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, advised by Professors Jordan Boyd-Graber and Rachel Rudinger. I conduct research with the goal of aligning, guiding, and interpreting LLMs, with a focus on **factuality** in text generation, **human-guided** frameworks, and **interpreting** the **safety** and **reliability** of LLMs. I am extremely grateful to be funded by the NSF GRFP and a Cohere for AI Research Grant.

Education

- **University of Maryland, College Park (UMD)** — College Park, MD
  Ph.D. Computer Science; GPA: 4.00/4.00 — *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

- Reinforcement Learning with Student Feedback: Verbalized and Outcome Preferences for Mnemonic Generation
  Under Review
  Nishant Balepur, Matthew Shu, Alexander Hoyle, ..., Seraphina Goldfarb-Tarrant, Shi Feng, Jordan Boyd-Graber
  TL;DR: We align an LLM with fine-tuning and DPO to generate mnemonic devices optimized on learning
- Is Your Large Language Model Knowledgeable or a Choices-Only Cheater?
  Under Review
  Nishant Balepur, Rachel Rudinger
  TL;DR: We investigate whether MCQA leaderboards are influenced by the choices-only abilities of LLMs
- The Prompt Report: A Systematic Survey of Prompting Techniques
  In Progress
  Sander Schulhoff*, Michael Ilie*, ..., **Nishant Balepur**, ..., Alexander Hoyle, Phillip Resnik
  TL;DR: We survey the current field and practices of prompt engineering in NLP (led LLM safety section)
- Plausibly Problematic Questions in Multiple-Choice Benchmarks for Commonsense Reasoning
  Under Review
  Shramay Palta, **Nishant Balepur**, Peter Rankel, Sarah Wiegreffe, Marine Carpuat, Rachel Rudinger
  TL;DR: We quantify the plausibility of answer choices in commonsense MCQA to uncover problematic data entries
- Artifacts or Abduction: How Do LLMs Answer Multiple-Choice Questions Without the Question?
  arxiv:2402.12483
  Nishant Balepur, Abhilasha Ravichander, Rachel Rudinger
  Best Paper Award (4%) and Oral Presentation (7%) at MASC-SSL 2024
  TL;DR: We discover that LLMs can obtain high accuracy without the question in MCQA, and analyze how
- It's Not Easy Being Wrong: Large Language Models Struggle with Process of Elimination Reasoning
  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
- KARL: Knowledge-Aware Retrieval and Representations aid Retention and Learning in Students
  arxiv:2402.12291
  Matthew Shu*, **Nishant Balepur***, Shi Feng*, Jordan Boyd-Graber
  TL;DR: We create a BERT-based retrieval-augmented flashcard scheduler to help students learn more effectively
- 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 transferring 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

Industry Experience

- **Adobe** — San Jose, CA
  Research Scientist Intern (Incoming) — *May 2024 - Aug 2024*
- **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, LLMs in Education
  First-authored paper under review

- **Jerry He** (HS Student), 2024-Present, Crossword Generation with LLMs

Professional Service

- **Conference Reviewer** — UMD
  Reviewer for: ACL 2023, ARR 2023-Present, LREC 2024 — *2022-Present*
- **Visiting Student Day Volunteer** — UMD
  Volunteer and ambassador for UMD's visiting student day — *Mar 2024*
- **Winter Storm LLM Workshop** — UMD
  Led a 5-day workshop on LLMs for non-CS graduate students — *Jan 2023*
- **Computer Science and Statistics Student Ambassador** — UIUC
  Mentor of new students and volunteer for computer science and statistics events — *Aug 2022 - May 2023*
- **SIGNLL** — UIUC
  President of Special Interest Group for Natural Language Learning — *Aug 2020 - May 2021*
- **Co-founder of Project: Code** — UIUC
  Co-founder of student organization to help students build computer science projects — *Aug 2019 - May 2021*

Honors and Awards

- **NSF Graduate Research Fellowship Program (GRFP)** — April 2023 - April 2028
  Provided $159,000 for 3 Years of Fully-Funded Ph.D. Support
- **Cohere for AI Research Grant Program** — April 2024
  Provided $1,000 from Cohere for AI to support the KARL research project
- **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