RESEARCH PROPOSAL : INTRINSIC EVALUATION IN MODELS TRAINED ON SOURCE CODE

Reshinth Adithyan
reshinth.adith@gmail.com

1 ABSTRACT

The proposal is to devise a set of probing tasks and classifiers to check how much of the implicit Code related Syntax/Semantics has the Pre-Trained Language Models for Source Code has learnt in contrast to Karmakar et al. (2021); Lu et al. (2021); Alon et al. (2018) where extrinsic evaluation is done . Intrinsic evaluation will play an important role in understanding the dexterity of the Language Models. I, hypothesize that using such intrinsic evaluation methods can help us choose Models for fine-tuning tasks which are heavily reliable on syntax/semantics heavy tasks.

Some tasks are as follows,

- **Lexical Probing(Token Level)** - Token level classification for lexeme identification Function, Method, Hardcoded, Variables, Operators.
- **Syntax** - AST based Structural Probing.
- **Semantics** - DFG based Semantic Probing(Use-Def Edge Probing).

To be precise, Building a Universal Dependency kind of **Intrinsic Evaluation for Models trained on Source Code**. These Probing Tasks/Datasets are aimed to capture the capability of the Model to evaluate from Syntactical, Semantic and Lexical Property of the Model.

REFERENCES

Uri Alon, Meital Zilberstein, Omer Levy, and Eran Yahav. code2vec: Learning distributed representations of code. *CoRR*, abs/1803.09473, 2018. URL http://arxiv.org/abs/1803.09473.

Anjan Karmakar, Julian Aron Prenner, Miltiadis Allamanis, and Romain Robbes. {GLUEC}ode: A benchmark for source code machine learning models, 2021. URL https://openreview.net/forum?id=5IqTrksw9S.

Shuai Lu, Daya Guo, Shuo Ren, Junjie Huang, Alexey Svyatkovskiy, Ambrosio Blanco, Colin B. Clement, Dawn Drain, Daxin Jiang, Duyu Tang, Ge Li, Lidong Zhou, Linjun Shou, Long Zhou, Michele Tufano, Ming Gong, Ming Zhou, Nan Duan, Neel Sundaresan, Shao Kun Deng, Shengyu Fu, and Shujie Liu. Codexglue: A machine learning benchmark dataset for code understanding and generation. *CoRR*, abs/2102.04664, 2021. URL https://arxiv.org/abs/2102.04664.