

**Research Interest** Machine Learning, Social Media, Information Extraction & Retrieval, Recommender Systems

**Education** **Carnegie Mellon University** 2010 - Present  
MS Student at Language Technologies Institute (**GPA: 4.0/4.0**) Advisor: Eric Xing  
**Courses:** Machine Learning for Ph.D. ( $A^+$ ), Algorithm for NLP ( $A^-$ )  
Language & Statistics, Social Media Analysis, Information Retrieval (in progress)

**Indian Institute of Information Technology & Management** 2003 - 2008  
Bachelor of Technology and Masters of Technology in Information Technology

**Work Experience** **Carnegie Mellon University** (with Eric Xing) Graduate Research Assistant  
Pittsburgh, PA, USA Aug 2010 - Present  
Genome-Phenome Association using structured sparse regression algorithms. Algorithms to learn sparse graph structure. Implementing these algorithms to efficiently process large datasets.

**Microsoft Research** (with Raghavendra Udupa) Research Intern  
Bangalore, India Jan 2010 - June 2010  
Worked on a keyword extraction algorithm for effective email retrieval. (A patent application is being filed currently)

**IR Group, University of Glasgow** (with Joemon Jose) Research Assistant  
Glasgow, UK June 2008 - Dec 2009  
Worked on feature selection/reduction algorithm for Video Retrieval, algorithms to promote diversity in retrieval & Video segmentation techniques based on visual and textual features.

**University of Applied Sciences** (with Kai-Uwe Barthel) Research Intern  
Berlin, Germany May 2007 - July 2007  
Implemented Image Sorting & Searching algorithms.

**Other Projects** **Ranking Tweets in User's Feed**  
Concepts from Personalized Retrieval, Collaborative Filtering and Graph Analysis are integrated to rank tweets in the user's social stream.

**Multi-Instance Multi-Label Classification of Gene Expression**  
Addressed the problem to classify Gene Expressions Images having more than one bag of features, more than one labels and having no specific association between feature bags and labels.

**Topic Modeling For Multimedia Retrieval**  
Implemented LDA based algorithm to reduce dimensionality of visual features and to match visual query with video documents for retrieval.

**Language model for part-of-speech (POS) tags**  
Implementing algorithms for N-gram, Maximum Entropy, Decision-Tree and HMM language models for POS tagging. (ongoing)

**Qaafiyaa - Couplet generation using Twitter feeds**  
It uses syntactic & semantic similarity between twitter feeds to generate next lines of a poem rhyming with the given first line.

**Achievements** **National Winner (India) - Lord of the Codes** 2007  
Organized by IIT-Bombay and RedHat

**National Winner (India) - IBM Great Mind Challenge** 2006  
IBM's National Level Software Design Contest (7000 Teams Participated)

<b>Computer Skills</b>	C, C++, Java, C#, Perl Matlab, R, Weka, SPSS, MALLETT Hadoop	
<b>Certifications</b>	IBM Certified Database Administrator (DB2) IBM Certified J2EE Associate/Solution Developer (Rational) IBM Certified XML Solution Developer	2005 2005 2006
<b>Publications</b>	<p><b>Promoting Temporal Diversity in Video Search Result Lists (RIAO 2010)</b> Martin Halvey, Anuj Goyal, Robert Villa, Joemon M. Jose</p> <p><b>Feature Subspace Selection for Efficient Video Retrieval (ACM MMM 2010)</b> Anuj Goyal, Reede Ren, Joemon M. Jose</p> <p><b>TV News Story Segmentation based on Semantic Coherence and Content Similarity (ACM MMM 2010)</b> Hemant Misra, Frank Hopfgartner, Anuj Goyal, Punitha P. and Joemon M. Jose</p> <p><b>Topic Prerogative Feature Selection using Multiple Query Examples for Automatic Video Retrieval (SIGIR 2009)</b> Punitha P., Joemon Jose, Anuj Goyal</p> <p><b>Split and Merge based Story Segmentation in News Videos (ECIR 2009)</b> Anuj Goyal, Punitha P., Frank Hopfgartner, and Joemon M. Jose</p> <p><b>Diversity, assortment, dissimilarity, variety: A study of Diversity Measures using Low Level Features for Video Retrieval (ECIR 2009)</b> Martin Halvey, David Hannah, P. Punitha, Robert Villa, Frank Hopfgartner, Anuj Goyal and Joemon M. Jose</p> <p><b>Optimising Similarity and Diversity in Image Retrieval (ImageCLEF 2009)</b> G. Zuccon, T. Leelanupab, A. Goyal, M. Halvey, P. Punitha, J. M. Jose</p> <p><b>Glasgow University at TRECVID 2009</b> P. Punitha, Hemant Misra, Reede Ren, David Hannah, Anuj Goyal, Robert Villa, Joemon M. Jose</p> <p><b>Improved Image Retrieval Using Visual Sorting and Semi-Automatic Semantic Categorization of Images (WIAMIS 2008)</b> Kai Uwe Barthel, Sebastian Richter, Anuj Goyal, Andreas Follmann</p> <p><b>Glasgow University at TRECVID 2008</b> P. Punitha, T. Urruty, Y. Feng, M. Halvey, A. Goyal, D. Hannah, I. Klampanos, V. Stathopoulos, R. Villa, J. Jose</p> <p><b>Robust Watermarking in Transform Domain using Edge Detection Technique (IC-CIMA 2007)</b> Naman Agarwal, Anuj Kumar Goyal</p> <p><b>Robust Watermarking using Watermark Spread Canvass in DCT &amp; DHT Domain (SPPRA 2007)</b> A.K. Goyal, N. Agarwal, V. Bahri, S. Verma, A.K. Saxena</p> <p><b>Robust Watermarking in DCT and DHT Domain using Edge Detection Techniques (ICSIP 2006)</b> A. K. Goyal, N. Agarwal, V. Bahri, S. Verma</p>	
<b>References</b>	Available Upon Request.	