report

November 2, 2024

1 Imports

[6]:
```
import pandas as pd
import matplotlib.pyplot as plt
```

[14]:
```
import dataset_gen as dg
```

2 Description

This is a report to accompany Lennie's codebase for Owain Evans' selection exercise.

[3]:
```
df = pd.read_csv('cache/odd_maxint_40_N_200_n_egs_[4, 8, 16,␣
↪32]_seeds_50_gpt-4o_maxtokens10_summary.csv')
```

[4]:
```
df
```

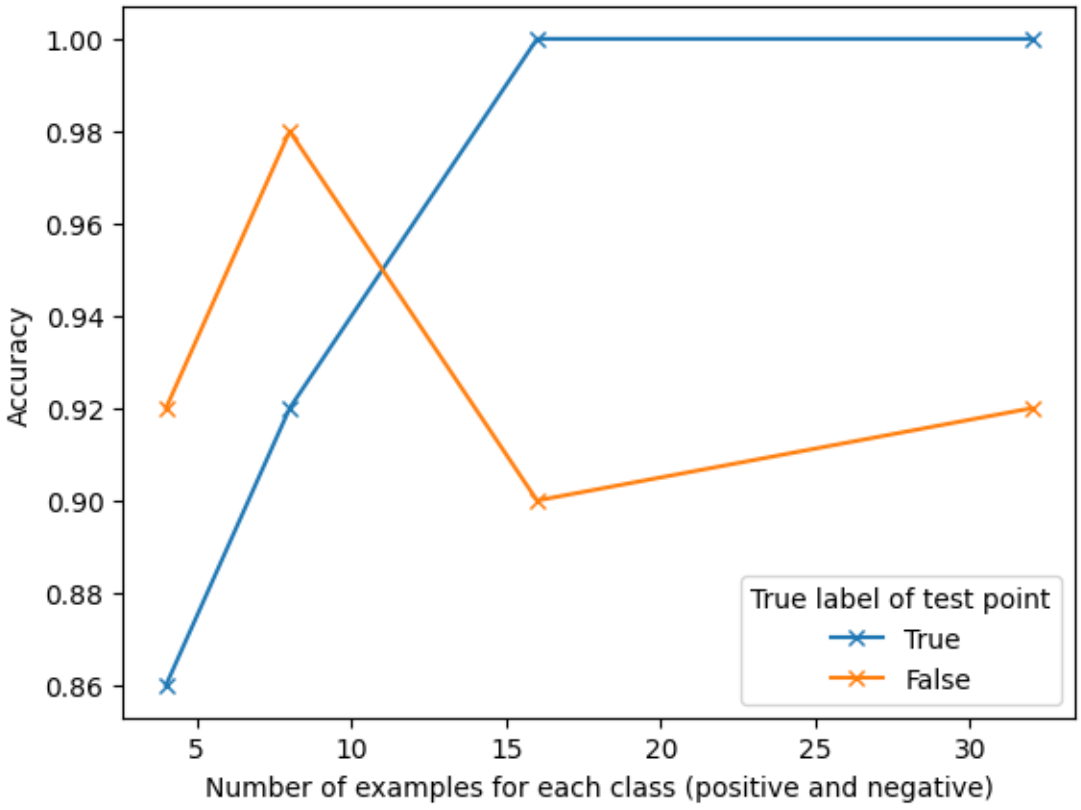
[4]:

	label	n_egs	correct
0	False	4	0.92
1	False	8	0.98
2	False	16	0.90
3	False	32	0.92
4	True	4	0.86
5	True	8	0.92
6	True	16	1.00
7	True	32	1.00

[12]:
```
fig, ax = plt.subplots()

df['accuracy'] = df['correct']
for label in [True, False]:
    subdf = df[df['label'] == label]
    ax.plot(subdf['n_egs'], subdf['accuracy'], label=label, marker='x')
ax.set_xlabel('Number of examples for each class (positive and negative)')
ax.set_ylabel('Accuracy')
ax.legend(title='True label of test point')
```

[12]:
```
<matplotlib.legend.Legend at 0x7f509ebab4f0>
```


1.00
0.98
0.96
0.94
0.92
0.90
0.88
0.86
Accuracy
5
10
15
20
25
30
Number of examples for each class (positive and negative)
True label of test point
True
False