

A standard approach to many natural language processing tasks is to take an existing, pre-trained large language model and fine-tune it for the given task. Such an approach leads to having a separate model for each task; furthermore, the fine-tuning must be repeated when upgrading to a new pre-trained model. This thesis explores the possibilities of using a single off-the-shelf model for three different tasks without fine-tuning. We present *Predictor*, a writing assistant that supports rewriting a sentence after replacing one of its words, suggesting continuations, and suggesting words that fit into a sentence. We design the system in a model-agnostic way, making it possible to upgrade to a new model with little effort. We also provide an extension that integrates the assistant into the text editor.