EMILY CHENG

emcheng@mit.edu

EDUCATION

Massachusetts Institute of Technology June 2021
Candidate for Masters of Engineering in Computer Science (2021) GPA: 4.7/5.0
Bachelor of Science in Computer Science and Engineering (2020)
Bachelor of Science in Mathematics (2020)

RESEARCH

Emergent Symbolic Communication in Humans and Machines Fall 2020 - Summer 2021
Master's Research: MIT Infolab *Cambridge, MA*

- Supervised by Boris Katz and Andrei Barbu.
- Designed and implemented a communication game to determine how and when sign communication develops between human players, and later machine players.

Few-Shot Text Classification with Meta-Learning Spring 2020
MIT Undergraduate Research: Natural Language Processing Group *Cambridge, MA*

- Supervised by Regina Barzilay. Directly supervised by Yujia Bao and Rachel Wu.
- Extended pipeline for few-shot documentation topic classification in PyTorch to include zero-shot classification baselines.

Reverse-engineering Nanophotonic Systems with BNNs Fall 2018 - Spring 2019
MIT Undergraduate Research: Soljacic Group *Cambridge, MA*

- Supervised by Marin Soljacic. Directly supervised by Sam Kim.
- Implemented a Bayesian neural network with multiplicative normalizing flows to reverse-design the hyperparameters of nanophotonic systems.

INDUSTRY EXPERIENCE

Palantir Technologies Summer 2020
Forward-deployed Software Engineering Intern *Remote*

- Developed insurance risk models using PySpark in Palantir Foundry in collaboration with external clients
- Architected and productionized backend features for map visualization.

Two Sigma Investments Summer 2019
Quant Research Intern: News Team *New York, NY*

- Designed and evaluated alpha models in Python and Groovy to forecast equity and options returns with news data.

Investment Technology Group (now Virtu Financial) January 2019
Algo Quant Research Intern *New York, NY*

- Developed cross-asset market impact models using Python for cash equities execution.

Goldman Sachs Summer 2018
Securities Research Intern: Equities Flow Vol, FICC SMM Execution Services *New York, NY*

- Developed alpha models in Python to forecast realized volatility for trading single stock options that is in production.
- Designed and built an order fill model for systematic trading simulation in Java to integrate submitted orders with historical market simulation data.

Avidyne Corporation Summer 2017
Avionics Systems Engineering Intern *Concord, MA*

· Built map sampling and networking tools in Python to automate and optimize flight planning for terrain alert testing.

PROJECTS

L2 Acquisition and Language Convergence in Neural Language Models Fall 2020
9.190 Group Project

· Conducted cross-lingual transfer and contact language experiments between monolingual French and English LSTM models using toy dataset.
· Found that utterances of monolingual models do not converge, but rather become mutually intelligible.

Cross-Lingual Text-to-Speech Transfer Learning for Low-Resource Languages Spring 2020
6.864 Group Project

· Performed cross-lingual transfer learning on text-to-speech synthesis using German to English single-speaker datasets.
· Designed and conducted Mean Opinion Score tests, finding evidence for optimal periods of transfer from partially trained systems.

Testing and Policy-augmented SIR for COVID-19 Spring 2020
6.435 Group Project

· Created modified Bayesian SIR models for COVID-19 that takes into account priors for testing capabilities and government stringency over time in order to predict ranges for true infection rates.
· Implemented models in PyMC, tested on synthetic data and actual data from Westchester County.

Automatic Image Colorization with Semantic Prior Fall 2017
6.867 Group Project

· Created a Keras/TensorFlow machine learning pipeline that predicts a colorized output image given grayscale input and a semantic tag.
· Designed, implemented, and trained the scene classifer and automatic colorizer CNNs, including data preprocessing and postprocessing.

Debtonator Fall 2016 - Summer 2017
Personal Project

· Designed and tested multiple algorithms for simplifying group IOU networks in Python, and built into a desktop application with tkinter.

AWARDS

Fulbright France Open Research Grant Semifinalist 2021
Meta-learning in low-resource multilingual generalization *ENS Ulm, Paris*

In collaboration with Thierry Poibeau at CNRS and École Normale Supérieure Ulm.

TEACHING

6.031 Software Construction Fall 2020
Graduate Teaching Assistant *Cambridge, MA*

· Held lab hours, graded assignments for students in MIT's intermediate Java software class.

Global Teaching Labs January 2020
Instructor *Grenoble, France*

· Taught middle school, high school, and preparatory school students concepts in math, physics, and computer science as part of an MIT STEM outreach program in Grenoble, France.
· Created and carried out lesson plans in both French and English for students aged 8th grade to prépa.

Math Learning Center Fall 2018 - Spring 2019
Tutor *Cambridge, MA*

· Held twice-weekly office hours for students in the math department taking Differential Equations (18.03), Linear Algebra (18.06), Probability and Random Variables (18.600), Physics (8.01/2) and Calculus (18.01/2).
· Reviewed lecture material and helped students with problem sets and code implementation.

MIT Math Department Fall 2017, Fall 2018
Grader *Cambridge, MA*

· Graded weekly problem sets and exams for Probability and Random Variables (18.600) and Statistics (18.650).

OUTREACH

Harvard MIT Mathematics Tournament Fall 2016 - Spring 2018
Finance Associate, Director *Cambridge, MA*

· Managed all finances for HMMT, a large-scale math competition for high schoolers, first as an associate and then as director.
· Recruited all corporate funding, created and managed the budget, and processed reimbursements and payments.

Florida Association of Mu Alpha Theta Fall 2016
Test Writer

· Wrote statewide math competition tests and solutions for high schoolers in geometry and algebra for Mu Alpha Theta, a widespread math competition circuit in Florida.

COURSEWORK

6.867 Machine Learning (G)	6.860 Statistical Learning Theory (G)	6.337 Numerical Methods (G)
6.435 Bayesian Inference (G)	6.031 Software Construction	18.615 Stochastic Processes (G)
6.864 Natural Language Processing (G)	6.046 Design & Analysis of Algorithms	6.436 Probability Theory (G)
6.884 Sensorimotor Learning (G)	24.933 Semantics & Pragmatics (G)	9.190 Comp. Linguistics (G)

SKILLS

Natural Languages	English (native), Mandarin (fluent), French (C1), Spanish (B1)
Computer Languages	Python, Java, C/C++
Software & Tools	Pandas/Numpy/Scipy, PyTorch/Keras/TensorFlow, Git, Linux, AWS