Winner, India Scholarship Award, Institution of Engineering and Technology (IET) *Out of 5000 participants*, Delhi, India

Honors

2016 Panelist, **IBM+Apache Spark Maker Community Event**, San Francisco, USA
2015 Grace Hopper Conference Scholar, Texas, USA
2013 Student Women Representative, Community Volunteers Conference, IET, Sri Lanka