

# CIS 419/519: Homework 1

Your name here

Although the solutions are my own, I consulted with the following people while working on this homework: {Names here}

1. (a) Show your work:

$$P(\text{play outside} = \text{yes}) =$$

$$P(\text{play outside} = \text{no}) =$$

...

$$IG_{\text{Snow}} =$$

$$IG_{\text{Sunny}} =$$

Snow/Sunny is better because...

- (b)

```
if color = blue:
    if size = small:
        if act = stretch:
            inflated = T
```

- (c) Yes/No because...

2. (a) See Figure 1 for the model performance in the Madelon dataset. (The figure was generated on a different dataset, so replace the `plot.png` file with your own.)

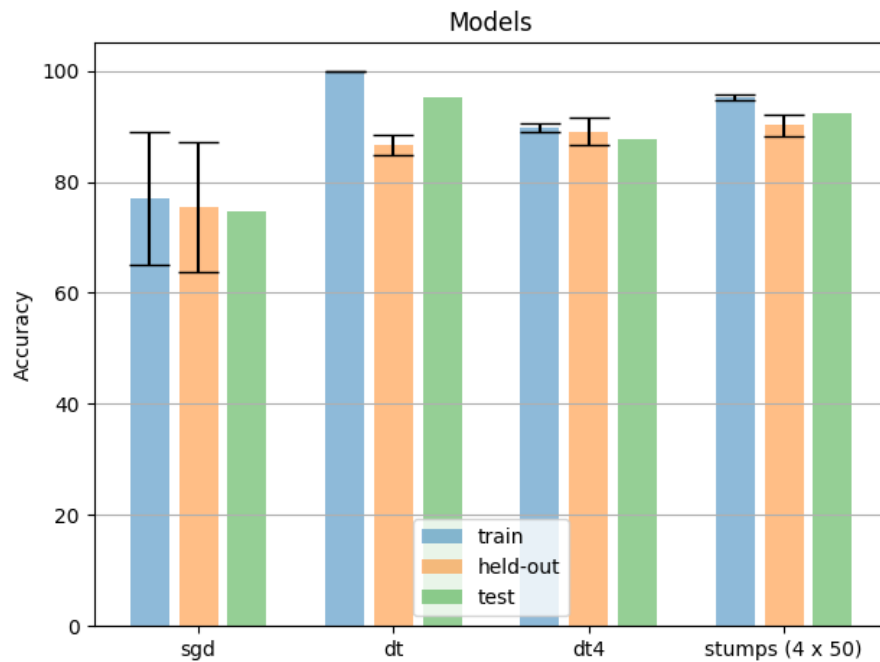


Figure 1: Model performance on the Madelon dataset

- 1.
- 2.
- 3.
- 4.

(b) The models' accuracies on the Badges dataset were as follows:

| Algorithm                     | Accuracy |
|-------------------------------|----------|
| SGD                           | a        |
| Decision Tree                 | b        |
| Decision Stump                | c        |
| SGD + Decision Stump Features | d        |

- 1.

### 3. Extra Credit

Please explain what features you used and the impact that they had on the model.