DS500: Does ChatGPT behave like a human?

By Marius Knipp, Benjamin Herzberger

Winter term 2023/2024

Eberhard Karls Universität Tübingen

School of Business and Economics

LLMs could act as surrogates in (market) research

- Type of AI that understands and produces human-like language
  - Generates most likely next sequence of text
- Models trained on vast amount of text data from the internet
  - Might include: Product reviews, messaging boards, brand forums, …
- Human individuals discuss consumer preferences online
- If training data includes this, answers of LLMs could reflect those preferences
- Market research tools time-consuming and expensive
  - Conjoint analysis
  - Focus group discussions
  - …

➢ LLMs could serve as a fast and low-cost alternative

Brand, James and Israeli, Ayelet and Ngwe, Donald, Using GPT for Market Research (March 21, 2023). Harvard Business School Marketing Unit Working Paper No. 23-062

Previous empirical findings are ambiguous

- Brand et al. (07/2023):
  - GPT-3.5 responds in ways consistent with economic theory
  - Downward-sloping demand curve
  - Estimated WTP for products and features of realistic magnitude
  - Also: State-dependence and human-like order bias

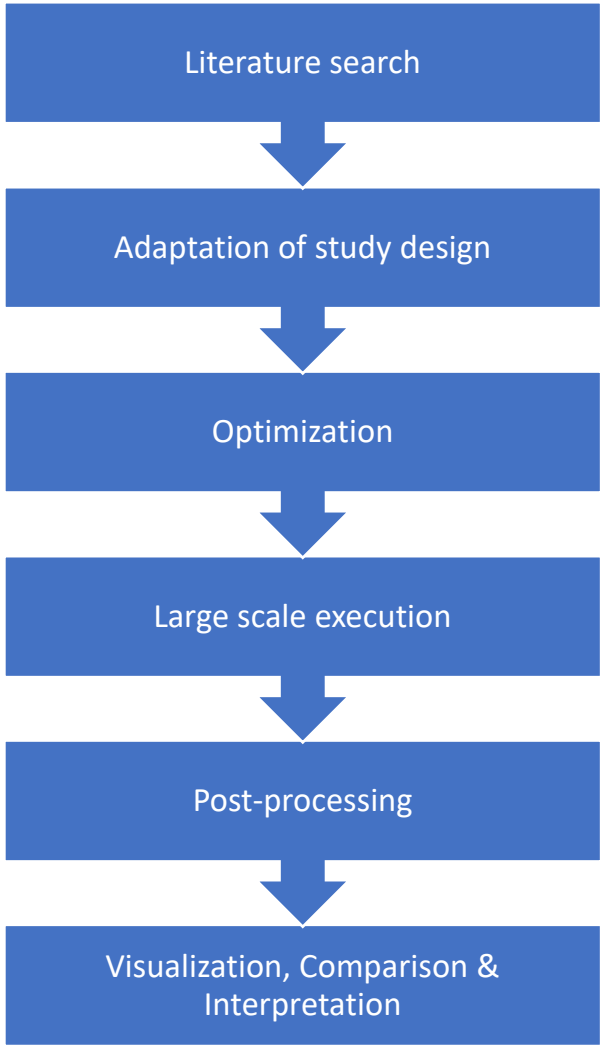
- Dominguez-Olmedo et al. (10/2023) focus on survey-like requests:
  - Models exhibit considerable position and labeling biases
  - Once corrected for this, responses trend to uniform distribution
  - Binary classifiers able to distinguish between model‘s and real responses to the ACS (American Community Survey) with very high accuracy

Brand, James and Israeli, Ayelet and Ngwe, Donald, Using GPT for Market Research (March 21, 2023). Harvard Business School Marketing Unit Working Paper No. 23-062
Dominguez-Olmedo, R., Hardt, M., & Mendler-Dunner, C. (2023). Questioning the Survey Responses of Large Language Models. ArXiv, abs/2306.07951.

Research approach for this project

- Regarded models:
  - OpenAI's GPT-3.5 & GPT-4
  - Meta's Llama-2-70b
- Access through API to issue thousands of requests
  - ➢ Distributional nature of responses
- Focus on *Temperature* parameter
  - Probability of choosing a less likely token
  - 0 for deterministic outcome, higher values induce randomness in response
- Priming & renaming/reordering answer options
- Purposely deviate from original study design
- Recreating consumer behavior patterns via experiments concerning:
  - Transaction Utility
  - Loss aversion
  - …

General workflow for an experiment



- Well-known or well-implementable experiments
- Multiple choice & open answer style

- Prompt engineering and instructions
- Resemble original as close as possible

- Small number of iterations
- Adapt phrasing for optimal results

- Large number of iterations (50 & 100)
- Loop over temperature values

- Identify valid answers
- Construct adequate functions

- Plot results against original findings
- Find similarities, key differences and possible reasons

Transaction Utility experiment

- Important determinant: Reference price
  - "Fair"
  - Seller's costs

- Example: overpriced water at the airport
  - Acquisition utility given
  - Transaction utility low

- Thaler (1985): Hockey ticket study
    - 3 anchors: Price on ticket, initial costs & current price
    - Buyer: Friend or Stranger
    - Expectation: Price individuals charge friends as estimate for „fair" price
  - ➢ Most participants charged friends their initial costs
  - ➢ Most participants charged strangers the current price