Understanding the Competitors in the Consulting Industry

Erin M. Swan-Siegel

Northwest Missouri State University, Maryville MO 64468, USA
`S201660@nwmissouri.edu`

Abstract. Keywords: business consulting · employee satisfaction

1 Introduction

1.1 Module 01 Questions

1. (2 points) What domain you are planning to choose and why?

Business Consultancy. I have been a Business Technology Consultant for nearly 6 years. In any business, it's important to understand the industry leaders and companies of similar size and differing geography, and I've enjoyed applying what I've learned so far in this degree program.

2. (2 points) Where are you planning to search for data sources? If you have them already, begin describing them. Include links. **My data source will be search result data from** https://www.Glassdoor.com

I will use the instruction/guidance provided by Ken Jee. Not only is this a rich resource for Data Science projects, it is particularly relevant to the study topic I eventually chose. [1]

I will use the guidance found in https://www.scrapingdog.com/ **to understand and counteract scraping limitations some website have implemented.**

I will modify the api code block found within a Kaggle dataset to use for this study. [2]

3. (3 points) What data problem do you want to analyze and why it is interesting and important to the world? Be specific and plan for a short but interesting and novel project. Do not duplicate any existing project.

The problem I would like to solve is to understand the attributes associated with consultant companies that are well-liked by their current and past employees. The importance of this study is to understand what elements are needed for a successful consulting company

4. (5 points) What are the steps or phases you may have in your project implementation? You may include a diagram to explain your steps.

Defining the Goal(s) of the study, obtain and prepare the data for analysis, explore the data and plan for the model that will best work for the study, build the model and assess if it is fitting the data in

order to be used for predictive analytics, then be able to communicate the results of the study

5. (2 points) What are the key components of your approach and results? Also, include any specific limitations.

I am planning to use web scraping to gain the initial set of data, SpaCy to classify and count-up key words. One limitation I believe to exist is the data available through the high-level html versus what would be found in full employee reviews, and its ability to be connected to the company's ratings

2   Associated Literature

2.1   Goals of this Study

2.2   Methodology

3   Data

3.1   Sources

3.2   Cleaning

3.3   Exploration

4   Model

4.1   Type Selection

4.2   Building

4.3   Training and Testing

5   Results

5.1   Summary

5.2   Visualizations

6   Discussion

6.1   Conclusion

6.2   Limitations

6.3   Future Exploration

References

1. Jee, K.: Data science project from scratch, https://www.youtube.com/@KenJee_ds
2. Shil,   J.:   https://www.kaggle.com/datasets/joyshil0599/glassdoor-company-insightsscraped-data-collection/code