**SRN6**

*In machine learning systems that we’ve discussed, we often use both positive and negative examples, sometimes large numbers of examples. This use of both positive and negative examples is in striking contrast to the case of language and number learning in children, which involves few or no negative examples. To attempt to model how children learn with only positive examples, we have to introduce constraints such as the Tolerance Principle in order to allow for learning and retreat from overgeneralization. Since using both positive and negative examples, and using lots of examples, turns out to be so powerful, what then is the benefit to AI research of examining human learning? Discuss both the potential positive and negative benefits to AI of studying human learning.*

Thoughout this semester, we have seen many examples of trying to understang how human learning occurs: Near-Miss, OS Learning, Prof. Boyden … Every one of these talks introduced some new AI concepts, directly inherited from the way humans learn and behave when exposed to new concepts. The new concepts we have seen allow to overthrow the needs of huge sets of examples in order for our models to learn. However, are there **other** benefits to understanding how humans learn ?

**Positive Benefits**:

* Understand Neural Networks. We have seen that, although the mathematical definition of the Neural Network is clear (a universal approximator), it is still unclear **why** do neural networks work that well ! Therefore, studying human learning might help us to first *understand*and then maybe improve the performances of Neural Network (*last part of Prof. Boyden speech)*
* Build true AI. During Prof. Boyden speech, we also tried to define what would *Artificial intelligence* mean. How could we build AI if we do not even know what is the I inside it.
* Define new concepts of learning: one major drawback of current AI models (overly-deep NN..) is that they require huge datasets in order to discover patterns in the data. However, studying human learning (like the way babies learn, or the Near-miss 🡪 OS learning concepts) helped us to defined new ways of learning (differing from the Bayesian Inference, or majority rules).
* Interpret the results: Maybe a better understanding of the way humans learn might be a gateway towards intepretable Neural Networks

**Negative Benefits:**

* This research is not oriented directly towards improving the capabilities of computational models, and therefore we may end up losing some time
* Until now, we know that computers are not equal to humans. What we have seen in OS learning might pose some ethical questions: what if computers become able to learn from a very small amount of data, better than humans do ?