Business Applications of Natural Language Analytics

Myles D. Garvey, Ph.D

Spring, 2022

1 Required and Recommended Materials

- R and R Studio
- Course Notes (To Be Handed Out)

2 Course Description

Natural Language Analytics (NLA) comprises of three areas: Natural Language Processing (NLP) , Natural Language Understanding (NLU), and Natural Language Generation (NLG). Over the prior decade, NLA has experienced exponential improvements in their capabilities and their applications to various areas within marketing, supply chain, finance, and public policy. In this course, we will survey through these three areas and their applications to the aforementioned areas within business. Students will have the opportunity to learn how to wrangle and analyze text data (NLP), apply foundational models to gain understanding of data contained within text (NLU), and have the opportunity to work with tools that will allow them to auto-generate language (NLG). Applications will showcase document classification, sentiment analysis, content generation, chat bot design, and question/answer. All applications will be showcased in R.

3 Learning Outcomes

- Identify the respective differences between Natural Language Processing, Understanding, and Generation.
- Apply each respective set of basic models and methods to chat bot design, QA sessions, and content generation.
- Interpret the results of each component of NLA to guide better decision making.

4 Prerequisites

Background in R is necessary. If you have not done work in R, then the minimum expectation is that you have worked in some other programming language in the past, such as Python, Java, C++, or Visual Basic. We will go through a brief review of R during our first session.

5 Course Format

This course will meet twice in person and once online. Our in person sessions will last for 4 hours, while our online session will be conducted synchronously at a pre-scheduled time via Zoom.

6 Evaluation

Lab 1 Participation	12.5%
Lab 3 Participation	12.5%
Lab 1 Exercise Set	25%
Lab 2 Exercise Set	25%
Lab 3 Exercise Set	25%

7 Labs

Each day will comprise of a Lab, which will be step-by-step instructions going through lines of R code. You will be expected to follow the code as we go line by line together. Parts of each lab, after demonstrating some of the basic functionalities in R, will comprise of solving problems using the code demonstrated throughout the first part of each lab. Each student is expected to work on their labs individually and in class. The same holds for our virtual session as well. Each lab will comprise of various mini-exercises that need to be solved using the topic material and previously demonstrated code in that day. Lab participation include the code types together that day along with the solutions to the mini-exercises performed in class. Students will also be responsible for undertaking 3 exercise sets, to be completing independently, outside of our meetings. These will be in the form of case-study responses, with code and solutions, as well as interpretations of results.

8 Course Schedule

- Day 1: Intro to NLA
  - Review of R
  - Overview of Natural Language Analytics
  - NLA Applications
  - Short History of Text Analysis

- Day 2 (Virtual Video): NLA Methods
  - What is Text Data?
  - Wrangling and Cleaning Text Data
  - Overview of Natural Language Processing Methods
  - Some Applications of NLP
- Day 3: Natural Language Understanding and Generation
  - Methods of Natural Language Understanding
  - Applications of NLU
  - Methods of Natural Language Generation
  - Applications of NLG