

Отчет по TTS (HW4)

Барановская Дарья

https://wandb.ai/darya_baranovskaya/hw4_hifi/runs/3l8skqsl?workspace=user-darya_baranovskaya

(Старт тренировки в логе https://wandb.ai/darya_baranovskaya/hw4_hifi/runs/1o8kn6g4?workspace=user-darya_baranovskaya)

<https://drive.google.com/drive/folders/1Z4UI20dvERE0jcPo4UdbbnQyUsVqempN?usp=sharing>

Для генерации wav мной была взята модель HiFi V2. V1 оказалось слишком большой, поэтому использовала вторую версию.

Переобучение батча

Батч переобучился очень быстро, всего за 10к итераций до отличного аудио. При чем оказалось, что, если не инициализировать веса в начале нормальным распределением, то лось генератора падает быстрее.

Эксперименты

Я проводила эксперименты с некоторыми гиперпараметрами стриминга. Так я меняла `lambda_wav` и `lambda_mel`. При равных значениях `lambda` обучение кажется идет не так хорошо, как при значении из статьи - 45, 2. Я пренебрегла шедулером, однако, скорее всего я сделала это зря. Также мне снова показалось, что без инициализации лоссы падают хуже.

Копипаста кода - плохо. Так я написала код шага дискриминатора и скопировала его в код шага оптимизатора. ИИИИ. Забыла убрать `detach` с прохода генератора вперед. Так что Ган не учился, лосс дискриминатора сразу ушел почти в 0, поэтому выявить ошибку оказалось не сложны. Из интереса я также пробовала шагать дискриминатором не на каждом батче, а раз в несколько батчей, ведь обычно дискриминатор быстро выигрывает. Но в данном случае дискриминатор достаточно сложно организован, поэтому в таком шагании не было необходимости.