Golam Mostaeen

Links
Email: golammostaeen@gmail.com
LinkedIn: https://www.linkedin.com/in/golam-mostaeen
GitHub: https://github.com/pseudoPixels
Blog: https://golammostaeen.wordpress.com/
Personal Web: https://pseudopixels.github.io/

Address
333 Lester St
Waterloo, Ontario
Canada
N2L 3W6
+1 (306) 717-9731

EDUCATION

Master of Science (Software Engineering), Computer Science — Sept. 2016 - Dec. 2018
University of Saskatchewan, Saskatchewan, Canada
Course Average : 90%
Thesis : Towards Collaborative Scientific Workflow Management System
Supervisors : Dr. Chanchal K. Roy & Dr. Banani Roy
Award : Best Graduate Student Award 2018, Dept. of CS, University of Saskatchewan, Canada
Nominated : Best Graduate Thesis Award 2018

Bachelor of Science, Computer Science and Engineering — Jan.2011 - Nov. 2014
Islamic University of Technology, Dhaka, Bangladesh
CGPA : 3.95/4.00 (First Class with Honors), Ranked 2nd in Class
Thesis : A Statistical Approach for Offline Handwritten Signature Verification Using SIFT Features
Supervisor : Dr. Md. Hasanul Kabir
Award : OIC Full Free Scholarship (three years of my undergraduate) for outstanding academic results

RESEARCH PAPERS

- Mostaeen, Golam, Banani Roy, Jeffrey Svajlenko, Chanchal K. Roy, and Kevin A. Schneider. 'A Machine Learning Based Framework for Code Clone Validation.' Journal of Systems and Software 169 (2020): 110686.
- Mostaeen, Golam, Jeffrey Svajlenko, Banani Roy, Chanchal K. Roy, and Kevin A. Schneider. 'CloneCognition: machine learning based code clone validation tool.' In Proceedings of the 2019 27th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, pp. 1105-1109. 2019.
- Mostaeen, Golam, Jeffrey Svajlenko, Banani Roy, Chanchal K. Roy, and Kevin A. Schneider. 'On the Use of Machine Learning Techniques Towards the Design of Cloud Based Automatic Code Clone Validation Tools.' In 2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation (SCAM), pp. 155-164. IEEE, 2018.
- Chakroborti, Debasish, Banani Roy, Amit Mondal, Golam Mostaeen, Chanchal K. Roy, Kevin A. Schneider, and Ralph Deters. 'A Data Management Scheme for Micro-Level Modular Computation-Intensive Programs in Big Data Platforms.' In Data Management and Analysis, pp. 135-153. Springer, Cham, 2020.
- Mostaeen, Golam, Banani Roy, Chanchal Roy, and Kevin Schneider. 'Designing for Real-Time Groupware Systems to Support Complex Scientific Data Analysis.' Proceedings of the ACM on Human-

Computer Interaction 3, no. EICS (2019): 1-28.

- Mostaeen, Golam. 'Towards Collaborative Scientific Workflow Management System.' Master of Science Thesis, University of Saskatchewan, 2019.
- Mostaeen, Golam, Banani Roy, Chanchal K. Roy, and Kevin A. Schneider. 'Fine-Grained Attribute Level Locking Scheme for Collaborative Scientific Workflow Development.' In 2018 IEEE International Conference on Services Computing (SCC), pp. 273-277. IEEE, 2018.
- Rahman, ABM Ashikur, Golam Mostaeen, and Md Hasanul Kabir. 'A statistical approach for offline signature verification using local gradient features.' In 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), pp. 1-4. IEEE, 2016.
- Ajwad, Rasif, Syed Nayem Hossain, Golam Mostaeen, and M. A. Mottalib. 'An optimized algorithm to find maximum parsimonious tree using PrimeNucleotide based approach.' In 2014 17th International Conference on Computer and Information Technology (ICCIT), pp. 127-131. IEEE, 2014.
- *Golam Mostaeen*, Banani Roy, Chanchal K. Roy, and Kevin A. Schneider. 'Consistency Management In Real-Time Collaborative Scientific Workflow Composition By Granular Attribute Level Locking.' *(to be submitted)*
- Golam Mostaeen, Banani Roy, Chanchal K. Roy, and Kevin A. Schneider. 'Collaborative Data Analysis With Scientific Workflow Management Systems in Theory vs Reality.' *(to be submitted)*
- *Golam Mostaeen*, Banani Roy, Chanchal K. Roy, and Kevin A. Schneider. 'SciWorCS: Towards Collaborative Scientific Workflow Management Systems.' *(to be submitted)*

RESEARCH PROJECTS

- **CloneCognition** is a machine learning based framework for automatic code clone validation - developed based on our recent research work. The method learns to predict user-specific code clone validation patterns. The current machine learning model has been build based on BigCloneBench - a collection of eight million validated clones within IJaDataset-2.0, a big dataset of software repositories containing 25,000 open source Java software application (*GitHub*: https://github.com/pseudoPixels/CloneCognition *Associated publication*: CloneCognition: machine learning based code clone validation tool).
- **SciWorCS** is a Collaborative Scientific Workflow management System. SciWorCS follows a plugin-based architecture for the scientific computational modules. It provides collaborative environment for scientific data analysis using efficient attribute level locking scheme. In addition to the collaborative data analysis, SciWorCS also provides real time monitoring of the computation steps. Developed as a proof of concept of proposition from my M.Sc. research works(*GitHub*: https://github.com/pseudoPixels/ /SciWorCS ; *Associated publication*: Designing for Real-Time Groupware Systems to Support Complex Scientific Data Analysis).
- **ML Clone Validation Framework** is a Machine Learning based framework for code clone labeling and validation. Works as a framework for code clone labeling with user-specific preferences. In addition to the clone labeling framework, this cloud based framework also supports the communication with any existing code clone detection tools for validation prediction responses using REST API (*GitHub*: https://github.com/pseudoPixels/ML_CloneValidationFramework; *Associated publication*: On the Use of Machine Learning Techniques Towards the Design of Cloud Based Automatic Code Clone Validation Tools).
- **iSeaborn** is an interactive version of the popular statistical data visualization library - Seaborn. iSeaborn is build on top of Seaborn and uses pandas for advanced statistical computation. iSearborn uses Bokeh as visualization library for interactive graphs and plots (*GitHub*: https://github.com/pseudoPixels/iSeaborn).

- **Dranslate** is a Deep learning based natural language translator for Bengali to English and vice-versa. Dranslate is an ecoder-decoder based LSTM model for the translation. Used Tensorflow and Keras for the development of this language model (*GitHub*: https://github.com/pseudoPixels/Dranslate).
- **Automathic** is sequence to sequence Deep Learning model using RNN to learn to do simple arithmetic operations such as addition or subtraction. The numeric operands are cast as string to break down the problem as sequence to sequence language modeling problem
(*GitHub*: https://github.com/pseudoPixels/ML_Playground).
- **DataGraphaite** is an AI based visualization library for advanced Business Intelligence data analysis. (*GitHub*: https://github.com/pseudoPixels/vizAI)
- **PyPoets** is a deep learning based poem generator. This is a language encoder-decoder based LSTM model for predicting the next verse of the poem for sequence to sequence modeling
(*GitHub*: https://github.com/pseudoPixels/ML_Playground).

RESEARCH TALK & POSTER DEMO

- Golam Mostaeen. 2019, 'Towards Collaborative Scientific Workflow Management System', M.Sc. Thesis Defence, University of Saskatchewan, Canada.
- Golam Mostaeen. 2018, 'A Framework for Collaboratively Building on-the-fly Scientific Pipelines for Plant Phenotyping and Genotyping', Poster Competition, Department of Computer Science, University of Saskatchewan, Canada.
- Golam Mostaeen. 2018, 'Fine-grained attribute level locking scheme for collaborative scientific workflow development', International Conference on Service Computing 2018, San Francisco, California, USA.
- Golam Mostaeen. 2017, 'A Framework for Collaboratively Building on-the-fly Scientific Pipelines for Plant Phenotyping and Genotyping', The 2nd Annual Plant Phenotyping & Imaging Research Centre Symposium, University of Saskatchewan, Canada.
- Golam Mostaeen. 2014, 'Statistical Approach for Offline Handwritten Signature Verification Using SIFT Features', B.Sc. Thesis Defence, Islamic University of Technology, Dhaka, Canada.

AWARDS & SCHOLARSHIP

- **Best Graduate Student Award 2018** - I was awarded the Best Graduate Award 2018 from the Department of Computer Science, U of S, Canada.
- **Best MSc Thesis Nomination 2018** - My M.Sc. thesis was nominated for the Best MSc Thesis Award 2018 by the Department of Computer Science, University of Saskatchewan, Canada.
- **SK Innovation Award 2018** - My M.Sc. research works received the prestigious SK Innovation, SK, Canada award 2018 for innovative use of technology towards data science.
- **Runner-Up, Programming Contest 2017** - Awarded Runner-Up in Programming Contest 2017, Department of Computer Science, UofS, Canada.
- **Champion, Game Jam 2016** - Awarded Champion, Game Jam 2016 in Audio-Visual Category, Department of Computer Science, UofS, Canada.

- **Graduate Travel Award 2017** - Awarded by University of Saskatchewan for SCC 2017 travel to San Francisco, California, USA. Award value: $550.
- **Runner-Up, Programming Contest 2013** - Awarded Runner-Up in Programming Contest 2013, Department of Computer Science, Islamic University of Technology, Dhaka, Bangladesh.
- **OIC Scholarship 2011** - Received OIC full free scholarship for three years starting from 2011 for outstanding academic results.
- **Government Scholarship 2008** - Received Government scholarship in Talent category for two years starting from 2008.

EMPLOYMENT HISTORY

- **Data Scientist - Electronic Arts (EA Sports), Ontario, Canada, (November 2020 – Present):** Researched and Developed highly scalable data visualization platform handling billions of data record efficiently. Developed and Maintained data pipelines with fast and efficient data Extraction, Transformation and Loading. Saved around 50% of the company production cost with data flow automation and scalable data visualization platform.

- **Data Scientist - Gameloft, Montreal, Canada, (April 2019 – November 2020):** Developed and deployed machine learning model for churn prediction; predicting from millions of users world wide. Added significant business impact for the industry in production. Worked directly on designing Paying User Prediction model that obtained promising results for the industry. Collaborated directly with brilliant Data Scientists, Analysts and ML/Data Engineer across the globe (Montreal, Toronto, Bucharest, Paris and Barcelona) on multiple ML projects. Developed data pipelines in full stack, worked on automation and scheduling, saving around 45% time.

- **Graduate Researcher (Data Science, ML, Data Engineering) - University of Saskatchewan, SK, Canada (Sept. 2016 – Feb 2019):** Researched and Developed platform for Collaborative Data Science which gained significant attention and in active R&D. Resulted in prestigious SK Innovation, Canada 2018 award. Published six research papers contributing in ML and Data Science in top international conferences and journals. Collaborated with Data Science Researchers of multiple domains (Image processing, Computer Vision, Software Eng., HCI and Bioinformatics).

- **Lecturer - Computer Science, Islamic University of Technology, Dhaka, Bangladesh (Mar. 2015 – Apr. 2018):** Organized and instructed CS courses to 250+ students (Data Structure, Algorithm, Software Development). Organized and tutored programming contests and teams.

- **Software Developer - Envato (Mar. 2015 – Apr. 2018):** Worked as full stack Android and AI developer. Developed Ten complete apps in full stack and collaborated with 150+ clients directly.

- **Internship - XeonBD, Dhaka, Bangladesh (October 2013 - November 2013):** Worked with software development team with major focus on SDLC and Scrum.

RESEARCH TOOLS EXPERIENCE

- **Machine Learning & Data Mining:** PyTorch, TensorFlow, Python, R, MATLAB, Weka, Scikit-Learn, Pandas, Numpy
- **Software Development & Maintenance:** Eclipse, PyCharm, IntelliJ, Visual Studio, ArgoUML, Doxygen, JUnit, JavaParser, Jsoup, PMD, FindBugs, CheckStyle, Ant, and Maven.

- **Software Version Control:** Git, GitHub, and GitLab.
- **Big Data Technologies:** Apache Spark 2.2, Hadoop 2.7.

REFERENCES

- Dr. Chanchal K. Roy
  Professor,
  University of Saskatchewan,
  Saskatoon, Saskatchewan, Canada
  Email : chanchal.roy@usask.ca
  URL : https://www.cs.usask.ca/faculty/croy

- Dr. Banani Roy
  Assistant Professor,
  University of Saskatchewan,
  Saskatoon, Saskatchewan, Canada
  Email : banani.roy@usask.ca
  URL : https://www.cs.usask.ca/people/faculty%20profiles/banani-roy.php

- Dr. Kevin Schneider
  Department Head, Professor,
  University of Saskatchewan,
  Saskatoon, Saskatchewan, Canada
  Email : kevin.schneider@usask.ca
  URL : https://clones.usask.ca/kas/