```python
output_dir = "Phi3-mini-LoRA-22kPython"
per_device_train_batch_size = 16
gradient_accumulation_steps = 8
optim = "adamw_torch"
save_steps = 10
logging_steps = 2
learning_rate = 5e-5
max_grad_norm = 0.5
max_steps = 100
warmup_ratio = 0.005
lr_scheduler_type = "cosine"

training_arguments = TrainingArguments(
    output_dir=output_dir,
    per_device_train_batch_size=per_device_train_batch_size,
    gradient_accumulation_steps=gradient_accumulation_steps,
    optim=optim,
    save_steps=save_steps,
    logging_steps=logging_steps,
    learning_rate=learning_rate,
    fp16=True,
    max_grad_norm=max_grad_norm,
    max_steps=max_steps,
    warmup_ratio=warmup_ratio,
    group_by_length=True,
    lr_scheduler_type=lr_scheduler_type,
    push_to_hub=True
)
max_seq_length = 512

trainer = SFTTrainer(
    model=peft_model,
    train_dataset=transformed_dataset,
    peft_config=peft_config,
    dataset_text_field="text",
    max_seq_length=max_seq_length,
    tokenizer=tokenizer,
    args=training_arguments,
)
trainer.train()
```