Deep Learning the Relationship between Visual and Textual Data

Patrick Poirson
University of South Alabama

Last Week

- Fine tuning parameters
  - i.e. Learning rate, number of layers, and number of nodes
  - Went from 348% MSE to 2%
- Can now go from vectors to images
- Setting up Aladdin computers for experiments
- Going through previous authors code
  - Ran into some trouble because I don't have experience with numpy or scipy

Error in Last Weeks Results

- Was using the author of the toolboxes method for calculating results
  - He jus compares the max of the two vectors
- We want to use Mean Average Precision (MAP)
- We were also using a bad learning rate

$$\mathrm{MAP} = \frac{\sum_{q=1}^{Q} \mathrm{AveP(q)}}{Q}$$

Next Steps

- Figure out how author computes MAP
- Write code for multimodal network
- Start using Aladdin computers to train with larger data sets