	spam	text
0	0	Subject: enron methanol ; meter # : 988291\r\n...
1	0	Subject: hpl nom for january 9 , 2001\r\n(see...
2	0	Subject: neon retreat\r\nho ho ho , we ' re ar...
3	1	Subject: photoshop , windows , office . cheap ...
4	0	Subject: re : indian springs\r\nthis deal is t...
...	...	...
5166	0	Subject: put the 10 on the ft\r\nthe transport...
5167	0	Subject: 3 / 4 / 2000 and following noms\r\nhp...
5168	0	Subject: calpine daily gas nomination\r\n>\r\n...
5169	0	Subject: industrial worksheets for august 2000...
5170	1	Subject: important online banking alert\r\ndea...

5171 rows × 2 columns

	text
0	[Subject, enron, methanol, meter, 988291, foll...
1	[Subject, hpl, nom, january, 9, 2001, see, att...
2	[Subject, neon, retreat, ho, ho, ho, around, w...
3	[Subject, photoshop, windows, office, cheap, m...
4	[Subject, indian, springs, deal, book, teco, p...

dtype: object

```
hello world hello hello world play

<Compressed Sparse Row sparse matrix of dtype 'int64'
       with 6 stored elements and shape (2, 5)>
  Coords        Values
  (0, 0)          3
  (0, 4)          2
  (0, 2)          1
  (1, 0)          1
  (1, 3)          4
  (1, 1)          1

bow4 shape: (2, 5)
```
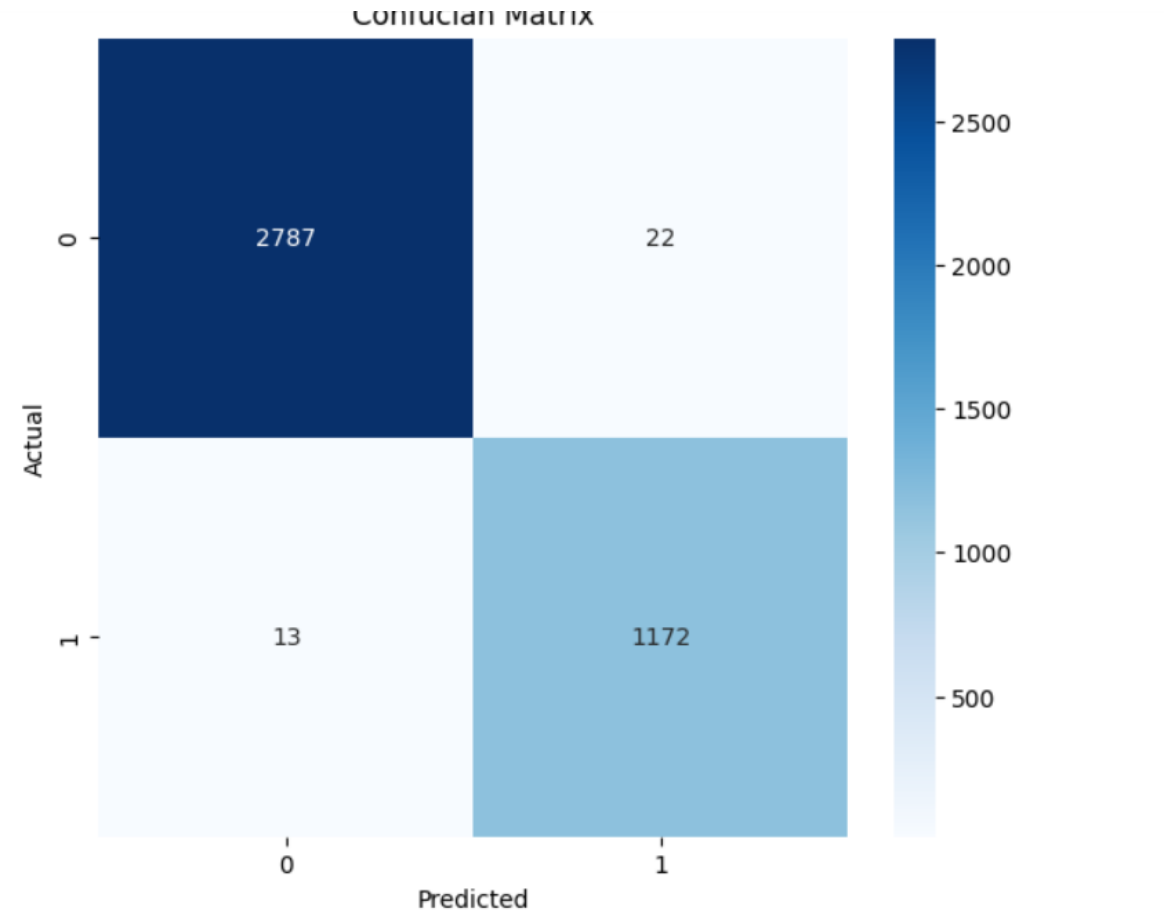
Predictions: [1 0 1 ... 0 0 1]

Y Train Values: [1 0 1 ... 0 0 1]

Classifcation Report (y-train ~ pred)

	precision	recall	f1-score	support
0	1.00	0.99	0.99	2809
1	0.98	0.99	0.99	1185
accuracy			0.99	3994
macro avg	0.99	0.99	0.99	3994
weighted avg	0.99	0.99	0.99	3994

Accuracy: 0.9912368552829244

Confusion Matrix:
 [[2787   22]
 [13 1172]]

Confucian Matrix

```
--------DF info--------
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5171 entries, 0 to 5170
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   spam    5171 non-null   int64
 1   text    5171 non-null   object
dtypes: int64(1), object(1)
memory usage: 80.9+ KB
-----------------------

Original Shape: (5171, 2)
Cleaned Shape: (4993, 2)

Dropped 178 duplicated rows

Missing Data:
```

	0
spam	0
text	0

dtype: int64