☆ Star 0

E

NEW FEATURES Lots of features were considered for training the model (including

use of smiles, dots, and various punctuation in a review)

• Length of the review Capitalization

- Use of punctuation (? and !)
- ngrams (word combinations)
- Words and ngrams frequencies

The following features were chosen:

- Checking for misspelled words
- "Bag of words" method

Multiple algorithms were considered including MultinomialNB, XGB, Bernoulli, RandomForest

ALGORITHMS

The following were chosen for the best model:

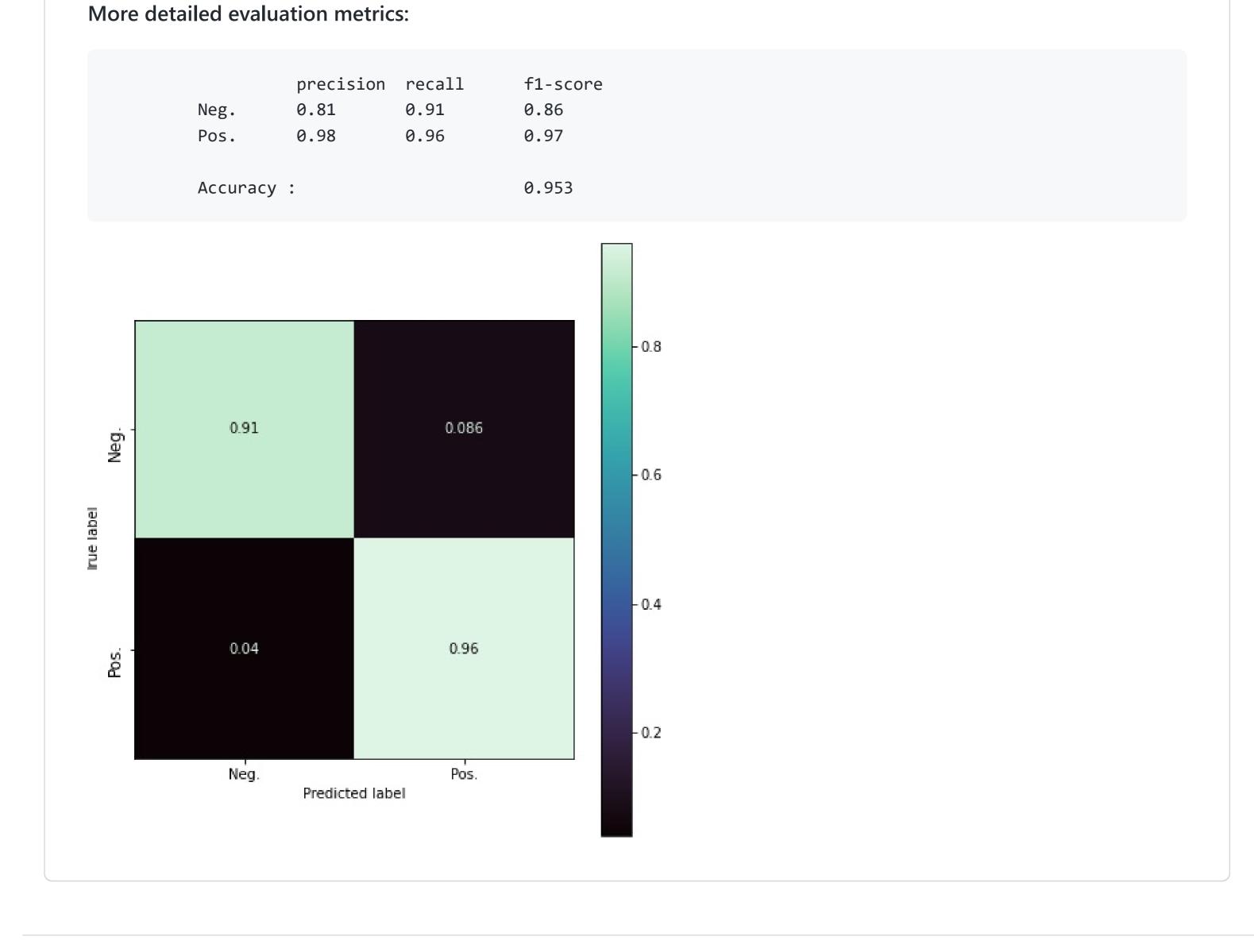
ComplementNB

Voting Ensemble of:

XGB boost

EVALUATION

• Accuracy of 95.3 was achieved



© 2022 GitHub, Inc. Terms Privacy Contact GitHub Security Pricing Training Status Docs About