MA678 Final Project Proposal

Denny (Ta-Hung) Chen

Project Name:

Facebook Post Comment Volume Prediction

Predict the volume of comments a post will receive in next H hours.

Personal Statement:

I aspire to become a data scientist with a specialization in statistical modeling and machine learning, targeting sectors such as technology, finance, retail, and media. The project at hand presents an interesting and novel challenge, distinct from the field of my past experiences, and intriguingly matches my industry interests.

Question:

1. What is the estimand?
2. What are the assumptions?
3. What are the key features for a Facebook post to gain comments?

Data Source:

UCI Machine Learning Depository: [Facebook Comment Volume Dataset]

Methodology:

The issue pertains to data that is quantified by counting, thus a Poisson regression approach will be taken into account. The "Page Category" attribute has the potential to introduce randomness in the multilevel regression analysis.

Proposed Timeline:

	Week	1	2	3	4	5	6
	Date	~ Nov 9	Nov 10 ~ Nov 16	Nov 17 ~ Nov 23	Nov 24 ~ Nov 30	Dec 01 ~ Dec 07	Dec 08 ~ Dec 12
Work Session	Proposal						
	Data Cleansing						
	EDA						
	Feature Engineering						
	Modeling, Validation						
	Write up						

References:

- Singh, Kamaljot, Ranjeet Kaur Sandhu, and Dinesh Kumar. "Comment volume prediction using neural networks and decision trees." IEEE UKSim-AMSS 17th International Conference on Computer Modelling and Simulation, UKSim2015 (UKSim2015). 2015.
- Singh, Kamaljot. "Facebook comment volume prediction." *International Journal of Simulation: Systems, Science and Technologies* 16.5 (2015): 16-1.