Devavrat Joshi

+1(650) 444-5277 | devavratj@gmail.com | [GitHub](#) | [LinkedIn](#) | Mountain View, CA

Enthusiastic and skilled recent graduate with a background in Natural Language Processing (NLP), Artificial Intelligence (AI), and Game Design. Seeking a challenging role that leverages my expertise in developing innovative solutions in NLP/AI and game design to contribute to cutting-edge projects and advance the field.

Education

Master of Science in Natural Language Processing, GPA: 3.98	Dec 2023
University of California, Santa Cruz (Silicon Valley Campus)	
Bachelor of Computer Science: Game Design, GPA 3.97	June 2020
University of California, Santa Cruz	

Technical Skills

Proficient in:	Familiar with:
AI/ML – SkLearn, Pytorch	WebGL, Assembly Language
NLP - Specifically RLHF & Conversational Agents	
Python, Java, JavaScript, C, C#, Pytorch	
Game Design, Unity, Phaser	

Experience

Programmer Intern – Llama chatbot with RAG   July– Sept 2023
Bizkonnect, Wilmington, DE
Worked on retrieval augmented generation with llama2 and milvus database using Langchain API.

Research Project - Dehumanization Target Identification   Sept 2022 – Present
UC Santa Cruz
Used Zeroshot NLI models and fine-tuned Roberta to identify the targets of dehumanizing language. With training on 947 examples, reached accuracy of 0.85.

Research Project - Hallucination Detection   Apr 2023 – Present
UC Santa Cruz
Created model to identify six types of hallucinations, faithfully explaining model reasoning based on CONFIT. Annotated 100 summaries with team annotation guidelines. Adapted baseline (Zhou 2021), trained BigBird, reaching F1 of 0.69 across 6 classes. Developed prompting approach with GPT4 to reach F1 of 0.71.

Research Project – History Dialog Corpus Chatbot   May– Jun 2023
UC Santa Cruz
Collected 2900 Reddit posts from r/AskHistorians, generated synthetic dialogs using ChatGPT, used these to train a GPT dialog agent with high Unieval groundedness of 0.8290 and coherence of 0.8775 and perplexity of 19.03.

Research Project – RLHF Data Bias Detection   May– Jun 2023
UC Santa Cruz
Investigated bias in Anthropic and REALTOXICITYPROMPTS using external toxicity metrics. Classified the types of bias using a dictionary classifier. Built a less toxic Reinforcement Learning with Human Feedback (RLHF) tuned T5 model based on Anthropic dataset.

Research Project – Clickbait Spoiling   Jan– Mar 2023
UC Santa Cruz
Question Answering model to detect main point of clickbait. Used BERT models to select sentences and then pass those to get answer spans.

Grader & Tutor – NLP Programming Assignments   Sept– June 2020
UC Santa Cruz
Grader & tutor for Natural Language Processing Class

Gaming Project – Polite Minotaur   July – Sept 2019
UC Santa Cruz
Worked as a Game Programmer & 3D Artist on team of three for maze strategy game, Polite
Minotaur. Created character & environment art in Blender 3D. Implemented AI techniques for enemy
behavior using JS & Phaser.

Research Project – Gene Data Cleanup   June – Aug 2019
UC Santa Cruz
Cleaned up read errors within bacterial genome & isolated viral genes for research paper using
Python.

Gaming Project – Number Puzzle   Jan– Mar 2023
UC Santa Cruz
Developed Number Puzzle Game using JavaScript & WebGL.

Intern Project – NER Resume Tagging   Jun– Aug 2018
BizKonnect, India
Worked as a programmer Intern on a classified gold standard training set for use in NLP model of
managerial resumes for sales intelligence company.