Behind the Screens: Unlocking Social Media Analytics – MKTG0005

Instructor: Dr. Vanessa Song
vsong19@wharton.upenn.edu
Office Hours: TBA

Course Description:

From TikTok trends to Formula 1 rivalries, from US Open tennis upsets to the latest must-have drop — social media shapes what's trending, what we talk about, and what we buy. This 7-week course introduces students to social media analytics, blending real-world business insights, data- driven analysis, and the technology behind today's big digital moments.

Students will explore how brands, influencers, and platforms harness data to make strategic decisions. Each week pairs real-world case studies with hands-on Python labs to teach how to collect, process, and analyze social media data. Industry examples from fashion, sports, and viral internet culture keep the content fresh and relatable. By the end of the course, students will have built a mini analytics project on a topic they care about — applying the tools professionals use to decode the digital world.

This course has **no prerequisites** and is designed to be accessible to all students. No prior coding experience is required; all necessary Python skills will be introduced gradually during class. Each activity will be guided step by step to ensure full participation and confidence, regardless of background.

For the assigned readings, students are asked to focus on selected sections, such as the introduction and data description, to gain a general overview. We will go through the more detailed parts together in class, so you can approach the readings without pressure and simply use them as a starting point for our discussions.

Learning Objectives:

- Explain how personalization algorithms shape the content we see.
- Collect and clean social media data from multiple platforms
- Apply sentiment analysis, topic modeling, and social network analysis to real-world cases.
- Analyze images and videos using basic computer vision tools.

1

- Evaluate ethical issues and emerging trends in social media.
- Present findings in a concise, visual, and data-driven story.

Readings:

The readings assigned for each class are integral to the course. Students are expected to have read the listed readings for each class *prior* to the start of the session. Additionally, students are expected to come prepared with questions and critical insights to discuss with classmates.

Requirements and Grading:

1. **Participation and Attendance – 15%**
   a. Attendance in each class is mandatory; see absence policies for excused absences. Evaluated based on thoughtful contributions to class discussions and proactive engagement with subject material. Students are expected to contribute to a positive, open learning environment.
2. **Weekly Assignments – 10% (Due 24 hours before each class)**
   a. Preparation for hands-on Python exercises (e.g., data collection).
3. **Weekly Reflections – 25% (Due 24 hours before the first class of each week)**
   a. 100-200-word reading responses, in-class activities, or case study-based projects.
4. **Discussion Leading – 20%**
   a. Students will help lead class discussions at least once, including summary slides and engaging peers with case questions.
5. **Final Project (Social Media Analysis, Individual or Group) – 30% (Due 7 days after the last class)**
   a. The final project requires students to apply class topics and readings to analyze a social media issue in the tech/AI space and provide strategic recommendations. A proposal is due at least one week before the final paper. Students may choose to complete the project individually or in groups of 2–4, and evaluation will be based on both the quality of the project and each student's contribution.

Academic Integrity:

Students are expected to know and adhere to all university policies and procedures regarding academic integrity (cheating, fabrication, plagiarism, improper collaboration, etc.) and ethical standards of behavior. The work you submit in this class is expected to be your own. If you submit work that has been copied without attribution from any published or unpublished source including the Internet, or that has been prepared by someone other than you, or that in any way

identifies somebody else's work as your own, you will be referred to Penn's Office of Student Conduct. For more information, please consult:

- **University Code of Academic Integrity** – http://www.upenn.edu/academicintegrity/
- **Center for Community Standards and Accountability** – https://csa.upenn.edu/
- **ChatGPT:** The unauthorized use of AI composition software violates academic integrity. This means that unless an assignment explicitly directs/allows you to use AI composition software (like ChatGPT), you may not use such tools to generate text that is subsequently submitted in your own name without attribution.

Course Schedule:

Introduction to ML & Personalization in Social Media

Covers the basics of machine learning concepts and their application in social media personalization algorithms. Students will explore how platforms like TikTok, Instagram, and YouTube recommend content and compete for user attention.

Readings:

- Harvard Business School. (2021). Social media war 2021: Snap vs. Facebook vs. TikTok (Case No. 59905). Harvard Business Publishing.
- Lu, S. (2023). Machine learning in marketing (1st ed., Pearson Originals), Ch. 7 Text Analysis. Pearson Learning Solutions.

Data Collection & Preprocessing for Social Media Text

Focuses on techniques for gathering social media data using APIs and web scraping. Students will learn how to preprocess and clean textual data for analysis, using examples from sports and entertainment events.

Readings:

- Elberse, A., & de Pfyffer, A. (2022). Spotify: Face the music (Update 2022). Harvard Business School.
- Liu, Y., Hu, Y. J., & Liu, Y. (2014). Competing for attention: An empirical study of online reviewers' strategic behavior. Management Science, 60(11), 2723–2737. https://doi.org/10.1287/mnsc.2013.1902

- Lu, S. (2023). Machine learning in marketing (1st ed., Pearson Originals), Ch. 7 Text Analysis. Pearson Learning Solutions.

Please note: This is a preliminary syllabus and is subject to change. It provides useful introductory detail and additional course insight as you prepare for the application process.

Sentiment Analysis & Viral Social Media Events

Introduces sentiment analysis to measure public opinion and emotional tone in social media posts. Students will examine viral events and discuss what makes content spread online.

Readings:
- Rosen, C. (2010). United breaks guitars (Case No. 510057). Harvard Business School.
- Lu, S. (2023). Machine learning in marketing (1st ed., Pearson Originals), Ch. 8 Topic Modelling. Pearson Learning Solutions.

Topic Modeling & Multi-Platform Social Conversations

Covers topic modeling methods to identify themes in large-scale social media discussions. Students will analyze conversations across different platforms and compare engagement trends.

Readings:
- Khanna, T., Varma, K., & Lane, D. (n.d.). Formula One motor racing. Harvard Business School.
- Elberse, A., & Moreno Vicente, D. (2022). Toto Wolff and the Mercedes Formula One Team (Case No. 522-075). Harvard Business School.
- Lu, S. (2023). Machine learning in marketing (1st ed., Pearson Originals), Ch. 8 Topic Modelling. Pearson Learning Solutions.

Social Network Analysis: Who's the Real Influencer?

Teaches students to map and analyze social media networks to identify influential users. Includes discussion on micro vs. macro influencers and the role of community structure.

Readings:
- Freberg, K. (2024). When it comes to influencers, smaller can be better. Harvard Business Review, Sep–Oct.
- Tan, B. C. Y., & Leong, C. (2023). Who is the right influencer? A social network analysis. Singapore Management University Case Publishing.
- Lu, S. (2023). Machine learning in marketing (1st ed., Pearson Originals), Ch. 9 Social Network Analysis. Pearson Learning Solutions.

Please note: This is a preliminary syllabus and is subject to change. It provides useful introductory detail and additional course insight as you prepare for the application process.

Stories Through Pictures: Image & Video Analytics

Examines how computer vision can be applied to marketing analytics, including image recognition and engagement prediction. Students will work with visual data from platforms like Instagram and TikTok.

- Zhang, S., Narayandas, D., Straaberg, S., & Lane, D. (2024). Loris (Case No. 524-010). Harvard Business School.
- Lu, S. (2023). Machine learning in marketing (1st ed., Pearson Originals), Ch. 9 Social Network Analysis. Pearson Learning Solutions.

Ethics, Future Trends & Capstone Project

Discusses ethical considerations in social media analytics, including privacy, misinformation, and algorithmic bias. Students will also explore emerging trends and present their final projects.

Readings:

- The Economist. (2018). Facebook–Cambridge Analytica data scandal. The Economist.
- Colicev, A., & Pauwels, K. (2024). Should your brand hire a virtual influencer? Harvard Business Review, May–June.