Detecting spam and non spam Amazon reviews

Problem Context:
Amazon reviews are becoming flooded with fake reviews in order to boost product ratings. This has removed some of the trust that users usually instill in product ratings. Many sub-par products have thousands of 5 star reviews with many done by fake accounts. This is creating a worse user experience for Amazon customers and reducing the trustworthiness of the brand

Criteria for Success
We are able to split reviews by spam vs. non-spam at a high enough rate to prevent further falsified reviews. We are also looking for key-words or phrases that spammers use.

Scope of Solution Space
We have 25 million labeled reviews spam and non-spam. This data will be used to create a model that I can then use on new data that I will scrape from the web.

Constraints
Spam vs non-spam data may not be a clear line in real life. There may be real reviews that look like spam or vice versa. Human input may vary widely depending on product and emotional responses. Also, my scraped data will be unlabeled, therefore any model used we will not know the accuracy unless we go line by line.

Stakeholders:
Amazon
Amazon Users
Amazon sellers

DataSource:
Kaggle
Amazon.com

Plan:
I will use machine learning focused around NLP datasets such as Naive Bayes. I will train several models, do cross validation andd hyperparameter tuning before deploying my model onto the scraped data. The scraped data will be acquired through beautifulsoup

Amazon reviews labeled spam or not spam
https://www.kaggle.com/naveedhn/amazon-product-review-spam-and-non-spam