


AMOS SS08 Project cloud-native-llm

Template

Please use the following template for creating your project report.

Project name	Amos2024ss08-cloud-native-llm
Project mission	<p>To create a Large Language Model on Cloud Native Technologies our work was focused mainly on the following topics:</p> <ol style="list-style-type: none">1. Automated Dataset Preparation: Develop an automated system that prepares structured datasets from CNCF landscape documentation.2. LLM Fine-Tuning: Enhance a Large Language Model specifically tailored to the prepared dataset for precise and relevant responses.3. Open-Source Contribution: Release and open-source the fine-tuned model and dataset preparation tools to the community, fostering collaboration and innovation.
Industry partner	Kubermatic
Team logo	
Project summary	<p>DeepCNCF is a specialized Large Language Model designed to provide context-aware answers to questions related to Cloud Native technologies. To create this model, we collected data from various sources, including documentation sites on Cloud Native technologies and Stack Overflow. This data was then processed into a Q&A format.</p> <p>Using this curated dataset, we retrained the gemma2 model to improve its ability to answer questions specifically about Cloud Native technologies. Evaluations indicate that our best-trained model, with 9 billion parameters, significantly outperforms the original gemma2 model in this domain.</p>
Project illustration	<div><div><div>Models5</div><div><div>Kubermatic/DeepCNCF2BAdapter</div><div>Updated about 23 hours ago • 30</div></div><div><div>Kubermatic/DeepCNCF7BAdapter</div><div>Updated 3 days ago</div></div><div><div>Kubermatic/DeepCNCF</div><div>Updated Jun 10 • 31</div></div><div><div>Kubermatic/DeepCNCF9BAdapter</div><div>Text Generation • Updated 1 day ago</div></div><div><div>Kubermatic/DeepCNCFQuantized</div><div>Text Generation • Updated Jun 10 • 82 • 1</div></div></div><div><div>Datasets5</div><div><div>Kubermatic/Benchmark-Questions</div><div>Viewer • Updated 15 days ago • 41</div></div><div><div>Kubermatic/stackoverflow_QAs</div><div>Viewer • Updated 15 days ago • 9.11k • 19</div></div><div><div>Kubermatic/cncf-raw-data-for-llm-training</div><div>Viewer • Updated 20 days ago • 40.7k • 14</div></div><div><div>Kubermatic/Merged_QAs</div><div>Viewer • Updated 15 days ago • 94.1k • 418</div></div><div><div>Kubermatic/cncf-question-and-answer-datas...</div><div>Viewer • Updated 20 days ago • 84.9k • 372</div></div></div></div>

	<div><div><div><div><div><div></div><div>User</div></div></div><div><div><div></div><div>Assistant (DeepCNCF_2b_F16)</div></div></div><div><div><div><div>What is Prometheus?</div></div><div><p>Prometheus is an open-source monitoring system that allows you to monitor your applications, servers, and infrastructure metrics. It collects data from various monitoring software and other sources and is available as a static file or a service.</p><p>Prometheus is split in two parts: Prometheus server: A scrape-based monitoring system that can be deployed as a static file or as a service. Prometheus metrics: An instrumentation layer that adds metrics to your applications, servers, and infrastructure.</p><p>How does Prometheus work? Prometheus data is scraped from various monitoring software and other sources at regular intervals. It collects data in an open standard format that is easy to process. The metrics are stored internally in Prometheus and are used to determine whether any problems have occurred.</p><p>What are the limitations of Prometheus? Prometheus is a tool for monitoring systems, not for analyzing data. For example, it can be used to track the number of web requests received by an application, but it cannot tell you what type of web requests are being received. Prometheus is a tool for collecting data, not for analyzing it. For example, it can be used to collect information about the number of web requests received by an application, but it cannot tell you what the values of that metric are.</p><p>What are the benefits of using Prometheus? Prometheus is a powerful and versatile monitoring tool that can be used to monitor your applications, servers, and infrastructure. It is available as an open-source tool and has a large and active community.</p><p>What are the drawbacks of using Prometheus? Prometheus is not a tool for analyzing data. It is a tool for collecting data. Also, a single Prometheus installation can only collect data from a single source of data. If you need to collect data from multiple sources, you will need to install multiple Prometheus instances.</p></div></div></div></div><div><div>Send a message...</div><div></div></div></div></div>
Team photo	<div><div><div><div><div><div><div><div></div><div><div>Christian Wielenberg</div></div></div></div><div></div></div></div><div><div><div><div><div></div><div><div>Li Jiao</div></div></div></div><div></div></div><div><div><div><div><div></div><div><div>Anosh Rezaei</div></div></div></div><div></div></div></div><div><div><div><div><div></div><div><div>Daniel Schaeffer</div></div></div></div><div></div></div><div><div><div><div><div></div><div><div>Thorben Wagner</div></div></div></div><div></div></div><div><div><div><div><div></div><div><div>Yashodhar Pansuriya</div></div></div></div><div></div></div></div><div><div><div><div><div></div><div><div>Julio Perez</div></div></div></div><div></div></div><div><div><div><div><div></div><div><div>Dominic Fischer</div></div></div></div><div></div></div></div></div></div></div></div></div></div></div>
Project repository	<div><div><div>https://github.com/amosproj/amos2024ss08-cloud-native-llm.git</div><div>https://huggingface.co/Kubermatic</div></div></div>