Abhishek Singh

Mahabirlane
Mussallahpur
Mehendru, Patna,
Bihar, India
✆ +91 9567071585
✉ abhishek.singh@research.iiit.ac.in

Areas of Interest

Information Retrieval, Natural Language Processing, Machine Learning, Deep Learning

Education

MS by Research in CSE (2015 - 2017)	**International Institute of Information Technology**, *Hyderabad*, India, CGPA: 8.67/10. **Advisers**: Dr. Vasudeva Varma, Dean R&D, IIIT-Hyderabad, Dr. Manish Gupta, Principal Applied Scientist, Microsoft, Hyderabad
Bachelor of Technology (2010 - 2014)	**B.Tech. in Computer Science and Engineering**, *Amrita Vishwa Vidyapeetham, Kerala, India*. CGPA: 8.31/10
Higher Secondary CBSE (2009)	**Central Academy**, *Rajasthan, India*. *Percentage: 81.2%.*
Secondary CBSE (2007)	**R.P.S Public School**, *Bihar, India*. *Percentage: 92.7%*

Experience

Full Stack Developer, Credii (July, 2015)

Implemented a full fledged domain lookup tool that returns domain of a company given name of the company as input. Information extraction through crawling and scrapping. Worked on product recommendation system.

Technologies: Ruby, Rails, Python, Redis

Teaching Experience

Teaching Assistant for Information Retrieval & Extraction Course (Aug, 2017 - Nov, 2017)

TA for a course of 120+ students. Conducted tutorial lectures, graded assignments and projects, held consistent offices hours to provide individual help to students.

Technical Trainer, Amrita University (July, 2014 - May, 2015)

Mentored undergraduate students to learn new programming languages and various FOSS technologies. Helped these students to make contributions to open source projects.

Learn IT Girl (Nov, 2014 - Feb, 2015)

Mentored a Romanian student to learn Python, Django and Postgres. She eventually created a web application. This web application help people find the best local dealer to repair their malfunctioning electronic gadgets.
Reference: https://sites.google.com/site/learnitgirl/home

Internship

Audio Fingerprinting Filters for FFmpeg (Dec, 2016 - April, 2017)

Implemented filter for FFmpeg which gives audio fingerprints for an audio which can be reused by other people for variety of applications. Some of its applications are: detecting duplicate songs, retrieving similar songs, can be used for recommending songs, clustering songs based on content/tags, identifying a fragment of a recording.

Environment: Linux, C

Reference: https://github.com/atana1/FFmpeg

Google Summer of Code(GSoC) 2016 (May, 2016 - August, 2016)

Implemented an Audio Information Retrieval (AIR) system for Peragro, that provides text based and content based search.

Environment: Linux, Python, C, C++, elastic search

Reference: https://github.com/AbhishekKumarSingh/pairs http://pairs.readthedocs.io/en/latest

Google Summer of Code(GSoC) 2014 (May, 2014 - August, 2014)

Worked for a virtual reality project of Italian Mars Society(IMS). Integrated the full body and hand gesture tracking of astronauts to the European MaRs Analogue Station(ERAS) virtual station, which enables an astronaut in the real world to control their avatar in the virtual world through body and hand gestures.

Environment: Linux, Python, C++, Kinect Xbox sensor, Blender, Tango, Sphinx, Visual Python

Reference: https://bitbucket.org/italianmarssociety/eras

Website and Pariksha application for Amrita FOSS Club

Developed a website for FOSS club at Amrita using Python, Django, Bootstrap, HTML, CSS. Also created a web application for conducting online exams.

SELECTED PROJECTS

Sandhan, CLIA, Jan 2017 - June 2017

Description   Cross Lingual Information Access System, a multi-institutional project funded by Department of Defence (Govt. of India) where my responsibilities are to contribute to the query processing pipeline which involves development of query translation and transliteration engine to enable cross language search in 9 different Indian Languages.

Wikimon: A Search Engine for Wikipedia

Description   Implemented a efficient and scalable search engine from scratch for the whole wikipedia dump. Supports normal as well as phrase based query.

Environment   *Python, Cython, NLTK*

References   open sourced at https://github.com/AbhishekKumarSingh/wikimon

Image classification

Description   Implemented a Random forest and ferns based image classifier on caltech dataset as a part of course project.

Person Recognition System(PRS)

Description   Implemented a PRS which is based on a 3D face recognition approach. The system recognizes a person by matching his/her features through a database. Solves many problems of 2D face recognition and has an accuracy of 94%.

Environment   *C#, OpenCV, Windows, Kinect sensor*

Page ranking based search result

Description   Implemented a mechanism for the search engine for a project used by Amrita Technologies for Business Incubator(ATBI) that uses page ranking method to return search results.

Environment   Python, Linux

Trident Shell

Description   Developed a shell compatible with Linux and Mac OSx. This open source software is currently under the Copyright process. Features implemented are tab auto-completion, history, support for the pipe command, Input/Output redirection, background processes, all external and shell built-in commands, added color to the output of some of the commands such as ls, grep, fgrep and egrep.

Environment   *C++, Unix*

References   open sourced at https://github.com/AbhishekKumarSingh/Trident

Open Source Contribution

Organization	**Fedora**
Description	Worked on writing **test suites** for **SSSD** that is a process which helps the system identify and authenticate remote users, local users and remote resources. Learned concepts of memory management library(**Talloc**) and event-driven programming (**Tevent**). Implemented tests for finding the user id and Input/Output module with **85** percent coverage. Also fixed some bugs.
Environment	*C, Cmocka unit testing framework, Unix, Git*
References	open sourced at https://github.com/AbhishekKumarSingh/sssd/

Organization	**Mozilla**
Description	Worked on expanding unit tests for **mozbase** with better code coverage. Mozbase provides utilities like Input/Output, basic networking, etc. Implemented unittest for the moznetwork module. Moznetwork is a utility that is a part of mozbase. It provides basic network functionality.
Environment	*Python, Unix, Pep8, Pyflakes, Git*
References	open sourced at https://bitbucket.org/abhisheksingh/mozbase

Organization	**Tor**
Description	Improved automated tests for Stem. Stem is a Python controller library that allows applications to interact with Tor. This project dealt with sanity checks and cleanup of the source code. Fixed and reported some major bugs.
Environment	*Python, Unix, Git*
Reported bug	https://trac.torproject.org/projects/tor/ticket/8157
	https://trac.torproject.org/projects/tor/ticket/8271
Fixed bug	https://trac.torproject.org/projects/tor/ticket/7926

Organization	**MoinMoin***(March 2012)*
Description	Improved the code for user settings and user registration by performing sanity checks and clean-up of the source code. Fixed and reported some major bugs.
Environment	*Python, Unix, Mercurial*
References	open sourced at https://bitbucket.org/abhisheksingh/moin-2.0

Organization	**Mozilla**
Description	Improved code quality by removing unsupported features. Conducted effective testing and sanity checks on the source code. Fixed bugs and submitted patches for the code.
Environment	*JavaScript, HTML, CSS*
References	Fixed bug: https://bugzilla.mozilla.org/show_bug.cgi?id=650322
	https://bugzilla.mozilla.org/show_bug.cgi?id=699041

Organization	**Ubuntu**
Description	Removed unwanted features from the code that were present in the bash-completion package.
Environment	*launchpad, Bazaar, Scripting*
References	sourced at https://bugs.launchpad.net/ubuntu/+source/bash-completion/+bug/818475

SKILLS & TOOLS

Language	C, C++, Python, Java (basic), Ruby, shell script
Web-dev	HTML, CSS, Javascript, Django, Ruby on Rails
Automation	Pytest, Python-unit-test framework, Cmocka-unit testing framework

Database	SQL, Elasticsearch, Redis
ML	Keras, Tensorflow, Pytorch, Scikit-Learn, Torch, Lua
VCS	Git, Mercurial, Bazaar

RELEVANT COURSEWORK

Machine Learning, Information Retrieval and Extraction, Topics in Information Retrieval, Topics in Natural Language Processing, Statistical Methods in Artificial Intelligence

SCHOLASTIC ACHIEVEMENTS

- Microsoft Research Travel Grant for presenting paper in CIKM'17 Conference.
- ACM SIGIR Research Travel Grant for presenting paper in CIKM'17 Conference.
- Selected for the prestigious **GSoC** (Google Summer Of Code)2014 & 2016 program.
- **Certificate of Appreciation** from my university for achievements in academics and FOSS (Free Open Source Software) during 2013-2014.
- Listed as one of the **top contributors** of SSSD (System Security Service Daemon) - one of the popular open source project.
  Reference: https://launchpad.net/sssd
- 1st rank in **Hack The Shell** (shell scripting) event conducted by **Vortex** at **NIT** (National Institute of Technology) Trichy, India.
- 2nd rank in **Crack The Shell** (shell scripting) event conducted by **CSI (Computer Society of India)** Kerala State Student Convention at Thangal Kunju Musaliar College of Engineering, Kerala, India.
- Awarded with 'Certificate of Merit' from the CIR (Corporate and International Relations) department at Amrita Vishwa Vidyapeetham. This certificate was awarded for an outstanding SSR (Student Social Responsibility) project.
- Our team ranked 3rd out of more than about 100 teams in **SCTF, Capture The Flag** hacking contest (August,2012).

DEMONSTRATED PERSONAL SKILLS

Communication skills	Addressed approximately 700 people during a health campaign organized as a part of 'Student Social Responsibility project'.
Leadership	Took up leadership roles in all academic and extra-curricular activities. I was one of the student co-ordinators for an event where a crowd of about 0.2 million was present. We managed the crowd successfully without any negative incidents.

CONFERENCE, PAPERS AND PUBLICATIONS

Title	**Hybrid MemNet for Extractive Summarization**
Authors	Abhishek Kumar Singh, Manish Gupta and Vasudeva Varma
	CIKM'17, November 6–10, 2017, Singapore, Singapore
Title	**Unity in Diversity: Learning Distributed Representation of Sentence for Extractive Summarization**
Authors	Abhishek Kumar Singh, Manish Gupta and Vasudeva Varma
	AAAI'18, February 2–7, 2018, New Orleans, Louisiana, USA
Title	**NASH: A Standard Neural Abstractive Summarizer for Hindi**
Authors	Abhishek Kumar Singh, Manish Gupta and Vasudeva Varma
	TACL, Journal (Under Review), 2017

ACTIVITIES

FOSS Club	An active member of Free and Open Source Software Club. Mentored many junior students learn new technologies.