SOHIL LAL SHRESTHA

712 Linda Vista Ave, Arlington, TX 76013 | +1-817-557-7646 | sohail.stha@gmail.com | https://50417.github.io/

EDUCATION

- **University of Texas, Arlington(UTA), Arlington, Texas** August 2017 - December 2022 (Expected)
  PhD Candidate in Computer Science GPA: 4.00
  - **Relevant Coursework:** Machine Learning, Design & Analysis of Algorithms, Software Engineering, Distributed Systems
- **Kathmandu University(KU)**, Dhulikhel, Nepal September 2012 - November 2016
  Bachelor in Computer Engineering GPA: 3.85
  - **Relevant Coursework:** Object-Oriented Design (OOD), Relational Databases, Data Structures, Cloud Computing, Operating Systems

TECHNICAL SKILLS

- **Language:** Java (Proficient), Python (Familiar), Matlab , C/C++
- **Databases & Frameworks:** SQL, Oracle, Vertica, REST, MVC, React, Django
- **Build/VCS Tools:** Maven, Git, SVN
- **Tools & Libraries:** TensorFlow, Keras, Numpy, MatplotLib, Eclipse, Pycharm

ACADEMIC PROJECTS

SLNET

- Built a tool to automatically mine Simulink models from open source repositories.
- Extract and analyzed the model metrics to visualize modeling practices and developed the largest corpus of Simulink models for future studies.

DeepFuzzSL

- Propose and develop to learn specification of dataflow programming language (aka MATLAB/Simulink) using LSTM architecture to automatically generate Simulink models.
- DeepFuzzSL consistently generated over 90% valid Simulink models and found 2 confirmed bugs

PROFESSIONAL EXPERIENCE

R&D Machine Learning Intern, Atos Syntel *May 2021 to August 2021*

- Prepared training dataset for object detection of 6 classes and trained Google AutoML model achieving 90% precision and recall.
- Applied BERT based contextual spell correction on optical character recognition's output along with heuristic based correction. The approach performed 95% good sentences based on metric using levenshtein ratio.
- Developed a web application using React and Django to visualize inference from models trained in Google AutoML along with Microsoft Azure Cognitive Services.

Intern/Associate Software Engineer

Cotiviti Nepal Pvt. Ltd. *July 2016 to May 2017*

- Performed root cause analysis of production issues (discrepancy in data, invalid sql queries) to reduce clients downtime in a Medical Intelligence application working
- Collaborated with senior member follow Agile development principles to reduce team's backlog by 50% involving Vertica SQL exception.

TEACHING AND RESEARCH EXPERIENCE

Graduate Teaching/Research Assistant, University of Texas at Arlington *August 2017 to Present*

- Design and grade assignments and projects for over 45 students each semester (Fall 2018, Spring/Summer 2019)
- Proposed an approach to integrate program analysis (dynamic symbolic execution) into an existing supervised machine learning pipeline in order to automatically produce additional labeled training samples.
- Currently working on using language modeling techniques to learn language specification of MATLAB/Simulink from existing Simulink sample models and use the resulting trained model to generate Simulink models to test Simulink compiler and toolchains.

Publications

- **Shrestha, Sohil L.**, and Christoph Csallner. "SLGPT: Using Transfer Learning to Directly Generate Simulink Model Files and Find Bugs in the Simulink Toolchain." Evaluation and Assessment in Software Engineering(EASE), 2021.
- Shafiul Azam Chowdhury, **Sohil Lal Shrestha**, Taylor T. Johnson and Christoph Csallner. "SLEMI: Equivalence modulo input (EMI) based mutation of CPS models for finding compiler bugs in Simulink." Proc. 42nd ACM/IEEE International Conference on Software Engineering (ICSE), 2020.
- **Shrestha, Sohil L.**, Shafiul Azam Chowdhury and Christoph Csallner. "DeepFuzzSL: Generating models with deep learning to find bugs in the Simulink toolchain." Proc. 2nd Workshop on Testing for Deep Learning and Deep Learning for Testing (DeepTest), 2020.
- **Shrestha, Sohil L.**, Saroj Panda, and Christoph Csallner. "Complementing Machine Learning Classifiers via Dynamic Symbolic Execution: Human vs. Bot Generated Tweets." 2018 IEEE/ACM 6th International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering (RAISE). IEEE, 2018.