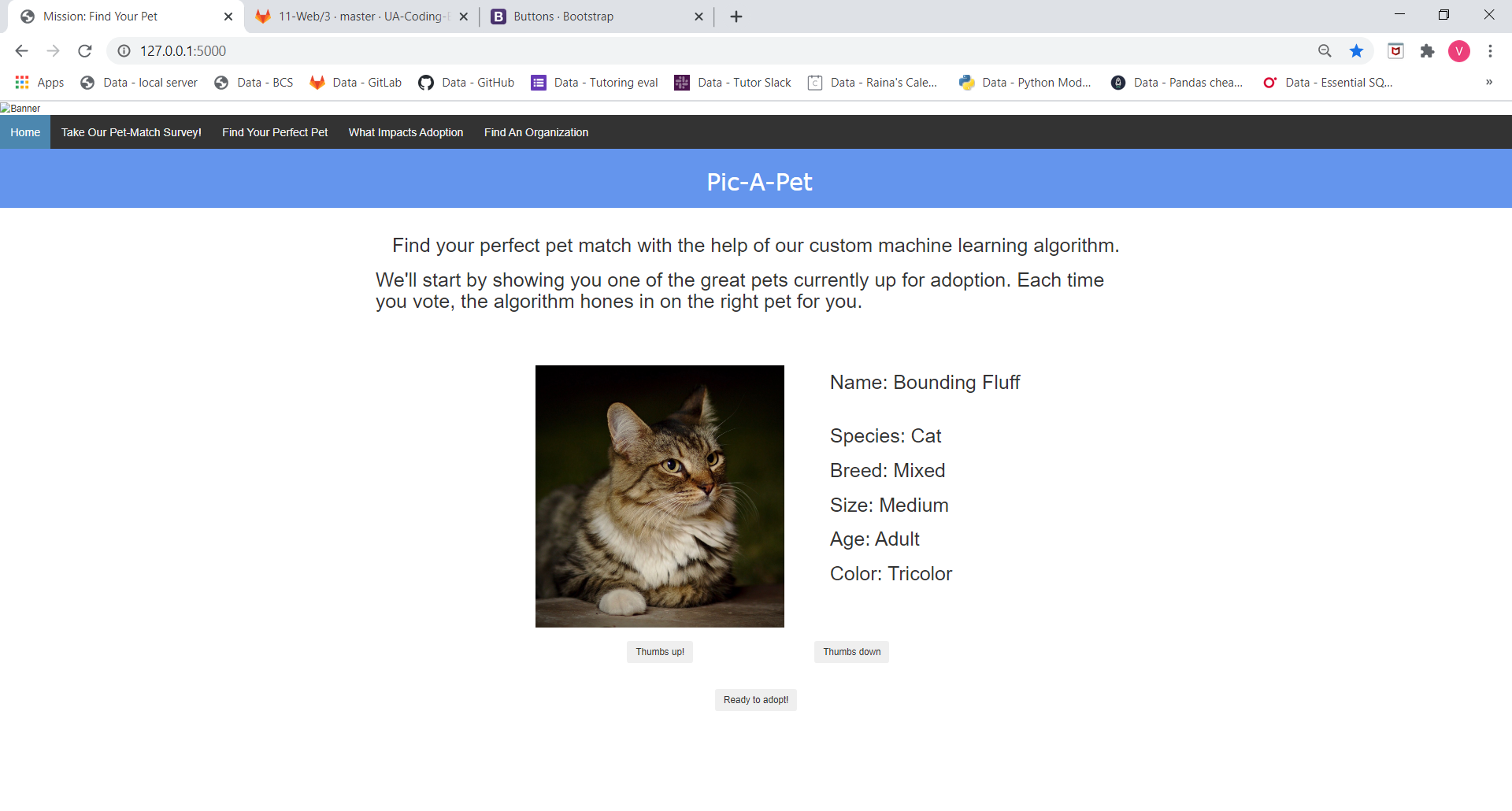
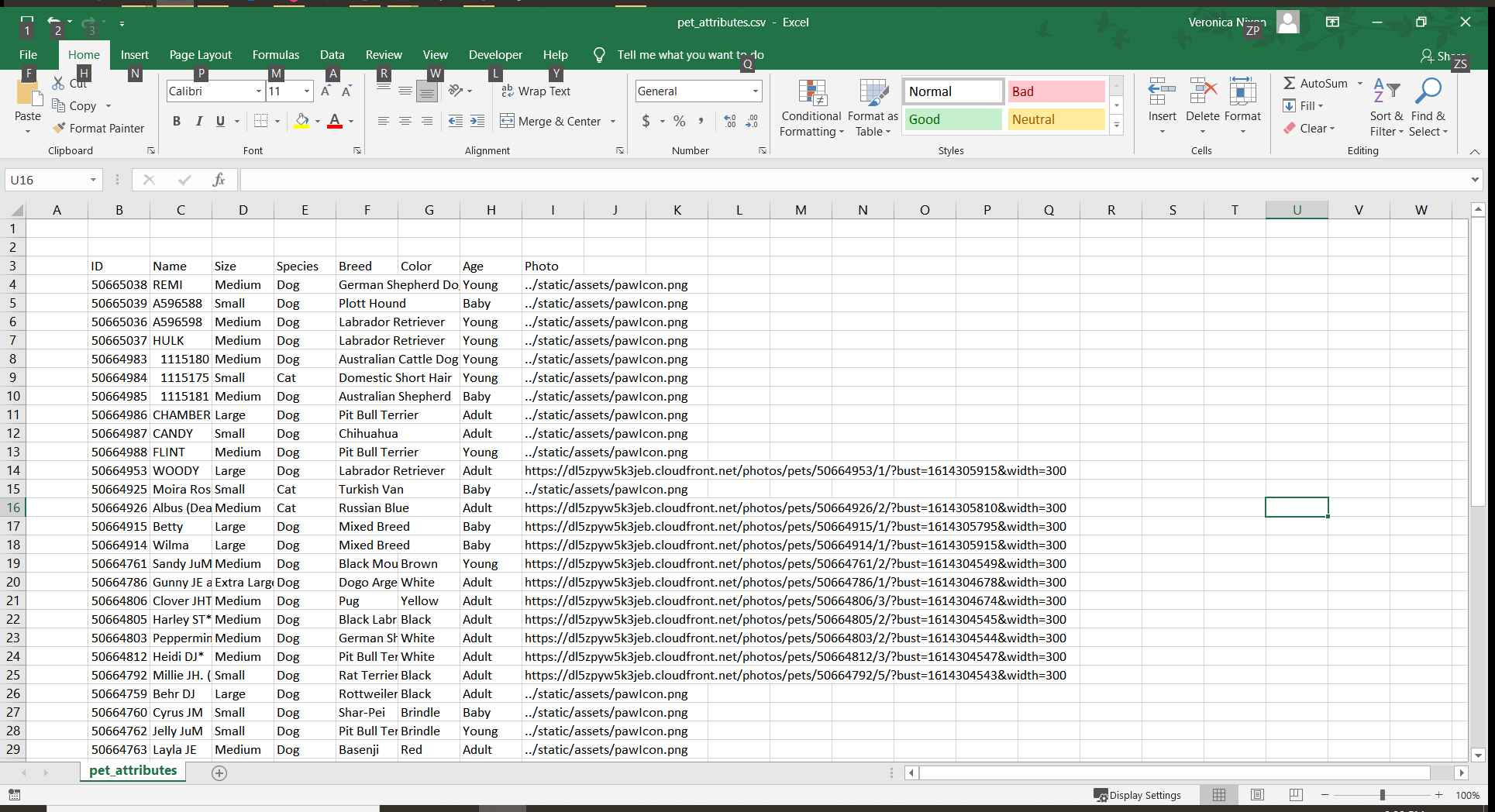
‘Pic-A-Pet’ Recommender System Model



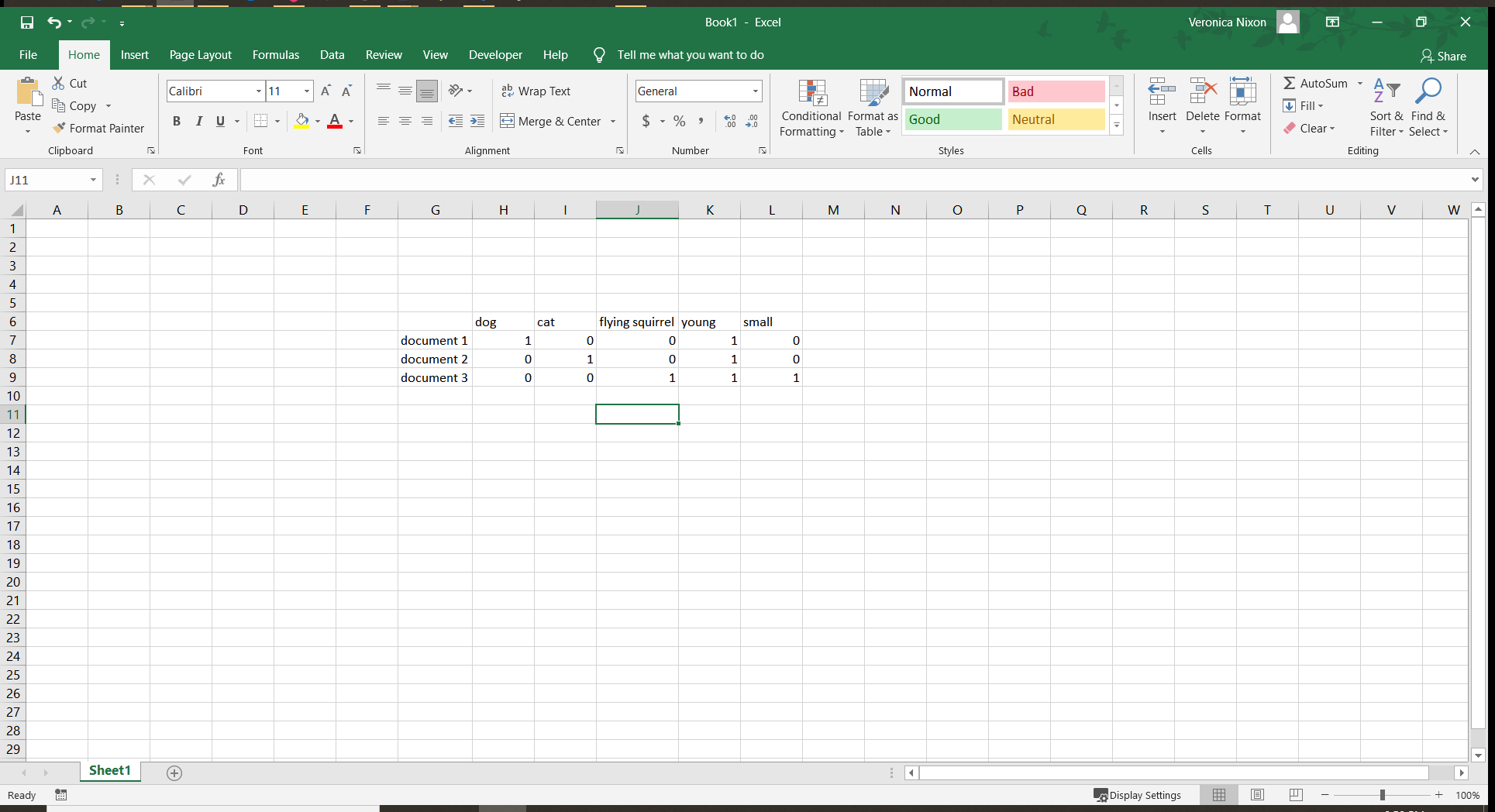
Preprocessing

Pandas is used to clean data downloaded from PetFinder and convert to ‘metadata soup’, e.g., “small cat mixedbreed young”.

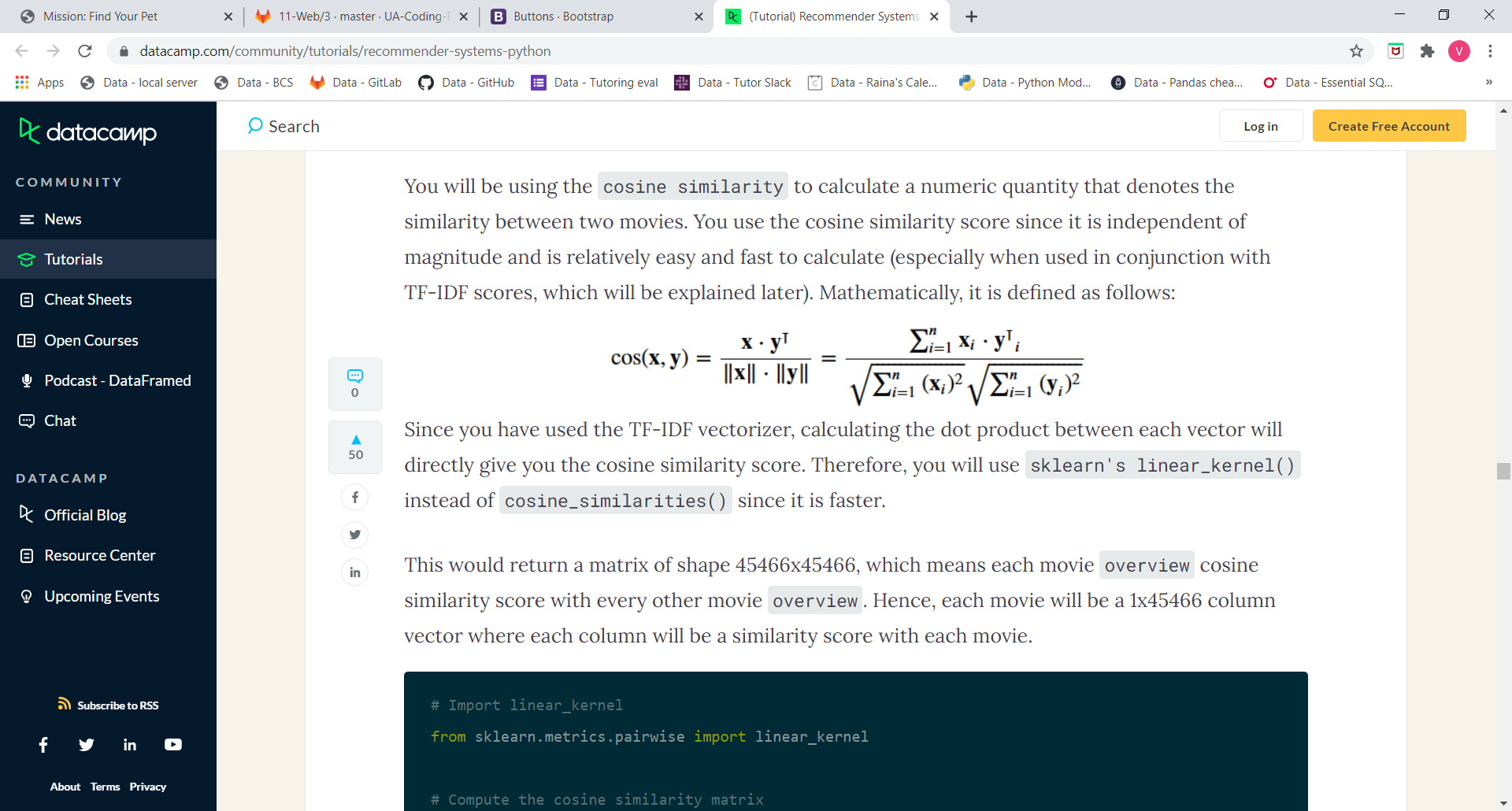


Training the model

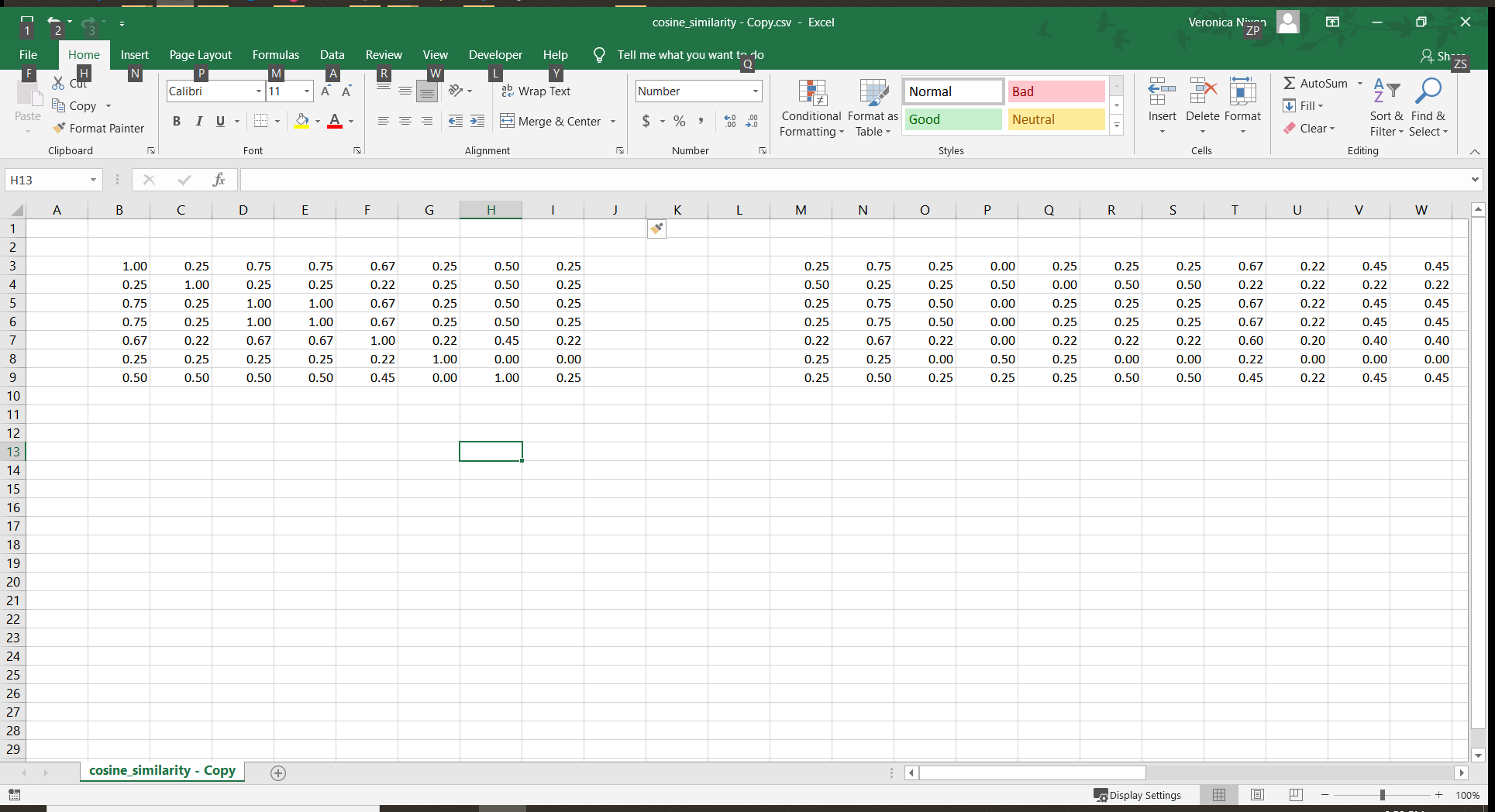
Python SciKitLearn’s CountVectorizer tool is used to compute the count matrix for the metadata soup. Each word present in the collection of metadata soups is listed along the columns. For each document, a 0 or 1 is assigned depending on whether it contains the word. Each document has now been converted to an n-dimensional vector.



Python SciKitLearn cosine\_similarity tool takes the above count matrix as input and computes the angle between each document.



The cosine\_similarity tool outputs a matrix containing scores between 0 and 1 representing the distance between each document.



Running the model

Each time the user votes a pet up or down, the current pet ID and the vote direction are used to select the next pet to display. A Flask app looks up the cosine similarity matrix distance between the current pet and every other pet and selects either the closest pet or the farthest pet depending on whether the user voted the pet up or down.