**Data Engineering Assignment – 25**

**CONTINUOUS INTEGRATION AND CONTINUOUS DEPLOYMENT:**

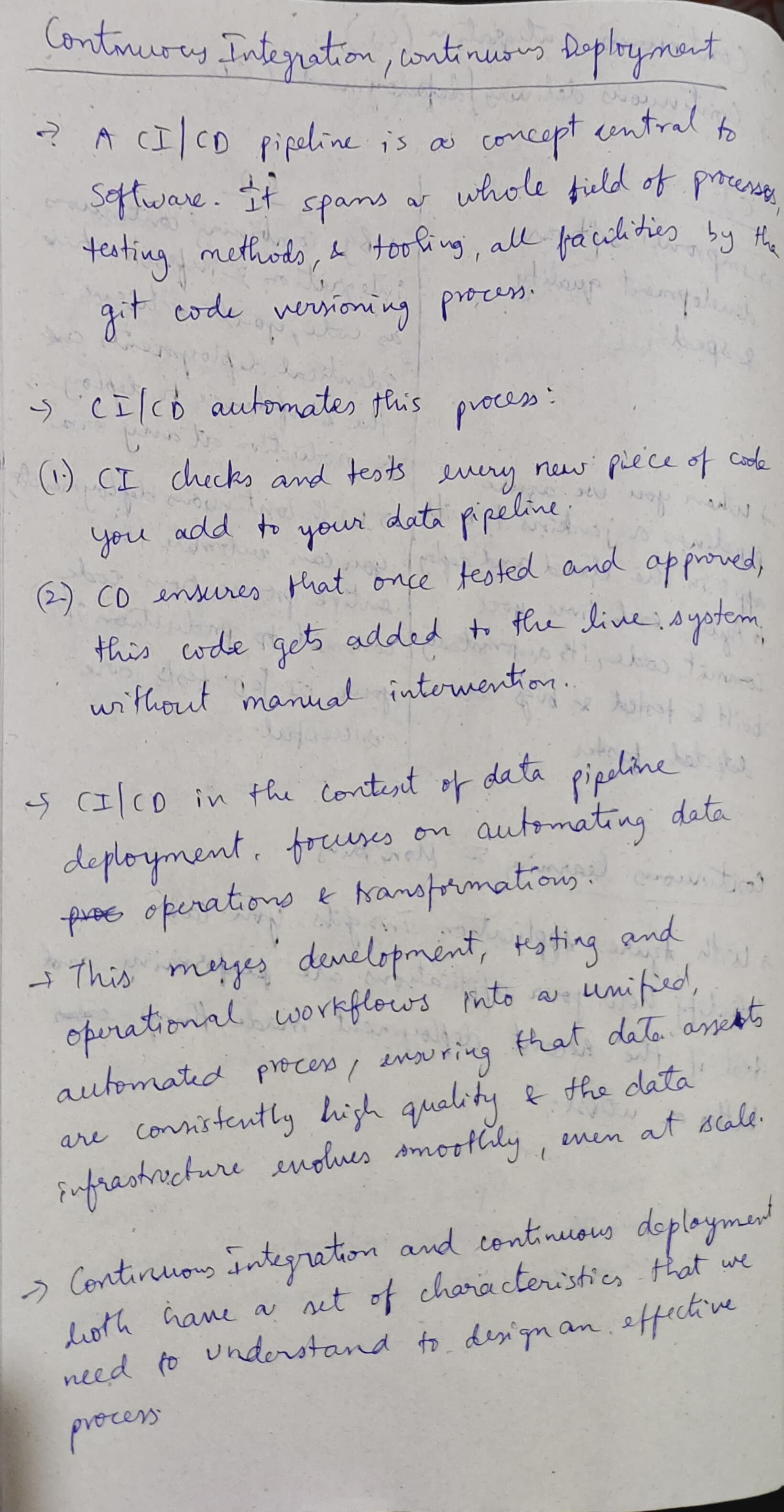
A CI/CD pipeline is a concept central to software. It spans a whole field of processes, testing methods, and tooling, all facilitated by the Git code versioning process. Since the terms “CI/CD Pipeline” and “Data Pipeline” are confusing, we will simply refer here to the CI/CD process.

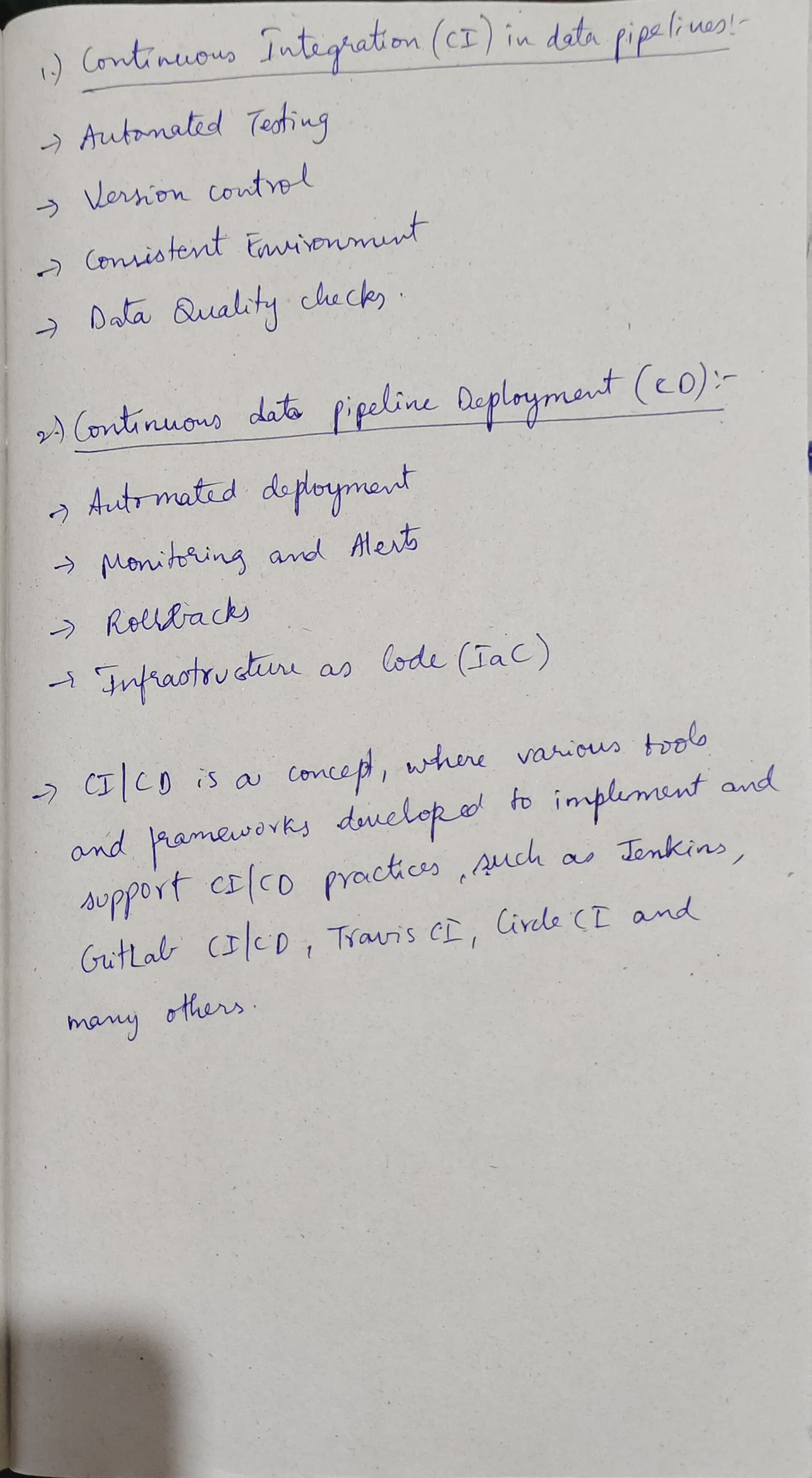
Imagine you're building a toy train track (your data pipeline). Every time you add a new piece (a code change), you want to ensure it fits perfectly and doesn't derail the train (break the pipeline).

**Continuous integration:** Every time you add a new track piece, you immediately test it by running the toy train (data) through it. This ensures that your new addition didn't introduce any problems. If there's an issue, you know instantly and can fix it before it becomes a bigger problem.

**Continuous deployment:** Once you've confirmed that your new piece fits and the train runs smoothly, you don't wait to show it off. You immediately let everyone see and use the updated track. In other words, as soon as your changes are verified, they're made live and functional in the main track (production environment).

**Hand Written Notes:**

****

****