

NLP – Fine tuning(LoRA) Projesi

3. Deneme

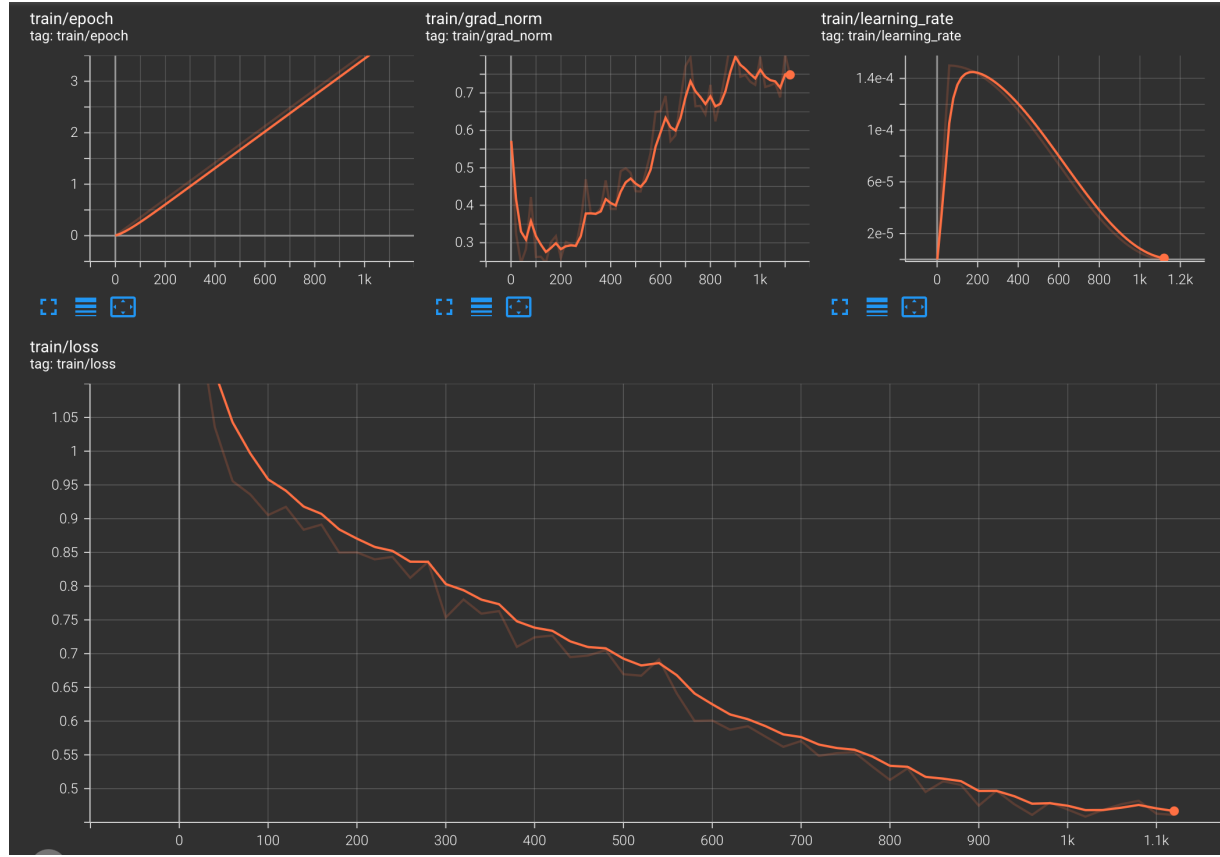
Bu pdf de son yapılan 3. Eğitim parametreleri ve sonuçları yer almaktadır. İlk iki eğitimde yapılan hatalı veri seti bölme(train-test-validation splitting) yapılmadan verisetinin kendi içindeki ayrıma göre eğitimler yapılmıştır. Ayrıca modelin eğitimi colab a100 çipi ile yapılmıştır

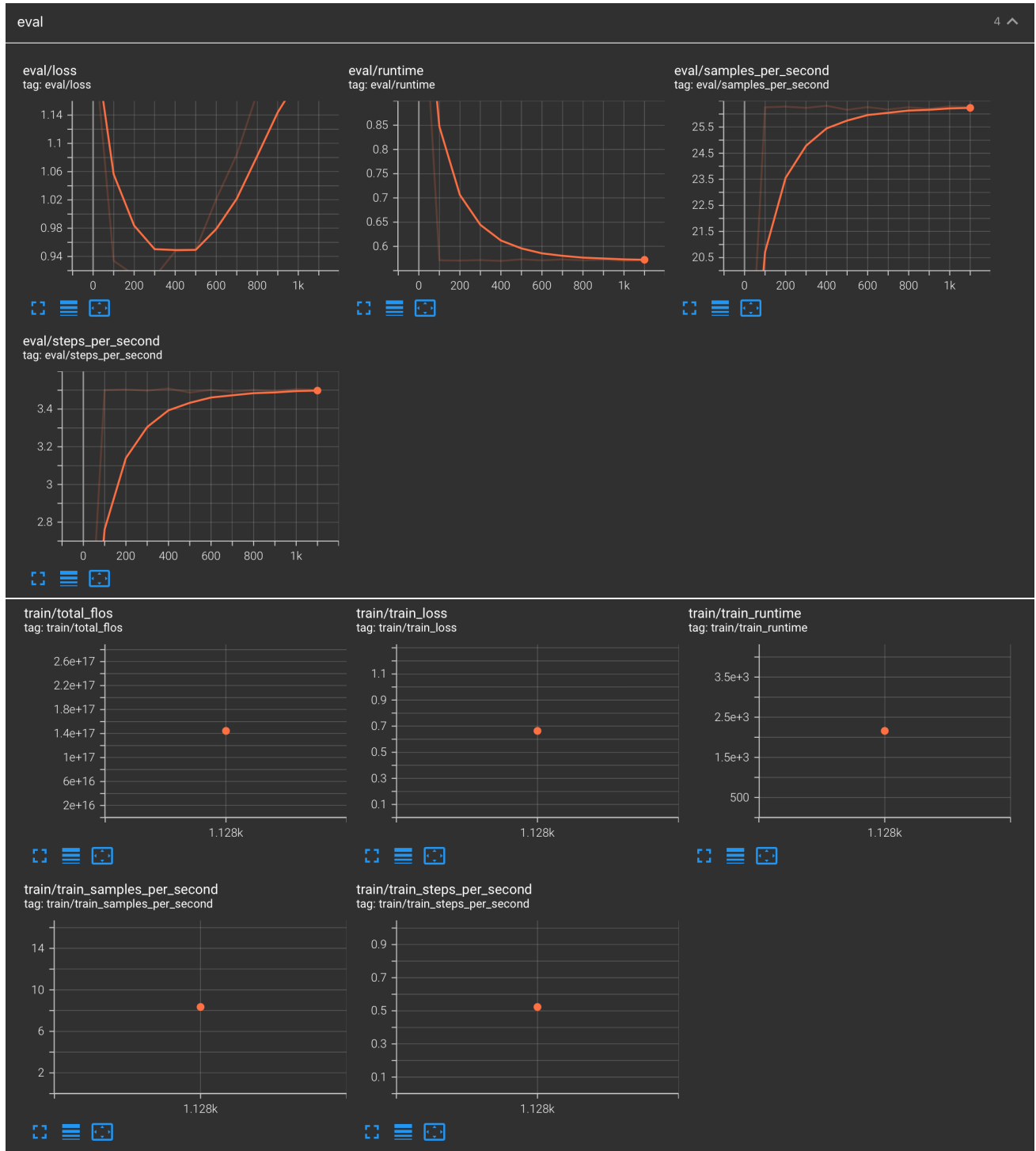
Her eğitim için kullanılan parametreler:

Parametreler	Deep - Diverse
MAX_SEQ_LEN	1024
BATCH_SIZE	8
LR	1.5e-4
NUM_EPOCH	4
WARMUP_RATIO	0.05
LORA_R	32
LORA_ALPHA	64
LORA_DROPOUT	0.08
SAVE_STEPS	200
LOG_STEPS	20
SEED	42

Yapılan LoRA eğitimlerinin sonucunda loglanan değerlerin eğitim boyunca oluşan grafikleri

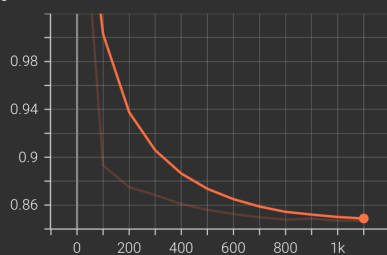
DEEP (aşağıdaki 3 resim);



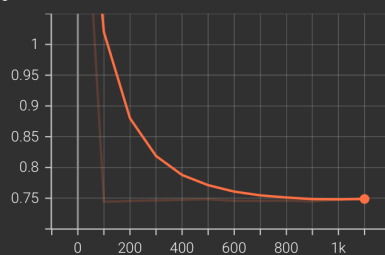


DIVERSE (aşağıdaki 3 resim);

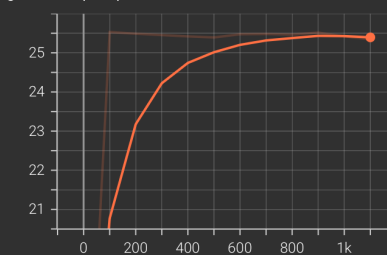
eval/loss
tag: eval/loss



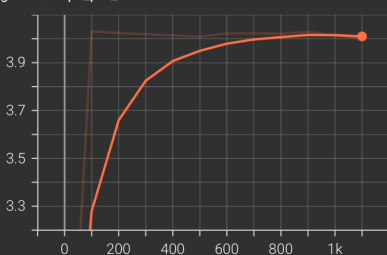
eval/runtime
tag: eval/runtime



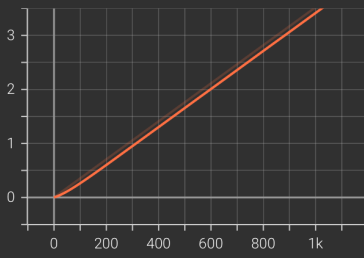
eval/samples_per_second
tag: eval/samples_per_second



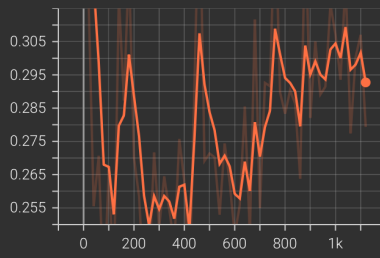
eval/steps_per_second
tag: eval/steps_per_second



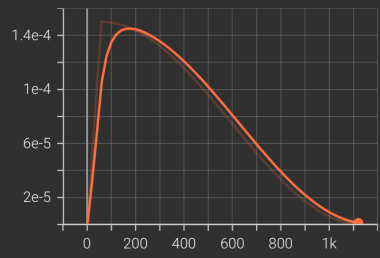
train/epoch
tag: train/epoch



train/grad_norm
tag: train/grad_norm



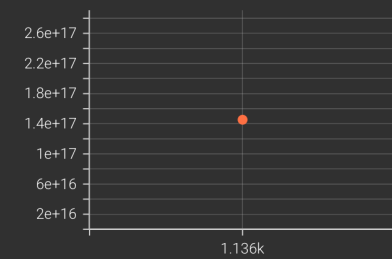
train/learning_rate
tag: train/learning_rate



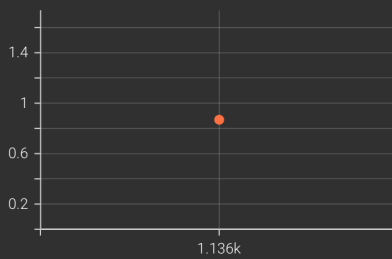
train/loss
tag: train/loss



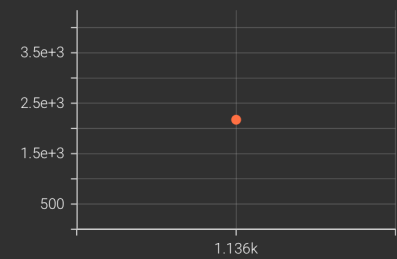
train/total_flos
tag: train/total_flos



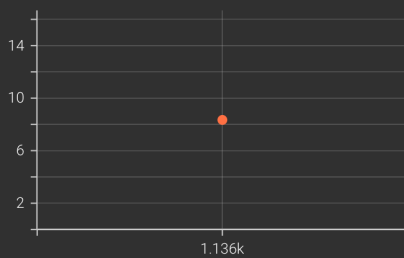
train/train_loss
tag: train/train_loss



train/train_runtime
tag: train/train_runtime



train/train_samples_per_second
tag: train/train_samples_per_second



train/train_steps_per_second
tag: train/train_steps_per_second

