

Contact Information	739 E El Camino Real Apt 112 Sunnyvale, CA-94087	<i>Email:</i> cvkkumar@cs.stanford.edu <i>Ph:</i> 650-384-9476
Interests	Large Scale Data Mining and Analysis, Machine Learning, Information Retrieval and Software Engineering	
Education	Stanford University, M.S. Computer Science	<i>Sep '08 - Apr '10</i> GPA : 3.86 / 4.0
	<ul style="list-style-type: none">• Recipient of the prestigious <i>BPCL Scholarship</i> awarded by Bharat Petroleum Corporation Ltd. awarded to the elite few for Graduate Study abroad.• <i>Relevant Coursework</i> : Machine Learning, Data Mining & E-Business (A+), Computational Advertising(A+), Information Retrieval & Web Search (A), Probabilistic Graphical Models(A), Data Mining(A), Transaction Processing and Distributed Databases, Network Analysis (A-), Biomedical Systems Design(A+), , Modern Applied Statistical Learning ,iPhone Application Programming	
	Birla Institute of Technology and Science, Pilani - Goa Campus M Sc.(Tech.) Information Systems	<i>Aug 04 - Jun 08</i> CGPA : 9.95/10.0
	<ul style="list-style-type: none">• Ranked No. 1 in Information Systems Department• Recipient of the Silver Medal awarded by BITS to the second rank holder in the graduating class of 2008.• <i>Relevant Coursework</i> : Artificial Intelligence, Data Structures & Algorithms, Software Engineering, Probability & Statistics, Discrete Structures for Computer Science, Data Mining.	
Experience	Apple Inc., <i>Software Engineer</i> <i>Internet Services Group</i>	<i>Apr '10 to Mar '13</i>
	<ul style="list-style-type: none">• Worked on Spam and Fraud Detection problems using Machine Learning and large scale data mining techniques.	
	Center for Study of Language and Information, Stanford University <i>Graduate Research Assistant</i> <i>Cognitive Assistant that Learns and Organizes (CALO)</i>	<i>Aug 08 to June 09</i>
	<ul style="list-style-type: none">• Member of a team involved in the addition of new features to CALO, a system that extracts decisions from multi-party meetings to enable the effective handling of feedbacks.• Also involved in evaluation of the Service Manager• Implementation used Java, Swing and libSVM	
	Apple Inc., <i>Summer Intern</i> <i>Internet Services Group</i>	<i>June 09 to Mar 10</i>
	<ul style="list-style-type: none">• Worked on automated sentiment analysis and opinion mining from Micro-blogs using a combination of sophisticated machine learning and data mining techniques.• Member of a two-member team that was responsible for the entire project from its conception and design to implementation and production deployment.	
	Hewlett Packard Labs-India <i>Research Intern</i> <i>STAIR : System for Topical and Aggregated Information Retrieval</i>	<i>Jan 08 to June 08</i>
	<ul style="list-style-type: none">• Developed the architecture and the prototype of STAIR - an IR system that applied of a combination of Collaborative analysis and Focused Crawling techniques on the web documents to provide personalized, consolidated information relevant to the used as an aggregated PDF document.• Implemented using Java, on top of Lucene. The semantics information were obtained from WordNet• Published as a HP Labs Technical Report in 2009.	
	Bhabha Atomic Research Center, Trombay <i>Project Intern at DRHR</i> <i>Image Processing and Software Development for Simplifying Robot Trajectory Generation</i>	<i>May 06 to July 06</i>
	<ul style="list-style-type: none">• Built a system that extracts information of a continuous path from any arbitrary raw image using graph-theoretic methods and provides input to the indigenous <i>Sensor-cum-Manipulator</i> ,a Parallel planar kinematic robot.• Implemented using C, Matlab(for image processing) and VB (as a wrapper GUI)	
Selected Projects	<i>Role Discovery in Social Networks using DMR Based Topic Modeling</i> <ul style="list-style-type: none">• <i>Instructor: Prof. Daphne Koller, Stanford University</i>	

Recommendation Systems based on Delicious and Twitter

- Implemented a People Recommendation System on Twitter (in Python) by a combination of several algorithms, that included collaborative filtering, network analysis and semantic consideration .
- Designed and implemented an URL Recommendation system by analysis of tags from *Delicious*

URL Recommendation Based on Asymmetric Tag Similarity and Diffusion-Based Grouping

- Implemented an URL Recommendation system by analysis of tag similarity using data from ShareThis and Delicious using MapReduce and Partition based Joins on top of the Aster Cluster.

Finding Answerers on Yahoo! Answers

- Designed and implemented a system for selection of most appropriate answerers in *Yahoo! Answers* based on textual, structural and auxiliary information. The result can be used to determine the routing for the new questions.

Analysis of Text Based Classifiers

- Implemented and analyzed the performance of different Naïve Bayes classifiers on the 20-Newsgroups dataset, using Java and Lucene

Comparison of Similarity Search Algorithms over Inverted Indexes

- Implemented and analyzed the performance of commonly used indexing similarity search algorithms - Term-at-a-Time and Document-at-a-Time. Also optimized the algorithms with efficient index compression.

RefMed - A Physician Referral and Review Service

- Designed and developed a physician referral and review service that enables patients to review and rate the physicians and facilitates the physicians to recommend other doctors to their patients.

An Ontology-based Automatic Staging system for Cancer

- Developed a Automatic Staging System for Breast Cancer over the NCI Thesaurus using SWRL, OWL and Protégé

Implementation Of a Search Engine for Personalized Information Retrieval by profiling of user data

- Designed and implemented a prototype search engine that incorporates an additional dimension of Personalization through User Profiling for Enhanced relevancy.

Time Table Generator

- Implemented a system for automating the process of timetable generation for a University with its constraints, using a sub-optimal graph coloring approach for Constraint Satisfaction.

An Expert System for selection of Polymer Composite Systems

- Implemented a novel method for evaluation and ranking of constituent materials for composite products using *TOPSIS* , an MADM (Multiple Attribute Decision Making) approach that ensures an optimum solution for the characteristics desired.

Publications

C.V.Krishnakumar and Dr.Krishnan Ramanathan, *STAIR : A System for Topical and Aggregated Information Retrieval* , Proceedings of the International Conference on Intelligent Human Computer Interaction (IHCI) 2009.

R.T. Durai Prabhakaran, B.J.C. Babu, V.P. Agrawal, C.V. Krishna Kumar, *A knowledge-based system for constituent material selection in polymer composite product design* - Proceedings of ISRS-2006, International Symposium for Research Scholars, IIT-Madras.

Skills

Languages: Java, Pig, Python, C, SQL
Frameworks: Hadoop (Cloudera Certified Hadoop Developer), HBase
Tools and Platforms: Mahout, Teradata, Weka, R, Eclipse, \LaTeX , Octave
Operating Systems: Unix based Systems, Microsoft Windows.

Honors and Achievements

Consistent recipient of the *Merit Scholarship* awarded by BITS Pilani to the top 10 students across the batch.

Secured the First Prize at *OpenSoft* - the software construction contest conducted as a part of QUARK-07, the national level technological fest at BITS-Pilani, Goa Campus.

Recipient of the Merit Certificate, awarded to the top 0.1% of students, for proficiency in English in AISSCE from the CBSE, 2004.