

# Vivek Nair

---

## CONTACT INFORMATION

3240 EBII, 890 Oval Drive  
Raleigh, NC, 27606  
*Github:* <https://github.com/vivekaxl>

*Phone:* (919) 523-9920  
*E-mail:* [vnair2@ncsu.edu](mailto:vnair2@ncsu.edu)  
*www:* <http://www.vivekaxl.com>

## EDUCATION

<b>North Carolina State University</b> Ph.D. candidate in Computer Science Advisor: Dr. Tim Menzies	Raleigh, NC May 2013 - Dec 2018 (Expected)
<b>National Institute of Technology, Durgapur</b> M.Tech in Information Technology	Durgapur, India 2011
<b>West Bengal University of Technology</b> B.Tech in Computer Science	Kolkata, West Bengal 2009

## IN PREPARATION

- **Vivek Nair**, Rahul Krishna, Tim Menzies, Pooyan Jamshidi. *Transfer Learning with Bellwethers to find Good Configurations*. 2018 → [http://tiny.cc/vivek\\_2018f](http://tiny.cc/vivek_2018f)
- Chin-Jung Hsu, **Vivek Nair**, Vincent W. Freeh, Tim Menzies. *Scout: An Experienced Guide to Find the Best Cloud Configuration*. 2018 → [http://tiny.cc/vivek\\_2018d](http://tiny.cc/vivek_2018d)
- **Vivek Nair**, Zhe Yu, Tim Menzies, Norbert Seigmund, Sven Apel. *Finding Faster Configurations using FLASH*. 2017 → [http://tiny.cc/vivek\\_2018b](http://tiny.cc/vivek_2018b)

## PUBLICATIONS

- **Vivek Nair**, Amritanshu Agrawal, Jianfeng Chen, Wei Fu, George Mathew, Tim Menzies, Leandro Minku, Markus Wagner, and Zhe Yu. *Data-Driven Search-based Software Engineering*. MSR (2018). → [http://tiny.cc/vivek\\_2018c](http://tiny.cc/vivek_2018c)
- Jianfeng Chen, **Vivek Nair**, Rahul Krishna, Tim Menzies. “Sampling” as a Baseline Optimizer for Search-based Software Engineering. IEEE TSE 2018. → [http://tiny.cc/vivek\\_2016b](http://tiny.cc/vivek_2016b)
- Jianfeng Chen, **Vivek Nair**, Tim Menzies. *Beyond Evolutionary Algorithms for Search-based Software Engineering* in IST 2017. → [http://tiny.cc/vivek\\_2017c](http://tiny.cc/vivek_2017c)
- **Vivek Nair**, Tim Menzies, Norbert Seigmund, Sven Apel. *Using Bad Learners to find Good Configurations* in FSE 2017. → [http://tiny.cc/vivek\\_2017b](http://tiny.cc/vivek_2017b)
- **Vivek Nair**, Tim Menzies, Norbert Seigmund, Sven Apel. *Faster Discovery of Faster System Configurations with Spectral Learning* in ASE Journal 2017. → [http://tiny.cc/vivek\\_2017a](http://tiny.cc/vivek_2017a)

## PROFESSIONAL EXPERIENCE

- Research Assistant** North Carolina State University (NCSSU)  
*Real-world Artificial Intelligence for Software Engineering (RAISE)* January, 2015 - present
- **Faster Discovery of Configuration Options of Software System**  
Performance optimization using machine learning techniques to discover system configurations to enhance the performance of software system. Spend 9 months collecting data by running benchmarks to find performance corresponding to the configurations of the system. Spend a month analyzing data to build a model which can be used for performance optimization.
  - **SWAY: Sampling WAY**  
Exploration of various alternatives to expensive evolutionary searches by intelligent sampling and population exploitation.

## Intern

*HIS Labs*

LexisNexis - Risk Solutions

*June, 2015 - Present*

- **FUSE Plugin:** Implementation of FUSE plugin, which can be used to mount HPCC cluster on to a local machine. Also included in the plugin is caching mechanism which does not require downloading large files on to a local machine. This is a proof-of-concept to which can reduce the time required for analysts to use data stored i  
Languages used: ECL, Python, Apache Spark, Apache Livy
- **Regression Suite:** Implementation of regression suite which includes both validity as well as scalability. The rig compares the performance scores of ecl-ml and sci-kit learn.  
Languages used: ECL, Python
- **ML-Plugins:** Implementation of Machine learning HIPIE plugins. Plugins make the library more accessible. The plugins developed are now a part of the Data Science Portal, an internal tool used across the organization.  
Languages used: DUD, ECL
- **Hyper-Parameter Tuner:** Implemented tuners, which would use algorithms like grid-search, differential evolution, etc. to tune the hyperparameters.  
Language used: ECL
- **Optimizing Random Forests:** Implementation and testing of Random Forest in the Enterprise Control Language (ECL) for use in the High-Performance Computing Cluster (HPCC).  
Languages used: ECL, Python

## Research Assistant

*Group on Autonomy, Resilience, Collaboration, and Energy (DANCE)*

North Carolina State University (NCSU)

*May, 2013 - December, 2013*

- **Detecting Contention in Data Center Environment:** Interference in a data center leads to performance variability, which causes losses to data center operators. Dynamically detecting and mitigating interference would save money for the operators. Use of hardware performance counters have proved useful in our initial experiments.  
Languages used: Python, Shell script, Hadoop, Web service (RUBiS and httpperf)

## Software Engineer

*File System & Memory Team*

Samsung Software Engineering Labs, India

*June 2011 - May 2013*

- Worked on various projects based on NOR Flash for Ultra Low-Cost cell phones.
- Developed tools to write data into One Time Programmable area of the NOR Flash.
- Worked on various filesystems like TargetFFS-NAND, Target-NOR, TargetZFS and Target FAT. Languages used: C

## TECHNICAL SKILLS

- Language: Python (4+ years), C (6+ years), Java (2+ years), ECL (2 year)
- Tools: scikit-learn, pandas, numpy, auto-ml
- OS: Linux, Windows
- Version Control: Perforce, git

## HONORS AND AWARDS

- Awarded for innovative use of HPCC Systems, 2017
- Awarded the 2<sup>nd</sup> prize in HPCC System Poster Competition, 2016.
- Awarded the Employee of the Month in January 2012 at Samsung Engineering Lab
- Awarded Scholarship by the HRD Ministry for pursuing M.Tech