Nishant Balepur

Emails: nbalepur@umd.edu, nishantbalepur@gmail.com
Website: nbalepur.github.io

RESEARCH INTERESTS

I am a first-year Ph.D. student in computer science at the University of Maryland, College Park, working with Professors Jordan Boyd-Graber and Rachel Rudinger. I conduct research with the goal of **aligning LLMs with human values**, with a focus on **factuality** in text generation, **human-centered** frameworks, and LLM **safety** and **intepretability**. I am extremely grateful to be funded by the NSF Graduate Research Fellowship Program.

EDUCATION

- **University of Maryland, College Park (UMD)** — College Park, MD
  Ph.D. Computer Science — *Aug 2023 - Present*
  Advisors: Professors Jordan Boyd-Graber, Rachel Rudinger

- **University of Illinois at Urbana-Champaign (UIUC)** — Urbana, IL
  B.S. Computer Science; B.S. Statistics (Dual Degree); GPA: 4.00/4.00 — *Aug 2019 - May 2023*
  Collaborators: Professors Kevin Chen-Chuan Chang, Jiawei Han, Hari Sundaram, Diyi Yang

PUBLICATIONS AND WRITTEN WORK

- It's Not Easy Being Wrong: Evaluating Process of Elimination Reasoning in Large Language Models
  arXiv:2311.07532
  Nishant Balepur, Shramay Palta, Rachel Rudinger
  TL;DR: We uncover a new weakness of LLMs – reasoning toward incorrect options on multiple-choice questions

- Expository Text Generation: Imitate, Retrieve, Paraphrase
  EMNLP 2023
  Nishant Balepur, Jie Huang, Kevin Chen-Chuan Chang
  TL;DR: We design a task and model to generate multi-sentence and stylistically consistent factual texts

- Text Fact Transfer
  EMNLP 2023
  Nishant Balepur, Jie Huang, Kevin Chen-Chuan Chang
  TL;DR: We propose a complement to style transfer, where models must preserve style while transfering facts

- DynaMiTE: Discovering Explosive Topic Evolutions with User Guidance
  Findings of ACL 2023
  Nishant Balepur*, Shivam Agarwal*, Karthik Ramanan, Susik Yoon, Diyi Yang, Jiawei Han
  TL;DR: We build a model to mine for topics evolutions in large corpora, leveraging user-provided seed guidance

- Mastering the ABCDs of Complex Questions: Answer-Based Claim Decomposition for Self-Evaluating LLMs
  arXiv:2305.14750
  Nishant Balepur, Jie Huang, Samraj Moorjani, Kevin Chen-Chuan Chang, Hari Sundaram
  TL;DR: We study whether LLMs can perform a fine-grained form of self-evaluation

- Aligning Language Models with Factuality and Truthfulness
  Undergraduate Senior Thesis
  Nishant Balepur, Kevin Chen-Chuan Chang

WORK EXPERIENCE

- **Meta** — Menlo Park, CA
  Software Engineering Intern — *May 2022 - Aug 2022*
- **HiMarley** — Remote
  Data Science Intern — *May 2021 - Aug 2021*
- **State Farm** — Champaign, IL
  Actuarial and Modeling Intern — *Aug 2020 - Dec 2020*
- **John Deere** — Remote
  Software Engineering Intern — *Jun 2020 - Aug 2020*

STUDENTS MENTORED

- **Matthew Shu** (B.S. Yale), 2023-Present, NLP in Education
  Submitted a first-authored paper to NAACL 2024

Honors and Awards

- **NSF Graduate Research Fellowship Program (GRFP)**   April 2023 - April 2028
  Provided 3 Years of Support over a 5 Year Period
- **Dean's Fellowship**   April 2023 - April 2025
  Awarded the Dean's Fellowship from UMD for outstanding academic achievement
- **UIUC Computer Science Graduation with Highest Honors**   May 2023
  Recommended by the UIUC computer science department to graduate with highest honors
- **C.W. Gear Outstanding Undergraduate Student**   May 2022
  Awarded to two seniors that have demonstrated excellence in research and service
- **James N. Snyder Memorial Award**   May 2021
  Awarded to three juniors based on academic merit