WEN XIE

Houston, TX, USA
(+1) 345 256 6096 ⋄ vaneshieh@gmail.com
linkedin.com/in/vincexie

EDUCATION

University of Houston
Ph.D. candidate in Electrical and Computer Engineering, *GPA: 3.85/4.00* *2018 - present*
Dissertation: Machine Learning and Unstructured Data Analytics for Digital Marketing
Selected courses: Deep Learning, Natural Language Processing, Optimization, Econometrics, Marketing Models, Financial Mathematics, Digital Image Processing
Advisor: Dr. Zhu Han

University of Electronic Science and Technology of China *2014 - 2018*
B.Eng. in Electronic Information Engineering, *GPA: 3.95/4.00*
B.Econ. in Finance, *GPA: 3.95/4.00*

RESEARCH INTEREST

Computer Vision, Natural Language Processing, Statistical Analysis, Causal Inference, Visual Marketing, Online Advertising, Social Media, Consumer Behavior

INDUSTRY EXPERIENCE

Research intern at Snap Research - Computational Social Science (CSS) team *Summer 2022*
- Users watch short videos or photos, shared by their friends (called stories on social platforms such as Snapchat and Instagram). They may encounter an ad between two consecutive stories.
- This project investigates the effects of story-ad congruence on ad viewing time by exploring two types of congruence: media content (17 types such as sports, games, and foods.) and format (video and image).
- Analyzed on-app behavioral data of over over 8 million users, extracted complete sets of covariates such as visual features (temperature and complexity). Employed propensity weighting to account for potential endogeneity, and the results revealed that the congruence increases ad viewing time by 11%.
- Collaborated with two external teams to acquire data for my research beyond the CSS team, and facilitated collaborations between Snap and UH. This project has a significant impact on ad-driven platforms.

Machine Learning Research Intern at Apple - Web Answers and Ranking Team *Summer 2021*
- Answering open-domain multi-hop questions needs multiple sources such as multiple paragraphs from different Wikipedia pages. It is challenging to answer such questions because the concatenation of several paragraphs is long while typical Transformer-based models have input length limits.
- Proposed to use text summarization to summarize multiple sources and extract key information for answering the question, thus avoiding the loss of important information due to the input truncation.
- Rranked Wikipedia data for picking important paragraphs with dense passage retrieval and Built Summarizer and Reader model using Ctrlsum, Pegasus, BART, and T5.
- Finetuned Summarizer and Reader on Hotpot-QA and NQ dataset, and improved Exact Match (by 30%) and F1 score (by 20%). Facilitated QA modeling for production in the industry.

Other experiences: Mentored two high school students on object detection projects; named entity recognition with conditional random fields; abusive language classification with LSTM; box-office prediction with textual and visual features using SVM, LightGBM, and hierarchical attention networks.

WORKING PAPERS

Wen Xie, Mi Hyun Lee, Ming Chen, and Zhu Han. "Understanding Consumers' Attention on Mobile Advertisements: An Ambulatory Eye-Tracking Study with Machine Learning Techniques," 3^{rd} *round of revision at* ***Journal of Advertising*** **Funded by 2021 Amazon Research Award**

- Understanding shoppers' attention to online ads is crucial for delivering messages effectively.
- Recruited 132 subjects for eye-tracking experiments and collected 235 videos recording shoppers' viewing trajectory when they book hotels on the Booking.com website or mobile app.
- Trained YOLO v3 object detection model to extract locations of four ad elements: hotel image, price, rating, and text from eye-tracking videos and tested its performance (mAP: 94.50% in PC shopping study and 87.5% in mobile shopping study).
- Proposed hypotheses based on theory of attention, quantified eye-fixation count and duration on each element during the shopping, and built regression models to test our hypotheses.
- Found that (1) textual ad elements both attract and keep more attention than pictorial ones, and such differences are more pronounced among mobile device users than PC users; (2) mobile ads attract and keep less attention than PC ads; and (3) online ads attract and keep more consumer attention when they are close to deciding on hotel choices, compared to when they search for hotels earlier in their shopping trajectory.

Wen Xie, Gijs Overgoor, Hsin-Hsuan Meg Lee, and Zhu Han. "Not a Black or White Matter: Auto-Detection and Perception of Skin Tone Diversity in Visual Marketing Communication," *in preparation for submission at* ***Marketing Science***

- Objectively measuring skin tone diversity from brand-posted images/videos is vital to researchers and practitioners: (1) brands gain insights into managing their and competitors' diversity positions; and (2) researchers have tools to address practical problems, e.g., colorism in marketplaces.
- Employed semantic segmentation to extract skin pixels from images, used K-means with Elbow method to find the dominant skin tone in an image, and quantified the skin tone brightness.
- Theoretically proposed diversity dimensions and indices, discovered that people perceive (1) darker skin-tone representations, (2) the presence of very dark or very light skin tones, (3) larger skin-tone richness and evenness as more diverse with pre-trained sentence transformer, and validated the proposed indices with t-tests through an online experiment.
- Crawled and analyzed over 70,000 images from fashion brands posted on Instagram and Twitter from 2019 to 2021 and found that (1) medium skin tones account for 51.9% of the whole representations on average over the two years, followed by light (30.8%), dark (10.2%), very light (4.8%), and very dark (2.3%) skin tones; and (2) fashion brands significantly featured more darker skin tones from May 2020 in response to the Black-out Tuesday, but (3) they did not significantly improve skin tone richness and evenness until August 2020, discovered by Bayesian change point detection.

Wen Xie, Ron Dotsch, Yozen Liu, Maarten Bos, and Zhu Han. "Congruence Affects Social Media Ad Engagement," *in preparation for submission at* ***Journal of Marketing Research***

WORK IN PROGRESS

"Advertising with High Quality Image Might Hurt: Insights from Airbnb Demand Analysis Leveraging Image Analytics" - with Sam Hui and Zhu Han.

PEER-REVIEWED CONFERENCE PROCEEDINGS

Wen Xie, Ron Dotsch, Maarten Bos, and Yozen Liu. 2023. "Improving Social Media Video Advertising Acceptance Using Priming: Evidence from Big Data Analysis." *Accepted at 2023 Academy of Marketing Science (AMS) Annual Conference, New Orleans, LA*

Wen Xie, Gijs Overgoor, Hsin-Hsuan Meg Lee, and Zhu Han. 2023. "Automated Detection of Skin Tone Diversity in Visual Marketing Communication." *In Proceedings of 2023 Hawaii international Conference on System Science (HICSS), Maui, Hawaii*

Wen Xie, Ming Chen, and Zhu Han. 2020. "How to Enhance Online Hotel Ad Effectiveness Based on Real-World Data: Mobile Eye-Tracking and Machine Learning Tell." *In Proceedings of 2020 American Marketing Association (AMA) Winter Academic Conference, San Diego, California*

- **Best Paper Award** in Market Research track

CONFERENCE PRESENTATIONS

2023: Hawaii International Conference on System Sciences, Marketing Science Diversity, Equity, and Inclusion (DEI) Conference, AMS Annual Conference (forthcoming)

2021: Annual ISMS Marketing Science Conference

2020: AMA Winter Academic Conference

HONORS and AWARDS

Cullen Graduate Student Success Fellowship (UH)

Best Paper Award in Market Research at AMA Winter Academic Conference

Excellent Student Leader Scholarship (UESTC)

WAC Scholarship (WAC Lighting CO.)

Grade A Certificate of Comprehensive Quality of Undergraduates (Sichuan Provincial Committee of the Communist Youth League of China)

National Encouragement Scholarship (UESTC)

The Provincial First Prize in China Undergraduate Mathematical Contest (Popularization Committee of the Chinese Mathematical Society)

SKILLS

Programming: Python, SQL, Matlab, BlueSky Stats, Eviews
Machine Learning Tools: PyTorch, Keras, TensorFlow
Others: Numpy, Pandas, Scipy, StatsModels, PyMC3, Matplotlib, GGplot
Languages: Fluent in Chinese and English

REFERENCES

Zhu Han
John and Rebecca Moores Professor
Electrical and Computer Engineering
and Computer Science Department
University of Houston
Phone: (+1) 713 743 4437
Email: zhan2@uh.edu

Gijs Overgoor
Assistant Professor of Marketing
Saunders College of Business
Rochester Institute of Technology
Phone: (+1) 704 493 8724
Email: gijsovergoor@gmail.com