

**MAI 475– Large Language Model**  
**IV MSAIM**  
**05-08-2025**

**Regular lab Question**

**Lab Exercise 7: To implement and analyze the performance of LLM Models on Code-Based Large Language Models**

**Problem Statement:**

**You are tasked with developing a code-specialized Large Language Model (LLM) for automatic code generation and understanding.**

**In this exercise, you have to:**

- 1. Fine-tune a pre-trained Transformer-based model (e.g., CodeBERT, PolyCoder, or GPT-Neo) on a given programming language dataset (Eg, Python, Java, or C++).**
- 2. Implement evaluation on a benchmark dataset for code tasks, code completion, and bug detection.**
- 3. Compare the performance of your model with a general-purpose LLM (such as GPT-2 or BERT, the same foundational model that you are using) trained on natural language.**

**Tasks:**

- Preprocess the provided code dataset (According to the requirement).**
- Fine-tune the chosen code-based LLM model.**
- Train a general-purpose LLM for the same task.**
- Evaluate both models using at least two evaluation metrics (CodeBLEU, Accuracy, or F1-score).**

**[Link for CodeBLEU - <https://arxiv.org/abs/2009.10297> ]**

- Present a comparative analysis of the results, highlighting which model performs better and why?**
- Compare the performance Evaluation Metrics of the foundation and code-based models.**

**NB: No marks will be credited for the pipeline implementation.**

## **Program Evaluation Rubrics**

<b>Model Selection and Implementation</b>	<b>6 Marks</b>
<b>Timely Submission</b>	<b>2 Marks</b>
<b>Viva</b>	<b>2 Marks</b>

## **General Instructions**

- The file you have to save with your name, last 3 digits of register number and program number "Anto\_501\_Lab1".
- The implemented code you have to download and upload to the Google Classroom in the given scheduled time.