EMPLOYEE SENTIMENT ANALYSIS

PROBLEM OVERVIEW

Sentiment analysis helps the business monitor team and individual performance and sentiment through feedback and reviews.

Sorts Data at Scale	Employee review data is comprised of unstructured text that can be overwhelming to process manually.
Efficiency at Scale	Qualitative data efforts are time-consuming to review manually; sentiment analysis speeds up the process and inform next steps more quickly
Reduce Bias	Tagging text by sentiment is highly subjective work and is unique to the individual doing the review. Sentiment analysis can improve accuracy and introduce standardization across an organization.

WHAT IS VADER

- Valence Aware Dictionary for Sentiment Reasoning
- Rule-based sentiment analyzer
- Natural Language Toolkit (NLTK) module
- Processes textual data

ABOUT THE DATA

- Collected with the help of Amazon Mturk workers and provided via Kaggle
- Dataset is comprised of main and additional columns

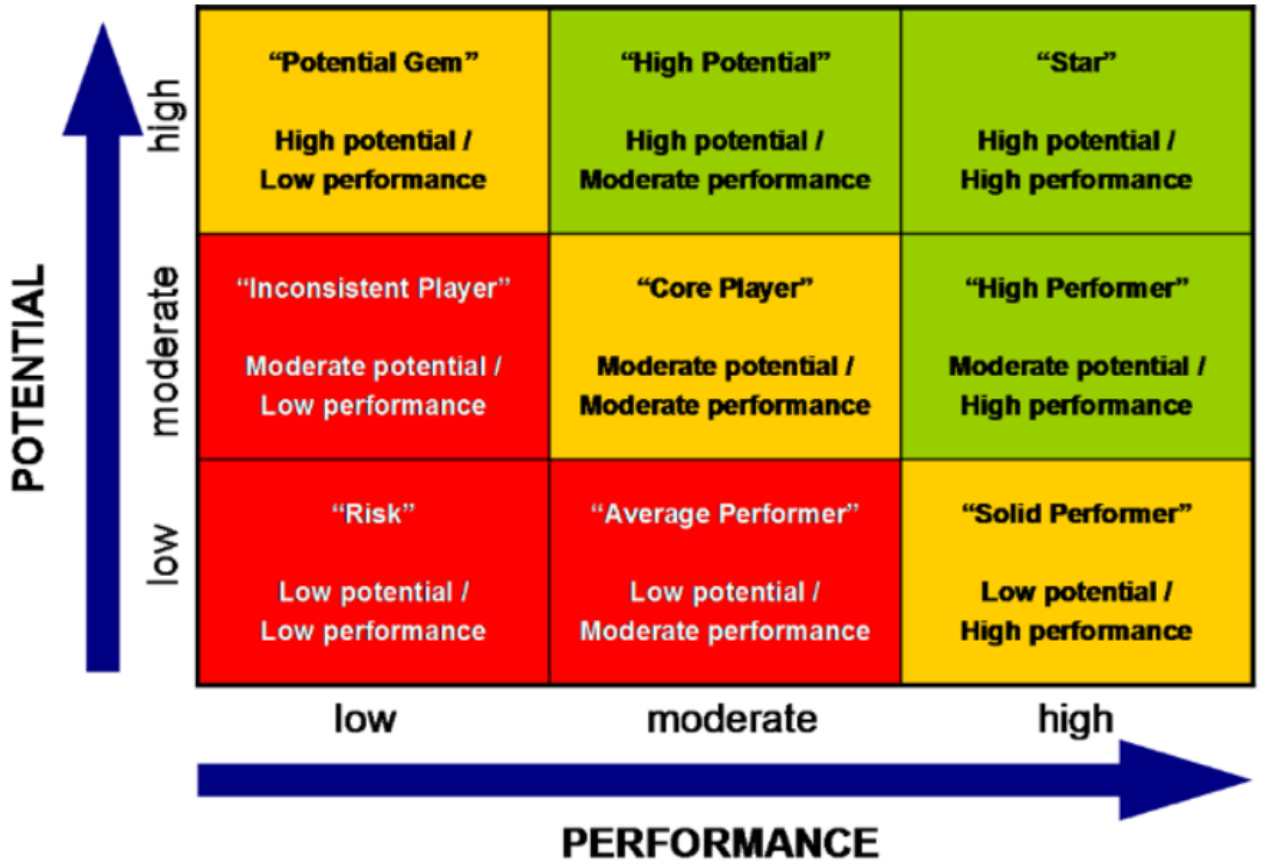
Main columns:

•**id** - unique identifier of the record
•**person_name** - imaginary employee name, for which feedback was given
•**nine_box _category** - human-readable 9-box category
•**feedback** - the actual review on the employee
•**updated** or **adjusted** - whether original category provided by MTurk employee was updated to properly match with the feedback (to sustain high degree of consistency)
•**reviewed** - flag that says whether this record was thoroughly reviewed or not with another pair of eyes

Additional columns:

•**label** - 0-based nine_box_category id
•**feedback_len** - length of the feedback
•**num_of_sent** - number of sentences in feedback
•**performance_class** - 0-based performance class id
•**potential_class** - 0-based potential class id
•**feedback_clean** - pre-processed feedback value

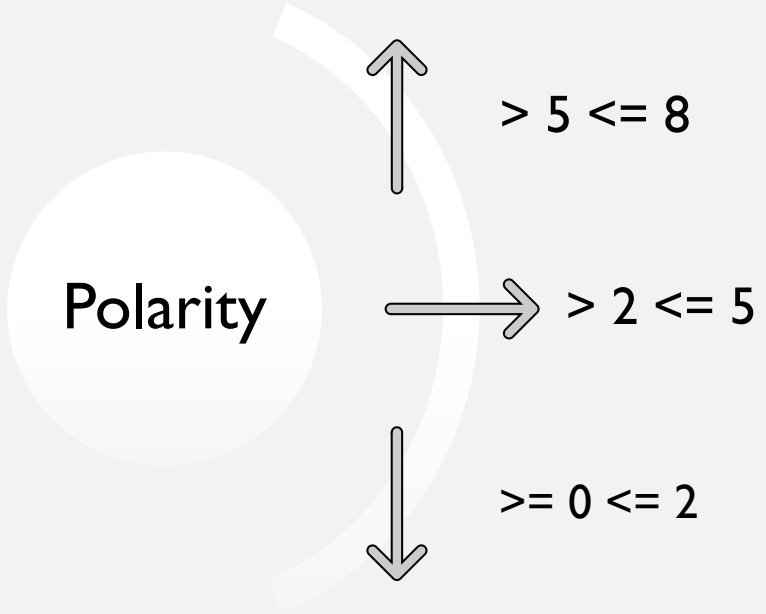
9-BOX GRID


POTENTIAL
high
moderate
low
"Potential Gem"
High potential / Low performance
"High Potential"
High potential / Moderate performance
"Star"
High potential / High performance
"Inconsistent Player"
Moderate potential / Low performance
"Core Player"
Moderate potential / Moderate performance
"High Performer"
Moderate potential / High performance
"Risk"
Low potential / Low performance
"Average Performer"
Low potential / Moderate performance
"Solid Performer"
Low potential / High performance
low
moderate
high
PERFORMANCE

DATA CLEANSING AND PREPARATION

- Validated dataset for any NaN/Null values
- Defined Polarity Score
- Defined VADER Score

Polarity Score



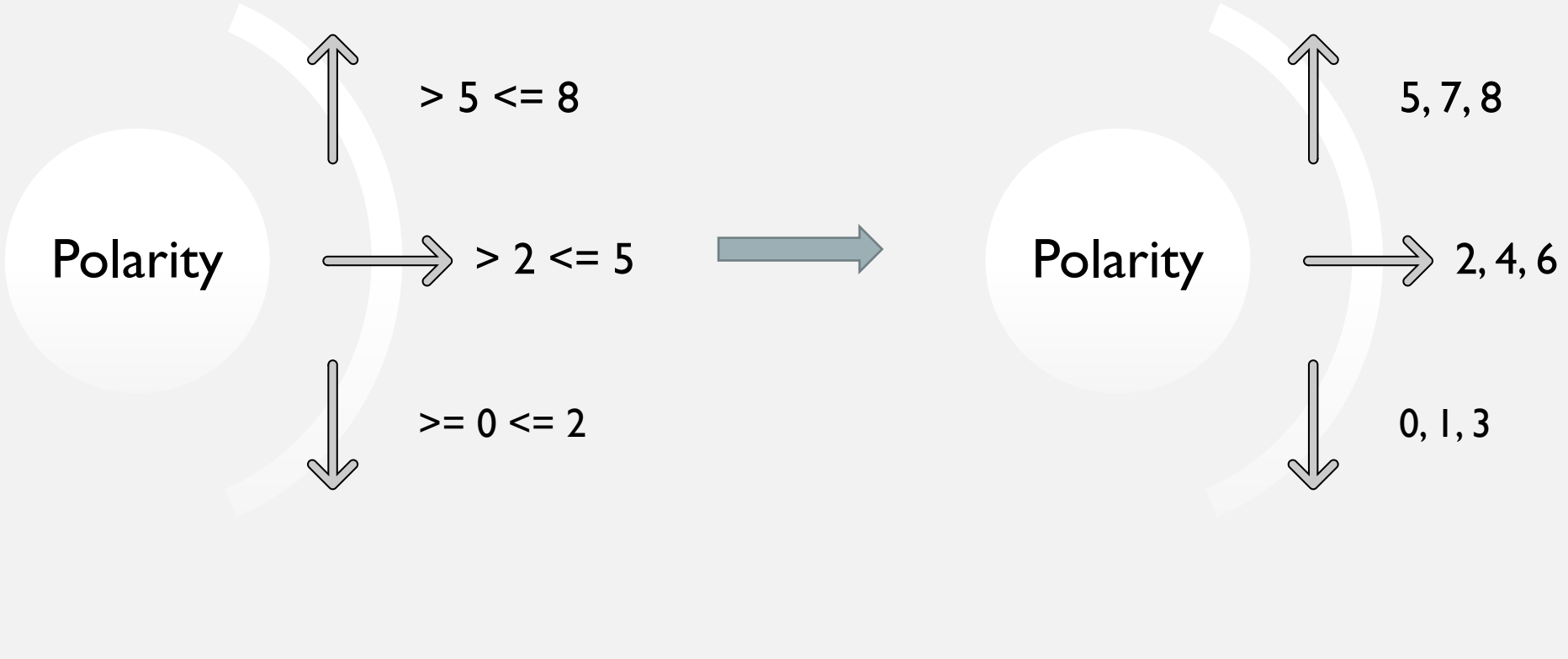
VADER Score

MODEL EVALUATION

Accuracy: 0.3706896551724138

	precision	recall	f1-score	support
negative	0.83	0.24	0.37	42
neutral	0.33	0.02	0.04	42
positive	0.32	1.00	0.48	32
accuracy			0.37	116
macro avg	0.49	0.42	0.30	116
weighted avg	0.51	0.37	0.28	116

CHANGING THE POLARITY SCORE

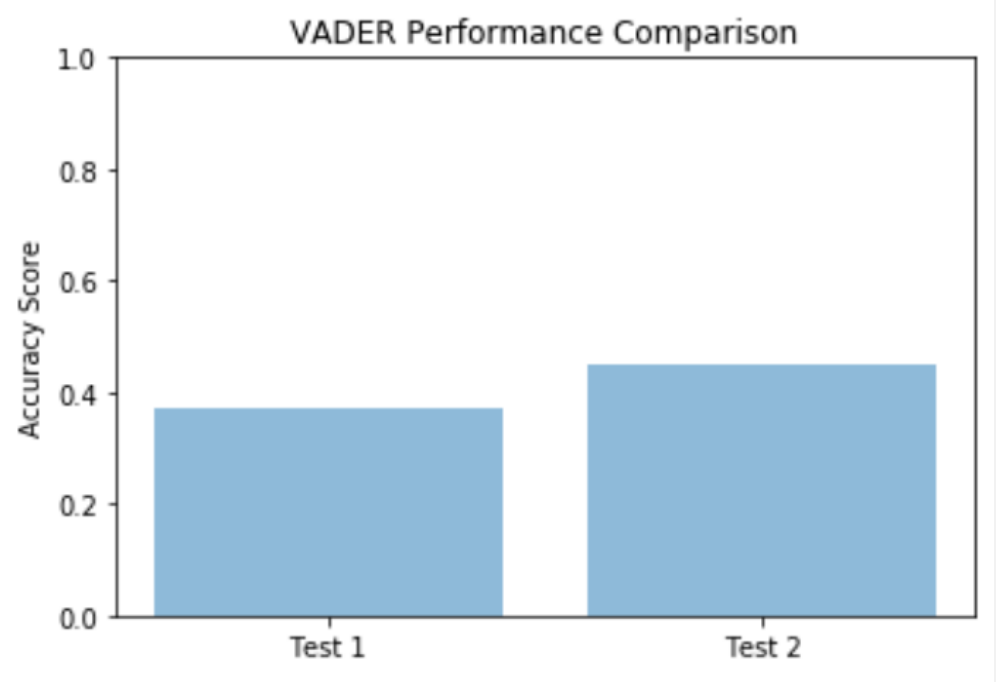


2ND MODEL EVALUATION

Accuracy: 0.4481707317073171

	precision	recall	f1-score	support
negative	0.85	0.31	0.45	249
neutral	0.21	0.02	0.03	189
positive	0.39	0.99	0.56	218
accuracy			0.45	656
macro avg	0.49	0.44	0.35	656
weighted avg	0.52	0.45	0.36	656

COMPARING RESULTS



2nd test performed 21% better than the 1st test

	Test	Accuracy Score	Average Precision	Average Recall	Average F1-Score
0	test 1	.37	.51	.37	.28
1	test 2	.45	.52	.45	.36

ETHICAL IMPLICATIONS AND LIMITATIONS

- Dataset required is sensitive in nature and information contained is usually only shared with pertinent HR professionals and direct leadership
  - Normal circumstances prevent this type of data from being shared outside of the organization, so the dataset used was collected explicitly for creator's use case
- Ethical data collection and analysis pertinent to prevent bias, manipulation, or influence
- There are limits to how well data can portray people and their actions, only meant to guide and inform organizations;

POTENTIAL FUTURE WORK

Different NLP

- Neither model resulted in better performance than random guessing (50% accuracy)
- Other sentiment analysis mediums such as TextBlob may be better suited for this dataset

Adjust the Polarity Score

- Original dataset creator used a complex 9-box grid that may not be aligned to sentiment analysis functionality; difficult to capture nuances in a trinary polarity score

REFERENCES

Ryzhykau, F. (2020). *Employee Review*. Retrieved from Kaggle: https://www.kaggle.com/datasets/fiodarryzhykau/employee-review