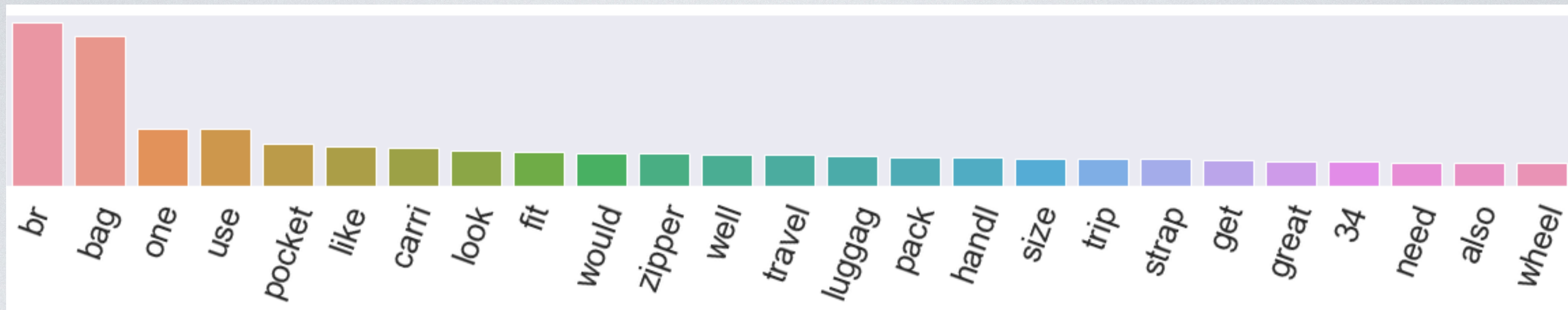


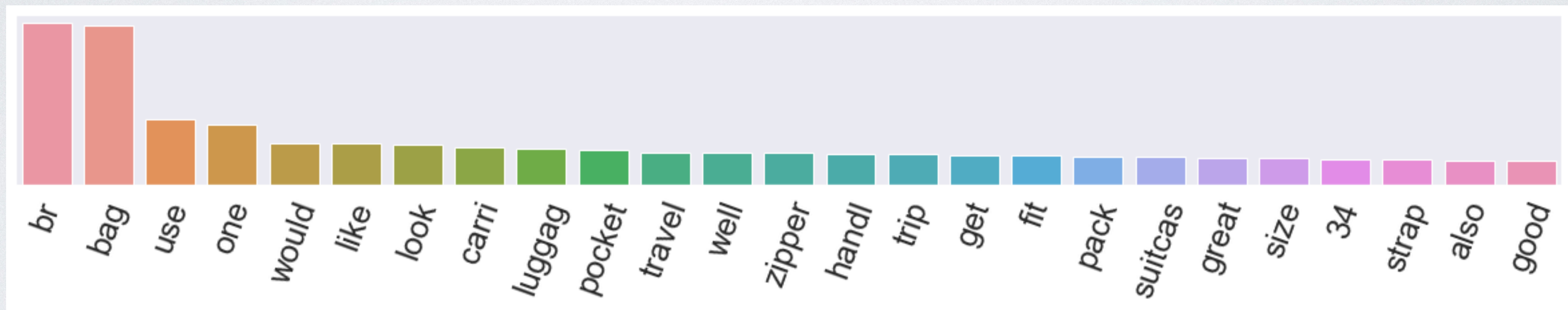


HELPING AMAZON REVIEWERS

Write More Helpfully



HELPFUL REVIEWS' TOP 25 MOST COMMON TOKENS



UNHELPFUL REVIEWS' TOP 25 MOST COMMON TOKENS

Helpful and **Unhelpful** Review Word Frequencies

Step 1:

Data preparation

Step 2:

EDA

Step 3:

Modeling

PROCESS

Random Forest

TRAIN

Accuracy: 0.6013

Recall: 0.7055

F1-Score 0.6365

TEST

Accuracy: 0.6040

Recall: 0.6951

F1-Score: 0.6383

XGBoost

TRAIN

Accuracy: 0.8370

Recall: 0.8415

F1-Score: 0.8363

TEST

Accuracy: 0.6135

Recall: 0.6049

F1-Score: 0.6114

RESULTS

BERT

Bidirectional Encoder Representations from Transformers

ERNIE

Enhanced Language Representation with Informative Entities

NEXT STEPS