* create capability to embed text (captions and queries) – Justin, Sydney
* create training and validation sets of triples – Justin, Sydney
* create function to compute loss, accuracy (in terms of triples correct) – Justin, Sydney
* create MyNN model for embedding images – Aparna, Nicole, Akash
* train model – Aparna, Nicole, Akash
  + embed caption
  + embed good image
  + embed bad image
  + compute similarities from caption to good and caption to bad
  + compute loss with margin ranking loss
  + take optimization step
* create image "database" by mapping whole set of image features to semantic features with trained model – Akash
* create function to query database and return top k images
* create function to display set of images (given image ids)
  + note that the image metadata (contained in captions\_train2014.json) includes a property called "coco\_url" that can be used download a particular image on demand for display
  + maybe display their captions, too (for debugging)
* create function that finds top k similar images to a query image
  + maybe give option for doing similarity search in either semantic space or original image feature space