Rishita Anubhai

+1 (650) 704 9220 • rishita@cs.stanford.edu

Education

Stanford University **Stanford, CA**
Master of Science, Computer Science *2011–2013*
Relevant Coursework: Machine Learning, Information Retrieval & Web Search, Mining Massive Datasets, Natural Language Processing

Institute of Technology, Nirma University **Ahmedabad**
Bachelor of Technology, Computer Engineering *2007–2011*

Experience

Baidu Research, Silicon Valley AI Lab **Sunnyvale, CA**
Machine Learning Engineer *September 2015–Current*

- Worked on Deep Speech 2 and its productization efforts..
- Currently investigating natural language understanding and dialog agents.

Microsoft **Redmond, WA**
Software Development Engineer *July 2013–September 2015*

- Data Platform Group - Proactive Analytics: Designed and implemented statistical and machine learning algorithms to detect anomalies in large scale telemetry datasets at near real-time latency. These key insights now drive livesite action for the performance and high availability of SQL Azure.
- Presented at Microsoft's internal machine learning conference to discuss Segmented Multiple Linear Regressions. This method was used to identify slow trends in time series signals that have adhoc reverse sawtooth patterns.

Facebook, Inc. **Menlo Park, CA**
Software Engineering Intern *June 2012–September 2012*

- Implemented a asynchronous and secure file upload component that worked across different user flow.
- Implemented a source code analyzer to support migrating to XHP from PHP. The analyzer informed clients once PHP code in a module had been converted to XHP, and enabled removal of the compatibility layer. (The compatibility layer wrapped XHP in strings and enabled engineers to incrementally migrate from string concatenations to XHP, the preferred and more secure language).

Stanford University **Stanford, CA**
Research Assistant to Professor Dan Boneh in the Stanford Security Laboratory. *September 2012–December 2012*

- Researched SSL certificate validation libraries for different languages and platforms. In addition to this, I also explored the use of GLSL in WebGL to see if cryptographic functions could be implemented in browsers.

Stanford University **Stanford, CA**
Teaching Assistant to Professor John Ousterhoust for Operating Systems (CS140) and Web Applications (CS142). *January 2013–June 2013*

- Conducted office hours, helping with exam question design and answering questions regarding the material taught in class.

Academic Papers and Projects

Deriving Patient-Similarity Metrics *Independent Study at Stanford Medical School* **April 2013–June 2013**

Worked on deriving effective patient similarity metrics based on medical records, for further analysis. I had to understand large datasets and explore various similarity metrics. This independent study was done under the guidance of Dr. Nigam Shah at the Biomedical Informatics Research (BMIR) center.

Event Extraction from Biological Text *CS 224U: Natural Language Understanding* **January 2013–April 2013**

Used Stanford NLP toolkits, feature engineering and experimentation to identify events and their agents in biological texts. This was done with the aim of gaining some semantic understanding of the facts in the text for eventual applications like question answering.

Predicting Rainfall in Rural India *CS 229: Machine Learning* **September 2012–December 2012**

We worked on rainfall and flood prediction for rural India in a group of three. This involved dataset exploration, feature engineering and experimentation with multiple algorithms ranging from basic ones like linear regression to non linear neural networks. Our method (Contextual window based neural networks) gave significant improvements over the out of the box methods. The final results were comparable to the state-of-the-art.

The most dangerous code in the world: validating SSL certificates in non-browser software
M. Georgiev et, al., Proceedings of ACM CCS '12, pp. 38-49, 2012

Awarded CSAW AT&T best paper in applied security.

Skills

Languages: C# (fluent), C/C++, R, Python, Java
Frameworks & Toolkits: Matlab, Stanford NLP
Data paradigms and stores: SQL Server, MySQL, MapReduce
Web technologies: PHP, HTML, CSS
Operating Systems: Linux/Unix, OS X, Windows