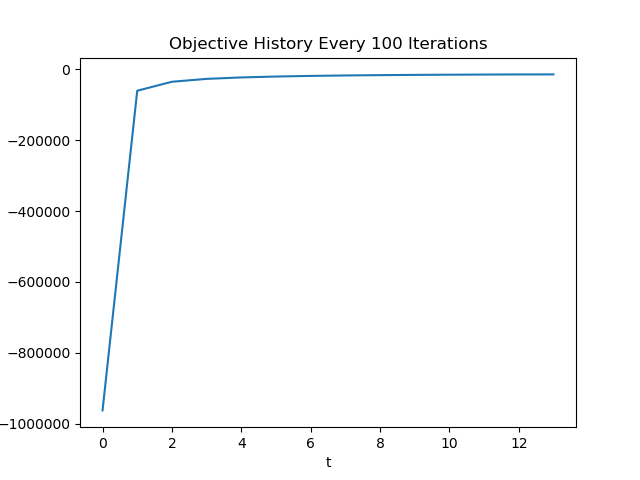
1. Record the performance of your top scoring submission on the validation and test sets.

- Accuracy on the validation set: 0.8068

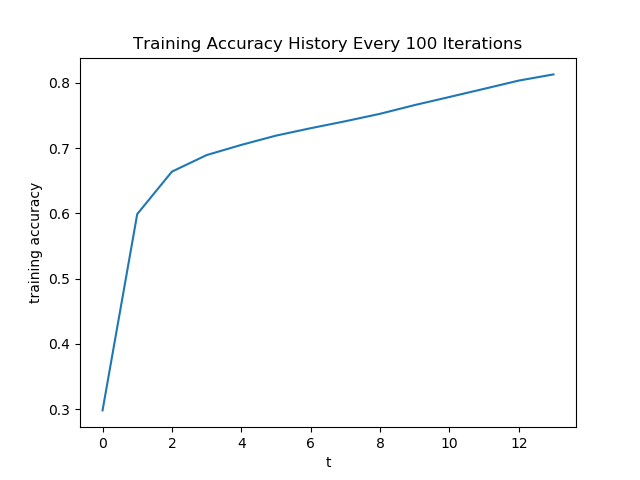
- Accuracy on the test set: 0.8001

2. Plot the objective function and accuracy, each as a function of the iteration t as your algorithm trains.

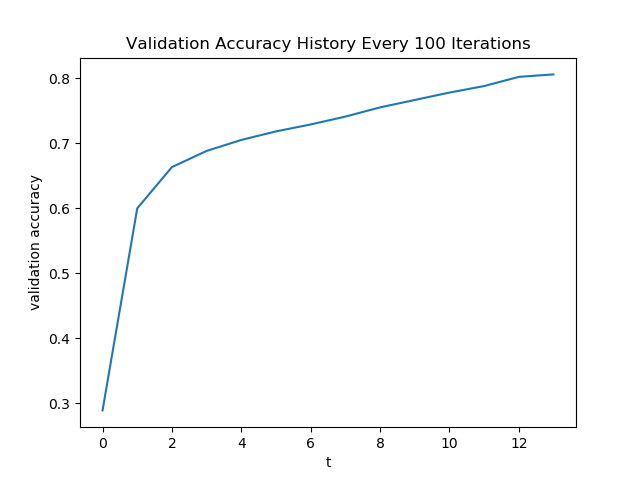
- Objective History every 100 iterations:



- Training Accuracy every 100 iterations:



- Validation Accuracy every 100 iterations:



3. Describe the features you added to boost the performance.

- Character n-gram features:

a. character unigram count

b. character bigram count

c. character trigram count

- Customized Boolean features:

a. number of words in the name string (=1, =2, =3, =4, >=5)

b. if name string contains digits

c. occurrence of frequent words for companies (INC, LTD, CO, TRUST, etc.)

d. occurrence of frequent words for drugs (GEL, CREAM, COUGH, LOTION, etc.)

e. length of the name string (<10, <20, <30, >=30)

f. pattern for people’s names:

Each word in the name string starts with an uppercase letter.