

TABLE I: List of the 31 topics with their top 10 keywords generated using BERTopic. The topics of the same cluster are grouped together

Topic Id	# of Posts	Topic Label	Top 10 Keywords
4	515	Maximum Context Length	length, context, tokens, context length, max, maximum, 2048, input, model, text
5	385	Tokenizer Implementation and Conversion	tokenizer, token, json, tokens, tokenizer model, model, tokenizer_config, tokenizer_config
7	349	Multilingual Translation Model	json, special, error
9	276	Prompt Format Template	language, english, languages, chinese, translation, model, french, support, multilingual, translate
12	224	Multilingual Audio Transcription Solution	prompt, prompt format, format, template, prompt template, assistant, user, instruction, prompts, use
18	141	Optimizing Chat Inference Time	audio, whisper, transcription, speech, transcribe, voice, language, file, output, model
28	66	Advanced Question Answering Models	electronic, music, inference, speed, time, td, slow, faster, inference time, inference speed
30	63	NSFW Censorship Analysis	question, answer, answering, question answering, context, answering model, prompt, model, answers, document
2,019		Cluster 1: Model Configuration and Performance Issues	
2	523	Loading Transformers Model Error	nsfw, uncensored, censored, scene, censorship, life, meaning, rules, oh, say
3	521	GPU Memory Allocation Issues	layers, model layers, model, transformers, self_attn, error, bias, py, file, from_pretrained
10	258	Inference API Error	memory, gpu, model layers, layers, self_attn, g_idx, g_idx model, bias model, model, vram
11	241	Model Loading Errors in WebUI	inference, api, inference api, 2023 07, 07, 2023, huggingface, 43, 42, endpoint
17	144	SageMaker Deployment Errors	webui, generation webui, text generation, generation, file, py, py line, text, line, ooba-booga_windows
1,687		Cluster 2: Model Load and Deployment Challenges	
0	843	Stable Diffusion Web UI Training	sagemaker, deploy, endpoint, aws, error, tgi, inference, aws sagemaker, instance, ml
8	310	License for Commercial Use	diffusion, stable, stable diffusion, image, ui, web ui, web, images, automatic1111, size
20	135	Model Download Issues	license, commercial, apache, mit, commercial use, use, license model, commercially, model, license hi
25	80	How to Use Model	download, download model, model, git, files, error, locally, run, load, access
26	72	DOI Citation Request	use, use model, example, model, app, usage, help, code, guide, example code
29	63	ONNX Model Conversion	doi, cite, request doi, paper, request, citation, model paper, doi hi, work, like cite
1,503		Cluster 3: Comprehensive Instruction Request	
6	353	Dataset for Training Model	onnx, onnx model, model onnx, export, convert, onnx version, onnxruntime, model, js, optimum
13	207	Sentence Embeddings for Semantic Similarity	dataset, data, bits, used, training, dataset used, int, set, std, binary
14	186	Model Evaluation and Benchmarking	embeddings, similarity, sentence, score, embedding, model, label, sentences, use, word
16	155	Fine-tune Model on Dataset	evaluation, score, leaderboard, scores, benchmarks, results, accuracy, model, benchmark, humaneval
			fine, tuning, fine tuning, tune, fine tune, tune model, model, model fine, tuning model, dataset

TABLE I: List of the 31 topics with their top 10 keywords generated using BERTopic. The topics of the same cluster are grouped together

Topic Id	# of Posts	Topic Label	Top 10 Keywords
19	136	Training Code Sharing	training, code, training code, share, share training, script, training script, train, scripts, thanks
22	102	Geneformer Perturbation Error	gene, cell, geneformer, perturbation, data, cells, dataset, insilicoperturber, silico, output_directory
24	87	LoRa Model Fusion	lora, loras, fine, lora model, lcm, training, peft, model, lora fine, tuning
1,226		Cluster 4: Model Training and Fine-Tuning Queries	
1	616	GGUF Model Version Request	gguf, version, ggml, gptq, quantized, quantization, model, bit, llama, gguf version
15	166	Neural Model Versions	13b, version, 7b, model, 3b, 70b, 33b, plans, 13b version, 30b
21	102	Model Merging Process	merge, merging, merged, models, mergekit, did, model, merge models, did merge, adapter
23	98	Model Discrepancy Investigation	llama, llama2, chat, teacher, llama 7b, older, 7b, old, years, mother
27	71	Mistral-7B Fine-Tuning Guide	mistral, mistral 7b, 7b, solar, v0, 7b v0, mistral model, 10 7b, solar 10, mixtral
1,053		Cluster 5: Updates and Requests for Model Versions	
Total	7,488		

TABLE II: List of the 3 topics with their top 10 keywords for the posts considered as outlier after clustering using BERTopic

Topic Id	# of Posts	Topic Label	Top 10 Keywords
0	3,510	Troubleshooting Model Loading Error	model, use, error, file, using, code, like, models, py, thanks
1	80	Choosing Between Bloom Versions	bloom, bloomz, use bloom, use, text, seed, like, generation, bigscience, model
2	64	Inference Methods for PEFT	checkpoint, checkpoints, method, model, training, intermediate, save, finetuning, inference, peft
3,654		Cluster 6: Efficient Model Fine-tuning and Deployment	
-1	136	Outlier	-
Total	3,790		