

2024

ServiceNow Alert Summarization LLM Model Card

ServiceNow Alert Summarization LLM Model Card

Intended Use

The ServiceNow Alert Summarization LLM Model ("the Model") is designed to support ServiceNow users with text generation use case such as summarization and analysis of alerts in 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 Llama2 large language model family and has been further fine-tuned on ServiceNow Platform-specific data.

The license for the Llama2 family of models includes explicit usage restrictions that preclude, among other things, generating or promoting fraud or disinformation, generating or disclosing sensitive personal or private information about individuals without rights and consents required by applicable laws, any use that violates the law or others' rights, engaging in or promoting activities that present a risk of death or bodily harm to individuals, and engaging in or facilitating the harassment or abuse of individuals or groups of individuals.

See [Llama2 Community License](#) for the license of the Llama2 family of models, and the linked acceptable use policy for details.

Model details

Developed by: ServiceNow
Research and Engineering
Released: Mar 2024, 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 and analysis use-case specific evaluation.

Metrics

F1, Rouge, BERTScore

Model Architecture

The Model is an instruction finetuned version of a model in the Llama2 model family.

The Model is an auto-regressive model, pre-trained and fine tuned with next token prediction. Details of pre-training are in the Llama2 paper.

Training data

This Model is based on a model in the Llama2 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/analysis, and question answering (Q&A).

Risks

The Model may generate text summaries and analyses 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 5.2 of the [Llama2 paper](#).

See Section 5.2 of the Llama2 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 Llama2 family of models](#).

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

