

# HENRIQUE SCHECHTER VERA

☎ 609-955-8911 ✉ [hvera@princeton.edu](mailto:hvera@princeton.edu)

## EDUCATION

**Princeton University B.S.E. in Computer Science**

Sep. 2019–May 2023

**Honors:** Summa Cum Laude, Phi Beta Kappa (early inductee), Tau Beta Pi (early inductee)

Minors in Applied & Computational Mathematics, Statistics & Machine Learning, Optimization & Quantitative Decision Science

**Overall GPA: 3.99, Departmental GPA: 4.0**

- **Graduate coursework:** Advanced Algorithm Design, Data Structures and Graph Algorithms
- **Undergraduate:** Natural Language Processing (NLP), Machine Learning Theory, Machine Translation, Optimization
- **Natural languages:** (native/fluent proficiency) English, Spanish, Portuguese, Guarani
- **Programming languages:** Java, Python, C, C++, Go, SQL, HTML, CSS, JavaScript, Assembly, R, MATLAB

## RESEARCH

**Gemini Embedding: Generalizable Embeddings from Gemini**

2024–2025

Google Deepmind ([Paper](#))

We introduce a state-of-the-art (SOTA) embedding model which produces generalizable embeddings for text spanning numerous languages and modalities. These representations can be precomputed and applied to downstream tasks including classification, similarity, clustering, ranking, and . Evaluated on the Massive Multilingual Text Embedding Benchmark (MMTEB), which includes > 100 tasks across 250+ languages, we substantially outperform prior SOTA models. Our unified model achieves SOTA across MMTEB’s multilingual, English, and code benchmarks, surpassing specialized domain-specific models.

**Structured Pruning for Pre-Trained Models**

2022–2023

Adviser(s): Prof. Danqi Chen, Mengzhou Xia (Princeton Language and Intelligence) ([github.com/princeton-nlp/CoFiPruning](https://github.com/princeton-nlp/CoFiPruning))

We adapt and evaluate the novel structured pruning method CoFi (Coarse- and Fine- grained Pruning) on pre-trained models (e.g. GPT and OPT). The key insight of CoFi is to jointly prune coarse-grained (e.g. layers) and fine-grained (e.g. heads and hidden units) modules, which controls the pruning decision of each parameter with masks of different granularity, as well as a novel layerwise distillation strategy to transfer knowledge from unpruned to pruned models during optimization.

**Selling to a Sophisticated No-Regret Buyer**

2022–2023

Adviser(s): Prof. Matt Weinberg, Second Reader: Prof. Mark Braverman ([Paper](#))

Consider a repeated single item auction with a single buyer who has a value for the item randomly drawn from known distribution  $\mathcal{D}$  each round and bids according to a no-regret online learning algorithm. Say the buyer specifically applies recency bias or  $k$ -switching to their algorithm, which is adversarial to the seller strategy presented in “Selling to a No-Regret Buyer” [Braverman et al., 2017]. In both settings, we (1) find for what parameter values the buyer remains no-regret, and (2) provide an auction which provably yields near-maximal revenue (i.e. revenue arbitrarily close to the expected welfare).

**Extending Randomized Single Elimination Bracket to Multiple Prize Vectors**

2022–2023

Adviser(s): Prof. Matt Weinberg, Prof. Huacheng Yu ([Paper](#))

We extend Randomized Single Elimination Bracket (RSEB) [Schneider et al., 2017] to tournaments with multiple prizes, creating two new tournament ranking rules: *Randomized Complete Bracket* and *Randomized Recursive Bracket*. We prove various guarantees for fairness and manipulability under these rules.

**Exploring the Role of Contextualization in Dynamic Contextualized Word Embeddings (DCWEs)**

2022

Adviser(s): Prof. Karthik Narasimhan ([Paper](#))

DCWEs incorporate both linguistic and extralinguistic context, combining a contextual language model with a dynamic model of time and social space. We evaluate the effect of the dynamic component on different contextual (e.g. GPT) and non-contextual models (e.g. Word2Vec) on a downstream sentiment analysis task. We find that the dynamic component improves performance on BERT and on non-contextual models, but not other contextual models.

**Cannabis Analysis: Understanding Strain Names through NLP**

2021–2022

Adviser(s): Prof. Christiane Fellbaum, Dr. Jérémie Lumbroso ([Paper](#)) We create and publish the largest public [cannabis strain dataset](#) and use NLP techniques (sentiment analysis, word embeddings) to find patterns in strain names using strains’ ratings, effects, flavors, and genealogy. We find that popular strains do not have distinct lexical distributions or sentiment values, strains’ parents account for their popularity, and strains named like their parents actually tend to differ in popularity from their parents.

**Lyric-Based Automatic Playlist Continuation Using Feature Engineering**

2021

Adviser(s): Prof. Srinivas Bangalore ([Paper](#))

We develop a novel automatic playlist continuation system. We establish a new method of featurizing songs based on vocabulary, structure, orientation, style, and semantics, which we validate using dimensionality reduction (t-SNE). We then train a neural network to determine a pair of songs’ compatibility using Spotify’s 1M playlist dataset, which we compare to cosine similarity. Our final system was 4x more accurate than our benchmark in both recommending one song and in recommending five songs.

## EXPERIENCE

---

- Google, Machine Learning Engineer III (NLP)** *July 2023–Present*
- Researching, designing, implementing, and deploying the query understanding model for Google Search (XXXB searches/day)
  - Conducting fundamental research in Natural Language Understanding (neural semantic parsing, representation learning, and information retrieval) and landing state-of-the-art innovations across Alphabet products, such as:
    - Search autocomplete predictions and web retrieval, Ads recommendations, Android sticker and emoji search
- Google, Software Engineering Intern (NLP)** *May 2022–August 2022*
- Completely re-implemented team's flagship neural semantic parsing model
  - Designed architecture alongside researchers, implemented the model, and built training and evaluation infrastructure
  - Led initiative and communication with research teams to open-source our model's classifier and new debugging tools
- Google, Software Engineering Intern** *May 2021–August 2021*
- Developed end-to-end features for electric vehicles in Google Maps, using NLP to create voice queries for Google Assistant
  - Served as direct liaison between two teams involved, driving communications with UX, PM, and SWE correspondents
  - Went beyond assigned tasks: Successfully lead work with a fellow intern to jumpstart a related project
  - Fixed a year-old bug directly affecting user experience. Answered several intern questions on Q&A site and groups
- Landed (Y Combinator W19), Software Engineering and Product Management Intern** *June–Aug. 2020*
- Created and shipped standalone, customer-facing pricing tool using JavaFX and CSS. Determines cost based on amount, location, urgency, and job of the hires. Continues to be used in meetings with prospective clients
  - Led team through product development life cycle for various projects, from ideation to deployment to analysis
  - Introduced successful features to platform with 10k+ users, such as a preference-based jobs page and filming overlay tips
  - Created data analytics dashboards and tracked key metrics to identify growth opportunities, evaluate product performance, and deliver concrete insights. Used SQL (PostgreSQL), Excel, and Google Data Studio
- Princeton Computer Science, Research TA, Head TA, Instructor, Grader, Tutor, Interactor** *Feb. 2020–2023*
- Analysis of Succinct Data Structures (Research TA), Machine Translation (Head TA), Machine Learning (TA + grader), Reasoning About Computation (Tutor), Algorithms and Data Structures (TA + grader), Programming Systems (TA), Introduction to Computer Science (Instructor + TA + Grader)
  - Research TA: help students with CS research projects, give feedback for proposals and papers (with Prof. Huacheng Yu)
  - Head TA: Created and led part of curriculum for intro NLP course, held office hours, monitored discussion forum
  - Instructor: Led an introductory CS class twice a week, teaching 15 Princeton undergraduates. Evaluated applications for future Undergraduate Course Assistants
  - Interactor: Worked with a professor as academic advisors for 1st-year CS majors
- Po Paraguay, Product and UX Intern** *June 2017*
- Conducted user interviews and tested prosthetic performance to refine product designs
  - Created public product pages for our 3D-printed prosthetics, with multilingual assembly guides and videos of prosthetic functionality, increasing reach and accessibility. Recruited and managed 60 volunteers for 2 years afterward

## ACTIVITIES

---

- Princeton Data Science, Officer** *2021–2022*
- Coordinated educational student events and speaker talks with experts, and established formal partnerships with companies
- Princeton Puzzle Hunt, Web Developer** *2020–2022*
- Learned HTML, CSS, and JavaScript to help reconstruct and manage the entire website for this fully digital event
- Jane Street ETC; Facebook; HackPrinceton; ReHack.co, Competitor; Fellow; Coder; Team Captain** *2019–2021*
- Created an algorithmic trading bot. Learned finance and socket programming. Placed 6th out of 22 team
  - Learned advanced CS over 10 weeks with FB Software Engineers
  - Built privacy app that utilizes ML-based facial recognition to protect one's computer from unwanted users
- Tigers For Nassau, Global Research and Consulting Group, Technology Strategist and Consultant** *2020–2021*
- Created digital solutions (delivery platforms, ordering systems) to adapt businesses to the COVID-19 crisis
- TechTogether New York, Organizer** *2020*
- Coordinated and carried out corporate deals and partnerships. Designed our sponsorship package
- Student Government and Student Council, Class Representative, Class Officer (Treasurer)** *2016–2019*
- Supervised marketing, partnerships, finances, staffing, and logistics for events with 40,000+ attendees and 1,500+ athletes

## AWARDS

---

- |  |   |
|--|---|
| <b>Phi Beta Kappa</b>                  | Early Inductee, Sep. 2022                                     |
| <b>Tau Beta Pi</b>                     | Early Inductee, Sep. 2021                                     |
| <b>Paraguay National Math Olympiad</b> | (2x) Silver Medal and Departmental Champion (2017–2018)       |
| <b>International Math Kangaroo</b>     | 1st, 3rd, 3rd, 5th, 5th, and 7th place nationally (2010–2019) |