Siddha Ganju

+1 (412) 789 6469
siddhaganju@gmail.com
http://sidgan.me/

Current appointments

2017-Present **Data Scientist**, *Deep Vision Inc.*, Palo Alto, CA, USA.

Education

2015-2016 **M.S. in Computational Data Science**, *Carnegie Mellon University (CMU), School of Computer Science*, Pittsburgh, PA, USA.
Capstone: Open Advancement of Question Answering using Deep Learning.

2011-2015 **B. Tech. in Computer Science and Engineering**, *National Institute of Technology (NIT), Department of Computer Science and Engineering*, Hamirpur, HP, India.
Major Project I: Deep Learning for Audio Recognition.
Major Project II: Missing Data Prediction using Data Mining for Wireless Sensor Networks.

Speaker

2016 **Strata+Hadoop World**, *San Jose, USA*.

2016 **IBM+Apache Spark Maker**, *San Francisco, USA*.

2015 **Talk by Team OpenCosmics**, *Mozfest, London, England*.

Achievements

2016 **Invited Member, Open Leadership Cohort, Working Open Workshop**, *Mozilla Science Lab, Berlin, Germany*.

2015 **Grace Hopper Conference Scholar**, *Texas, 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, U.A.E.*

2013-2014 **Women Ambassador**, *The Institution of Engineering and Technology, U.K. (IET)*.

2013 **Winner, India Scholarship Award**, *IET*, New Delhi, India.

2013 **Student Representative, Community Volunteers Conference**, *IET, Sri Lanka*.

2013 **Finalist, Nokia Do Good Hackathon**, *New Delhi, India*.

2011-2015 **Ambuja Scholarship**, *NIT*.

Media

Apache Spark for Atom-Smashing experiments, *O'Reilly Media*, June 9, 2016.

CERN seeks to predict new and popular data sets, *O'Reilly Media*, March 22, 2016.

Publications

Peer-reviewed conference articles

CVPR 2017 **What's in a Question: Using Visual Questions as a Form of Supervision**, *S Ganju, O Russakovsky, A Gupta*, Computer Vision and Pattern Recognition, 2017.

CERN Research

Zenodo 2015 **Evaluation of Apache Spark as an Analytics framework for CERN's Big Data Analytics**, *S Ganju, V Kuznetsov, T Wildish, M Martin Marquez, A Romero Marin*, 10.5281/zenodo.3186, 2015.

Research Experience

2016 **Research Extern**, *CMU*, Mentors: Dr. Olga Russakovsky, Prof. Abhinav Gupta.
Research focused on weak supervision: Utilizing supervision from visual questions asked about images. Accepted as a poster presentation at the **IEEE Computer Vision and Pattern Recognition conference, 2017.**

2016 **Capstone**, *CMU*, Mentors: Prof. Eric Nyberg, Dr. Matthias Grabmiar.
Open Advancement of Question Answering Consortium: Developed Question-Answering systems based on an ensemble of Deep Learning and Rule-based systems.

2016 **Course Project, Deep Learning (10807)**, *CMU*, Instructor: Prof. Ruslan Salakhutdinov.
What's in the Future? Generating Videos with Motion Sensitive Adversarial Networks: Used optical flow and GAN's to generate future frames using our **FlowGAN** architecture. Transferred the learned representations for Action Recognition and Static Image Editing.

2016 **Request for Research**, *OpenAI*.
Jokes Entity Recognition (JER): Collected 16031 joke-urls licensed under fair use of data. Trained a character-level LSTM language model on collected data and developed **JER**

2015 **Course Project, Machine Learning with Large Datasets (10805)**, *CMU*, Instructor: Prof. William Cohen.
Qualitative Evaluation of Word2Vec for Recommendations: Used Word2Vec to gain semantic relations from text reviews.

2015 **Openlab Research Intern**, *CERN*, Mentors: Dr. Valentin Kuznetsov, Dr. Tony Wildish, Manuel Martin Marquez, Antonio Romero Marin.
Evaluation of Apache Spark as an Analytics framework for CERN's Big Data Analytics: Used Apache Spark to streamline different predictive prototypes by gathering information from the Compact Muon Solenoid experiment, 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. Talk presented at **Strata+Hadoop World, 2016**.

2015 **Major Project II**, *NIT*, Mentor: Dr. Narottam Chand.
Missing Data Prediction using Data Mining for Wireless Sensor Networks: Used data imputation methods to predict missing data in wireless sensor networks.

2014 **Major Project I**, *NIT*, Mentor: Dr. Narottam Chand.
Deep Learning for Audio Recognition: Used deep learning to recognize and transcribe audio in the ambient surroundings to help the hearing impaired become aware of their surroundings.

2014 **Summer Internship**, *NIT*, Mentor: Anirudh Koul, Microsoft.
Automated Pipeline for Machine Learning Problems: 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 our trained 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. Poster presentation at **Grace Hopper, 2015**

2014 **Course Project, Natural and Artificial Intelligence**, *ShanghAI Lectures, University of Zurich, Switzerland*, Mentor: Prof. Martin F. Stoelen.
Swiss Robots with Adaptive Morphology: Attributed vision and perception to Swiss Robots to help them navigate the environment.

Professional Activities

Reviewer and organization

2016 **Women in Machine Learning Conference (WiML)**.

2016 **Language Technology Institute Student Research Symposium**, CMU.

Workshop and tutorial organization

2016 **Open Source Day**, Grace Hopper Conference.

2014 **International Workshop on Machine Learning Algorithms and Data Analytics**, Thapar University, Patiala, PB, India.