Vincent J.
Hellendoorn
vjhellendoorn
@gmail.com
+1 530 979 4600
vhellendoorn.github.io

Vincent Hellendoorn

Assistant Professor of Computer Science

About me I am an Assistant Professor of Computer Science at Carnegie Mellon University and visiting researcher at Google. In my research, I study how developers write and understand code using empirical software engineering methods and deep learning. This allows me to build intelligent tools, advance deep learning methods, and work towards my broader goal of making good software development more accessible.

Top Publications

ICLR 2020: Global Relational Models of Source Code (*first author*)
We show that present models of code poorly capture global information and propose alternative models that greatly improve automated repair.

ICSE 2019: When Code Completion Fails: a Case Study on Real-World Completions (*first author*)
By running various top code completion models on real developer data, we expose systemic biases in synthetic training data.

FSE 2018: Deep Learning Type Inference (*first author*)
We leveraged large corpora of TypeScript code to learn to automatically and reliably annotate code with types.

FSE 2017: Are Deep Neural Networks the Best for Modeling Source Code? (*first author*)
A criticial re-evaluation of language modeling for software leads to the most powerful language model to date.

ICSE 2016: On the "Naturalness" of Buggy Code (*shared first author*)
We demonstrated that bugs often have remarkable entropy footprints that can be discovered with appropriate models.

MSR 2015: Will they like this? Evaluating Code Contributions With Language Models (*first author*)
Open-source projects have unique stylistic preferences that influence their decisions on accepting external contributions.

Industry Experience

Visiting Faculty at Google Brain (since September 2020)
Part-time at Google Brain, studying how developers produce code and how better to support their efforts.

Intern at Google Brain (summer 2019)
Worked with Google Brain on improving models of source code, setting new state-of-the-art performance in automated bug repair.

Intern at Microsoft Research (summers 2017 & 2018)
Improved bug-triaging with neural nets, built a learned type inference engine for JavaScript, and trained models of program invariants.

Software Developer at Milvum (Sep'13 - June'14)
Developed tailored applications for mobile & web. In the final 6 months (before moving to Davis, CA), I was the lead developer on a new project.

Vincent J.
Hellendoorn
vjhellendoorn
@gmail.com
+1 530 979 4600
vhellendoorn.github.io

Education

2016 - 2020, PhD. Computer Science at UCDavis

Focus on combining Empirical Software Engineering Research with Machine Learning, and (psycho-)linguistic aspects of reading code.

2013 - 2015, MSc. Computer Science at Delft University

Completed MSc Cum Laude, with a specialization in Software Engineering. Awarded faculty best Master's Thesis.

2010 - 2013, BSc. Computer Science at Delft University

Completed BSc Cum Laude, with Honors, with a Minor in Applied Mathematics at the National University of Singapore.

Academic Experience

Academic Service

- FSE'21 Organizing Committee, online co-chair
- **Program committee** for: ICSE, FSE, NeurIPS, ICML, MSR, ICSME
- **Journal reviewer** for: TSE, EMSE, TOSEM, CACM, and JSS.

Teaching Experience

- **Deep Learning for Software Engineering**: a 7-11 week seminar or (grad level) course covering the fundamentals of deep learning as well as recent research in ML4SE. Students are provided with modeling scaffolding and work in teams on class-long projects to replicate (and exceed) state-of-the-art results. Taught at UC Davis & CMU.
- **Software Engineering Toolchains**: covers the full suite of tools commonly used in SE (think: git to cloud) and their theoretical and historical background. Taught at UC Davis.

Delft University of Technology: Honours Programme (Nov'11 - June'13)

I conducted research in the area of Natural Language Processing as part of the Honours Programme for BSc. students.

National University of Singapore: Minor Abroad (Aug'12 - Dec'12)

I participated in Delft's 'Study Abroad' program by traveling to Singapore, where I completed a minor in Applied Mathematics.

Emphasis

General Research Interests

- Empirical Software Engineering
- Deep Learning

Interdisciplinary Interests

- Natural Language Processing
- Psycho-Linguistics Research

Published Projects

- ICLR-GREAT: a simple toolkit to compose arbitrary neural layers to construct models for a variable misuse bug detection problem. `https://github.com/VHellendoorn/ICLR20-Great`
- SLP-Core: a fast, dynamic n-gram language modeling toolkit for source code. `https://github.com/SLP-team/SLP-Core`