Decoding Consumer Preferences: Reverse-Engineering LLMs for Personalization

MPhil Thesis Update, Tentative Pre-Defense Date and Committee

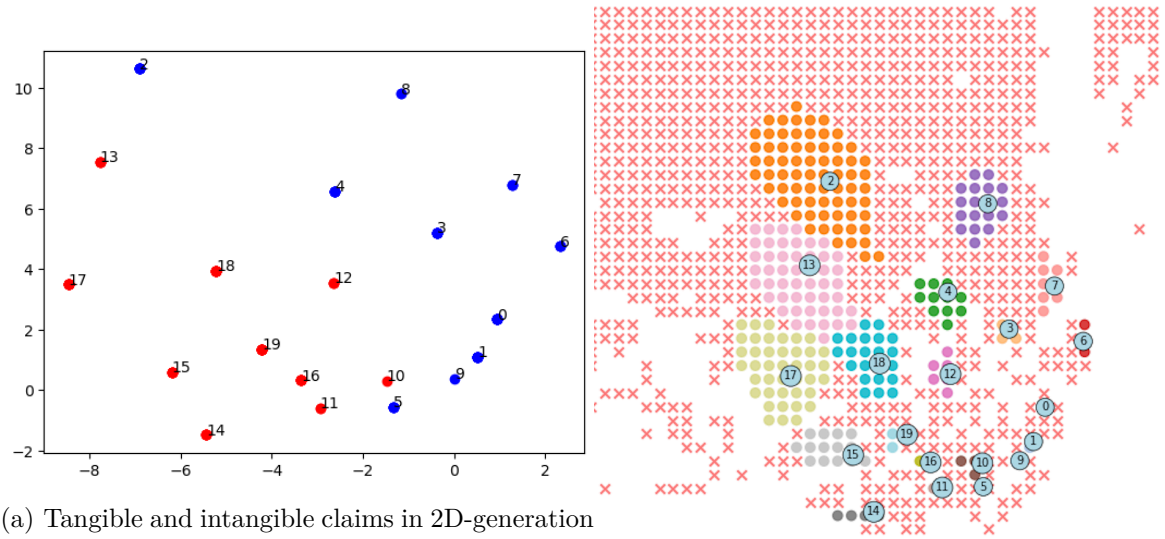
Update

- **Application:** Generating advertising claims with an LLM, based on consumer preferences.
- **Method:** Development of a new method. Find input-embeddings for LLM, that generate a specific advertising claim (we call these "summary embeddings"). We model these "summary embeddings" with an autoencoder, which gives us a "generation space" for advertising claims. These summary embeddings maximize the likelihood to generate a specific advertising claim.
- **Data:** Advertising claims from market research company, different brands, different products. Ratings by respondents on different dimensions, e.g. fit with brand, and preference over other claims.
- **Intended contribution:**
  - Introduction of reverse-engineered "summary embeddings"
  - Use of generative models for personalization in marketing
  - Explore the role of brands in the generation space
  - Learning about consumer preferences and linking them to a "generation-space", from which we can generate new advertising claims. Possibly, relate this to the idea of perceptual maps, i.e. mapping the different offerings on the market and identifying the "gaps" in-between.
- **Key references:**
  - Radford et al. (2018)
  - Devlin et al. (2018)
  - Mullainathan and Rambachan (2023)
  - Pang et al. (2023)

  - Morozov and Tuchman (2024)
  - Burnap, Hauser, and Timoshenko (2023)
  - Li et al. (2024)
  - Schmalensee and Thisse (1988)

- **Progress:** Data acquisition, development of algorithm, validation on toy-data, exploration of generation space.

  - Summary embeddings generate target sequences.
  - Separation of tangible and intangible claims in generation space (Figure 1).
  - Grid-search exploration reveals "candidates" for new claims (red crosses) and "islands" regenerating training data claims (color-coded circles) (Figure 2).



(a) Tangible and intangible claims in 2D-generation space.

(b) Grid-search across 2D-generation space.

Figure 1: Two preliminary results on the toy-data. The toy-data are a collection of ChatGPT generated advertising claims for a hair-shampoo product, where the first half of claims is tangible and the second half is intangible in their wording.

- **Next steps:** Improve training, introduce LLM judge for newly generated claims, incorporate e.g. brands in training process, perform analysis on market research data, explore relation to perceptual maps.

Pre-Defense Admin

- **Expected defence date:** Fri, 28th June 2024

- **Thesis Committee**: Prof. Meike Morren (VU), Prof. Jonne Guyt (UvA); Supervisors: Prof. Fok (EUR), Prof. Donkers (EUR)

References

Burnap, Alex, John R. Hauser, and Artem Timoshenko. 2023. "Product Aesthetic Design: A Machine Learning Augmentation." *Marketing Science* 42 (6): 1029–56. https://doi.org/10.1287/mksc.2022.1429.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." https://doi.org/10.48550/ARXIV.1810.04805.

Li, Peiyao, Noah Castelo, Zsolt Katona, and Miklos Sarvary. 2024. "Frontiers: Determining the Validity of Large Language Models for Automated Perceptual Analysis." *Marketing Science* 43 (2): 254–66. https://doi.org/10.1287/mksc.2023.0454.

Morozov, Ilya, and Anna Tuchman. 2024. "Where Does Advertising Content Lead You? We Created a Bookstore to Find Out." *Marketing Science*, April. https://doi.org/10.1287/mksc.2023.0138.

Mullainathan, Sendhil, and Ashesh Rambachan. 2023. "From Predictive Algorithms to Automatic Generation of Anomalies." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4443738.

Pang, Chao, Jianbo Qiao, Xiangxiang Zeng, Quan Zou, and Leyi Wei. 2023. "Deep Generative Models in De Novo Drug Molecule Generation." *Journal of Chemical Information and Modeling*, November. https://doi.org/10.1021/acs.jcim.3c01496.

Radford, Alec, Karthik Narasimhan, Tim Salimans, Ilya Sutskever, et al. 2018. "Improving Language Understanding by Generative Pre-Training."

Schmalensee, Richard, and Jacques-Frangois Thisse. 1988. "Perceptual Maps and the Optimal Location of New Products: An Integrative Essay." *International Journal of Research in Marketing* 5 (4): 225–49. https://doi.org/10.1016/0167-8116(88)90003-1.