Natural Language Processing

Research conducted and validated by third-party Data Science team, in partnership with the Prodoscore Research Council

Published May 19, 2022

prodoscore.com/prodoscore-research-council/

Prodoscore™
Productivity Intelligence

What is NLP?

Natural Language Processing

- A field of study that combines linguistics, mathematics, and computer science to generate technological solutions that concern teaching computers to automatically understand, interpret, and generate human language (Cambria & White, 2014)

- A subset of machine learning, a family of algorithms and techniques that give computers the ability to learn from data and make data-driven predictions

- Sentiment analysis, or opinion mining, is a data mining technique used to extract emotions, opinions, points of view, and other emotional expressions from a document (Liu, 2012)

- In recent years, the field of NLP has adopted the use of pre-trained word embeddings that leverage transfer learning in the form of transformer multilayered encoder- decoder architecture as an innovative approach towards analyzing natural language (Ruder et al., 2019)
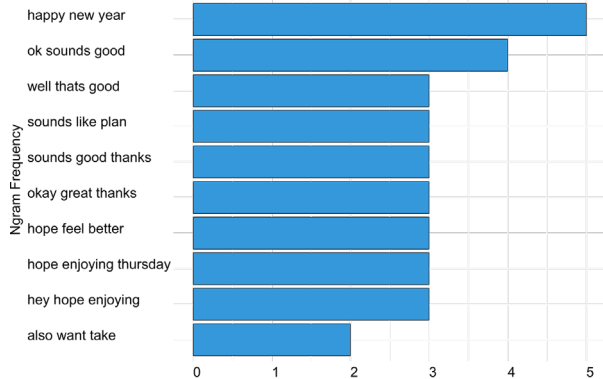
Employee Wellness and Culture Indicator

- As the lived experiences of employees at work are of interest to business and human resource leaders, we have created an **Employee Wellness and Culture** indicator that leverages natural language processing (NLP) and employee chat data to help leaders uncover the artifacts and potential threats to their business and company culture.

- Google's Natural Language Platform uses a transformer encoder-decoder architecture to perform NLP tasks (specifically BERT: Bidirectional Encoder Representation of Transformers) to uncover meaning from employee text. The benefit of using a transformer-based algorithm comes from its ability to quantify and manage the interdependence between data inputs (i.e., words in our example). BERT's attentional mechanism is what manages and quantifies input interdependencies by assigning weights to input features based on their importance.
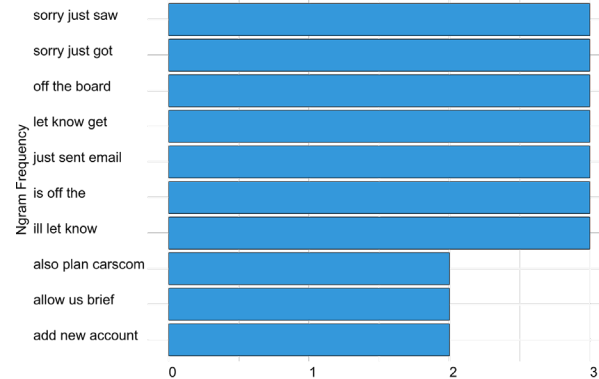
N-Grams

Positive Messages

Ngram Frequency

Phrase	
happy new year	5
ok sounds good	4
well thats good	3
sounds like plan	3
sounds good thanks	3
okay great thanks	3
hope feel better	3
hope enjoying thursday	3
hey hope enjoying	3
also want take	2

The n-grams include the most common 3-term pairs used by Prodoscore employees. The n-grams are separated by sentiment capturing a sample of phrases and terms that are indicative of an employee's happiness.

Neutral Messages

Ngram Frequency

Phrase	
sorry just saw	3
sorry just got	3
off the board	3
let know get	3
just sent email	3
is off the	3
ill let know	3
also plan carscom	2
allow us brief	2
add new account	2

Negative Messages

Ngram Frequency

Phrase	
money back incorrect	2
going stay meeting	2
going drive crazy	2
get pissed calling	2
feel last part	2
dont telling many	2
doesnt make sense	2
didnt get chance	2
calling mobile phones	2
actually going stay	2