There are four stages of understanding a paper, from shallow to deep. However, no matter the stage of understanding, the understanding itself must be structured into a coherent summary where each step in the story is clearly motivated by the previous steps. It is also best to use your own languages for the summary rather than copy-paste from the paper (rephrase the paper basically). Otherwise, we will forget about the paper.

Stage 0: only know the high level.

1. Most papers we read will fall into this category.
2. Know what the paper is trying to achieve, how it is different from previous works and key insights from the experimental sections.

Stage 1: understand every sentence in the paper

1. Cho (NYU) makes a nice point, in which it is important to read every sentence in very relevant papers carefully.
2. The strength and weakness of the approaches proposed in the paper might be contained in a few or even a single sentence.

Stage 2: understand every line in the code

1. Since research in ML is driven to a large extent by experimental science methods, a lot of important knowledge about a proposed method is in the code and not the paper.
2. What I like to do is to use a debugger and step through the code execution step-by-step when the command to reproduce the paper is run.
3. Often, an algorithm might seem complicated from the description in the paper, but is actually very simple when presented in code.
4. VSCode and Pycharm are good option for this.
5. When stepping through the code, pay attention to any implementation details that is either not discussed in the paper or surprising.
6. This should be done for extremely relevant papers to the current research direction (e.g. SOTA paper, paper that sheds understanding into SOTA results, highly cited paper, etc.)
7. Also, we should be able to reproduce the results in the paper. If that is not possible, we should know why.

Stage 3: know what the authors might not themselves know

1. Sometimes, it is possible to know more about a proposed algorithm than even the original authors of the algorithm.
2. This knowledge, for example, can be obtained by running an ablation study that was not run in the paper.
3. Or questioning the assumptions made by the author and running experiments to validate their assumptions.
4. Research is by definition producing new knowledge. So when we get to this stage, we are ready to publish.