

Research

- **Datasets available as a benchmark**

- a. [CommitBench](#)
 - Newest Dataset 2024
 - Multilingual Dataset
 - Filtered
 - [Paper link](#)
- b. [MCMD](#)
 - Dataset 2022
 - Multilingual
 - Not filtered
 - [Paper link](#)

- **Different Approaches that papers implement**

- a. [LLMs via In-Context Learning](#)
 - Use prompt engineering on different language models and evaluation on MCMD dataset
- b. [RACE: Retrieval-Augmented Commit Message Generation](#)
 - Train T5 model with code diff along with relevant code to code diff and its commit message from vector database.
- c. [From Commit Message Generation to History-Aware Commit Message Completion](#)
 - Auto complete commit message using CMG models and gpt 3.5.
- d. [Fine tuning CodeBert](#)
 - First commit generation model.