The solutions I provided are quite straightforward and clear. I personally didn't like the implementation of RoPE using query_imag and query_real, but the assignment is the assignment.

Also, I ran into a problem that sometimes sanity check fails due to uncertainty of computations or hardware problem.

What I mean is that my sanity check failed one time when there was the difference in one number in 4th digit after the decimal separator. And on the next launch the problem disappeared and never happened again.