Vinay Kumar Kola

251 10th St NW, THF 313A, Atlanta, GA-30318

http://vinaykola.com

Phone: +1-404-490-6361

Email: vinaykola@gatech.edu

Education:

M.S. in Computer Science   expected December 2014

Georgia Institute of Technology, Atlanta, GA   GPA: 4.0 (after 2 semesters)

B. Tech. (Honors) in Computer Science and Engineering   March 2013

Indian Institute of Technology, Kharagpur, West Bengal, India   CGPA: 7.88/10.0

Professional Experience:

Internship at Facebook   May '14 – July '14

Data Engineer - Data Engineering Tools team

- Built the integration for Vertica (a column-oriented database) into the data discovery search engine at Facebook, which consisted of:
    - Designing the ETL pipeline to query Vertica system metadata for the purposes of the data discovery search engine
    - Building the front-end using PHP and Javascript (using React framework)
- Hence increased the productivity of all data analysts and data scientists within the company by making it easier to find, annotate and collaborate on data stored in Vertica tables

Graduate Teaching Assistant   August '14 – present

- Computational Science and Engineering Algorithms

Internship at Amazon.com Inc.   May '12 – July '12

Software Development Engineer - eCommerce Platform team, specifically the Authentication team

- Designed and built a real-time metric aggregator, a system to aggregate and perform calculations on software metrics (such as the number of successful sign-ins) using Java
- Applied concepts from shared-memory multiprocessor design to build a scalable and efficient task queue model for the aggregator, showing innovation
- Made the task of computing ratios, differences, etc. of multiple metrics easier for other engineers across all teams in Amazon and would have prevented a major outage in Amazon China which was the incentive for my project

Internship at Ozonetel Systems Pvt. Ltd.   May '11 – July '11

- Worked on a cloud telephony web framework called KooKoo and built the KooKoo - Google Places integration using PHP and Google Places API, making it accessible to low-end phones too instead of only smartphones

Technical Skills:

Programming Languages: Python, PHP, Javascript, SQL, C, Java   Software: NumPy, SciPy,Hadoop, Git, Mercurial

Academic Projects:

Politeness classification using Support Vector Machine (SVM)   Dec '13 – May '14

Under the guidance of Prof. Eric Gilbert

- Extracted a set of 20 features from a sentence that model politeness using Stanford CoreNLP tools and Python to train an SVM to give near-human accuracies
- Built an API using Flask that returns the politeness score of a queried sentence.

Machine learning based approach to GPU/CPU Program Partitioning   July '12 – May '13

B.Tech. Thesis, under the guidance of Prof. Pabitra Mitra

- Used machine learning algorithms for program partitioning between GPU and CPU executions
- Involved extraction of code features from a program implemented in C and Nvidia CUDA, by static and dynamic analysis, and using them as an input feature vectors to an SVM

JARVis   Jan '14 – May '14

- Created an interactive website using NodeJS to allow comic book fans to explore the comic universe by finding characters similar to their favorite ones and recommending popular story arcs as well.
- Key features included developing a similarity metric for characters (Python) and developing a UI using D3 that doesn't overwhelm the user. We used Neo4J to model the character graph for performance reasons.

Banshee   Jan '14 – May '14

- We attempted to check the hypothesis that music listeners listen to gloomier songs on a rainy day. Used the Million Musical Tweets Dataset, weather data from NCDC and song data from Echo Nest as data sources.
- Also performed topic modelling on the lyrics of the song using Latent Dirichlet Allocation to uncover differences in the themes of songs played during different weather scenarios.

Stack Overflow Tagger using Latent Dirichlet Allocation (LDA)   August '13 – Dec '13

- Analysed the Stack Overflow dataset using LDA to model each question as a tag distribution and each tag as a word distribution using Python and used these distributions to predict possible tags for new questions