# Nikunj Yadav

Mobile: 512-201-0676
Email: nikunj@cs.utexas.edu
Web: www.nikunjyadav.com

## Objective

Research and innovate in the field of computer science and produce tangible results which can be used widely.

## Educational Qualification

Bachelor of Technology (Information Technology) 2006-2010
*Indian Institute of Information Technology Allahabad*
*Cumulative Grade Point Average: 9.4/10.0*

All India Senior School Certificate Examination 2005-2006
*Central Board of Secondary Education*
*Aggregate Percentage: 90%*

## Research Interests

Pattern Recognition, Information Retrieval, Natural Language Processing, Distributed and Workflow Systems

## Research Experience

**Research Intern at Intel** ***(December 2009 – June 2010)***

At Intel I worked on high level modeling and validation of processors. I was involved in developing models which could be served for the validation of processor at the architectural stage itself. During this period, I had insight into several software design choices and different concepts of a processor. I was part of a team in co-writing two models, one for validation of power algorithm and another for cache coherency of a processor. I was also involved in writing a decoupled test bench for these models as well and writing some tests which could run on regression.

**Research Assistant at Natural Language Processing Lab** ***(June 2007 – November 2009)***

I worked at lab called TDIL (Technology Development of Indian Languages Lab, http://tdil-dc.in), funded by Ministry of Communications and Information Technology. This is a national group of labs situated across several premier colleges of India and I was part of this lab during my undergraduate study at Indian Institute of Information Technology Allahabad. My work included majorly semantic document classification, and grammar exchange for translation from English to Hindi, another paper that I am currently working on.

## Publications

1. Nikunj Yadav, Yanu Gupta, Manish Kumar, and Ratna Sanyal.
   *Semantic Document Classification and Keyword Spotting in Digital Repositories.*
   Presented at ACM MEDES'09 Lyon France
   Published in proceedings of International Conference of Management of Emergent Digital Ecosystems Pg. 157-161, ACM, New York 2009
2. Nikunj Yadav, Yanu Gupta, Manish Kumar and Ratna Sanyal (2010)
   *Semantic Classification, Keyword Mining and Search Space Optimization for digital ecosystems*. Journal of Multimedia Processing 1(2): P. 131.

## Work Experience

**Software Development Engineer at Amazon.com (November 2010 – Present)**

I work in the Global Payments group at Amazon, as part of a team of 4 engineers called CROW which is responsible for developing and maintaining the refund workflow for the whole of Amazon. I have been working on service oriented architecture, innovating solutions which improve the availability and scalability of a service. Being involved in CROW, I have obtained in depth exposure to working of a stateful workflow system, handling large requests volume, host optimizations, database optimizations and other details involved in maintenance of service. My major work includes delivery of refund scenarios of several Amazon world-wide impacting projects, integration with new services, several scaling and latency optimizations, development of new features, innovating to improve experience of the customers, optimized data base and host configurations. I also have complete ownership of UI of our refund system which is used within Amazon.

### Member of Technical Staff at Oracle (July 2010 – October 2010)

In my short stint with Oracle I worked on a performance evaluation tool called Hermes for a team collaboration project called OnTrack. This tool was basically a multi-threaded demon which filed different types of requests to different API on the service, evaluated response times and calculated several statistics. My work was inclusion of several on track api calls in the Hermes suite, understanding how Ontrack works, and developing small tools like web crawler, news feed reader which helped in gathering data for Hermes.

## Projects

**High Level Modeling and Validation of Microprocessors**

This project was carried out at Intel India. In this project my work was to write high level model for a particular concept of intel processor which was still in development stage. I had written model for power algorithm of a processor, power algorithm for a multi core processor with a master slave combination among them, and cache coherency. I was also exposed to writing a de-coupled test bench which could run tests on the model. Apart from skills of software design, other skills acquired were computer architecture and microprocessors concepts.

**Project Guide**: Raj Bahadur Singh Rajput, Tech Lead at Intel

**Semantic Document Classification and Keyword Spotting using Probabilistic Semantic Analysis**

The Project focused on classifying the documents semantically, and finding keywords in the document which make the query retrieval efficient. Research employed modified PLSA by starting at a trained point on the probability function curve so that the subsequent EM iterations lead at an optimum point, thus giving a classification into labeled classes.

**Project Guide :** Dr. Ratna Sanyal, Coordinator(UDL,TDIL), IIIT-Allahabad

**An efficient and fast parser for translation from English to Hindi using TAG Formalism**

Developed parsing technique which used patterns specific to English –Hindi grammar, thus making the parsing of TAG much more efficient O(n) and the tree-bank shorter (n is the length of sentence, number of trees were 10 because of the patterns observed, superfluous trees were removed). During this work, I designed an efficient clause breaker, thus to make the translation at the clause level rather than the phrase level, which gave better results.

**Project Guide :** Dr. Ratna Sanyal, Coordinator(UDL,TDIL), IIIT-Allahabad

**Query Retrieval using LDA**

The project focuses on using LDA to extract semantic knowledge from the documents. Each word in document is related to a topic and LDA is used to relate the documents to concepts. Based on the semantic knowledge the documents are classified and further query retrieval is done aiming at higher efficiency with respect to the semantics.

**Project Guide :** Dr. Ratna Sanyal, Coordinator(UDL,TDIL), IIIT-Allahabad

**Proto-Symbol Manipulation for HOAP-2 using HMM**

The project aimed at mimesis theory, the learning of segmented motion patterns through HMM and then extrapolating or interpolating them together to make new patterns for the movement of HOAP-2 robot.

**Project Guide** : Dr. G.C. Nandi, Professor, IIIT-Allahabad

**Virtual 3D Tour of IIITA**

The project used the OpenGL library with VC++ to make a virtual tour of IIIT-A. The view was 3D and could be seen from any orientation, the viewer camera was keyboard controlled.

**Project Guide :** Dr. Pavan Chakraborty, IIIT-Allahabad

## Achievements

- Among top 10 (ranked 6$^{th}$) at Indian Institute of Information Technology Allahabad.
- Among the Top 0.1% students (in India ), All India Sr Secondary Exam in Mathematics, 2005-06
- Silver Medal in Science stream in School in All India Sr Secondary Certificate Exam, 2005-06
- First runner up in Web Designing Competition, Effervescence 2007, IIIT, Allahabad
- Indian Association of Physics Teachers awarded certificate for being among top 10% in National Standard Examination in Physics, 2005-06
- Silver Medal in School in All India Secondary Certificate Exam, 2003-04
- Among Top 50 students selected in India for ACM – ICPC 2009.

## Leadership Skills

Very Actively Involved in Hiring, taken more than 50 Technical Interviews at Amazon

Lead Website Designer of Effervescence 2008 (Annual Cultural & Technical fest, gets international level technical participation, http://effervescence.iiita.ac.in/2008/flash_site/)

Technical Events Coordinator and Network Administrator, Effervescence 2008. (Managed Sponsors for Technical Events like Microsoft and Intel)

Organizer of Effervescence 2007 as an Executive member, PMP club (Program Management and Publicity Club, consist 7 members only), IIIT- Allahabad

Treasurer, Music Club 2007-08, IIIT - Allahabad

Website Designer, Program Management and Publicity Club 2006-07 , IIIT-Allahabad

## Social service

Financial aid for education and other necessities of a child through an NGO Action Aid 2011.

## Hobbies

Vocal Music, Guitar and Drums.
Awarded Best Bassist in Rock Band Competition at Effervescence 2009
Reading books, deep interest in Indian mythology, swimming.

## References

Dr. Sudeep Sanyal
*Professor*
*Indian Institute of Information Technology Allahabad, India*
Phone: +91- 9415235180
E-mail: ssanyal@iiita.ac.in

Dr. Ratna Sanyal
*Coordinator, Technology Development of Indian Language Laboratory*
*Indian Institute of Information Technology Allahabad, India*
Phone: +91- 9793403758
E-mail: rsanyal@iiita.ac.in

Praful Mattoo
*Senior Software Development Manager,*
*Amazon.com, Bangalore, India*
Phone: +91-9986018363
E-mail: praful@amazon.com