jupyter **second** Last Checkpoint: 8 minutes ago

File   Edit   View   Run   Kernel   Settings   Help

Markdown ∨

1. Dataset overview

1.1 Brief overview of the Amazon product review dataset. ¶

The dataset contains the following columns:

1. `reviewText` : The review text.
2. `Positive` : The target variable. It is a binary variable indicating whether the review is positive or neg

The dataset contains 20000 rows and 2 columns.

1.2 Describe columns

```python
import pandas as pd
df = pd.read_csv('amazon.csv')
df.head()
```

```python
df.shape
```

2. Data Preprocessing