# 5. Temperature

We will now explore the effects of the temperature parameter in our algorithm.

**You will be working in the files `part1/main.py` and `part1/softmax.py` in this problem**

## Effects of Adjusting Temperature

1/1 point (graded)

Explain how the temperature parameter affects the probability of a sample $x^{(i)}$ being assigned a label that has a large $\theta$. What about a small $\theta$?

- [ ] Larger temperature leads to less variance
- [x] Smaller temperature leads to less variance
- [ ] Smaller temperature makes the distribution more uniform

✔

Submit You have used 3 of 3 attempts

✔ Correct (1/1 point)

## Reporting Error Rates

2.0/2.0 points (graded)

Set the temperature parameter to be 0.5, 1, and 2; re-run `run_softmax_on_MNIST` for each one of these (add your code to the specified part in **main.py**).

$\text{Error}|_{T=0.5}$ = 0.08399999999999996 ✔

$\text{Error}|_{T=1}$ = 0.10050000000000003 ✔

$\text{Error}|_{T=2}$ = 0.1261 ✔

Submit You have used 1 of 20 attempts

**Return the `temp_parameter` to be 1 before moving on to the next section.**

## Discussion

Hide Discussion

**Topic:** Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks):Project 2: Digit recognition (Part 1) / 5. Temperature

Add a Post

All Posts

### What is "Effects of Adjusting Temperature" question asking

question posted 4 days ago by **Cool7** (Community TA)

Anybody get what this question is asking?

Basically higher temperature will make all outcomes more equiprobable and lower temperature causes a greater difference. But which distribution is the question refer to?

This post is visible to everyone.

**Mark_B2** (Community TA)
4 days ago - marked as answer 4 days ago by **Cool7** (Community TA)

$p\left(y\left(i\right)=j|x\left(i\right),\theta\right)$

Thanks, that means "variance" in the question is not refer to this distribution's variance. Then it must be the variance of probability of a sample $x^{(i)}$ being assigned a label that has a large $\theta$.

This feels more like a word puzzle then a ML question. 😒

posted 4 days ago by **Cool7** (Community TA)

"*"variance" in the question is not refer to this distribution's variance.*"
It does. It returned by `compute_probabilities` function and temperature is one of parameters. Would temperature change the returned variance?

posted 4 days ago by **Mark_B2** (Community TA)

Oh! I don't know what was I thinking, confused myself. I really should stop doing things at night,embarrassing. Thanks very much.

posted 4 days ago by **Cool7** (Community TA)

I'm definitely confused on this one. If a higher temperature makes all outcomes more equiprobable then why is a decrease in variance for larger temperature apparently not correct?

posted 4 days ago by **synnfusion**

The variance of uniform distribution on [0, 1] is 1/12, the variance of delta function is zero.

posted 3 days ago by **Mark_B2** (Community TA)

Might be helpful if I remember that variance depends on distance from mean and not how much the pdf is going up and down.

posted 3 days ago by **synnfusion**

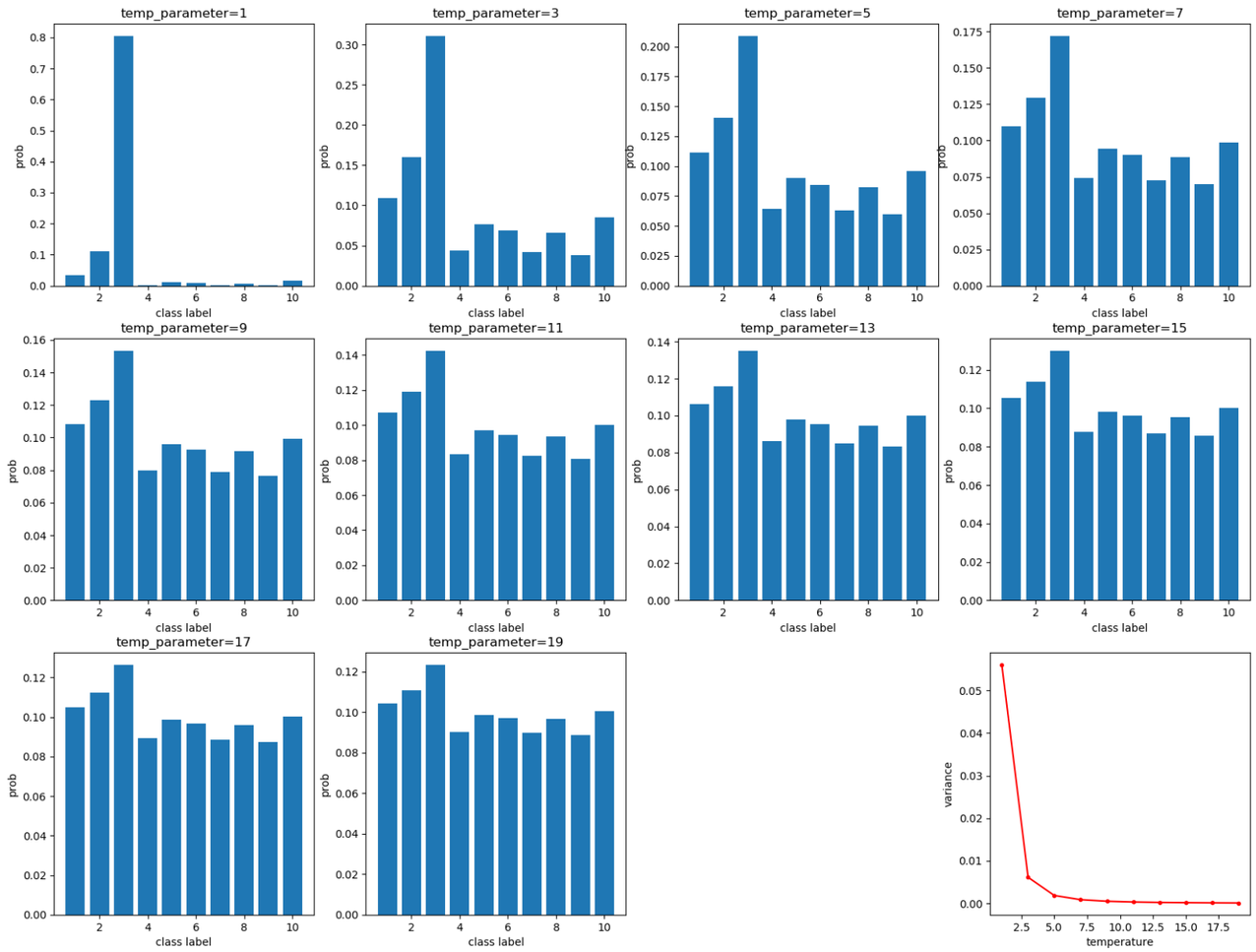
Add a comment

Add a Response

3 other responses

**sandipan_dey**
2 days ago

Already used 2 attempts and yet to get the correct answer. Used a random theta matrix and plotted the class-prob-distribution and temperature-vs-variance with the probabilities obtained using compute_probabilities() ,for a test datapoint. The plot supports the intuition, although the grader does not accept the answer and i am not sure what i am doing wrong. Any help? Thanks in advance.




I think from the earlier comments, this question is asking a different thing, to what I think it's asking!

posted 2 days ago by **robweatherston18**

Check your variance function, note that np.var is not appropriate to calculate variance in this case.

posted about 23 hours ago by **nontawat**

Why is np.var inappropriate @nontawat? I thought the question is asking the variance of the computed probabilities (one for each class). Even if we use ddof=1, the result will not change.

posted about 14 hours ago by **sandipan_dey**

Hint: look at your plot and figure out which kind of distribution it is.

posted about 6 hours ago by **nontawat**

Add a comment

**sakimarquis**
a day ago

same here, i use all 3 attempts just guessing, it's really confusing.

Add a comment

**dkhachatrian**
about 14 hours ago

What @Mark_B2 described is what the grader is asking for. Staff should probably make this more explicit. Especially because usually, looking at the distribution of the *ID number corresponding to a predicted label for a classification problem* doesn't have much meaning. For example, what if the label-encoding were {0: "cat", 1: "dog", 2:"other"} rather than our current example of {0:'0', 1:'1', ..., 9:'9'}? (I imagine a more "meaningful" distribution to look at could be the 1-D distribution of probabilities -- to see whether/how often the classifier is "confident" vs "ambivalent".)

Add a comment

Showing all responses

## Add a response:

Preview

**Submit**

Learn About Verified Certificates