Siddha Ganju

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

Current appointments

2017-Present **Deep learning 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.

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*.

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

2014 **Winner, Grace Hopper Conference Hackathon**, *Bangalore, India*.

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*.

Speaking Engagements

2018 **Being smarter than dinosaurs: How NASA uses Deep Learning for Planetary Defense**, *Strata Data Conference, San Jose, USA*.

2017 **Embedded Deep Learning: Deep Learning for Embedded Systems**, *O'Reilly Artificial Intelligence, San Francisco, USA*.

2017 **Embedded Systems and Deep Learning**, *Global Big Data Conference, Santa Clara, USA*.

2016 **Atom smashing using machine learning at CERN**, *Strata+Hadoop World, San Jose, USA*.

Panel Discussions

2017 **The road to becoming a Data Scientist**, *Global Big Data Conference, Santa Clara, USA*.

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

Media

Contouring learning rate to optimize neural nets, *O'Reilly Media*, August 17, 2017.

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 Spotlight **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.

Peer-reviewed conference workshops

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

CVPR 2017 **What's in a Question: Using Visual Questions as a Form of Supervision**, ***S Ganju***, *O Russakovsky, A Gupta*, VQA Challenge Workshop , 2017.

NIPS 2017 **In Search of Long-Period Comets with Deep Learning Tools**, *S Zoghbi*, M DeCicco*, A P Stapper*, A J Ordonez*, J Collison*, P S Gural,* ***S Ganju****, J L Galache, P Jenniskens*, Women in Machine Learning (WiML) Workshop, 2017.

NIPS 2017 **Searching for Long-Period Comets with Deep Learning Tools**, *S Zoghbi*, M DeCicco*, A P Stapper*, A J Ordonez*, J Collison*, P S Gural,* ***S Ganju****, J L Galache, P Jenniskens*, Deep Learning for Physical Science Workshop, 2017.

GHC 2015 **Automated Python Pipeline for Machine Learning Problems**, ***S Ganju****, A Koul*, Grace Hopper Conference, 2015.

Talks

IMC 2017 **Artificial Intelligence Techniques applied to Automating Meteor Validation and Trajectory Quality Control to Direct the Search for Long Period Comets**, *M DeCicco*, S Zoghbi*, A P Stapper*, A J Ordonez*, J Collison*, P S Gural,* ***S Ganju****, J L Galache, P Jenniskens*, International Meteor Conference, 2017.

GTC 2017 **In Search of Long-Period Comets: Deep Learning tools for Automatic Meteor Classification and Shower Characterization**, *S Zoghbi*, M DeCicco*, A P Stapper*, A J Ordonez*, J Collison*, P S Gural,* ***S Ganju****, J L Galache, P Jenniskens*, GPU Technology Conference, 2017.

MozFest 2015 **Open Cosmics: Cosmic-Ray Physics for Everyone**, *S Gupta*,* ***S Ganju****, *A Bose*, G Azzopardi*, H Urhan**, Mozfest, 2015.

*=Equal Contribution

Research Experience

2017 **Deep Learning Coach**, *NASA*, Frontier Development Lab.

Mentored deep learning researchers to develop machine learning and deep learning tools to discriminate meteors from other detections. Designed Python scripts that automate the data collection, run the CAMS software, and compare the calculated orbit with a meteor shower template file based on showers assigned by Peter Jenniskens. Developed a new web tool that displays each shower radiant in sun-centered ecliptic coordinates. Talks and posters presented at **International Meteor Conference 2017**, **GPU Technology Conference 2017** and **Neural Information Processing Systems 2017**.

2016 **Research Extern**, *CMU*, Mentors: Prof. Olga Russakovsky, Prof. Abhinav Gupta.

Research focused on weak supervision: Utilizing supervision from visual questions asked about images. Spotlight presentation & poster at the **IEEE Computer Vision and Pattern Recognition conference, 2017.**

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**.

2014 **Summer Intern**, *NIT*, Mentor: Anirudh Koul, Senior Data Scientist, 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**

Professional Activities

Volunteer 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.