# andrew j. mcgehee

andrew@andrewjmcgehee.com
https://andrewjmcgehee.com
+1 (205) 300-0030


## EDUCATION

**Auburn University** // Bachelor of Software Engineering 2020
Cumulative GPA: 4.0 / 4.0

## RESEARCH EXPERIENCE

### Auburn University

*Undergraduate Research Assistant* May 2018 - Dec 2020

**bioinformatics**

- Investigating generative machine learning approaches to improve the quality of noisy or shallow multiple sequence alignment (MSA) inputs in ways which are evolutionarily meaningful.
- Developed an open source 3D protein folding simulator (PolyFold) in Java which allows users to visualize distance-based protein folding processes through highly convergent stochastic optimizations — namely gradient descent and simulated annealing.

**evolutionary computing & reinforcement learning**

- Applying evolutionary methods to the automated design of novel adaptive resonance theory (ART) algorithms which are highly tailored to specific problem classes.
- Developed integrations of evolutionary computing techniques and deep reinforcement learning (RL) agents which improve training stability for benchmark control tasks.

**deep learning**

- Investigated and compared the learned representations of common image classification models in order to more intuitively understand the benefits of including adversarial examples in training sets.

### National Water Center Comet Cooperative Proposal

*Software Consultant* Jan 2018 - Jan 2020

**optimization & data mining**

- Designed and implemented an efficient pipeline in Cython for interpolating the stages of ~2.7 million water catchments with a given discharge for a given range of stream orders
- Built a web scraper in Python for automatically downloading nationwide stream data within a given time interval

## PUBLICATIONS & PRESENTATIONS

- Real-time Visualization of Distance-based Protein Folding with PolyFold
  Journal of Visual Examples // Video Paper†
  *Andrew J. McGehee, Sutanu Bhattacharya, Rahmatullah Roche, Debswapna Bhattacharya*

- EDM-DRL: Toward Stable Reinforcement Learning through Ensembled Directed Mutation
  Workshop Paper†
  *Michael H. Prince, Andrew J. McGehee, Daniel R. Tauritz*
- PolyFold: an interactive visual simulator for distance-based protein folding
  PLOS One // arXiv // Journal Paper
  *Andrew J. McGehee, Sutanu Bhattacharya, Rahmatullah Roche, Debswapna Bhattacharya*
- PolyFold
  ACM Bioinformatics & Computational Biology 2020 // Best Poster
  *Andrew J. McGehee, Sutanu Bhattacharya, Rahmatullah Roche, Debswapna Bhattacharya*
- PolyFold: augmenting human intuition with machine learning for protein folding
  Auburn Student Research Symposium 2019 // Oral Presentation
  *Andrew J. McGehee, Debswapna Bhattacharya*

† *under review*

# TEACHING EXPERIENCE

## Auburn University

*Intro to Evolutionary Computing // Teaching Assistant* Aug 2020 - Dec 2020

- Led bi-weekly two hour lab sessions for ~30 students
- Graded 7 programming intensive assignments for 6 students

*Intro to Algorithms // Recitation Leader* Aug 2020 - Dec 2020

- Led bi-monthly 30 minute interactive recitations for ~100 students
- Designed 6 interactive Colab notebooks to enforce key concepts and introduce students to typical programming interview questions

# INDUSTRY EXPERIENCE

## Google

*Software Engineer Intern* Summer 2020

WORK

- Designed and implemented a general infrastructure for joining arbitrary public datasets (e.g., Kaggle, NOAA, COVID-19) with client provided BigQuerry data in C++ and SQL in order to augment AutoML training sets and improve demand forecasting seq2seq model performance.
- Implemented the specific use case for public holiday data augmentation. Augmentations consider geographic location of clients and various user data time granularites in order to properly match holidays with provided datasets.
- Designed a data conversion tool which validates and converts user BigQuery datasets into properly formatted and typed AutoML datasets for model training. Conversions allow users to override traditionally categorical data (e.g., strings) into AutoML strings, enabling the usage of natural language understanding predictions as input signals to the forecasting model.

### Google

*Software Engineer Intern* Summer 2019

WORK

- Integrated black-box hyperparameter optimization tools in C++ into an evaluation pipeline for deep learning computer vision models.
- Increased F1 score by 1.5% in initial proof of concept.
- Wrote hypothesis testing tools in Python to allow insignificant difference of means to trigger early exit in hyperparameter optimization.
- Wrote visualization tools in Python to intuitively demonstrate the progress and relative success of the automated hyperparameter tuning process.

PROFESSIONAL DEVELOPMENT

- Delivered live demo and oral presentation to > 100 senior engineers and peers
- Attended weekly mentor meetings with a staff engineer discussing career trajectory, graduate studies, technical skills, and soft skills

## HONORS

- Auburn University College of Engineering Student Marshal (1 / 1500+ selected)
- Auburn University Undergraduate Research Fellow
- Auburn University Engineering Alumni Council Student Panel (4 / 6000+ selected)
- Auburn University CSSE Business Advisory Council Student Panel (20 / 1000+ selected)
- Auburn University Board of Trustees Scholarship
- Barbara Drummond Thorne Scholarship
- Estelle and Dame Hamby Endowed Scholarship
- Thomas Goode Jones Scholarship

## RELEVANT SKILLS

### General Programming Languages

| | |
|---|---|
| C / C++ | ●●●●● |
| Java | ●●●●● |
| JavaScript | ●●●●○ |
| Python | ●●●●● |

### Markdown / Scripting Languages

| | |
|---|---|
| Bash / Zsh | ●●●●○ |
| GitHub Markdown | ●●●●● |
| HTML / CSS | ●●●●● |
| LaTeX | ●●●●○ |

### Foreign Languages

| | |
|---|---|
| German | ●●●○○ |
| Russian | ●●●○○ |
| Spanish | ●○○○○ |

## RELEVANT COURSEWORK

- Intro to Deep Learning
- Intro to Evolutionary Computing
- Research Methods of Evolutionary Computing

## LEADERSHIP & SERVICE

### Auburn A.I. Club
*Founder, President* Aug 2019 - Dec 2020

RESPONSIBILITIES & ACHIEVEMENTS

- Devised and taught a 12 week beginner friendly curriculum covering topics including: regression, classification, clustering, neural networks, reinforcement learning, and evolutionary algorithms
- Created interactive, take-home Jupyter notebook exercises as well as "from scratch" (Python and Numpy only) implementations of each topic covered to reinforce concepts
- Delivered 45 minute lectures weekly; Average weekly attendance: 50

### Auburn ACM
*President* Aug 2018 - Aug 2019

RESPONSIBILITIES & ACHIEVEMENTS

- Oversaw team of 6 officers responsible for ACM and all sub-clubs: A.I. Club, Ethical Hacking Club, and Auburn Competitive Programming Team
- Increased weekly attendance across sub-clubs from < 50 members to ~150 members

### Auburn Competitive Programming Team
*Coach, Member* Aug 2017 - Aug 2019

RESPONSIBILITIES & ACHIEVEMENTS

- Delivered biweekly lectures covering problem solving strategies for technical interviews, common algorithms, and data structures to ~30 members
- Placed 3rd out of 86 teams in 2019 ICPC Southeastern Regional — Division II
- Placed 8th out of 82 teams in 2018 ICPC Southeastern Regional — Division II

# REFERENCES


## Debswapna Bhattacharya, Ph.D.

**Auburn University**
*Computer Science & Software Engineering, Biological Sciences*
Assistant Professor
bhattacharyad@auburn.edu

## Daniel Tauritz, Ph.D.

**Auburn University**
*Computer Science & Software Engineering*
Interim Director and Chief Cyber AI Strategist, Auburn Cyber Research Center
Associate Professor
dtauritz@auburn.edu

## Dean Hendrix, Ph.D.

**Auburn University**
*Computer Science & Software Engineering*
Director of Computer Science and Software Engineering Undergraduate Programs
Associate Department Chair
Associate Professor
hendrtd@auburn.edu