

# Kevin Berry

(678) 237-3418 · Alpharetta, GA

[kpberry11@gmail.com](mailto:kpberry11@gmail.com) · [linkedin](#) · [github](#) · [kpberry.github.io](http://kpberry.github.io)

## EDUCATION

---

**Georgia Institute of Technology**, B.S. Computer Science

Aug 2015 - Dec 2018

**Selected Coursework:** Machine Learning, Linguistics, Computer Vision, Compilers, Linear Algebra

**Extracurriculars:** Communications officer for The Agency (ML/AI club) and Big O (Theoretical CS club)

**GPA:** 3.96/4.0

## SKILLS

---

**Programming Languages:** Python, Rust, JavaScript, C

**Tools:** PyTorch, pandas, networkx, scikit-learn, Docker, Git, Azure, SQL, MongoDB, L<sup>A</sup>T<sub>E</sub>X

**Human Languages:** English (fluent), Portuguese (intermediate), Spanish (intermediate)

## EXPERIENCE

---

**Machine Learning Engineer**

Jun 2017 - Present

*Worthix*

*Alpharetta, GA*

Here, my goal was to help companies understand customer feedback. I have been responsible for designing, developing, deploying, monitoring, and maintaining a wide variety of tools to accomplish this goal. I recruited and currently direct a small team which works on these tasks with me. The following are some highlights of my individual contributions:

- Developed LLM-based API for large scale multi-level feedback analysis and summarization
- Developed and trained NLP models and API for real-time survey response topic classification
- Developed novel algorithms for large-scale unsupervised graph classification
- Developed queries and statistical analyses for client-facing dashboards
- Deployed and monitored APIs and applications as scalable microservices in Azure
- Conducted scientific experiments to improve survey data and model quality
- Developed fast Monte-Carlo Shapley value approximation for calculating response topic importance
- Developed reverse-geolocation API capable of resolving millions of point-in-region queries per second
- Contributed Python code to networkx which massively improved performance for graph unions and intersections
- Contributed Rust code to polars dataframe library which solved infinite loop in core groupby operations

**Teaching Assistant**

May 2016 - Dec 2017

*Computer Organization and Programming (CS 2110)*

*Atlanta, GA*

- Led recitations on C, Assembly, CPU datapaths, and digital logic
- Wrote software to automate grading of Java programs and circuits

## PROJECTS

---

**image-to-ascii (Rust):** Uses computer vision techniques to convert images and gifs to ASCII art. Capable of converting 120+ images per second. Uses hand-coded SIMD instructions for improved performance. ([Github](#)) ([Crates.io](#)) ([Github](#))

**Rolling a  $d_7$  in Finite(ish) Time:** Paper analyzing the problem of simulating dice with any number of faces given only a standard die set. Proves some efficiency results for interesting cases. ([Paper](#))

**detect-duplicates (Rust):** Recursively finds all of the duplicated files in a directory. Reads each file at most twice and, in expectation, only has one file loaded at a time. Can check the entire Linux kernel repository (roughly 5GB, over 70,000 files) in about 5 seconds. ([Github](#)) ([Crates.io](#))

**Readability Analyzer (JavaScript):** Assigns average text grade level based on standard readability metrics. Uses MLP with hand-engineered features to count syllables. Can analyze 2,000,000+ characters per second. ([Web App](#))

**Vertex Cover Algorithms (Python):** Implemented algorithms to solve minimum vertex cover problem. Made a new approximation algorithm which achieved state of the art accuracy on 10 out of 11 real-world datasets. ([Paper](#))