8. Parameter Tuning

Project due Oct 7, 2020 05:29 IST *Completed*

You finally have your algorithms up and running, and a way to measure performance! But, it's still unclear what values the hyperparameters like T and λ should have. In this section, you'll tune these hyperparameters to maximize the performance of each model.

One way to tune your hyperparameters for any given Machine Learning algorithm is to perform a grid search over all the possible combinations of values. If your hyperparameters can be any real number, you will need to limit the search to some finite set of possible values for each hyperparameter. For efficiency reasons, often you might want to tune one individual parameter, keeping all others constant, and then move onto the next one; Compared to a full grid search there are many fewer possible combinations to check, and this is what you'll be doing for the questions below.

In **main.py** uncomment Problem 8 to run the staff-provided tuning algorithm from **utils.py**. For the purposes of this assignment, please try the following values for T: [1, 5, 10, 15, 25, 50] and the following values for λ [0.001, 0.01, 0.1, 1, 10]. For pegasos algorithm, first fix $\lambda = 0.01$ to tune T, and then use the best T to tune λ

Performance After Tuning

7/7 points (graded)

After tuning, please enter the best T value for each of the perceptron and average percepton algorithms, and both the best T and λ for the Pegasos algorithm.

Note: Just enter the values printed in your main.py. Note that for the Pegasos algorithm, the result does not reflect the best combination of T and λ.

For the **perceptron** algorithm:

$T =$ 25 ✔

With validation accuracy =

0.7940 ✔

For the **average perceptron** algorithm:

$T =$ 25 ✔

With validation accuracy =

0.8000 ✔

For the **pegasos** algorithm:

$T =$ 25 ✔

$\lambda =$ 0.01 ✔

With validation accuracy =

0.8060 ✔

Submit You have used 1 of 20 attempts

Accuracy on the test set

1/1 point (graded)
After you have chosen your best method (perceptron, average perceptron or Pegasos) and parameters, use this classifier to compute testing accuracy on the test set.

We have supplied the feature matrix and labels in `main.py` as `test_bow_features` and `test_labels`.

Note: In practice the validation set is used for tuning hyperparameters while a heldout test set is the final benchmark used to compare disparate models that have already been tuned. You may notice that your results using a validation set don't always align with those of the test set, and this is to be expected.

Accuracy on the test set :

0.802 ✔

Submit You have used 4 of 20 attempts

The most explanatory unigrams

10/10 points (graded)
According to the largest weights (i.e. individual i values in your vector), you can find out which unigrams were the most impactful ones in predicting **positive** labels. Uncomment the relevant part in `main.py` to call `utils.most_explanatory_word`.

Report the top ten most explanatory word features for positive classification below:

Top 1 : delicious ✔
Top 2: great ✔
Top 3: ! ✔
Top 4: best ✔
Top 5: perfect ✔
Top 6: loves ✔
Top 7: wonderful ✔
Top 8: glad ✔
Top 9: love ✔
Top 10: quickly ✔

Also experiment with finding unigrams that were the most impactful in predicting negative labels.

Submit You have used 2 of 20 attempts

Discussion

Hide Discussion

Topic: Unit 1 Linear Classifiers and Generalizations (2 weeks)\Project 1: Automatic Review

Add a Post

Show all posts by recent activity

[Staff] can you share the steps when "show Answer"	[Staff] can you share the steps when "show Answer" I didn't get the right answers, but with only answers showing here, It's hard to k...	1
[STAFF] Grading Technical Error?	Hi, so it turns out that my responses to the last question on this page, 'The most explanatory unigrams', are essentially the same as t...	3
Again, same answers as submitted (but worng) after using given algorithms	I still do not understand, why after introducing in project1.py the given solutions, I still get the same answers as previously provided,...	1
Tuning	Can anybody explain how to perform parameter tuning program using main.py i'm not getting any output when I uncomment 8th q...	1
[STAFF] Please Help !	Dear staff, I keep getting the accuracy of the perceptron algorithm wrong, both in the "Baseline accuracy" in problem 7 and the "Per...	4
Help: most_explanatory_word, best_theta	Traceback (most recent call last): File "main.py", line 124, in <module> sorted_word_features = utils.most_explanatory_word(best_the...	12
Query regarding last	Should it not be mentioned in the question which training model we are using , as theta as depends on the model ?	4
"Could not format HTML for problem. Contact course staff in the discussion forum for assistance."	I get this error in the section where we are supposed to enter the 10 words with the most weight. How can I resolve this problem? It ...	5
Could not format HTML for problem. Contact course staff in the discussion forum for assistance.	Staff Kindly Help, i could not answer this part and it has greater proportion of marks i.e. 10 Marks. I have all the answers of this part ...	2
confused ☹ need help!!!	under "The most explanatory unigrams" section in project-1 i'm not able to understand the question . how do we get best_theta vect...	4
please help. after i run The most explanatory unigrams for 2 hours i get the following message.	I mention that till now everything went OK File "F:\MTI-ML\resources_sentiment_analysis\sentiment_analysis\utils.py", line 167, in m...	1
Most_Explanatory_Words?	I am getting only two words, and they are also not correct?? Any Idea about how should I select my best_theta?	1 new 18
Last question - rearrange the values?	My last four words were marked wrong. I tried some of the words form the next 10 words from the list. Then I tried re-arranging the...	2

Previous Next

edX for Business
Open edX
Careers
News

Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap

Connect

Blog
Contact Us
Help Center
Media Kit
Donate

/