

# LLM-for-code-intelligence

---

Fine Tuning Large Language Models for Code Intelligence - Msc Project

## Todo List 19/06 - 25/06

---

- [x] Document Current Challenges, Progress and Next Steps to progress\README.md
- [x] Identify Relevant Literature
- [x] Read Identified Literature and summarise to lit-rev\README.md (Ongoing)
- [x] Identify Relevant Datasets
- [x] Document Identified Datasets to lit-rev\README.md
- [x] Identify Relevant Evaluation Metrics
- [x] Document Identified Evaluation Metrics to lit-rev\README.md
- [x] Setup the code generation environment
- [x] Document the code generation environment setup
- [x] Setup the functional evaluation environment
- [x] Document the functional evaluation environment setup
- [x] Setup the similarity evaluation environment
- [x] Document the similarity evaluation environment setup
- [x] Setup the Full Fine Tuning environment
- [x] Document the Full Fine Tuning environment setup
- [x] Setup the Parameter Efficient Fine Tuning environment
- [x] Document the Parameter Efficient Fine Tuning environment setup
- [ ] Setup the Knowledge Distillation environment
- [ ] Document the Knowledge Distillation environment setup

## Todo List 26/06 - 02/07

---

- [x] Try to experiment with the Learning Rate
- [ ] Try to experiment with the Number of Layers to Freeze
- [x] Try to experiment with using a sample of the dataset processed and downloaded from the cloud
- [x] correlation between evaluation metrics and the number of fine-tuning steps
- [x] Similar Setup for the Parameter Efficient Fine Tuning and Full Fine Tuning
- [ ] Documentation of progress and next steps

# Fine-tuning CodeGen on TheStack

---

Colab:  [Open in Colab](#)

Huggingface:  [Share to community](#)

Weights & Biases:  [Visualize in W&B](#)