Sentiment Analysis

Business Problem

This project seeks to perform a sentiment analysis on employee review data. Talent retention and acquisition has historically involved a lot of manual and human effort from hiring efforts to yearly performance reviews to offboarding. Introducing a social monitoring such as sentiment analysis can help the organization gain not only a broader perspective on their workforce but also better understand individuals.

Background/History

Sentiment analysis is a natural language processing (NLP) technique that identifies whether something is positive, negative, or neutral. Focusing on the polarity of the text, sentiment analysis can be tailored to an organization's needs from social data and brand reputation to employee review and experience.

Sentiment analysis on employee review data is especially interesting as it can help the organization sort their qualitative and unstructured data at scale efficiently. Humans also struggle with sentiment analysis manually, so introducing machine learning modeling can reduce bias and improve accuracy overall.

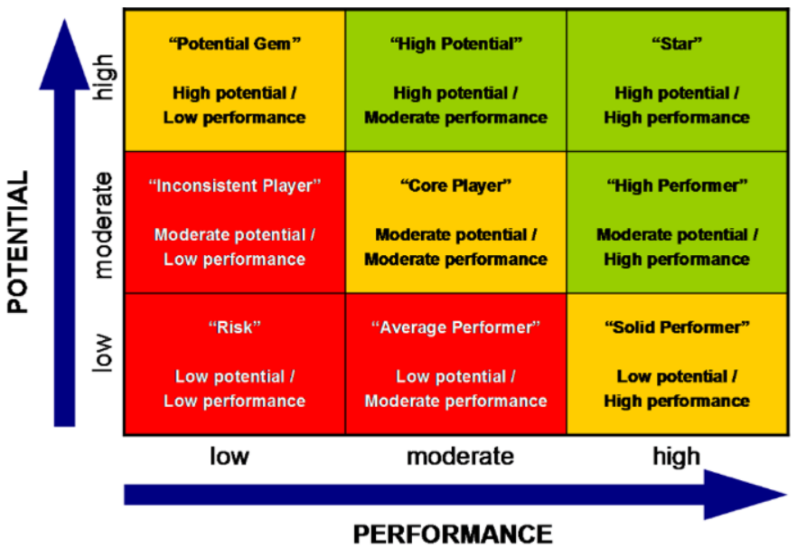
Data Explanation

The dataset used was a part of a research project that sought to leverage BERT for multi-dimensional sentiment analysis. Given that the nature of this data is sensitive, this dataset was collected explicitly or generated in partnership with Amazon MTurk workers (Ryzhkau, 2021). Participants were asked to generate a free-form review of an imaginary colleague that was no shorter than 4 sentences length. From there, the review was tagged based on a 9-box grid, which graded employees across two dimensions—performance and potential (Ryzhkau, 2021).

The author provided two distinct datasets—the core dataset and the test dataset—which do not overlap. The test data set has 225 rows that have been reviewed and included a balanced distribution of classes, unlike the core dataset.

Methods

Initially, TextBlob and VADER were considered for this analysis effort. Ultimately, VADER was selected as the NLP medium as it didn't require previous training before set-up. Both TextBlob and VADER use a lexicon-based method, which is where mapping between words and sentiment is created. As the author of the original dataset created the 9-box grid, the polarity score was defined based off the unique categorization.



Analysis

The initial VADER model defined the polarity score was positive being a score of five and above, neutral being between two and five, and negative being anything less than two. The VADER score was based off general consensus, of positive being anything greater than or equal to 0.05 and negative being anything less than negative 0.05. Neutral was anything else.

Unfortunately, the first model resulted in an accuracy score of approximately 37%, which is significantly worse than just random guessing alone (50%). Additional research was conducted before moving onto additional testing.

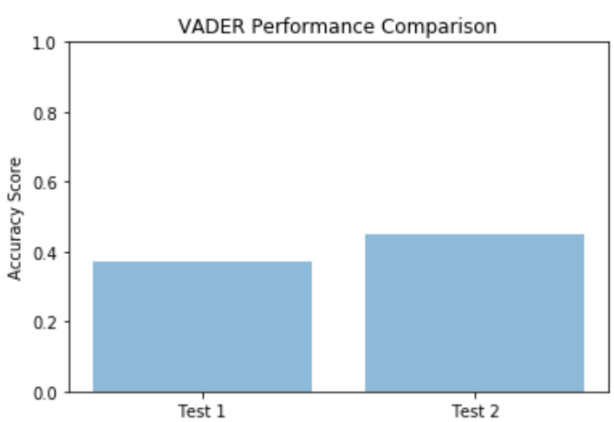
It was identified that the polarity score could have been misattributed to the original author's 9-box grid, as the scoring was not linear as initially thought. Manual matching of numbers to the polarity score was done in order to better align to the 9-box grid. This resulted in a polarity score of:

- Positive: 5, 7, 8
- Neutral: 2, 4, 6
- Negative: 0, 1, 3

The second testing resulted in an accuracy score of approximately 45%, which was a 21% score increase from the initial testing.

Conclusions

Unfortunately, neither model outperformed random guessing, so the sentiment analysis model is not ready for deployment across an organization. Additional work needs to be done on the foundational polarity score in order to see marked improvement of VADER's performance.



Assumptions

Given the qualitative nature of this project, assumptions had to be made to align the reviews to a positive, neutral, or negative scoring. It is not possible to start from a bias-free place of analysis when working with unstructured textual data as this.

Limitations and Challenges

There were several potential barriers to a successful sentiment analysis, as this data is very subjective in nature given the nuanced nature of language. Usually, this type of data would only be accessible by human resources professionals or an employee’s direct leadership, which resulted in the analysis dataset being from a collected exercise.

Generally speaking, there are also limits to how well data can translate to people and their actions or sentiments.

Future Uses and Additional Applications

As natural language processing continues to evolve, organizations will be able to integrate sentiment analysis tools more confidently and accurately. This type of integration will bring a positive impact in reducing bias, manipulation of results, and human interference and influence.

Recommendations

Future work on this project could include further evaluation of the polarity score defined, as there was a significant improvement from the first test to the second test. Alternatively, other natural language processing mediums should be explored, such as TextBlob. One concern was that VADER is more suited for social media analysis and not formal workplace review. It would be interesting to see if TextBlob is able to more accurately identify positive and neutral employee reviews.

Implementation Plan

If this type of work were to be implemented to an organization, it is recommended that it be limited to human resource professionals and key leadership as an informational and guidance tool only. Significant work needs to be done in this space before it is more widely deployed.

Ethical Considerations

As discussed earlier, employee data is sensitive in nature. It is pertinent to conduct any analysis with privacy and legal considerations top of mind, as typically, it would be unlikely that any employee review data would ever be released outside of a direct leadership team or employees with specific security access.

As such, the goal of this project, as related to any real world applications, was to inform direct leadership, human resources business partners, and executive leadership of the sentiments of their organization.

Reference

Ryzhkau, F. (2021). *Employee Review*. Retrieved from Kaggle: https://www.kaggle.com/datasets/fiodarryzhykau/employee-review