

Machine Learning CS 453X

Homework 2

Mean Square Error

Formula used:

$$f_{\text{MSE}}(\mathbf{y}, \hat{\mathbf{y}}; \mathbf{w}) = \frac{1}{2n} \sum_{i=1}^n \left(g(\mathbf{x}^{(i)}; \mathbf{w}) - y^{(i)} \right)^2$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\text{Predicted}_i - \text{Actual}_i)^2}{N}}$$

Method 1:

Training MSE: 39.24296298929068

Testing MSE: 206.79647485406986

Method 2: (T = 5000 and learning_rate = 0.001)

Training MSE: 87.89921317334458

Testing MSE: 97.00998215844058

Method 3: (T = 5000, learning_rate = 0.001, alpha = 0.1)

Training MSE: 88.05593100785939

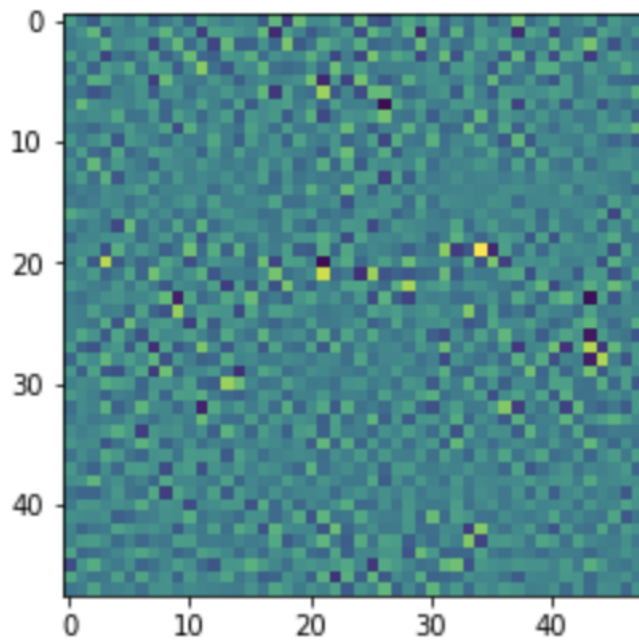
Testing MSE: 96.73951901294978

Testing RMSE: 13.925709329510624

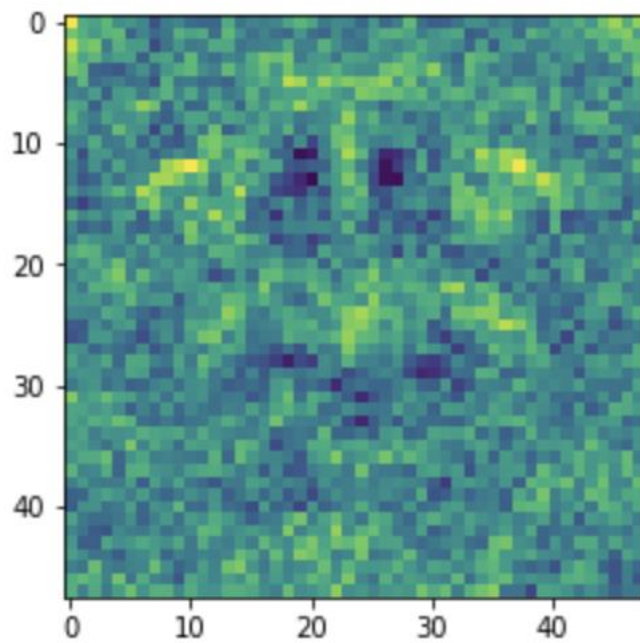
Note: On method 2 and 3, when I set learning_rate = 0.003, my gradient descent “overjumped” and the weights became infinity. Hence, I changed my weights to 0.001. I already asked the professor and he said the results are reasonable.

Visualizing Weights

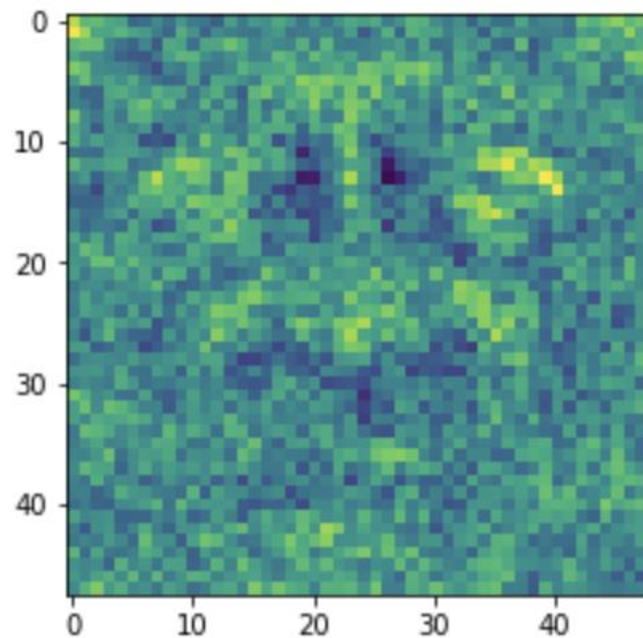
Method 1:



Method 2:



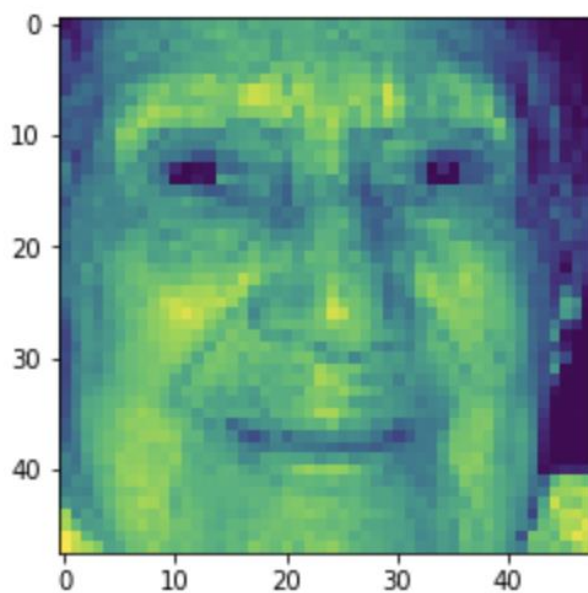
Method 3:



The images of weights from method 3 and method 2 are somewhat similar. However, both of them are visually different from the image of weights from method 1.

Top 5 Egregious Error

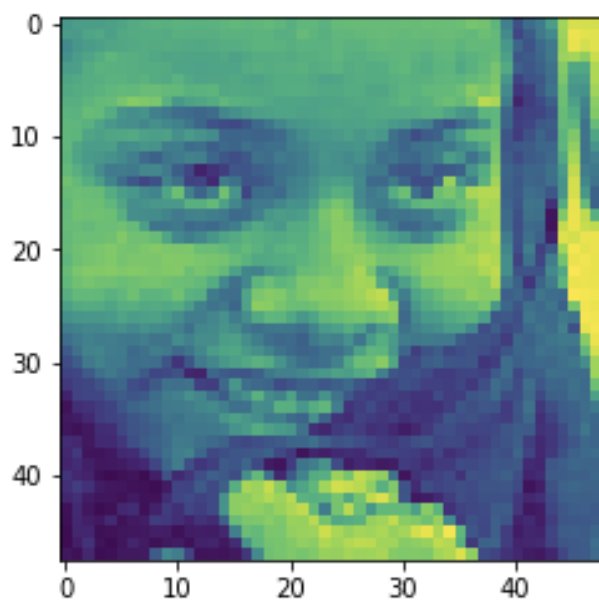
Image at index: 581



Predicted: 32.07881314858461

Actual: 80.0

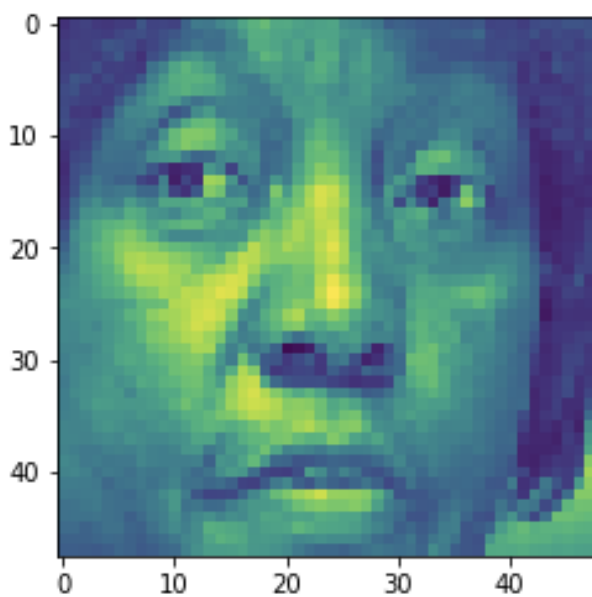
Image at index: 884



Predicted: 59.85149905268884

Actual: 10.0

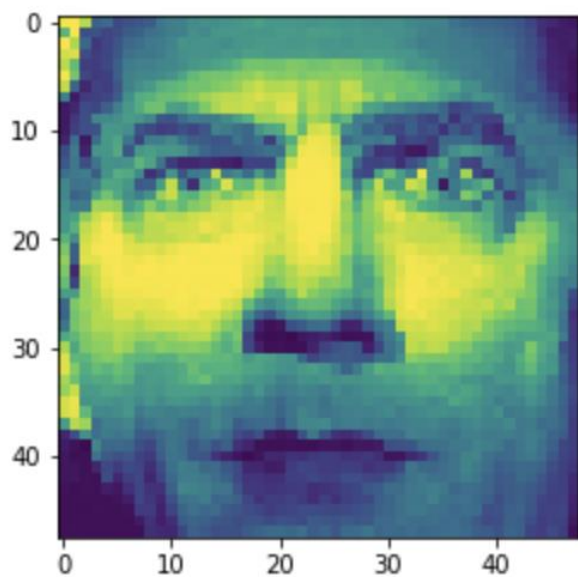
Image at index: 830



Predicted: 40.16579593130702

Actual: 89.0

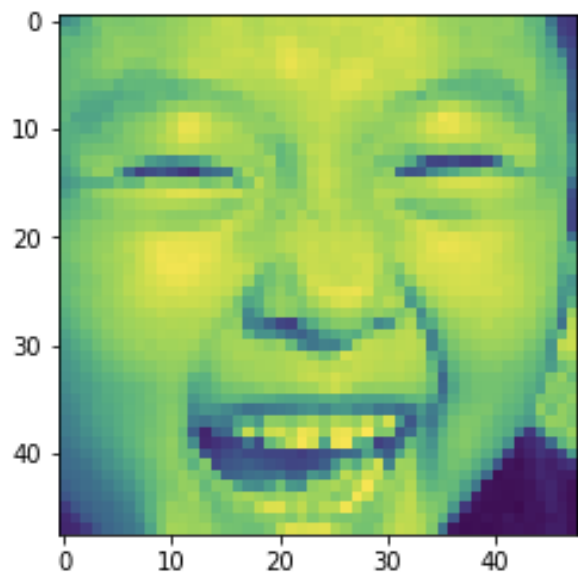
Image at index: 1120



Predicted: 52.9035657047763

Actual: 99

Image at index: 1640



Predicted: 52.73122484883238

Actual: 4.0