

Machine Learning Attractiveness

Author: Christopher MacLellan
Advisor: Jeff Van Baalen

Questions to be Answered

- How do we determine if a given face is attractive?
- Can a machine learning algorithm predict how attractive we find a particular face?
- What are the application of such an algorithm if it was successful?

Applications

- Online dating websites
 - Facebook
 - Match.com
- Online marketing and advertising
 - Personalized Ads (where you think the sales person is amazingly attractive)

Facial Attractiveness

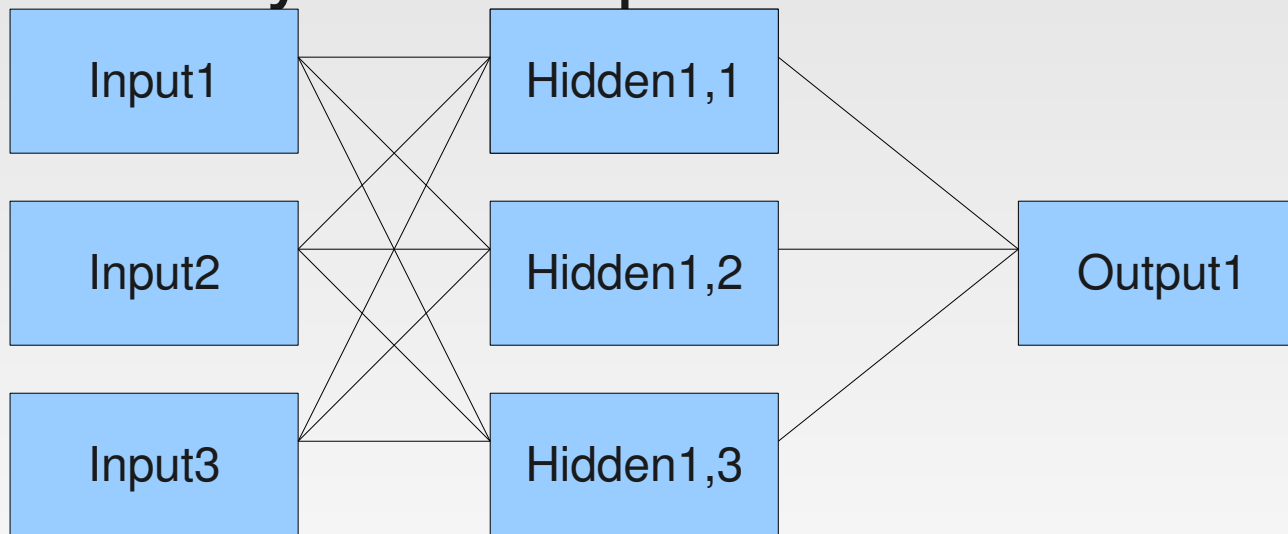
- Image quality (when rating images on the web)
- Facial symmetry
- Feature preferences
 - Specific type of nose, mouth, eyes, hair, etc...

How to Learn Attractiveness

- Two approaches were used.
 - Neural Networks
 - Support Vector Machines

Neural Network

- Statistical approach which simulates neural activity using link weights and nodes
- One layer example

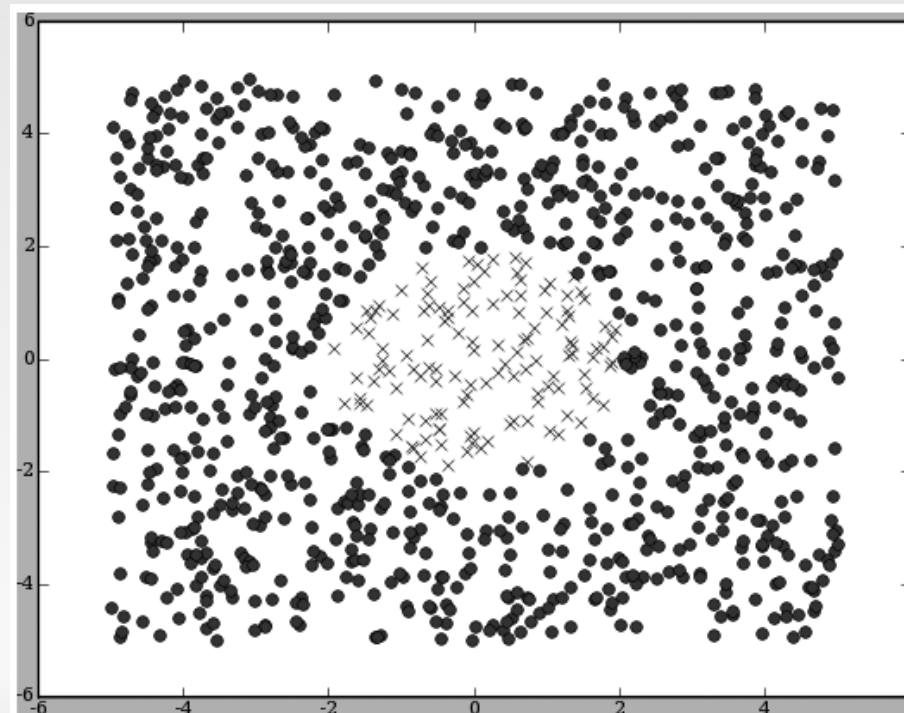


Neural Network (cont.)

- Forward propagation to make a prediction.
 - This consists of summing all of the links into each node and entering that into a sigmoid function to determine how much the node should be activated.
- Back propagation is used to train the network.
 - This consists of adjusting the links into each node based on the amount of error that each node had with the actual value.

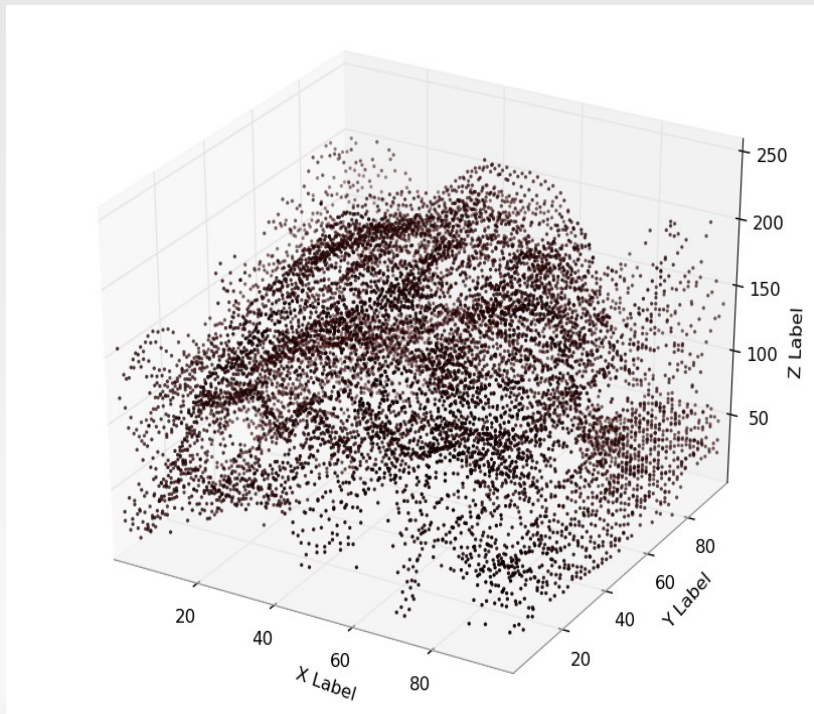
Support Vector Machine

- A classification technique which analytically calculates boundaries between data points and uses these boundaries to make future predictions.

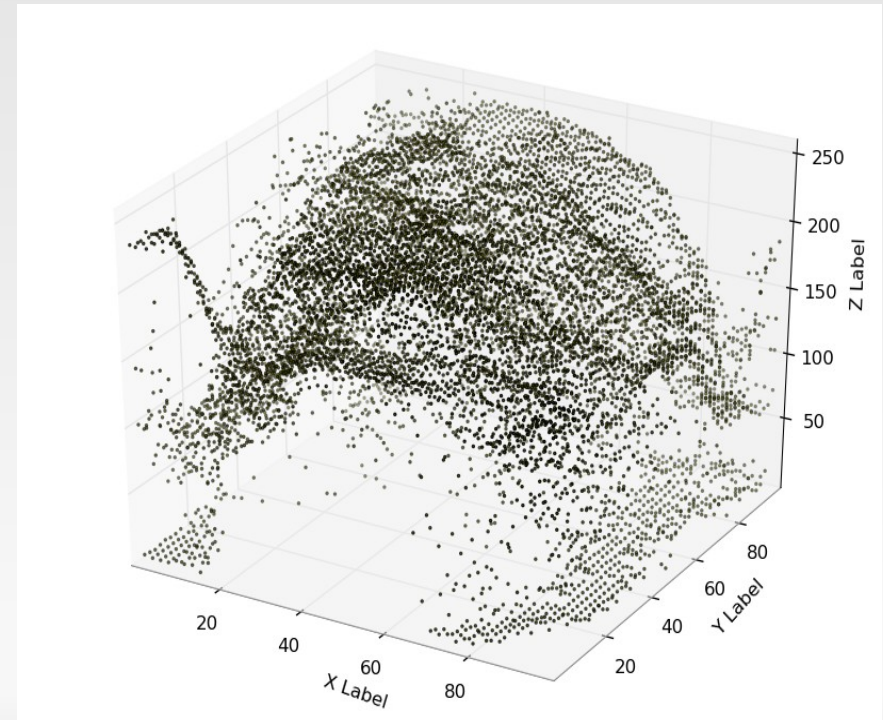


Applying these techniques

- Used the pixel (rgb) intensities of the images as input into the neural network / support vector machine.



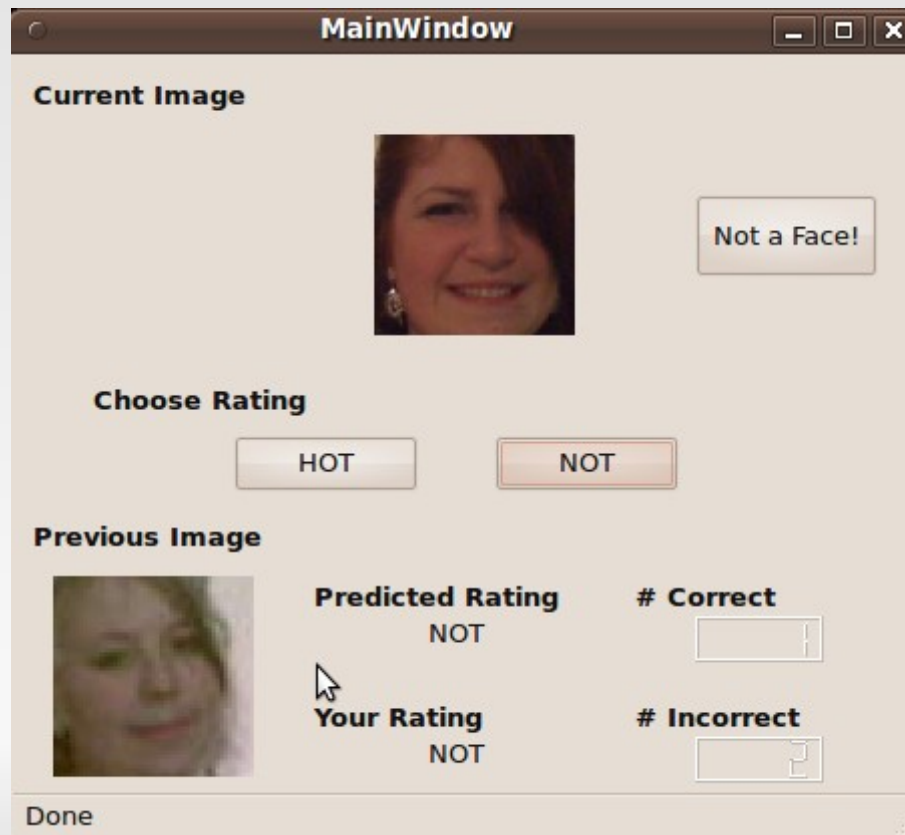
Four HOT



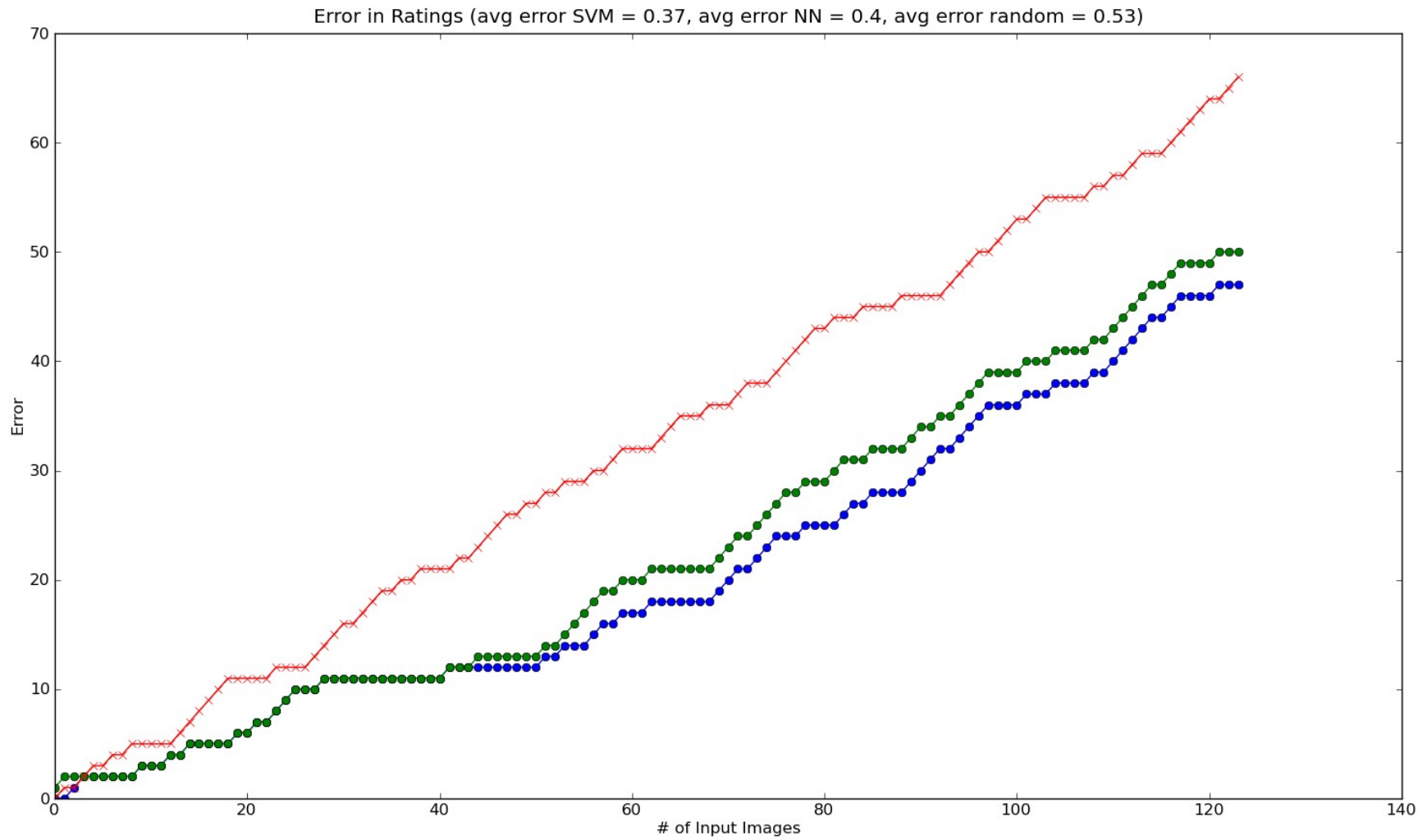
Four NOT

Collecting Faces

- www.hotornot.com
- Rating Images HOT or NOT

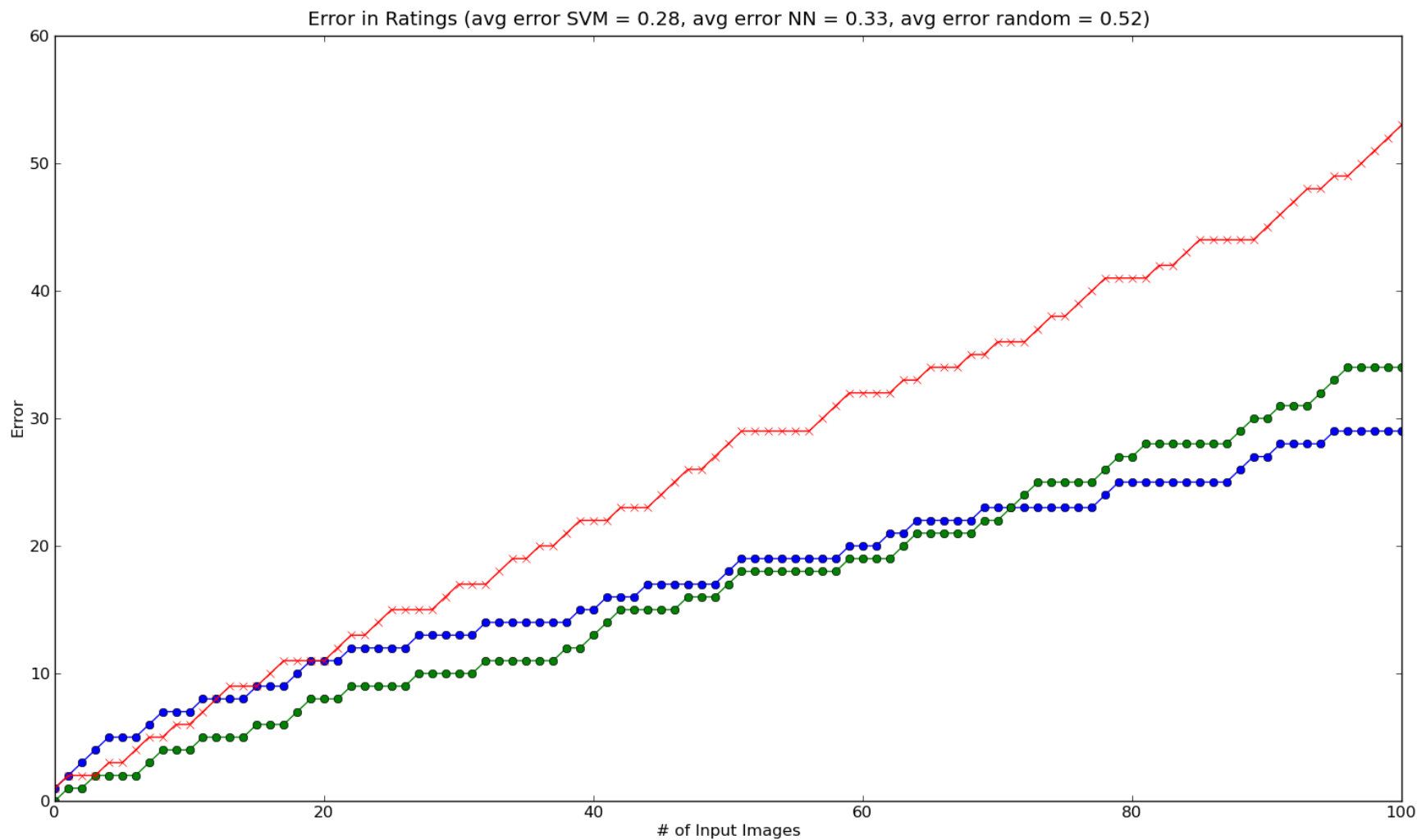


Results



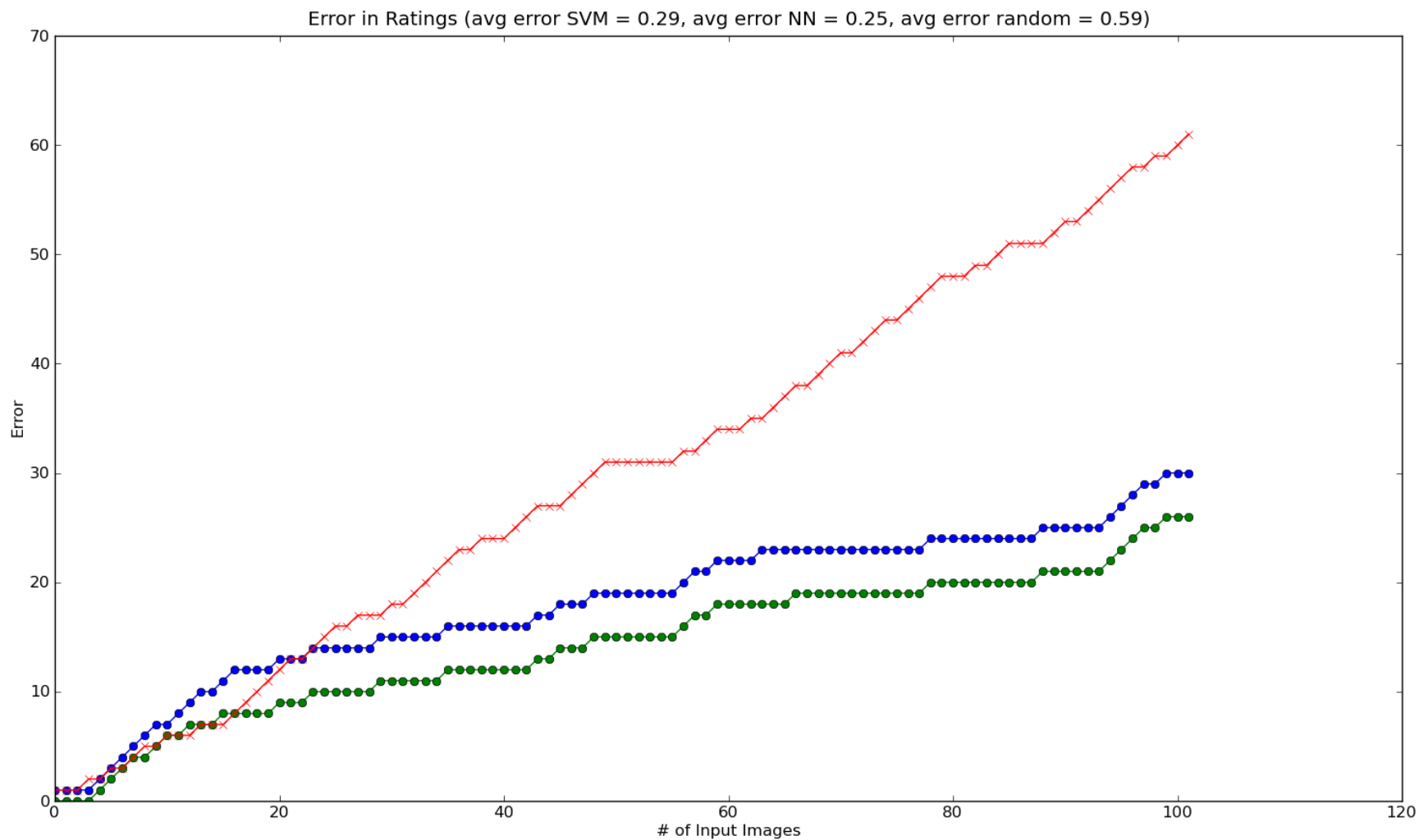
Chris

Results (cont.)



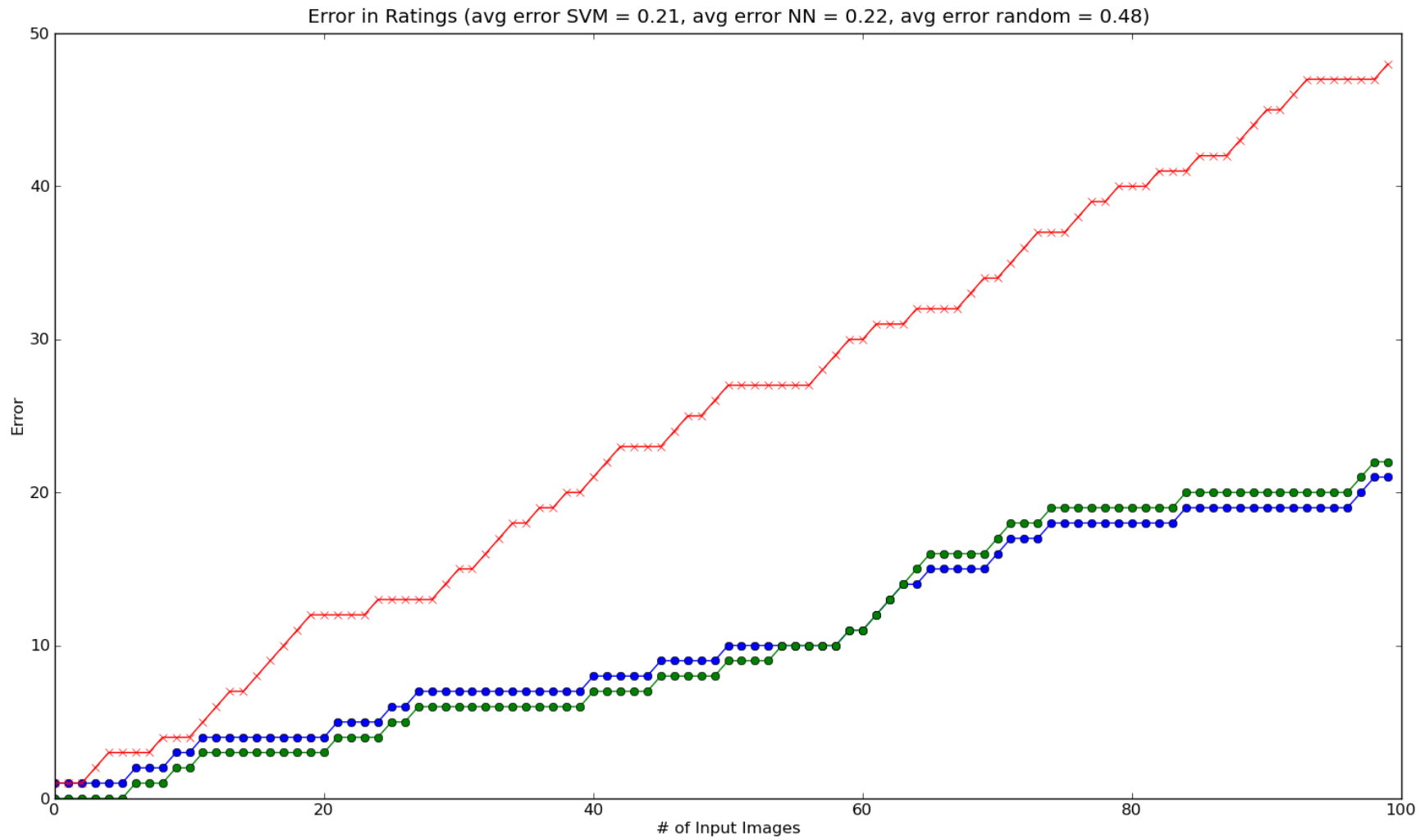
Joe

Results (cont.)



Nema

Results (cont.)



Pat

Result Summary

Subject	Average Error		
	Neural Net	SVM	Random
Me	0.4	0.37	0.53
Joe	0.33	0.28	0.52
Nema	0.25	0.29	0.59
Pat	0.22	0.21	0.48
Average	0.3	0.29	0.53

What Does All It Mean?

- It means that the algorithms consistently do ~20% better than guessing.
- It means there is some type of correlation between pixel data and attractiveness ratings that the machine is able to learn.
- It hints at future research to improve results by preprocessing the images to reduce noise and possibly extract features (facial symmetry information, hair / eye / skin color, type of nose or mouth structure, etc...)

Additional Thoughts

- I believe that there is a significant amount of noise in these images that can be reduced.
- I am curious to see what would happen with 1000+ ratings as opposed to the 100 I asked my subjects for.

Thank You

- Dr. Jeff Van Baalen