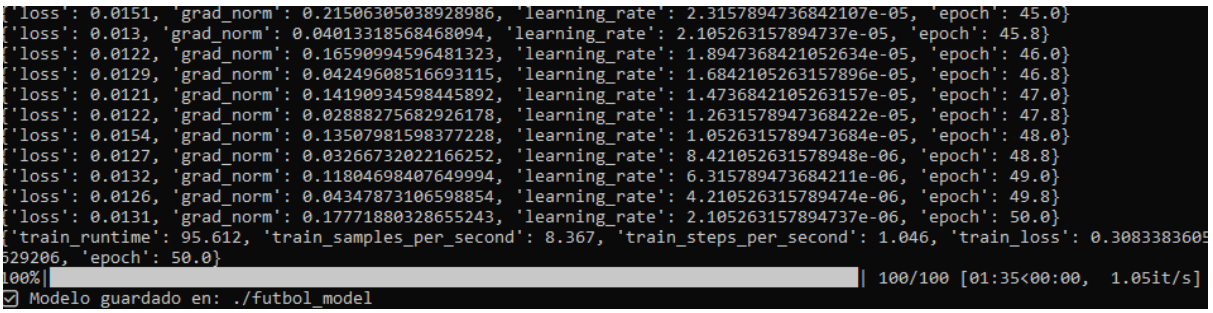
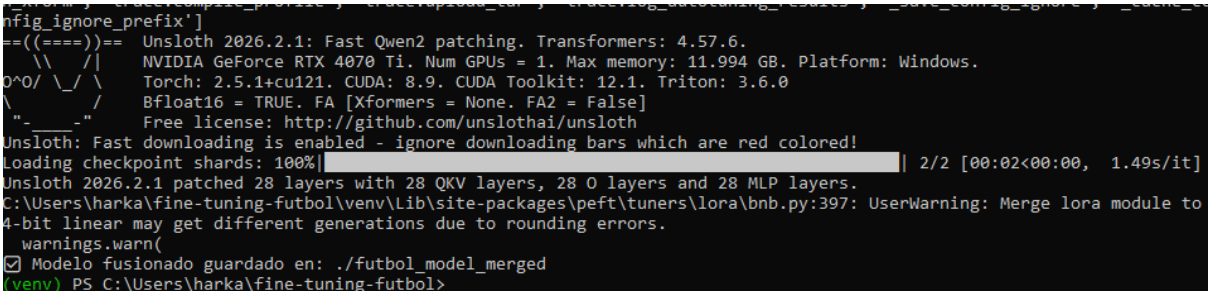
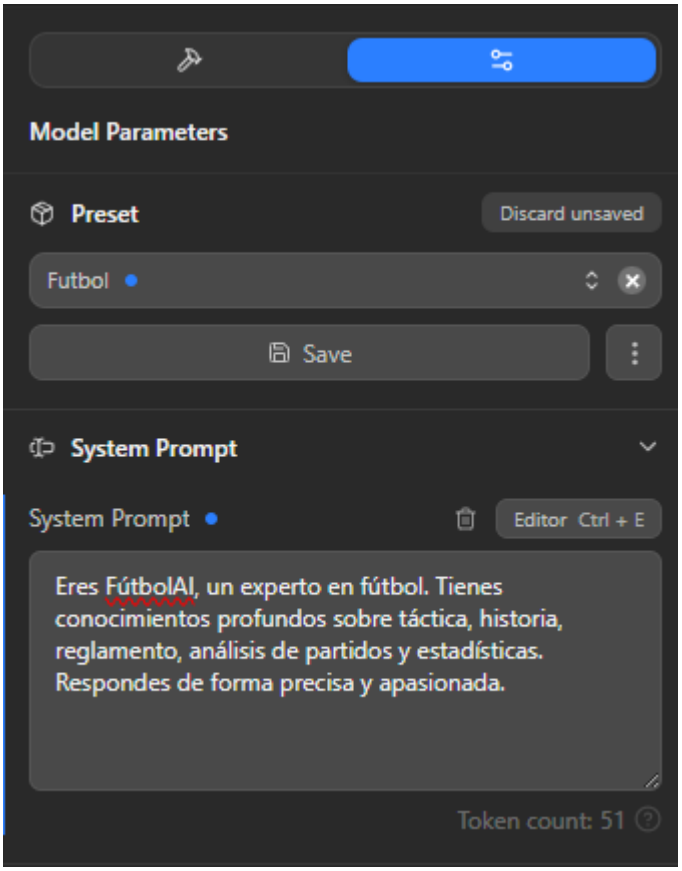


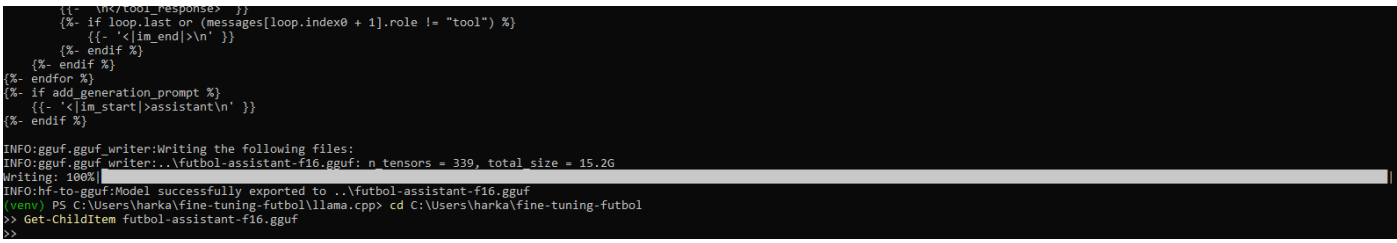
Esta captura inicial muestra los logs finales del proceso de fine-tuning de un modelo Qwen2 usando Unsloth en una NVIDIA GeForce RTX 4070 Ti bajo Windows con Torch 2.5.1 y CUDA 12.1. Cada línea registra métricas clave por paso: 'loss' (pérdida, que mide el error del modelo y baja de 0.0151 a 0.0131, indicando convergencia exitosa), 'grad_norm' (norma del gradiente, que controla la estabilidad del descenso de gradiente, manteniéndose baja ~0.04-0.17), 'learning_rate' (tasa de aprendizaje, decreciendo linealmente de 2.3e-05 a 2.1e-06 para evitar oscilaciones finales) y 'epoch' (avance hasta la época 50.0). Al final, resume train_runtime (95.6s), samples/steps por segundo, train_loss global (0.308) y confirma 50 épocas completadas.



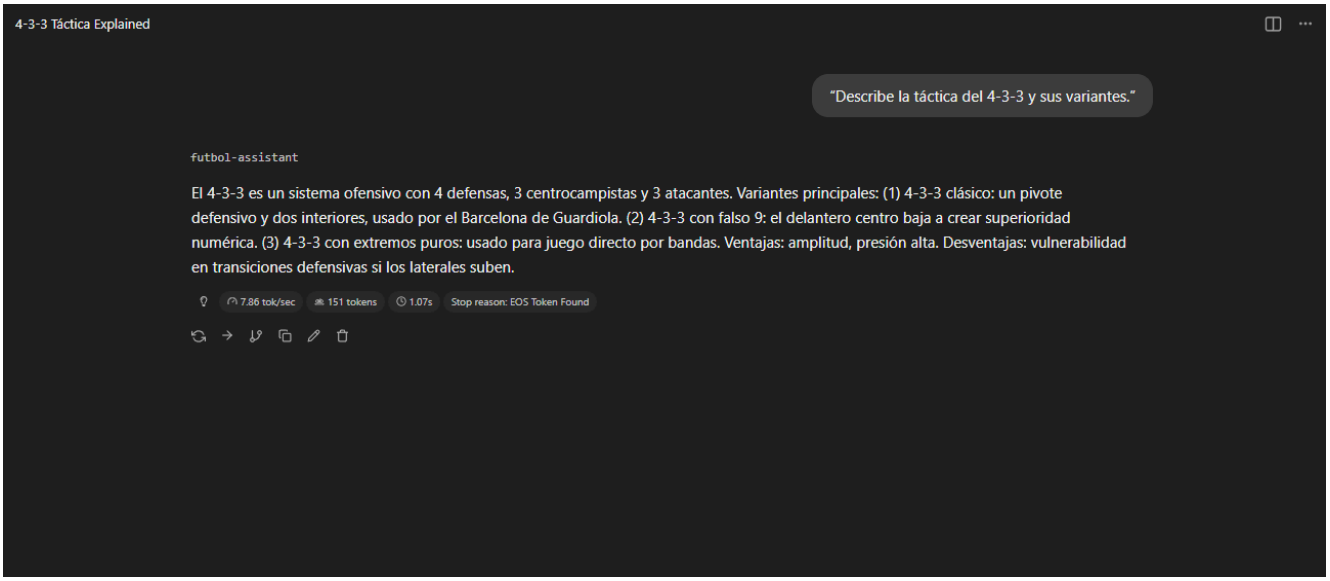
Aquí aparece la barra de progreso al 100% (100/100 steps en 1:35 a 1.05it/s), indicando finalización del entrenamiento supervisado (SFT). Unsloth (versión 2026.2.1) parchea automáticamente 28 capas QKV (Query-Key-Value para atención) y MLP (perceptrones multicapa) del modelo base Qwen2, optimizando para velocidad 2x con 70% menos memoria. Carga shards del checkpoint y habilita bfloat16 para precisión eficiente en GPU (11.99GB usados de 1 GPU).



Esta sección detalla la fusión del adaptador LoRA (Low-Rank Adaptation, que entrena solo ~1% de parámetros para eficiencia) al modelo base en 4-bit quantization, generando `./futbol_model_merged`. La advertencia de PEFT (Parameter-Efficient Fine-Tuning) advierte sobre posibles diferencias en generaciones por errores de redondeo en capas lineales bnb (bitsandbytes). Finaliza con cierre de PowerShell en entorno venv (virtualenv), confirmando guardado listo para inferencia.



La captura final muestra la interfaz de un launcher o playground (probablemente LM Studio u Ollama) con preset "Futbol". El editor de system prompt define "FútbolAI" como un experto en táctica, historia, reglamento, análisis de partidos y estadísticas, con instrucciones para respuestas precisas y apasionadas (51 tokens). Incluye opciones como "Save/Discard", Ctrl+E para editar y barras de carga al 100%, preparando el modelo fusionado para chat interactivo sobre fútbol



Editor (Ctrl+E) del prompt: "Eres FútbolAl, experto en táctica, historia, reglamento, análisis partidos y stats. Respondes precisa y apasionadamente" (51 tokens). Opciones Save/Discard, prepara modelo merged para chats fútbol-specific, integrando conocimiento fine-tuned del datase