Supervised Fine-tuning Vs Unsupervised Fine-tuning

References:

https://klu.ai/glossary/supervised-fine-tuning

Constraints taken into account:
- Hardware: NVIDIA A30 GPU (24GB VRAM) on Titan
- Timeline: 3 Weeks
- Team size: 4 people
- Goal: Beat the RAG implementation in having Indian cultural nuanced responses

A: Supervised Fine-tuning:
- Since it teaches on Q & A pairs giving examples on how exactly you want to the model to answer questions
- More efficient with limited data, you get more learning progress per data point even 500-1000 labelled examples can show significant learning improvement
- Fast results: You see training results within hours, allows for quick iterations over the 3 weeks (really good for our timeframe and team size)

D: Supervised Fine-tuning:
- Training data must be carefully handpicked, labelled and potentially manually annotated for the model to like it
- You need to create Q & A pairs created manually by hand, using python automation or with the help of LLMs
- Overfitting risk, they might reduce the generalization level of the model and become biased too much towards Indian cultural nuances
- Coverage limitations, it doesn't learn what you don't teach it (Coverage limitations)

A: Unsupervised Fine-tuning:
- Can leverage vast amounts of easily obtainable data to fine-tune the model (You can just scrape Indian books, news articles, websites etc.
- No time used creating Q & A pairs you just throw in the tokens into the model
- Broader knowledge Acquisition, gets a general understanding of the language, better generalization to unseen quesitons
- Picks up nuances that we might miss labeling or adding questions for
- May discover patterns and nuances not typically picked up by humans

D: Unsupervised Fine-tuning:
- Requires 10,000 - 100,000+ examples to start seeing noticeable results
- Very computationally expensive, takes longer to train so no chance for iterations when we make mistakes in setting up the parameters | Because of this the 3 week timeline may not be sufficient to come up with a working solution
- The model may learn patterns but not use that knowledge to answer the questions so we may still need to use SFP after to see some differences
- May learn irrelevant or negative patterns because we don't have full control what it's learning

Conclusion:
- Due to time, team size and computer resources constraints the only realistic option would be to use Supervised Fine Tuning where the main focus would be to get really good high quality Q & A pairs somewhere around 1000-1500 of them and do some fine tuning iterations on Titan and since training is in the form factor of hours in terms of the time taken, if the model is not performing as well it is easy to tweak parameters easily and un it again within the time that we have but with UFT, all this would not be possible.