

Manisha Verma

(+44) 7823890600, manisha.verma.13@ucl.ac.uk
http://www0.cs.ucl.ac.uk/staff/M.Verma/

Education	University College London (UCL) <i>PhD/MPhil in Computer Science</i>	2013 - Present Advisor: Dr Emine Yilmaz
	Search and Information Extraction Lab (SIEL) <i>MS (by Research) in Computer Science</i>	2009 - 2012 CGPA: 8.4/10
	Jaypee Institute of Information Technology <i>Bachelor of Technology in Computer Science</i>	2005 - 2009 CGPA: 8.1/10

Research Interest Information Retrieval and Information Extraction

Work Experience **ContextScout**
Data Science Intern *June 2016 - August 2016*
Was responsible for building analytics infrastructure to mine potential customers and training models to detect active users of context scout chrome extension.

Google Ads Team
Phd Software Intern *June 2015 - August 2015*
Was responsible for developing system to extract commercial user search tasks for ad targeting. It involved evaluating and testing several classification models with user, query and session oriented features.

Naukri.com Search team
Senior Software Engineer *July 2012 - Sept 2013*
Was responsible for developing models for information extraction from resumes and job description pages. We employed supervised approaches that exploit both html dom structure and its content to extract phrases and fields like title, location, experience etc to convert semi-structured data into structured data.

Microsoft R&D Ad Sciences team
Research Intern *July 2011 - Feb 2012*
We developed models to suggest bid phrases for an advertisement in sponsored search. We extracted features from landing pages, query and click logs to train models to identify candidate phrases. Our project won the MS-IDC innovation award in March 2012.

Search Engine and Information Extraction Lab (SIEL) IIIT, Hyderabad
Research Assistant *Jan 2010 - May 2011*
We developed a customized cross-lingual search engine for Patents in English, German and French. Our system with key phrase extraction and patent meta data stood 1st in CLEF-IP 2011 track. Query expansion techniques were explored to improve retrieval performance. Assisted Btech students to construct cross-lingual dictionary for patents using Wikipedia.

Publications	Search Costs vs User Satisfaction on Mobile, ECIR'17 <i>Manisha Verma, Emine Yilmaz</i>
	Overview of TREC 2016 Tasks Track, TREC'16 <i>Manisha Verma, Evangelos Kanoulas, Emine Yilmaz, Ben Carterette, Nick Craswell and Rishabh Mehrotra</i>
	Going Beyond Relevance: Incorporating Effort in IR, SIGIR DC'16 <i>Manisha Verma</i>
	Characterizing relevance on mobiles and desktop, ECIR'16 <i>Manisha Verma, Emine Yilmaz</i>
	Category based Task extraction, CHIIR'16 <i>Manisha Verma, Emine Yilmaz</i>
	On obtaining effort based judgments for Information retrieval, WSDM'16 <i>Manisha Verma, Emine Yilmaz, Nick Craswell</i>
	Overview of the TREC 2015 Tasks Track, TREC'15 <i>Emine Yilmaz, Evangelos Kanoulas, Manisha Verma, Ben Carterette, Nick Craswell and Rishabh Mehrotra</i>

Entity oriented Task Extraction from Query Logs, CIKM'14

Manisha Verma, Emine Yilmaz

Bringing Head Closer to the Tail with Entity Linking, ESAIR'14

Manisha Verma, Diego Ceccarelli

Relevance and Effort: An Analysis of Document Utility, CIKM'14

Emine Yilmaz, Manisha Verma, Filip Radlinski, Nick Craswell, Peter Bailey

Patent search using IPC Class vectors, PaIR'11

Manisha Verma, Vasudeva Varma

Exploring Key phrase extraction and IPC Vectors for Prior Art Search, CLEF'11

Manisha Verma, Vasudeva Varma

Applying Key Phrase Extraction to aid Invalidity Search, ICAIL'11

Manisha Verma, Vasudeva Varma

Projects

Characterizing relevance on mobiles and desktop

We collected and compared how relevance of search result pages change with device size. We conducted two user studies on mobile and desktop with same search tasks. We analyzed different aspects of these judgments: judging time, label distribution and features important to predict relevance on both devices.

Characterizing user effort in search result examination

Designed two user studies to explore and evaluate user effort involved in understanding a web page given a search query. We investigated several content and visual factors that impact cognitive effort required to evaluate a search result web page for an input query.

Evaluating impact of search tasks on Session search

We used entity oriented tasks to improve session search. We experimented with Markov models to expand search queries based on session context.

Extracting and Ranking Phrases for Patent Retrieval

Implemented a supervised approach to extract and rank candidate key phrases from a patent. The goal is to maximize a retrieval score like MAP or Recall depending upon the nature of search task. Query performance prediction and patent specific features were used to represent a key phrase.

IR and Infobox Suggestion System for Wikipedia

A search engine was developed for Wikipedia articles in English. The application also suggested Infobox attribute values for the pages without an Infobox. Rule based attribute ranking and slot filling were implemented to extract information from the input page.

Data Mining Algorithms for Mahout

Conversion of Data Mining/Machine Learning algorithms Linear Discriminant Function (LDF) and SPRINT into Map-Reduce paradigm. Both the algorithms were implemented and integrated in Mahout.

Skill Set

Programming Languages: Python, Java

Softwares: Lucene, Nutch, Hadoop, Weka, Apache UIMA

Research Talks

Entity oriented Task Extraction from Logs (Cambridge, Microsoft Research Reading Group)

Binary Task Extraction from Query Logs (Google, Zurich)

Exploring Mobile Centric Challenges and Opportunities in Pro-IR (Pro-Active workshop, ECIR'16)

Other

Responsibilities

Reviewer at AAAI'17, WWW'16, WWW'15 and TOIS journal

ISD Training Assistant at UCL

Python instructor for Scipy course at UCL

Python instructor for CodeFirst Girls (nonprofit) programming course at UCL

Ambassador for Post graduate women in Computer Science at UCL

Teaching Assistant (3 terms) for Information retrieval course at IIIT and UCL

Track coordinator in TREC Tasks track 2015 and 2016

Achievements

Our cross lingual patent retrieval system stood 1st at CLEF-IP 2011

Won Best Hack in Yahoo Hack-U at IIIT-H, Jan 2010