

# Hayden Housen

[haydenhousen.com](https://haydenhousen.com)  
[github.com/HHousen](https://github.com/HHousen)  
[linkedin.com/in/HHousen](https://linkedin.com/in/HHousen)

[hayden@haydenhousen.com](mailto:hayden@haydenhousen.com) • (845) 518-1380 • Ithaca, NY

## Education

**Cornell University** **Computer Science, BS** **Aug 2021 – May 2025**

- GPA: 3.865, Dean's Honor List, Rawlings Presidential Research Scholar, Cybersecurity Club President
- Coursework: Machine Learning, Artificial Intelligence, Functional Programming, Linear Algebra, Probability and Statistics

## Work Experience

**Software Engineering Intern** **Vocode (YC W23)** **May 2023 – Aug 2023**

- Collaborated directly with founders as a pivotal engineer to shape Vocode's pioneering AI-driven call automation solutions.
- Demonstrated versatility as a fullstack developer by building the FastAPI backend, open-source self-hosted service (Python and asyncio), and frontend dashboard (Next.js), while also resolving issues and reviewing pull requests from the community.
- Evaluated the latencies of multiple LLMs, speech-to-text APIs, and synthesis services to enhance Vocode's performance.

**Undergraduate Researcher** **Cornell University** **Sept 2021 – Present**

- Worked with Dr. Kevin Ellis on unsupervised object discovery using Slot Attention and investigated applications to options learning.
- Overcame bias in paraphrase identification by using transformers & out-of-distribution detection techniques: "[GAPX: Generalized Autoregressive Paraphrase-Identification X](#)." Published in **NeurIPS 2022** (3rd author). Advised by Dr. Sernam Lim at Meta AI.

**Machine Learning Intern** **Ada Support (Remote)** **May 2022 – Aug 2022**

- **Improved Ada's production accuracy by 8%** using only 3% of production data by developing a novel intent classification pipeline.
- Conducted **>60 experiments** and trained >110 models to determine the most accurate methodology.
- Experimented with knowledge transfer, unsupervised learning of sentence embeddings, multi-task learning, and contrastive losses in the context of transformers and support vector machines.

**Machine Learning Intern** **Ada Support (Remote)** **May 2021 – Aug 2021**

- Led the discovery and experimentation phases of a project to enable Ada chatbots to better understand non-English languages.
- Wrote a data processing pipeline to efficiently clean and analyze **9 billion** chat messages for machine learning models.
- Researched novel techniques in multilingual intent prediction and cultivated skills in PyTorch, transformers, and pandas.

## Projects

**AI Music Workstation** **Sonauto** **Jan 2023 – Present**

- Training diffusion models to generate long-form pop songs through a collaboration with StabilityAI researchers.
- Developing a diffusion autoencoder using PyTorch, achieving 64x music compression to a continuous latent space, by testing various diffusion objectives, loss functions, and attention mechanisms.
- Creating [an app](#) with React Native and Firebase to streamline hit pop song creation using generative AI.

**AI Lecture Notes Generation** **lecture2notes** **Sept 2019 – Jan 2022**

- Created a state-of-the-art system to summarize classroom lectures using PyTorch, transformers (BERT), optical character recognition, speech to text, and convolutional neural networks. Source on [GitHub](#). Learn more in the [research paper](#).
- Named a **top 300 scholar in the 2021 Regeneron Science Talent Search**, the nation's oldest and most prestigious science and math competition for high school seniors.
- Deployed ML pipeline in production via a [full-stack website](#) powered by Docker, Flask, Celery, Bootstrap, and Stripe.

**Neural Summarization Library** **TransformerSum** **Mar 2020 – Oct 2020**

- Furthered research in neural-network text summarization with a focus on long document summarization. **340+ stars on GitHub**.
- 4.45x smaller than the state-of-the-art model but 94% as accurate at release. 10+ pre-trained models available.
- Rewrote researchers' code with enhanced performance and a focus on code readability and [thorough documentation](#).

## Technologies and Languages

### Languages

Python, Java, JavaScript, HTML/CSS, SQL, OCaml, C, Bash

### Machine Learning

PyTorch, transformers (BERT), scikit-learn, NumPy, Lightning, pandas, OpenCV, Spacy

### Web

Flask, FastAPI, React Native (Expo), Bootstrap, jQuery, web scraping, API design

### DB and DevOps

MongoDB, PostgreSQL, MySQL, Docker, AWS, Firebase, Supabase, CI/CD, Git