

RAHUL KRISHNA

http://rkrsn.us ◊ Github: github.com/rahlk
119 Karen Ct, Cary, NC 27511

Email : rkrish11@ncsu.edu
Mobile : +1-919-396-4143
LinkedIn: linkedin.com/in/rkrns

EDUCATION

PhD in Computer Science

North Carolina State University

Jun. 2015 – May 2019

Raleigh, NC

MS in Electrical Engineering

North Carolina State University

Aug. 2013 – May 2015

Raleigh, NC

BE in Electronics & Communication

Ramaiah Institute Of Technology

Aug. 2009 – May 2013

Bengaluru, India

TECHNICAL SKILLS

General Expertise: Statistics, Machine Learning, Natural Language Processing, Program Analysis, Optimization;

Programming: Proficient: Python (5+ years), MATLAB (4+ years), C/C++ (2+ years). Working Knowledge: R, Javascript, Scala;

WORK EXPERIENCE

Phase Change LLC

Research intern, Financial Data Mining

Golden, CO

June 2018 - Aug. 2018

- Developed an ontology based vector embedding model to automatically extract financial domain knowledge from COBOL.
- Achieved confidence scores $\geq 80\%$ on determining banking concepts in COBOL code as measured by surveying subject matter experts and active COBOL developers.

Languages: Python; **Frameworks:** Apache Jena, Gensim; **Skills:** Static Code Analysis, Ontology development, SPARQL.

LexisNexis Legal & Professional

Data Science Intern, Natural Language Processing and Information Retrieval

Raleigh, NC

May 2017 - Aug 2017

- Developed and deployed a text summarization framework for generating “headnotes” in more than 1 Million legal documents.
- Designed a scalable clustering and a classification scheme with *doc2vec* to categorize documents into specific legal topics;
- Demonstrated that the above framework can reduce document review time by $\geq 50\%$ according to in-house user surveys.

Languages: Python, Scala; **Frameworks:** Tensorflow, Spark, Livy, AWS-EMR, Lambda, CloudFormation; **Skills:** ML, NLP.

LexisNexis Legal & Professional

Data Science Intern, Natural Language Processing and E-Discovery

Raleigh, NC

May 2016 - Aug 2016

- Designed a sandbox application for automated document retrieval in Technology Assisted Review (TAR) and E-Discovery.
- Improved the classification accuracy of SVM in highly skewed samples by $\approx 20\%$ using active reinforcement learning.

Languages: Python, Javascript; **Frameworks:** Spark, Elasticsearch, Logstash, Flask; **Skills:** ML, NLP, Web design.

RESEARCH PROJECTS

Planning in Software Engineering

NSF funded project in the RAISE Lab

Raleigh, NC

Sept 2015 - Present

- Developed a novel planning algorithm called XTREE to assist developers in software maintenance and defect reduction.
- Demonstrated that XTREE can generate succinct plans (75% smaller plans than other approaches).
- Demonstrated that XTREE can reduce defects by more than 80% in several cases.

Transfer Learning in Software Engineering

NSF funded project in the RAISE Lab

Raleigh, NC

Sept 2015 - Present

- Demonstrated the existence of a “Bellwether Effect” in several domains within software engineering.
- Demonstrated that the *bellwether* projects outperformed state-of-the-art transfer learners with Recall and False Alarm $\geq 75\%$.

IBM Software Group (Research Collaboration)

Assessing Development Practices in opensource and in-house projects.

Research Triangle Park, NC

March 2017 - May 2018

- Developed time-series models to forecast issues, bugs, and enhancements in 800+ opensource & IBM in-house projects.
- Demonstrated effectiveness of moving-window ARIMA models in forecasting 4-weeks of future issues, bugs, and enhancements in 66% of the projects with an accuracy of 85-95% (IBM in-house) and 75-85% (opensource).
- Studied the effect of CI/CD and developer collaboration in reducing bugs in 800+ opensource & IBM in-house projects.
- Demonstrated that earlier adoption of CI/CD stimulates developer collaboration and reduces bugs in IBM in-house projects.

LexisNexis Legal & Professional (Research Collaboration)

Software Engineering for Big Data and Industrial Text Mining

Raleigh, NC

Sept 2015 - May 2017

- Validated large scale natural language processing pipelines empirically for technology assisted review at LexisNexis.
- Developed a quantitative 6-point metric set for assessing best-practices in industrial text mining.
- Demonstrated the effectiveness of the 6-point metric in alleviating effort in making design choices for industrial text mining.

TEACHING ASSISTANTSHIPS

CSC417: Programming Languages

Spring 2019 (ongoing)

Senior Level Undergraduate Course

- Helping teach a class on programming languages to a class of 40 undergraduate students.
- Offering guest lectures on some popular abstractions in various commonly used programming languages.
- Aiding in the design and evaluations of homework assignments and examinations.

CSC591: Foundations of Software Science

Fall 2018

Special Topics in Computer Science

- Assisting in teaching the foundations of software engineering principles to a class of 30 graduate students.
- Offering guest lectures on intermediate and advanced functional programming in python with a focus on best practices.
- Assisting in designing and grading homework assignments and mid-semester examinations.

CSC591: Automated Software Engineering

Fall 2016

Special Topics in Computer Science

- Aided in teaching the principles of software engineering and applied data science to a class of 49 graduate students.
- Assisted in designing and grading homework assignments, mid-semester examinations, and final project reports.
- Offered guest lectures on the state-of-the-art in software engineering, software testing, and maintainability research.
- Assisted in mentoring over 12 distinct class projects on applying data mining, metaheuristic search, and natural language processing to software engineering.

SELECTED PUBLICATIONS

- [1] Krishna, R. & Menzies, T. "Bellwethers: A Baseline Method For Transfer Learning". In **IEEE Transactions on Software Engineering**, 2018. Online: <https://arxiv.org/abs/1703.06218>;
- [2] Krishna, R., Menzies, T., & Layman, L. "Less is more: Minimizing code reorganization using XTREE". In **Information and Software Technology**, 2017. Online: <https://arxiv.org/abs/1609.03614>;
- [3] Krishna, R., Menzies, T., & Fu, W. "Too much automation? The Bellwether Effect and its Implications for Transfer Learning." In **Intl. Conference on Automated Software Engineering**, 2016. Online: <https://doi.org/10.1145/2970276.2970339>;
- [4] Krishna, R., Agrawal, A., Rahman, A., Sobran, A., & Menzies, T. "What is the Connection Between Issues, Bugs, and Enhancements? (Lessons Learned from 800+ Software Projects)". **Intl. Conf. Software Engineering, 2018 SEIP**. Online: <https://arxiv.org/abs/1710.08736>;
- [5] Krishna, R., Yu, Z., Agrawal, A., Dominguez, M., Wolf, D. "The 'BigSE' Project: Lessons Learned from Validating Industrial Text Mining. ". In **Intl. Workshop on Big Data Software Engineering (BIGDSE), 2016**. Online: <http://tiny.cc/BIGDSE16>;
- [6] Krishna, R. & Menzies, T. "Actionable=Cluster+Contrast?". In **Intl. Conference on Automated Software Engineering Workshop, 2015**. Online: <http://tiny.cc/ase15action>;
- [7] Ganesh, M.R., Krishna, R., Manikantan, K. & Ramachandran, S., "Entropy based Binary Particle Swarm Optimization and classification for ear detection". In **Engineering Applications of Artificial Intelligence**, 2014., Online: <http://tiny.cc/EBPSO14>;
- [8] Chen, J., Nair, V., Krishna, R., & Menzies, T. "Sampling as a Baseline Optimizer for Search-based Software Engineering". In **IEEE Transactions on Software Engineering**, 2018. Online: [arXiv:1608.07617](https://arxiv.org/abs/1608.07617);
- [9] Chen, D., Fu, W., Krishna, R., & Menzies, T. "Applications of psychological science for actionable analytics". In **Intl. Conference on Foundations of Software Engineering**, 2018. Online: [arXiv:1803.05067](https://arxiv.org/abs/1803.05067);
- [10] Agrawal, A., Rahman, A., Krishna, R., Sobran, A. & Menzies, T. "We Don't Need Another Hero? The Impact of 'Heroes' on Software Development". **Intl. Conf. Software Engineering, 2018 SEIP**. Online: [arXiv:1710.09055](https://arxiv.org/abs/1710.09055);
- [11] Rahman, A., Agrawal, A., Krishna, R., Sobran, A. & Menzies, T. "Characterizing The Influence of Continuous Integration. Empirical Results from 250+ Open Source and Proprietary Project". **Intl. Conference on Foundations of Software Engineering, SWAN 2018**. Preprint: <https://arxiv.org/pdf/1711.03933>;