

# Saikat Chakraborty

saikatc.info

PHONE: +1 (434) 242-1306

E-MAIL : saikatc@cs.columbia.edu

## INTEREST

---

**AI driven software systems** - I work on the intersection of Artificial Intelligence and Software Systems. My interest is on designing *intelligent* tools and models to extract information from big sources of software information (e.g. Codebase), design and develop analysis tool to aid developer for better maintenance of softwares. I am interested in developing intelligent systems that help reduce human effort to detect, locate and fix *bugs* through intelligent program analysis and machine learning.

## EDUCATION

---

January 2019 - Present	<p>Ph. D in COMPUTER SCIENCE COLUMBIA UNIVERSITY, New York, NY Area : Data driven software Engineering <b>Current Work</b> : Working with Professor Baishakhi Ray towards developing a <i>Variational Autoencoder</i> based hierarchical language model for code to capture the semantic structure of the developer written code bases.</p>
AUGUST 2016 - December 2018	<p>Ph. D in COMPUTER SCIENCE UNIVERSITY OF VIRGINIA, Charlottesville, VA Area : Data driven software Engineering CGPA : 4.00/4.00 <b>Graduate Courses Taken</b> : Software Engineering, Natural Language Processing, Defense Against Dark Arts (Software Security), Cyber Physical System, Design and Analysis of Algorithm, Mobile Security. <b>Projects:</b> <b>1. Natural Language Processing</b> : We built a Language Models for Text with Named Entities. novel and effective approach to building a language model which can learn the entity names by leveraging their entity type information. This work has been accepted to be published in <b>ACL'18</b> [C1]. <b>2. Software Engineering</b> : In this project, we leveraged the naturalness of code to localize faults in demonstrated by failing test cases in real software systems. <a href="#">arXiv link</a>. <b>3. Code Translation</b> : In this project, we are building a tree-to-tree machine translation model to capture the frequent code change pattern and bug fix patterns [A1].</p>
MARCH 2009 - JULY 2014	<p>B.Sc. in COMPUTER SCIENCE AND ENGINEERING BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY Dhaka, Bangladesh. Thesis : Optimization algorithms in Multi Depot Vehicle Routing Problem [J1]. Supervisor: Dr. Md. Monirul Islam. CGPA : 3.80/4.00</p>

## WORK EXPERIENCE

---

SUMMER 2017	<p><b>Research Intern</b> at Fujitsu Laboratories of America, Sunnyvale, CA Supervisor : Dr. Ripon K Saha Worked with AI based program repair tool. Extracted subtle information from auxiliary sources to improve the performance of fault localization and program repair.</p>
JUNE'15 - AUGUST'16	<p><b>Quantitative Researcher</b> at Stochastic Logic Ltd, Dhaka, Bangladesh Supervisor : Dr. Arif Dowla Designed a robust model to estimate Value-at-risk(VAR) from stock market data. Also Developed an intelligent model to calculate Value-at-risk(VAR) from stock data and manage stock portfolio.</p>
OCTOBER'14 - AUGUST'16	<p><b>Lecturer</b> at Ahsanullah University of Science and Technology Taught undergraduate course on digital system design and compiler. Supervised undergrad research project. [C2]</p>

JULY'14 -  
OCTOBER'14

**Jr. Software Engineer** at Kona Software Lab (KSL)

Designed Proof of Concept of Host Card Emulation System with VISA based payment networks presented by KSL at CARTES SECURE CONNEXIONS 2014 .

## SKILLS

---

Programming Languages : Java, Python, C/C++, SQL, PHP, Matlab, R, Assembly  
Deep Learning Tools : Theano, Pytorch

## PUBLICATIONS

---

- [A1] Tree2Tree Neural Translation Model for Learning Source Code Changes, **Saikat Chakraborty**, Miltiadis Allamanis, Baishakhi Ray, [arXiv pre-print](#).
- [C1] Building Language Models for Text with Named Entities, Rizwan Parvez, **Saikat Chakraborty**, Baishakhi Ray, Kai-wei Chang, 56th Annual Meeting of the Association for Computational Linguistics (ACL) 2018.
- [C2] FP-ANK: An improvised intrusion detection system with hybridization of neural network and K-means clustering over feature selection by PCA, NA Biswas, FM Shah, WM Tammi, **S Chakraborty**, 18th International Conference on Computer and Information Technology (ICCIT), 2015
- [C3] PlssGA: An ultra fast meta-heuristic approach to solve protein inference problem, SKK Santu, **S Chakraborty**, S Rahman, MS Rahman 16th International Conference Computer and Information Technology (ICCIT), 2013.
- [J1] A Heuristic Initialized Stochastic Memetic Algorithm for MDPVRP With Interdependent Depot Operations, AS Azad, MM Islam, **S Chakraborty**. IEEE Transactions on Cybernetics, 2017
- [W1] An SMS based Smart System for Remote Control of Utility Services During Earthquake. **Saikat Chakraborty**, Tanmoy Sen, Madhusudan Basak, Mir Tazbinur Sharif and Ahmed Al Muzaddid, In Proceedings of Workshop on Design and implementation of Emerging Computing Systems (WECS), 2013.
- [P1] Which similarity metric to use for software documents?: a study on information retrieval based software engineering tasks. Md Masudur Rahman, **Saikat Chakraborty**, and Baishakhi Ray in Proceedings of the 40th International Conference on Software Engineering: Companion Proceedings. ACM, 2018.

## AWARDS AND HONORS

---

<b>Champion</b> : 1st Undergraduate Thesis Poster Presentation.	2014
<b>Best Paper Award [w1]</b> : Workshop of Emerging Computing System (WECS), Bangladesh.	2014
<b>University merit and deans award</b>	2009-2014

## OTHER ACTIVITIES

---

Volunteering :	Vice President, Badhan (A community blood donor organization), Suhrawardi Hall unit, BUET zone, Dhaka.
Leadership :	Founding Vice President, Engineering Students' Association of Bangladesh, Dhaka Bangladesh (2013- 2015). Organizer, International Engineering Innovation Summit Bangladesh, 2015