**Bloomberg LP**
*731 Lexington Ave,*
*New York, NY 10022*
*USA*

January 19, 2018

Dear Sir or Madam,

I am writing to apply for a full-time position at Bloomberg. I am a Final year Ph. D. student majoring in computer science from North Carolina State University. I am being advised by Dr. Tim Menzies. My research focuses primarily on empirical software engineering, computational linguistics, and data driven decision making.

I believe that I have the industrial and research experience required to be an effective team member at your group. In terms of industrial experience, I have worked extensively with machine learning, natural language processing, and their applications to legal text mining. Following is a brief summary of my work:

- Over the summer of 2017, I worked on deploying large scale computational linguistics and machine learning algorithms for processing over 1 million of legal documents. Specifically, I developed tools for automated text summarization of very large legal documents using TensorFlow. With these I was able to summarize large legal documents into so-called *headnotes*. Subject matter experts opined that these headnotes were very similar to human-generated analogs. Further, I was also responsible for developing additional machine learning algorithms so that they can be deployed to production. To achieve this, I worked extensively with the devops pipeline within the AWS ecosystem to enable seamless integration of machine learning algorithms with the current product.
- In the summer of 2016, I was tasked with "opening the black-box" that is SVM. For this, I designed a sandbox app for e-discovery. This sandbox was very useful improving the classification accuracy of SVM by approximately 20%. My Contributions include: (1) Translating internal mechanisms of SVM into human comprehensible format. For this, I extracted the support vectors and mapped them to actual words, phrases, and paragraphs in the text. These were then presented to users so they may get some insights on how SVM made its prediction; (2) By soliciting feedback from the end users (human-in-loop) regarding the quality of these support vectors, I was able to incorporate these feedback with the help of active learning to improve the text classification accuracy of SVM.

As to my research work, several of my recent papers show that I have done work in areas that may be of further interest to you. Much of my work studies how machine learning applications can be applied to streamline software engineering practices. I have 3 papers under review at this time for the ICSE'18 software engineering in practice track.

- 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)"*. (Under review) **ICSE 2018 SEIP**. Pre: arXiv:1710.08736;
- Rahman, A., Agrawal, A., Krishna, R., Sobran, A., & Menzies, T. *"Continuous Integration: The Silver Bullet?"*. (Under review) **ICSE 2018 SEIP**. Preprint: arXiv:1711.03933;

**Rahul Krishna**
*119 Karen Ct – Cary, NC 27511 – USA*
*+1 (919) 396-4143 • rkrish11@ncsu.edu • rkrsn.us*

- Agrawal, A., Rahman, A., Krishna, R., Sobran, A., & Menzies, T. *"We Don't Need Another Hero? The Impact of Heroes on Software Development"*. (Under review) **ICSE 2018 SEIP**. Preprint: arXiv:1711.03933;

As to my accepted work, I have published an article at *Automated Software Engineering* 2016, Singapore. I also have two accepted papers premier journals such as *IEEE Transactions in Software Engineering* and and *Information and Software Technology* journal. My other work at *IEEE Transactions in Software Engineering* journal is awaiting a second round review. These are listed below:

- Krishna, R., Menzies, T., & Layman, L. *"Less is more: Minimizing code reorganization using XTREE"*. In **Information and Software Technology**, Volume 88, 2017, Pages 53-66;
- Krishna, R., Menzies, T., & Fu, W. *"Too much automation? The Bellwether Effect and its Implications for Transfer Learning."* **31st Intl. Conference on Automated Software Engineering, Sept. 2016**;
- Krishna, R. & Menzies, T.. *"Bellwethers: A Baseline Method For Transfer Learning"*. In **IEEE Transactions on Software Engineering** (pending revisions), 2017;
- Chen, J., Nair, V., Krishna, R., & Menzies, T. *"Sampling as a Baseline Optimizer for Search-based Software Engineering"*. In **IEEE Transactions on Software Engineering** (accepted, to appear), 2017;

I strongly believe that I would be an asset to your organization. A full-time position here would provide me with the ideal opportunity to assist your organization and to further expand my skills.

I have attached my resume for your reference. If you feel that my qualifications seem to be a match for positions in your organization, I should greatly appreciate an opportunity to schedule an interview at a mutually convenient time.

Thank you very much for your consideration.

Yours Sincerely,

**Rahul Krishna**

*Attached: Résumé*

# 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/rkrsn

## EDUCATION

**PhD in Computer Science** — Jun. 2015 – Dec. 2018 (expected)
*North Carolina State University* — *Raleigh, NC*

**MS in Electrical Engineering** — Aug. 2013 – May 2015
*North Carolina State University* — *Raleigh, NC*

**BE in Electronics & Communication** — Aug. 2009 – May 2013
*Ramaiah Institute Of Technology* — *Bengaluru, India*

## TECHNICAL SKILLS

**General Expertise**: Machine Learning, Natural Language Processing, Multiobjective Optimization, Empirical Software Engineering;
**Data Analytics**: Spark, Hadoop, Elasticsearch, S3, Weka, Sklearn, JMetal. Visualization: Kibana, D3JS, Matplotlib;
**Cloud Computing**: AWS ecosystem: EMR, Apache Livy, Cloud Formation, AWS Lambda, Chalice;
**Programming**: Proficient: Python, Javascript, & R. Also familiar with: Java, Scala, C++, & Lua;
**DevOps**: Ansible, Vagrant, Travis, Jenkins, Docker.

## WORK EXPERIENCE

**Data Science Intern** — May 2017 - Aug. 2017
*LexisNexis* — *Raleigh, NC*

- Worked on deploying computational linguistics and machine learning algorithms for processing 1M+ legal documents.
- Contributions include: (1) Clustering more than 1 million documents based word2vec and doc2vec to identify representatives for specific legal topics; (2) Developing automated text tools to summarize very large legal documents into legal "headnotes";

**Software Engineering Intern** — May 2016 - Aug. 2016
*LexisNexis* — *Raleigh, NC*

- Designed a sandbox app for e-discovery. Sandbox was used to improve the classification accuracy of SVM by $\approx$ 20%.
- Contributions: (1) Translating internal mechanisms of SVM into human comprehensible format; (2) Improved text classification accuracy of SVM by modifying support vectors using active learning and feedback from human-in-loop.

## SELECTED RESEARCH PROJECTS

**Planning in Software Engineering** — Sept 2015 - Present
*NSF funded project in the RAISE Lab* — *Raleigh, NC*

- Developed a novel planning algorithm called XTREE to assist developers in software refactoring and code reorganization.
- Showed that XTREE can generate succinct and effective plans. Experiments showed that XTREE can reduce defects by more than 80% in several cases.

**Transfer Learning in Software Engineering** — Sept 2015 - Present
*NSF funded project in the RAISE Lab* — *Raleigh, NC*

- Demonstrated the existence of a *"Bellwether Effect"* in several domains within software engineering.
- The *bellwethers* were shown to be a very effective baseline for transfer learning. Also showed that they are very easy to discover and usually outperform several state-of-the-art transfer learners in software engineering.

**Validating Industrial Text Mining** — Sept 2015 - May 2017
*Industrial collaboration with LexisNexis* — *Raleigh, NC*

- Worked on validating large scale natural language processing pipelines for technology assisted review at LexisNexis.
- Demonstrated the usefulness of context specific ensemble learners and active learning for document classification.
- Demonstrated the effectiveness of several data preprocessing techniques such SMOTE for enhancing information retrieval.

## SELECTED PUBLICATIONS

[1] Krishna, R., Menzies, T., & Layman, L. *"Less is more: Minimizing code reorganization using XTREE"*. In **Information and Software Technology**, Volume 88, 2017, Pages 53-66. DOI: 10.1016/j.infsof.2017.03.012;

[2] Krishna, R., Menzies, T., & Fu, W. *"Too much automation? The Bellwether Effect and its Implications for Transfer Learning."* **31st Intl. Conference on Automated Software Engineering, Sept. 2016**. DOI: 10.1145/2970276.2970339;

[3] Krishna, R. & Menzies, T.. *"Bellwethers: A Baseline Method For Transfer Learning"*. In **IEEE Transactions on Software Engineering** (pending revisions), 2017. Preprint: arXiv:1703.06218;

[4] Chen, J., Nair, V., Krishna, R., & Menzies, T. *"Sampling as a Baseline Optimizer for Search-based Software Engineering"*. In **IEEE Transactions on Software Engineering** (accepted, to appear), 2017. Preprint: arXiv:1608.07617;

[5] Krishna, R. *"Learning effective changes for software projects"*. **32nd Intl. Conference on Automated Software Engineering Doctoral Symposium, October 2017**. Available: http://dl.acm.org/citation.cfm?id=3155562.3155695;

[6] 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)"*. (Under review) **ICSE 2018 SEIP**. Pre: arXiv:1710.08736;

[7] Rahman, A., Agrawal, A., Krishna, R., Sobran, A., & Menzies, T. *"Continuous Integration: The Silver Bullet?"*. (Under review) **ICSE 2018 SEIP**. Preprint: arXiv:1711.03933;