Market Reseach & Business Opportunities Based on Yelp Data

Team #105:

Michael A Fronda, Hal Zhang, Vladimir Orlov, Shivani Narahari, Abhishek S Mishra, Anthony J Mazza

Yelp dataset

Tips

User tips for businesses

Check-ins

Check-ins per business

Business

Business Features such as ratings and location

Users

User statistics and friends

Reviews

User reviews for businesses

Q1 What are you trying to do?

- Study a geographical area for current supply of businesses
- Create demand metric based on customer preferences, *Glaeser et al. [1]*
- Measure delta between supply and demand for business opportunities

Q2 How is it done today; what are the limits of current practice?

- Business Opportunities are not a part of the Yelp product
- Limited to user and business

Q3 What's new in your approach? Why will it be successful?

- Creating a Demand metric based on customer reviews along Business data
- Large enough dataset to calculate a meaningful metric

Q4 Who cares?

- Businesses, Local Entrepreneurs, Investors in small businesses

Q5 If you're successful, what difference and impact will it make, and how do you measure them?

- Creating data driven decisions for businesses
- Natural Experiment
  - Business using the demand metric
  - Business without demand metric

Q6 What are the risks and payoffs?

- Risk: difficulty to calculate actual demand limited to Yelp data
- Payoff: Easier start up for a business

Q7 How much will it cost?

- No monetary costs for usage

Q8 How long will it take?

- 2 Months

Q9 What are the midterm and final "exams" to check for success? How will progress be measured?

- Midterm
  - Data preprocessed
  - Majority of modeling finished
- Final
  - Supply and Demand Calculated
  - Interactive Visualization
- Progress measured via Gannt Chart

Plan of Activities

Project Literature Survey	2/21/2024	2/26/2024	All
Proposal Slides and Video	2/26/2024	2/29/2024	Anthony, Michael
Data Processing	3/1/2024	3/9/2024	Vladimir, Shivani
Modeling	3/9/2024	3/27/2024	Abhishek, Hal
Visualizations	3/27/2024	4/10/2024	Shivani, Michael, Anthony
Final Deliverables	4/10/2024	4/19/2024	All

Innovation: Modeling

- Demand metric
  - Star Rating
  - Number of Reviews
    - Quality of Review, *Sihombing et al. [5]*
    - Text Embedding, *Oscar et al. [4]*
  - Sentiment Score, *Multinda et al. [3]*
  - Time Series short term prediction, *Pereira et al. [2]*

Innovation: Visualization

- Geographic Visualization, *Wang et al. [6]*
  - Business Opportunities
    - Answers: If I want to place a business in this area will there be enough demand?
  - Heatmap with various tooltip information
  - Supply and Demand graphs

Citation

1. Glaeser, E. L., Kim, H., & Luca, M. (2017, November). Nowcasting the Local Economy: Using Yelp Data to Measure Economic Activity (Working Paper No. 24010). National Bureau of Economic Research. http://www.nber.org/papers/w24010.
2. Luis Nobre Pereira & Vitor Cerqueira (2022) Forecasting hotel demand for revenue management using machine learning regression methods, Current Issues in Tourism, 25:17, 2733-2750, DOI: 10.1080/13683500.2021.1999397.
3. Mutinda, James, Waweru Mwangi, and George Okeyo. 2023. "Sentiment Analysis of Text Reviews Using Lexicon-Enhanced Bert Embedding (LeBERT) Model with Convolutional Neural Network" Applied Sciences 13, no. 3: 1445. https://doi.org/10.3390/app13031445
4. B. Oscar Deho, A. William Agangiba, L. Felix Aryeh and A. Jeffery Ansah, "Sentiment Analysis with Word Embedding," 2018 IEEE 7th International Conference on Adaptive Science & Technology (ICAST), Accra, Ghana, 2018, pp. 1-4, doi: 10.1109/ICASTECH.2018.8506717
5. A. Sihombing and A. C. M. Fong, "Fake Review Detection on Yelp Dataset Using Classification Techniques in Machine Learning," 2019 International Conference on contemporary Computing and Informatics (IC3I), Singapore, 2019, pp. 64-68, doi: 10.1109/IC3I46837.2019.9055644.
6. Wang, Y., Haleem, H., Shi, C., Wu, Y., Zhao, X., Fu, S. and Qu, H. (2018), Towards Easy Comparison of Local Businesses Using Online Reviews. Computer Graphics Forum, 37: 63-