

ABSTRACT

A data processing method and device, a model optimization method and device are provided. The data processing method includes: performing sampling for multiple datasets, and determining target sample data based on a sampling result; inputting the target sample data into multiple language models for processing, to obtain target prediction data outputted by each of the language models; performing text annotation processing on the target prediction data by a preset text annotation model, to obtain text annotation information corresponding to the target prediction data; and constructing a sample data array based on the target sample data, the target prediction data and the text annotation information, where the sample data array is used to perform a model training task and a model validation task.

(Figure 1)