* What's done.
-all the main blocks of the model are written separately as new classes
-data processing is done via custom collate_fn and dataloader
-some tries to combine everything altogether

* What didn't work.
-positional encoding function
-alignment module
-total model is not completed and not properly tested :(

* Explain some worst bugs that you had to fix.
Dataloader struggles and understanding how collate_fn can help with it:)
Tries with positional encoding and not finishing due to the time limit; alignment block as well.

* For you personally, what were the ****top 3 hardest things**** in this assignment? (name anything, even "**waiting to be unblocked on Colab***")
-you are right about "unblocked Colab", even took Pro version to prevent this issue
-to make the assignment on somewhat blank list (notebook) - no pre-written helpful functions, just some details to the picture and internet to complete the model "picture"
-the process of understanding theory itself - a lot of components, non-trivial blocks, just stucked for much time on almost every issue. For example, alignment model - first think that it is an ordinary upsampling, then dived into a notebook, researched the paper, and tried to implement it in some way.
PS (one more thing) it is hard to make smth, when there is no one right answer (as in previous assignments) and organize yourself with an idea "if you would like to make smth, just make an argument for it and make this thing works"

* What have you learned while doing this task?
-that real life deep learning problems take much more time than classical homework assignments
-usage of torch transformers
-work with audio data

* What questions of yours has this task provoked? E.g. "I was curious if Transformers are used in X" or "I'm wondering if real papers really don't mention Y".
-that papers not provide all the details to reimplement the paper (in my ideal world all the details should be mentioned in some place)
-how to work with audio data - phonemes, duration and all this stuff (this was the first experience, images are more natural for work with data)
-in what tasks can transformers can be used, except classical NLP

* How would you improve this assignment?

-test all the blocks properly to retrieve all the bugs

-combine all the blocks together

-make tests on audio quality of the model