Bharath **Venkatesh**

Research Engineer · Artificial Intelligence and Data Analytics

MALE · 08-05-1988 · INDIAN · Sint-Jansbergsesteenweg 101, 3001 Heverlee, Belgium

(+32) 496798363 | bharath.venkatesh85@gmail.com | bhatturam.github.io | bhatturam | bhatturam

Work Experience

SAP

Bangalore, India

Research Associate - Data Science — *Jul. 2013 - Jul. 2015*

I was a junior scientist at a next-generation applications team, which reported directly to the CTO at Palo Alto CA. My primary role was to identify relevant literature and make proof-of-concept implementations. I worked on various Data Mining and Analytics projects in the domains of Bioinformatics, Health and Internet of Things. I also contributed to software architecture and development as a secondary activity

Education

KU Leuven

Belgium

Advanced Masters in Artificial Intelligence — *Sep. 2015 - PRESENT*

- *Aggregate* 76% (15.2/20)
- *Courses* - Programming for Big Data (18/20), Machine Learning, Data Mining, Uncertainty in AI, Robotics (19/20), Computer Vision, Artificial Neural Networks, Support Vector Machines.
- *Thesis* - Exploring Deep Kernel Machines for Unsupervised Learning

IISc (Indian Institute of Science), Bangalore

India

Master of Science (engineering) in Computer Systems — *Aug. 2010 - Jul 2013*

- *Aggregate GPA* 6.2/8
- *Courses* - High Performance Computing (7/8), TCP/IP Networking (7/8), Probability and Statistics, Linear Algebra
- *Thesis* - Fast identification of structured P2P botnets using community detection algorithms.

NIT (National Institute of Technology), Tiruchirappalli

India

Bachelor of Technology in Mechanical Engineering — *Aug. 2006 - May 2010*

- *Aggregate GPA* 7.87/10
- *Thesis* - Design and Simulation of a MEMS Microgripper

Projects

User Profiling and Item Recommendations

SAP

In the Context of a Health and Fitness smartphone application, I worked on the design and proof-of-concept implementation of the recommendation component. The system combined collaborative filtering and content based approaches. I worked on designing the attributes that were collected from the user for building profiles. In this regard, I tackled the problem of identifying important places for a given user from his GPS traces.

Research Project on Topological Pathway Analysis

SAP

This was a bioinformatics research project aimed at ranking biological pathways based on how much they were affected in a disease state (cancer). The approach combined differential gene expression data from microarrays and gene-gene network graphs using PageRank. This work has been submitted for publication and the R package that I developed has been open sourced and is available on Github.

NGS Data Simulator

SAP

This was a bioinformatics based application designed for bioinformatics research. The aim of the application was to generate realistic Next-Generation DNA Sequencing data and introduce mutations. I contributed to the drafting of the specifications and came up with the architecture and developed the application from scratch (over 5000 LoC) in Java, SQL and Javascript (Server side REST APIs).

Rural Child Health Tracking and Analytics *SAP*

This was an IoT corporate social responsibility project in collaboration with an NGO which was interested in tracking child health and combating malnutrition in rural areas in Karnataka. I conceived, designed and lead the implementation efforts of a low-cost system to track the height and weight of children in Anganwadis(child shelters). I designed the data acquisition system using an Arduino, ultrasonic and strain sensors. I developed an android smartphone client application and REST services for data collection and transfer to the cloud for analytics.

Large Scale Graph Based Botnet Detection *IISc*

This project aimed at detecting Botnets at the Internet infrastructure level, where the data velocity is very high. The approach relied on detection of nearly regular subgraphs of a large IP-IP graph. I devised a linear time modified graph clustering approach in order to perform the detection. This work was published in a good computer security journal. During the course of this work, I developed a network science library in C++ which can be found at Github. It was used extensively in our lab at IISc

Malware Classification and Clustering *IISc*

I contributed to a machine learning based malware analysis project. In this work, system call traces of application were used to classify them into known malware families using Hidden Markov Models (HMM). Unknown samples were clustered using the outputs of these models to enable labeling and perform further analysis. This work was presented at a good security conference.

Deep Kernel Architectures for Unsupervised Learning *KU Leuven*

As a part of my thesis I am working on an exploratory research project where I am working on deep architectures consisting of kernel principal component analysis units to learn good representations for clustering. The implementation is in MATLAB.

Incisor Segmentation using Active Shape Models *KU Leuven*

For the Computer Vision course project, I implemented Model-based segmentation of incisors from dental radiographs used for forensic studies. I implemented a Point Distribution Model based on Prinicipal Component Analysis to model tooth shapes, and an iterative fitting algorithm that searches for the best position and pose parameters of the shape model. This was implemented in Python using the OpenCV libraries.

Large Scale Machine Learning and Data Mining Algorithms *KU Leuven*

As the main requirement of the Programming for Big Data course at KU Leuven, I implemented online classifiers- Hoeffding trees and stochastic gradient descent based algorithms, approximate nearest neighbor search using the Locality Sensitive Hashing and various approximate counting and sampling techniques for large datasets.

Publications

Preprint

Patil, SS., Venkatesh, B, and Singh, R. "From Differentiated Genes to Affected Pathways." bioRxiv (2016): 038901.

International Journal

Venkatesh, B., Choudhury, S. H., Nagaraja, S., Balakrishnan, N. (2015). *BotSpot: fast graph based identification of structured P2P bots.* Journal of Computer Virology and Hacking Techniques, 11(4), 247-261.

International Conference

Ravi, S., Balakrishnan, N., Venkatesh, B. (2013). *Behavior-based Malware analysis using profile hidden Markov models.* In 2013 International Conference on Security and Cryptography (SECRYPT)(pp. 1-12). IEEE.

Software Skills

Programming Languages	C/C++, Java, PERL, Python, R, MATLAB, SQL, PROLOG, Javascript
Parallel Programming	MapReduce, CUDA, OpenMP, MPI
Frameworks	Scikit-learn, WEKA, OpenCV, Arduino, Android, SAP HANA
Miscellaneous	Git, SVM, Linux System Administration