Read the Reviews: Analyzing NLP Signals of Wayfair Products
Alex Spiride, Erin Yang, Jason Zhang, Sangyoon Park

ABSTRACT
E-retailers need to predict future return rates for quality control and pricing applications. We investigate NLP methods for feature extraction from free-text reviews. These signals meaningfully improve the prediction of future product return rates.

GOAL
Do product reviews contain signals that improve the prediction of future return rates?

DATA
Product-level
Product-level variables include market categories (e.g. sofas, rugs, lamps), average weight, and historical annual return rate.

Review-level
Reviews feature star ratings, whether the product was returned, text, and purchaser information.

NLP FEATURE EXTRACTION

Sentiment Analysis
We use the VADER rule-based sentiment extraction approach originally developed for social media and find it performs better than more general methods.

BERT Returnability & Embeddings
Transfer learning with BERT encodes the probability that a reviewer returned their order. We also extracted word embeddings and analyzed what features they encode.

MODELS
Wayfair wants to avoid incorrectly removing items from their catalog. We built an asymmetric loss function that penalizes overprediction more than underprediction. Our goal is to determine whether the Full model outperforms the Non-NLP model with this loss function.

Regression: Predicting Return Rates
We fit Random Forest and Neural Network regressors to product data.

Classification: Outlier Detection
To help Wayfair identify low-quality products, we define outliers as the products in the highest 10% returned for each market category. The Reg+Rank classifier ranks the RF regressor’s raw predicted rates to determine outliers. A precision metric encodes the false positive penalty.

CONCLUSION
We successfully extracted interpretable NLP features from the review text data. These features improve the prediction of return rates in both outlier classification and regression settings. This increase is robust to changes in the test set and different metrics.

REFERENCES

Certain details, including certain figures and numbers, have been transformed in advance of this analysis.