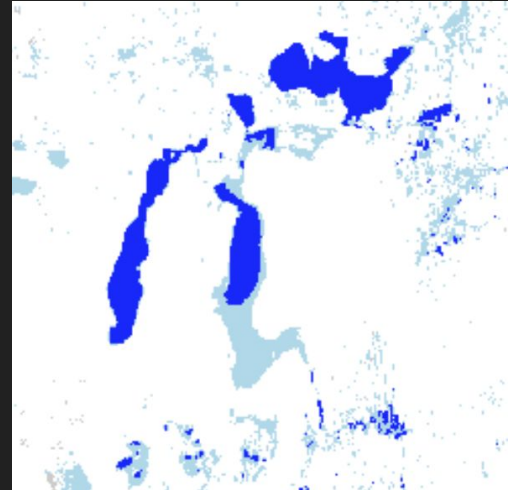# Machine Learning for Waterbody Identification

## Team Lakeside View

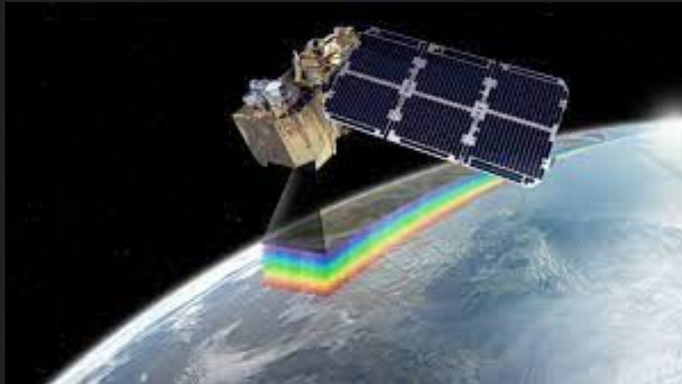James Kervin, Adam Subora, Benjamin Maves, Gabriel Humbert

# Question

What is the best method to classify water using machine learning?

# Data Available

1. 12-band Sentinel-2 Imagery:
   - [Blue, Green, Red, NIR 1 - 5, SWIR 1 - 4]
2. 4-band Michigan Public Imagery
   - [Blue, Green, Red, NIR]
3. 3-band Sentinel-2 Imagery:
   - [Blue, Green, Red]
   - Classified with Normalized Water Difference Index (NWDI)

# Accuracy and Possible Methods

## Accuracy

- Compare to number of correct pixels
    - "Correct" Labels found through NWDI classification system.
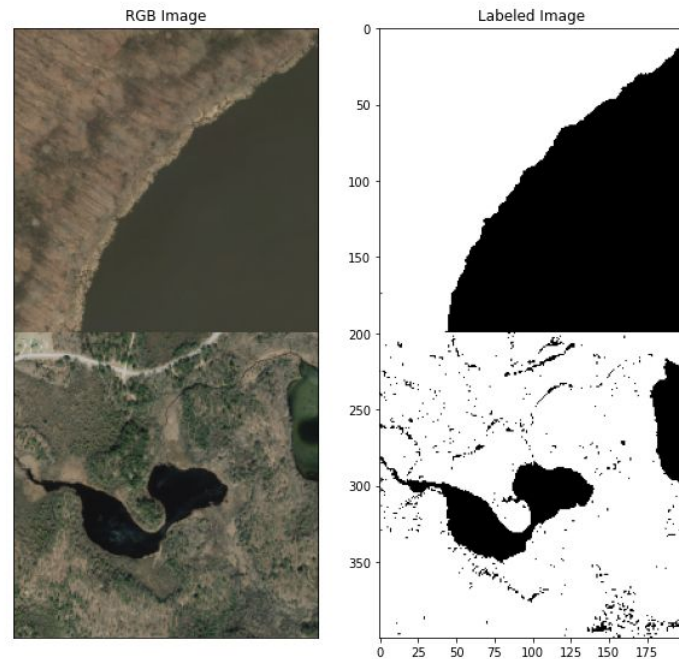- Down sample the code
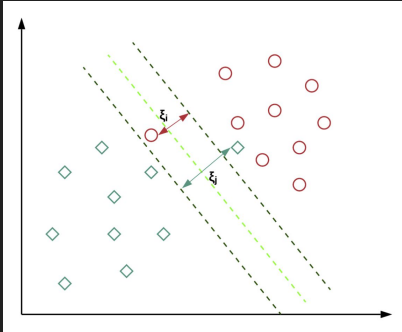


## Methods

- Perceptron
    - RGB Class
    - Near Infrared Class
- Support Vector Machines (SVM)
    - RGB Class
    - Near Infrared

# Perceptron Near Infrared

- Features:
    - Michigan Public Imagery
    - Red, Green, Blue, NIR
- Labels:
    - NWDI
    - (Green - NIR) / (Green + NIR) > 0.1
- Split:
    - 2 Training Images (80,000 px / band)
    - 4 Testing Images
- Parameters:
    - Iterations: 2000
    - Learning Rate: 0.001
    - Downsampling Rate: 0.5
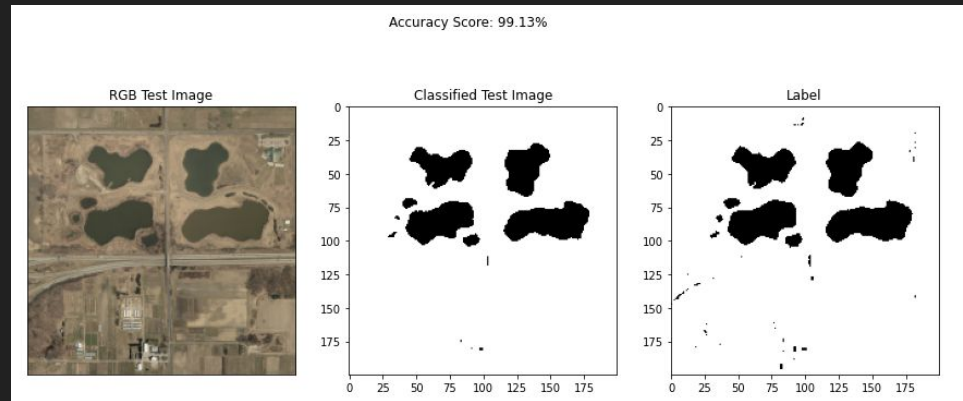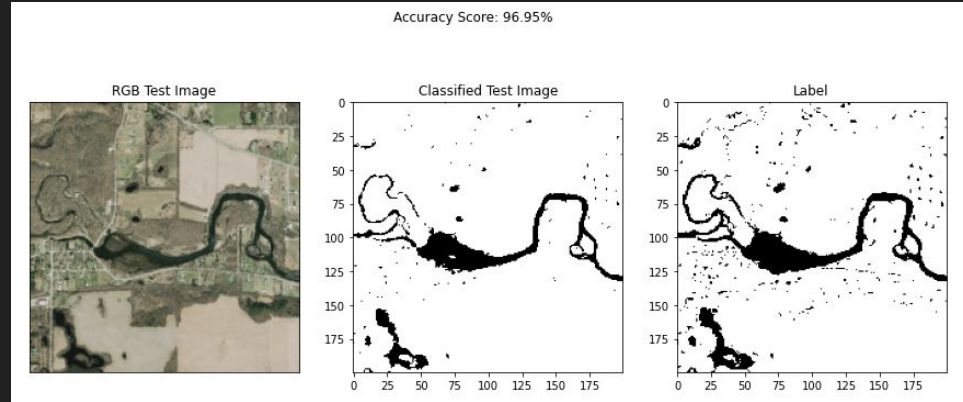
# Perceptron Near Infrared Results

- Very Accurate
  - Min. Acc.: 96.95%
  - Max. Acc.: 99.13%
  - Does not account for inaccuracy of NWDI labels
- Very Slow
  - Over 24 minutes with 2 training images
  - Would be over double without downsampling

```
Weights:
 [-0.00488776  0.03046392  0.00269948 -0.0390093 ]

Bias Weight: 0.002999999999999216

Time Elapsed: 1470.995376110077 Seconds
```
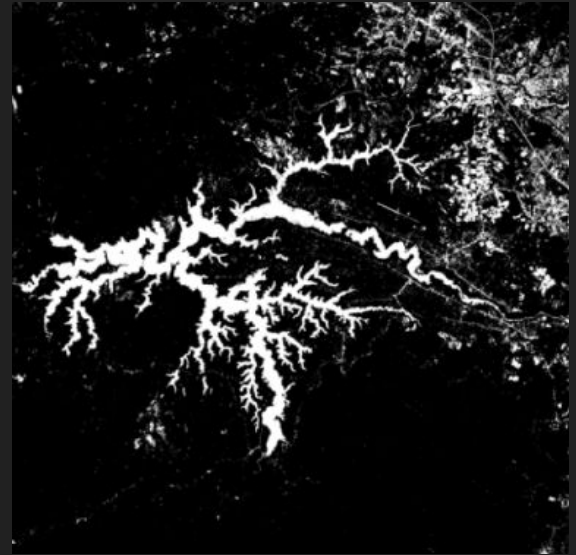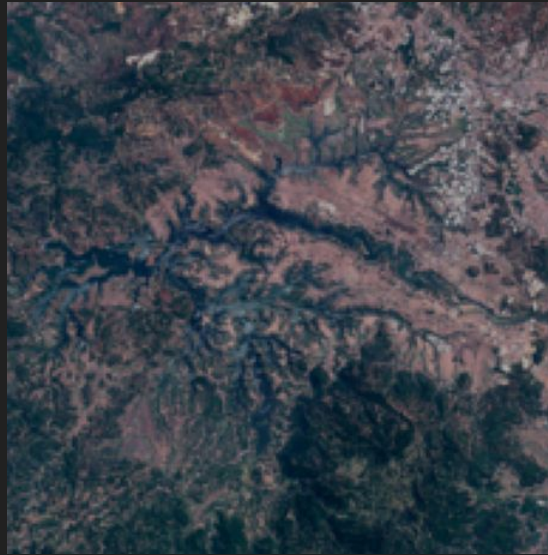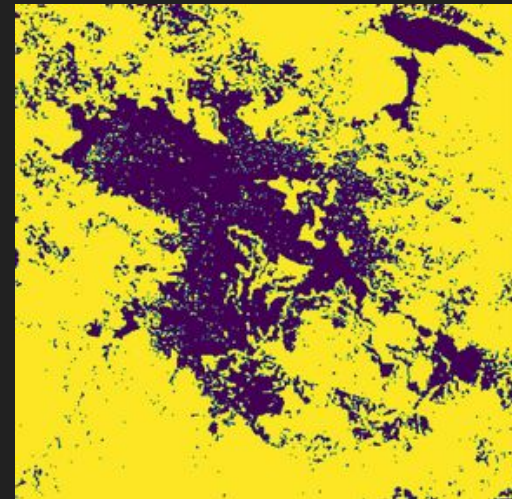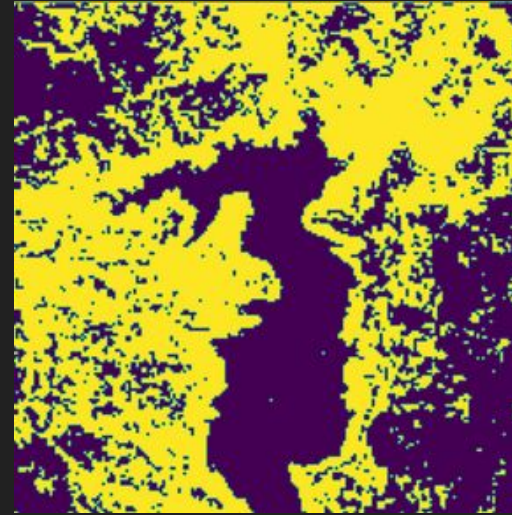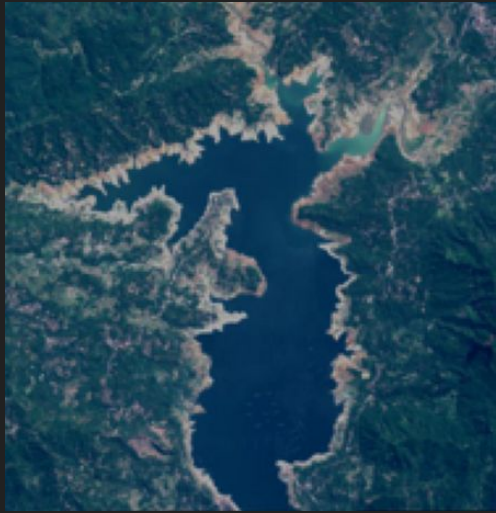
# Perceptron RGB

Features:

- Red
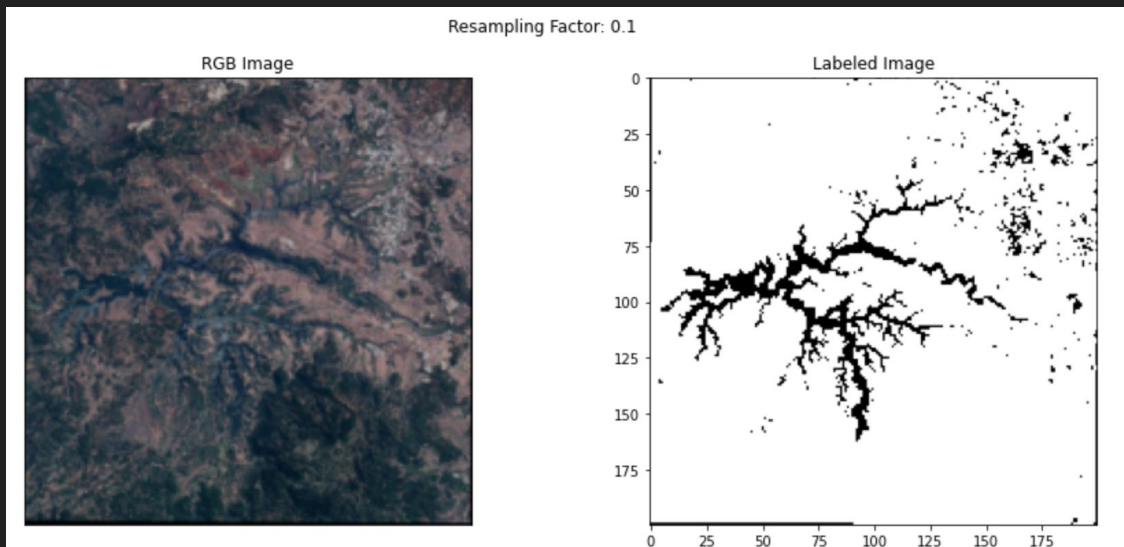- Green
- Blue
- White

# Perceptron RGB results

- Not very accurate.
- Takes a very long time to train
- Very quick analysis

# RGB SVM method

Features:

- Red
- Green
- Blue
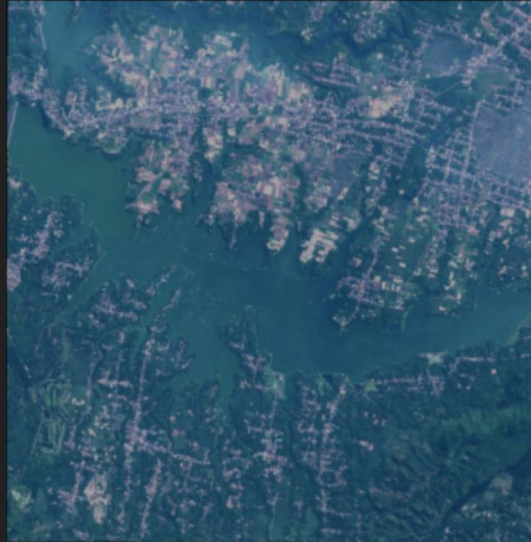- White



Best estimator found by grid search:
SVC(C=100000.0, class_weight='balanced', gamma=0.1)
Best parameters found by grid search:
{'C': 100000.0, 'gamma': 0.1, 'kernel': 'rbf'}

# RGB SVM results
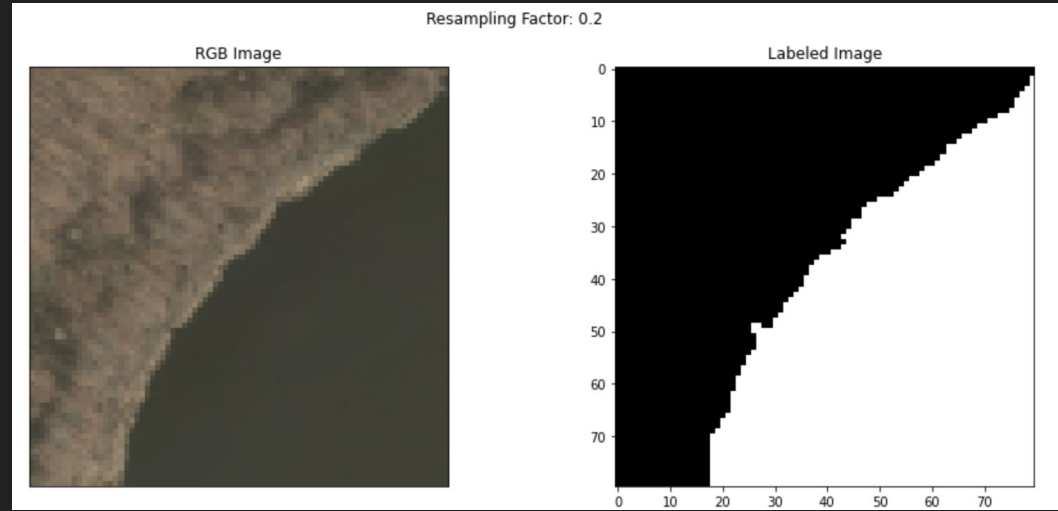
- Very inaccurate.
- Takes a very long time to train
- Quick analysis

# NIR SVM

Features:

    -Red

    -Green

    -Blue

    -White

    -Near Infrared



Best estimator found by grid search:

SVC(C=50000.0, class_weight='balanced', gamma=0.1)

Best parameters found by grid search:

{'C': 50000.0, 'gamma': 0.1, 'kernel': 'rbf'}

# NIR SVM results

- Fairly accurate (>.98)
- Relatively quick training
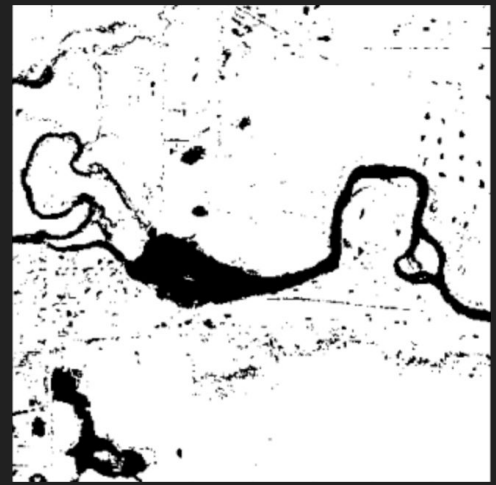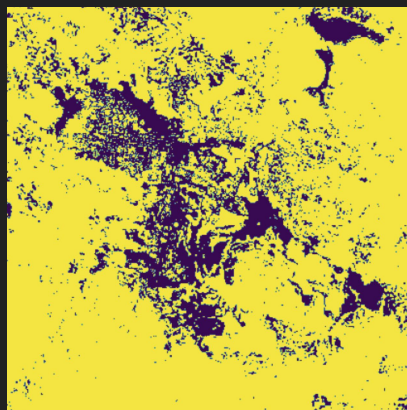- Very quick analysis

Image accuracy: 0.9808375



Image accuracy: 0.9873625

# Comparison



| Perceptron RGB | Perceptron NIR | SVM RGB | SVM NIR |

# Conclusion

- RGB data is easy to find, lots with labels, but hard to model around
- NIR data is very difficult to find and many of the labels we had to make ourselves
- Inclusion of NIR features enhances both the Perceptron and SVM predictive capabilities.
- SVM and Perceptron have similar predictive capabilities when using same features
- NIR SVM ran much faster than NIR Perceptron

# References

https://www.kaggle.com/datasets/franciscoescobar/satellite-images-of-water-bodies