

# L3Masking: Multi-task Fine-tuning for Language Models by Leveraging Lessons Learned from Vanilla Models

Yusuke Kimura<sup>†</sup>, Takahiro Komamizu<sup>‡</sup>, and Kenji Hatano<sup>†</sup>

<sup>†</sup>Doshisha University, <sup>‡</sup>Nagoya University, Japan



## Eliminating the gap between pre-training and fine-tuning using Multi-task Fine-tuning

### Designing an auxiliary task leveraging lessons learned from vanilla models

**L3Masking** (Idea: The lower likelihood, the more masking, vice versa)

