**D214 Task 3**

Jon Isak

D214: Capstone

Contents

[Executive Summary 3](#_Toc155656417)

**Building a Logistic Regression Model to Detect Fake Reviews**

# **Executive Summary**

A.

The proliferation of fake reviews within online platforms poses a substantial threat to both consumers and businesses. This study aims to construct a logistic regression model capable of accurately distinguishing between fake and real reviews in a dataset comprising 20,000 reviews for each category. The research hypothesis asserts that the identified logistic regression model cannot achieve an accuracy rate exceeding 70% in identifying fake reviews.

The dataset obtained from Kaggle comprises 40,433 rows and 4 columns, encompassing continuous, and textual data (Mexwell, n.d.). The 'text\_' column, containing the reviews, is a pivotal predictor for identifying fake reviews. Python, executed in a Jupyter notebook, served as the primary tool for data analysis. This data underwent preparation which included cleaning the text data by removing special characters, punctuation, and stop words, followed by lemmatization. The label column was transformed into a binary variable to indicate a fake or real review. Natural Language Processing (NLP) techniques were employed for text data cleaning, using libraries like NLTK for efficient text preprocessing. Feature extraction was conducted using TF-IDF Vectorization, transforming the text data into numerical features suitable for the logistic regression classifier. The logistic regression model was trained and tested on this prepared dataset. The logistic regression model exhibited promising results within the original dataset, achieving an accuracy of 87% in distinguishing between real and fake reviews.

A screenshot of a computer code

Description automatically generated

However, when applied to an external dataset, the model's accuracy dropped to 50%, indicating challenges in generalizing across diverse review contexts.

A screenshot of a computer

Description automatically generated

Despite achieving high accuracy within the original dataset, the model's inability to generalize well across different review domains poses a significant limitation. This limitation highlights potential issues with dataset variations and the model's adaptability to diverse review contexts. Additionally, while logistic regression is effective in binary classification, it might struggle with capturing nonlinear relationships and adapting to different data distributions.

To address the observed limitations, proposed actions include further analysis of dataset variations between the original and external datasets, exploration of domain-specific models tailored to specific review categories, and refinement of the model using ensemble learning or transfer learning techniques.

Enhancing the model's generalizability through these proposed actions is anticipated to yield substantial benefits. Refining the model through diverse dataset analysis and techniques can result in an approximate 20% increase in model accuracy on external datasets and would establish the model's reliability in distinguishing between fake and real reviews across diverse contexts. This improvement will bolster consumer trust in online reviews and significantly reduce the potential impact of deceptive reviews on businesses, thereby fostering a more transparent and trustworthy online marketplace. Accurate identification and removal of fake reviews can safeguard businesses from reputational damage caused by malicious reviews, ensuring fair competition.

B. https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=3b087d8c-ef11-42bf-a3d7-b0f2006b19f7

C.

Mexwell. (n.d.). *Fake Reviews Dataset*. Retrieved January 2, 2024, from https://www.kaggle.com/datasets/mexwell/fake-reviews-dataset