Machine Learning and I Spy Accuracy

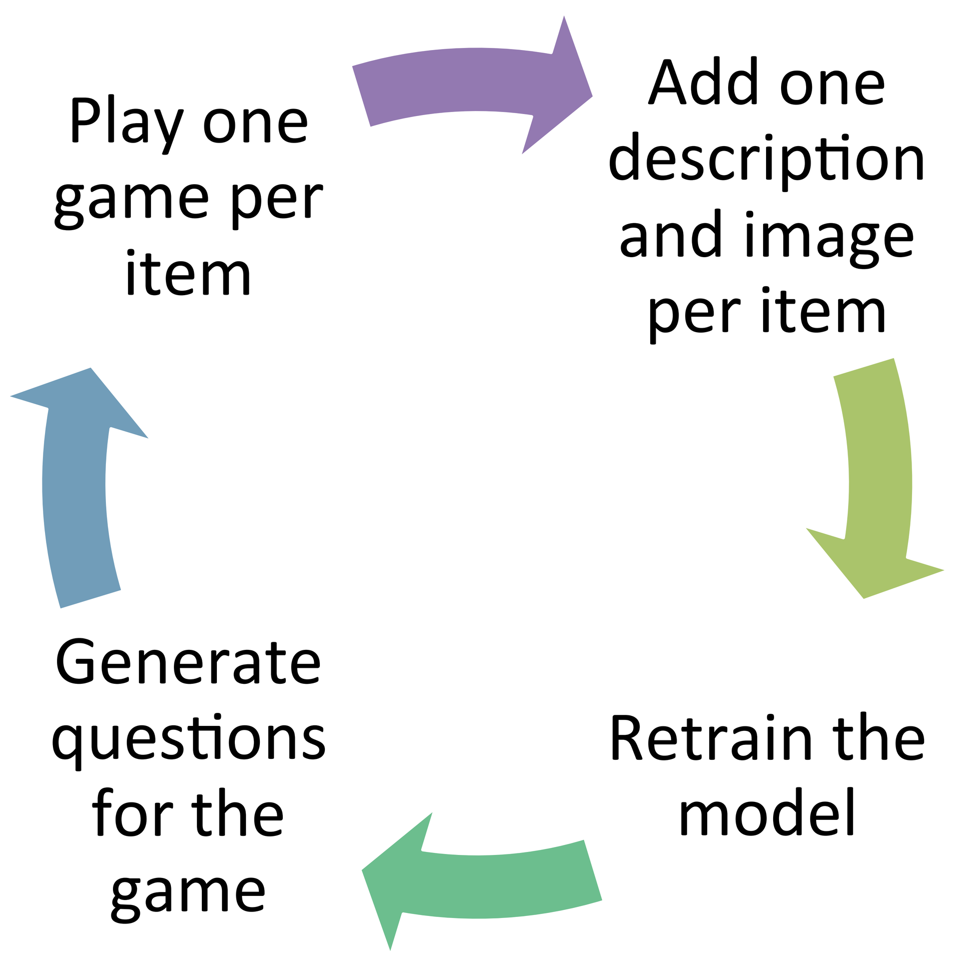
Background:

Changes in social norms, limitations of healthcare systems, and an increasing elderly population have resulted in situations where human companions are not always available. Companion robots (Companionbots) offer one means of complementing available care. The HILT Lab is constructing Companionbots utilizing advanced machine learning to interact with clients and communicate patient needs to medical professionals. The lab is creating an implementation of the game I Spy to improve machine learning techniques and integrate the learning naturally into the participant's environment. The system learns to recognize common objects based on a few photos and participant descriptions. Based on this initial training, the system automatically generates descriptive questions to determine the object selected by the user. Once the system is confident in a specific answer, it will guess the user's object.

Training the system takes a significant amount of time and collecting valid descriptions can be costly in terms of time and money. So, there is value in finding the minimum amount of training required to generate a certain level of accuracy.

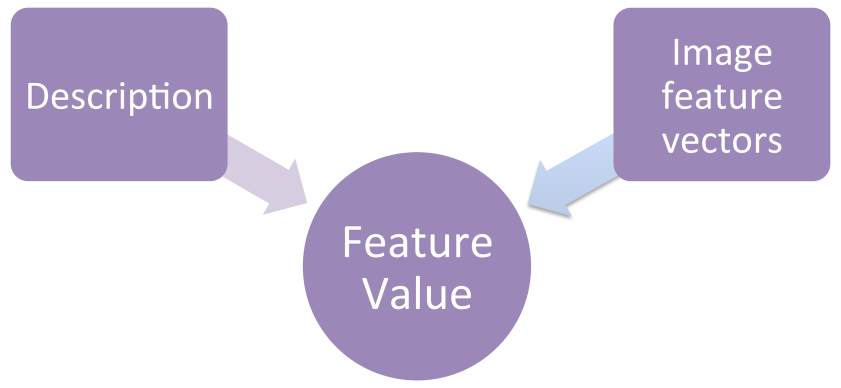
Example of Human-Generated Description Used:

This is a coffee mug. The mug is black and has a handle on the side. It has the words "No, I will not fix your computer" printed on the side.

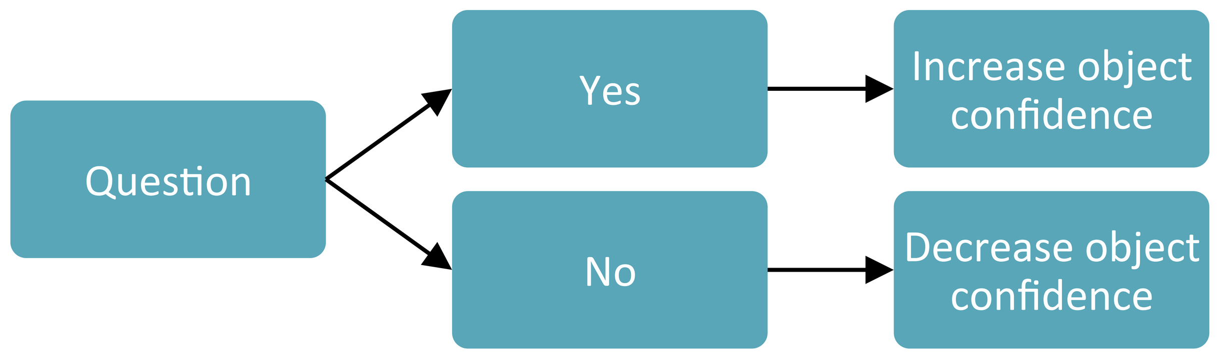


Methods:

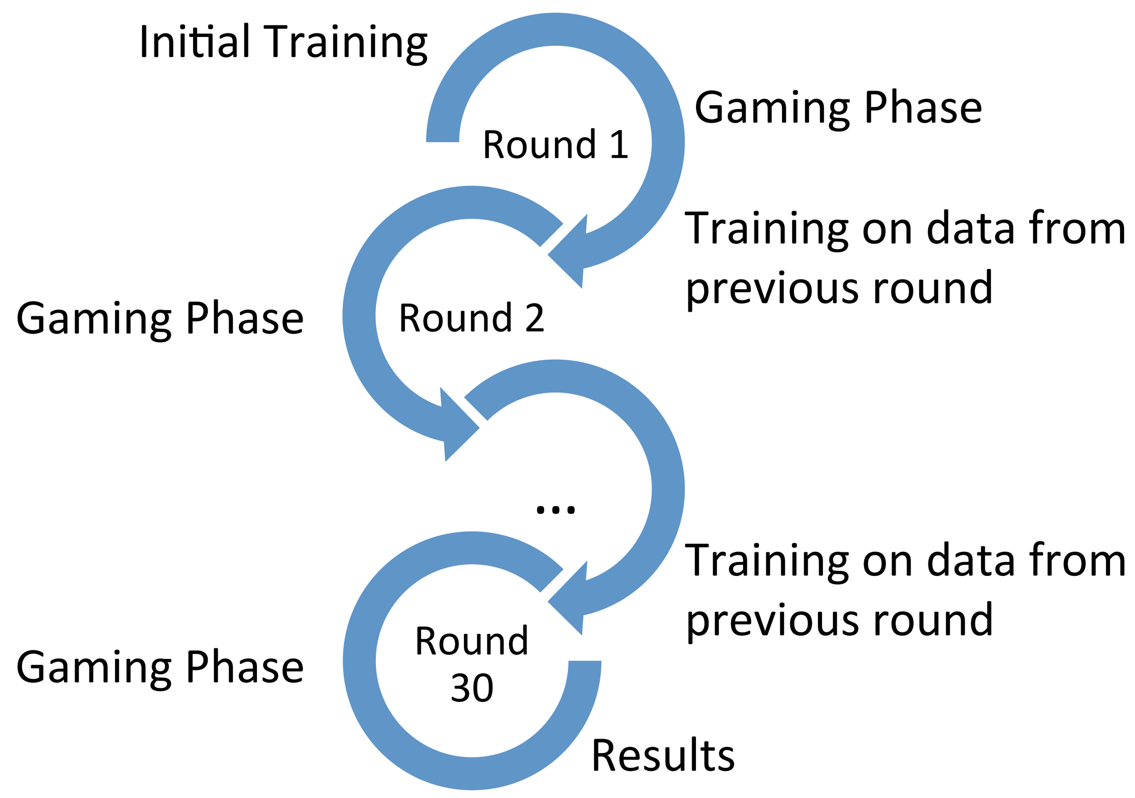
- 30 rounds of 17 games
- Purpose of each game is to correctly guess one object
- Each tag/object ID pair receives a numeric value based on how important the system thinks the tag is to describing the object



- Descriptions for rounds 1-6 are complete sentences
- Descriptions for rounds 7-30 are manufactured
- Gaming system asks the 'player' questions until it is confident enough to guess the 'player's' object



- If the system guesses correctly, it wins that round
- The learning system then uses the images and descriptions from the previous round to retrain the model

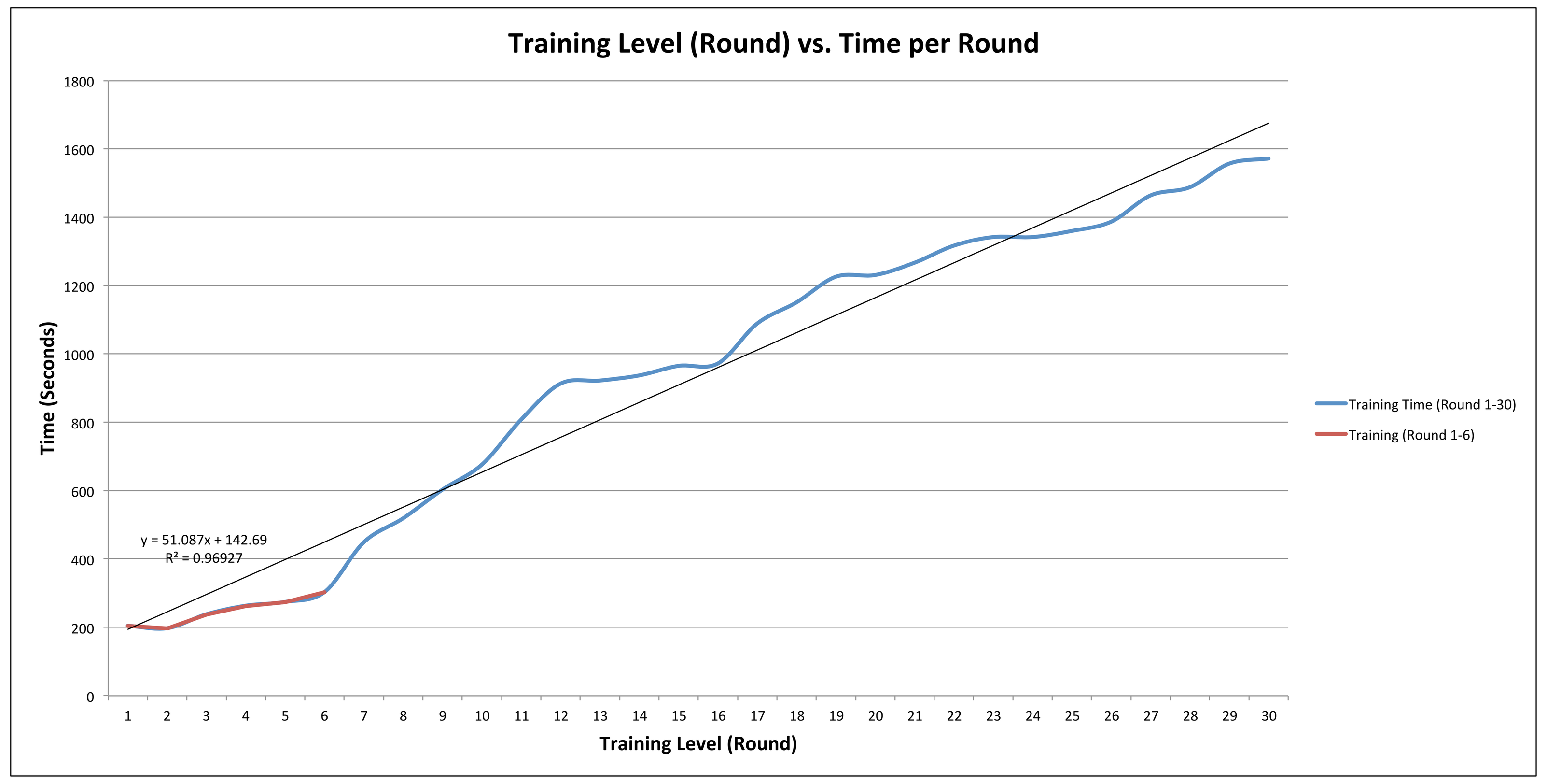


Objects Used:

1. Digital Clock
2. Analog Clock
3. Red Soccer Ball
4. Basketball
5. Football
6. Yellow Book
7. Yellow Flashlight
8. Blue Soccer Ball
9. Apple
10. Black Mug
11. Blue Book
12. Blue Flashlight
13. Cardboard Box
14. Pepper
15. Green Mug
16. Polka-Dot Box
17. Scissors

Results:

The entire 30-round process took almost 8 hours to complete. The accuracy fluctuates with each level of training and does not follow any practical trend. This could be due to the quality of descriptions used, the way the system was trained, or the way the game selects which questions to ask. The maximum accuracy reached was 59% and the minimum was 24%. The time used in each level of training increases at each training level and follows an approximately linear curve. The first training level took about 3.4 minutes to finish where the last took 26.2 minutes to complete.



The training time increases with each training level due to the increased number of tags in the database each round. In rounds 1-6, between 10 and 15 tags are added per object. In rounds 7-30, anywhere from 50 to around 200 tags are added per object.

Conclusions:

The gaming system could potentially make more accurate guesses if better descriptions were used. The tag-built descriptions are unfiltered, and even though each tag might be a correct description of the object, it may not be an appropriate training parameter. More concise object descriptions could also help reduce training time, because no time is wasted calculating the importance of extraneous tags. The system could also benefit from finding a way to utilize the 'no' answers in the tag-built descriptions; it could be trained on positive and negative values. Future versions of the gaming system will also have better decision logic in order to select the best tag query. This should lower the number of questions asked and improve accuracy. Eventually, Companionbots will be able to recognize and react to key objects in their environments, and long-term patient care will be transformed.

Dr. Rodney Nielsen, Associate Professor
Adam Hair, Undergraduate CS Student

Robot image comes from the Nao website: http://www.aldebaran.com/en/humanoid-robot/nao-robot