Dzhabaev Islam’s report about interview tests

Part 1

The first task turned out to be more difficult than I expected. The site is hard to scrape. I tried to use BeautifulSoap first. Generally no result. Returned empty responses, although the same requests worked well on other sites. Then I tried to use Selenium. After each request, the site blocked me. So, I had to take long breaks between requests. Then I bought a VPN subscription. And I have used almost every VPN that I have. But, even with all the precautions, such as making the script behave like a human, the site blocked me when I tried to click on any button or link. Then I tried through a simple request with JSON, it blocks instantly. And I had to manually download a couple of photos to complete the second task. But, regardless of the results of this interview, I will continue to try to scrape data from this site. This is an interesting problem.

Part 2

The second task was easier than the first. To convert photos to .jpg, I wrote a small script in which I simply renamed the file names. Since a full conversion with an indication of the format for saving did not work, because on my windows 11 .jpg is saved as .jfif. So I just renamed the files.

I then wrote a "preprocess" function that takes 2 images, a clothing image and a clothing mask image. All downloaded mask images from google colab turned out to be smaller than the image of the clothes, so the first thing I did was make them the same size. (I also cropped the images, because the downloaded images contained scales on the sides. If all mask images are the same, then this process can be automated). Then I processed the mask image a little with MinFilter (here I wrote a small "erode" function to apply MinFilter several times for the best effect, after several iterations I found that 4 iterations give the best result). Then I converted the mask to gray scale and blurred the image a bit. Then I created an empty image in blue. And in the end I glued it all together using the "composite" function. This is all.