servicenow.

ServiceNow summarization LLM Model Card

ServiceNow summarization LLM Model Card

Intended Use

The ServiceNow summarization LLM Model ("the Model") is designed to support ServiceNow users with text generation use cases, such as summarization as part of application workflows on the ServiceNow Platform.

The Model is designed for assistive text generation use cases, and it is important to have a human-in-the-loop or use human-AI collaboration to review the AI-generated output to ensure that the output is accurate and appropriate for the intended use case.

The Model is based on a model in the StarCoder large language model family, and has been further fine-tuned on ServiceNow Platform-specific data.

The license for the StarCoder family of models includes explicit usage restrictions that preclude, among other things, generating or disseminating verifiably false information and/or content, generating or disseminating personally identifiable information (PII) to harm others, any use that violates applicable laws, any use that results in automated decision making that adversely affects individuals' rights, and any use that is intended to discriminate based on legally protected characteristics.

See Attachment A of the Open RAIL-M v1 license for the StarCoder family of models and a list of usage restrictions.

Model details
Developed by: ServiceNow Research and Engineering
Released: Sep. 2023, v1.0

Evaluation data
The Model is evaluated on English datasets. Open-source benchmarking datasets were used for general NLP task evaluation, and ServiceNow Platform-specific data was used for summarization use-case specific evaluation.

Metrics
Rouge, BERTScore

Model Architecture
The Model is an instruction finetuned version of a model in the StarCoder model family. The Model is pre-trained using fill-in-the-middle and finetuned with next token prediction.

Training data
This Model is based on a model in the StarCoder model family, which was trained on data from publicly available sources.

Fine-tuning was done with ServiceNow Platform-specific text data.

User Benefits

The Model is trained and fine-tuned for several natural language processing (NLP) tasks including intent classification, summarization, and question answering (Q&A).

Risks

The Model may generate text summaries that may be inaccurate, omit key information, or include irrelevant or redundant text. The Model may, at times, produce socially unacceptable or undesirable text, even if the prompt itself does not include anything explicitly offensive. For more information, refer to section 10 of the [StarCoder paper](#).

See Section 10 of the [StarCoder paper](#) for a discussion of documented biases and harmful outputs that the Model can produce.

Factors and limitations

The Model has been primarily trained on English data. While the model may respond to non-English text, the responses are unlikely to be correct or consistent.

The Model needs sufficient context in a prompt to create an acceptable response. If the prompt is too generic, the Model will not be able to follow the instruction.

Ethical considerations

The Model has been tuned with the intention of limiting bias, toxicity, and hallucinations, although such limitations may still exist. Text LLMs can produce harmful text based on how it is prompted, and the Model is not free from such limitations similar to the behavior of other industry LLMs.

Users should not use the Model to generate PII with the purpose of harming others, to generate verifiably false information with the purpose of harming others, or in any other way prohibited by the usage restrictions for the StarCoder family of models.

Please report instances of hallucinations, malicious text, or PII in output so that we can evaluate for remediation.