LLMScan.online

AI Model Security Scan Report

B

Score 77/100

Report summary

GGUF v3 model | architecture: llama | quantization: F32 | 2 tensors | 1 blocking issue / 1 warning

Model

Filename	`zero-length-tensor-name.gguf`
Format	`gguf`
Architecture	`llama`
Size	`336 B`
SHA-256	`feb2404f9e891067f6f5a6fb9331754087ccfba51c66fd964084f8dede8605de`
Scan mode	`static-deep`
Scanned at	`2026-08-14T22:29:35+00:00Z`
Report ID	`ffe126d42a4d4a6ea09d4f0068ff2344`

Findings (2)

SEVERITY	CATEGORY	DETAIL
HIGH	Header Validation	**Tensor #1 has a zero-length name** GGUF tensor names are length-prefixed strings; a declared length of 0 is spec-valid but leaves the tensor unidentifiable by name and has caused parsers to mis-resolve tensor offsets or dereference a null/empty name pointer, leading to a crash (DoS) **Recommendation:** Reject this file; a production GGUF file should never contain an unnamed tensor
MEDIUM	Resource Limits	**Unusually high bits-per-parameter: 224.0** Values >128 bpp suggest most file content is not tensor data (padding, embedded files, or metadata bloat) **Recommendation:** Investigate what occupies the non-tensor file space

Runtime compatibility

RUNTIME	SUPPORTED	NOTES
llama.cpp	Yes	GGUF v3 container: supported | quant F32: supported
Ollama	Yes	GGUF via llama.cpp backend: container and quantization supported
LM Studio	Yes	GGUF container and quantization supported; GPU/CPU inference
vLLM	No	GGUF not supported; requires safetensors format
HuggingFace Transformers	No	GGUF not supported; requires safetensors or PyTorch checkpoint

Generated 2026-08-21 06:48 UTC by LLMScan.online (engine v2.0.0, report schema v1.0). Full interactive report: https://llmscan.online/report/ffe126d42a4d4a6ea09d4f0068ff2344 — Static analysis only; not a guarantee of safety. See llmscan.online/docs/how-it-works for methodology.