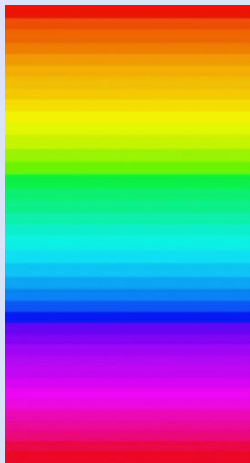
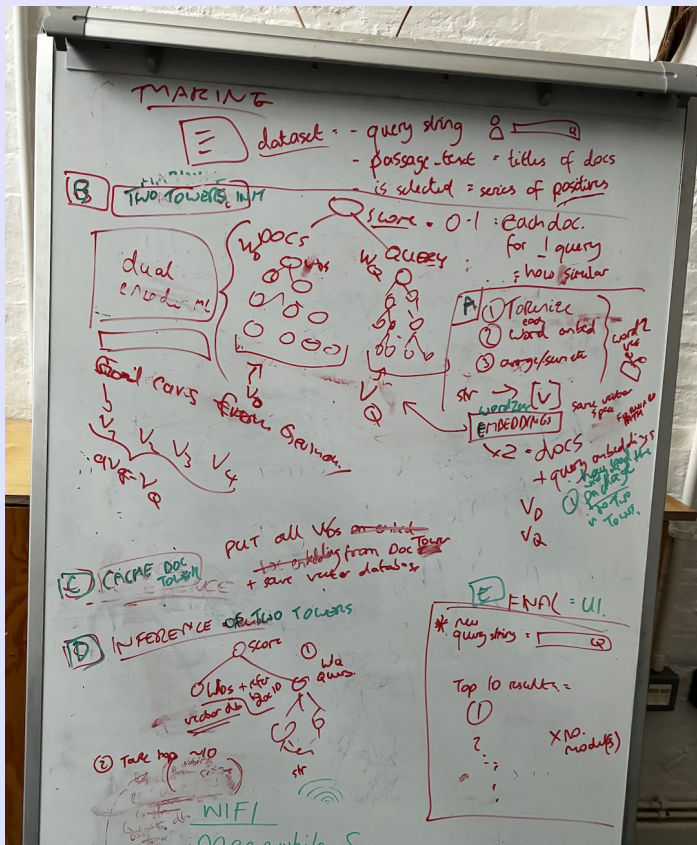


# Gradient Gigglers

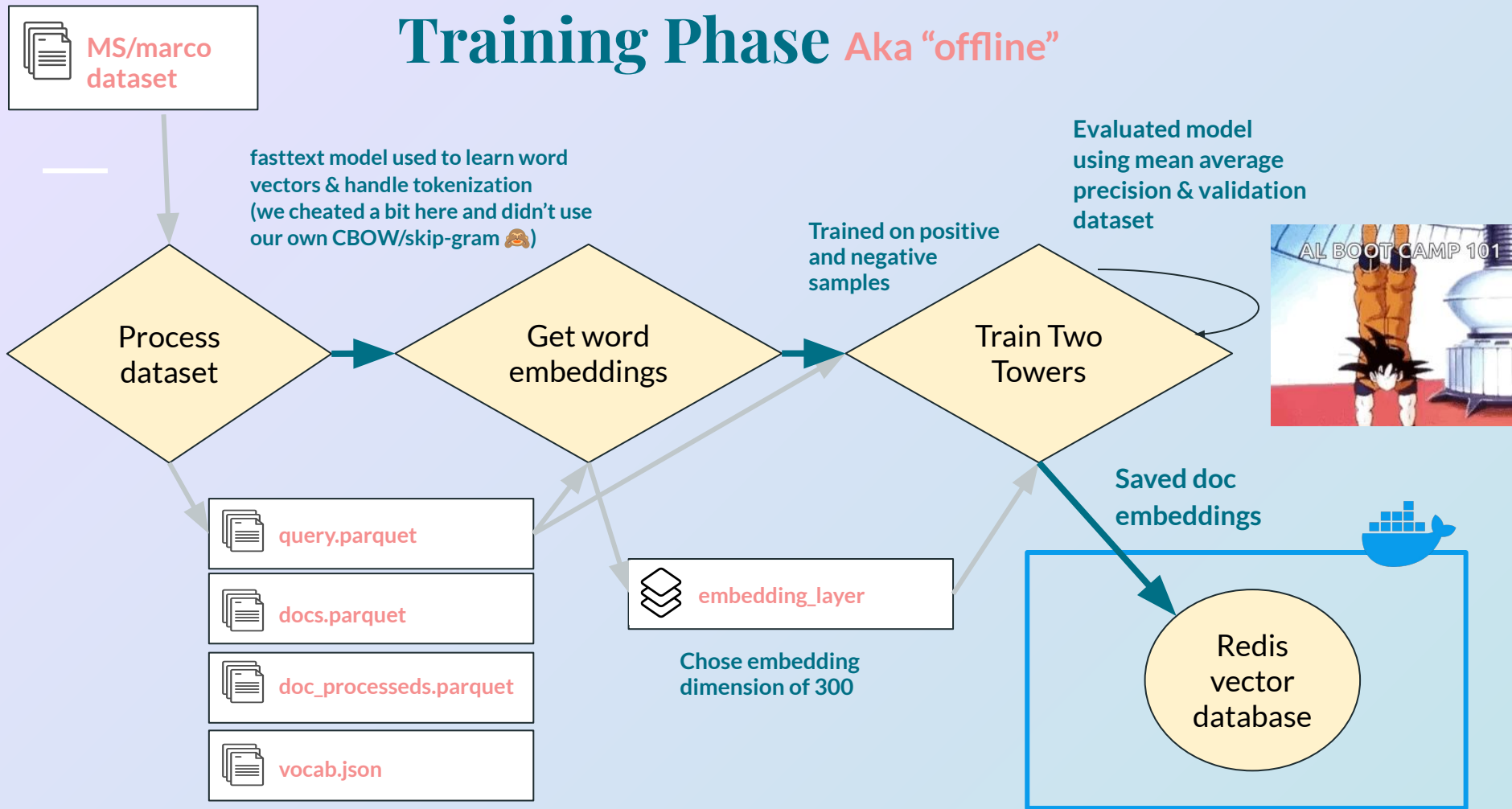
vs Two Towers Task. MLX week 2



# The Start



# Training Phase Aka “offline”



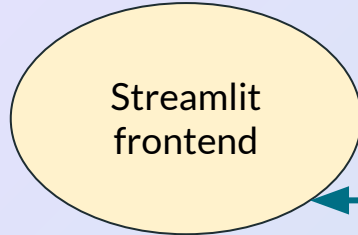
# Inference Phase Aka “online”



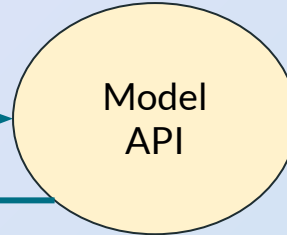
Pre trained model



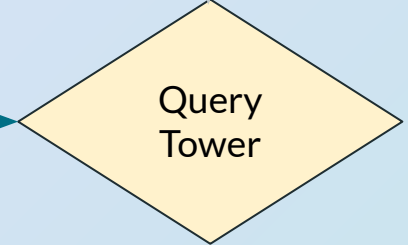
embedding\_layer



Calls /search  
with query

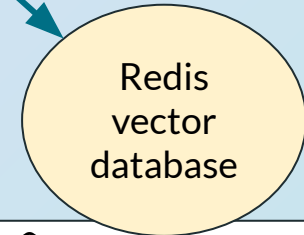


Gets query  
vector



Returns top docs

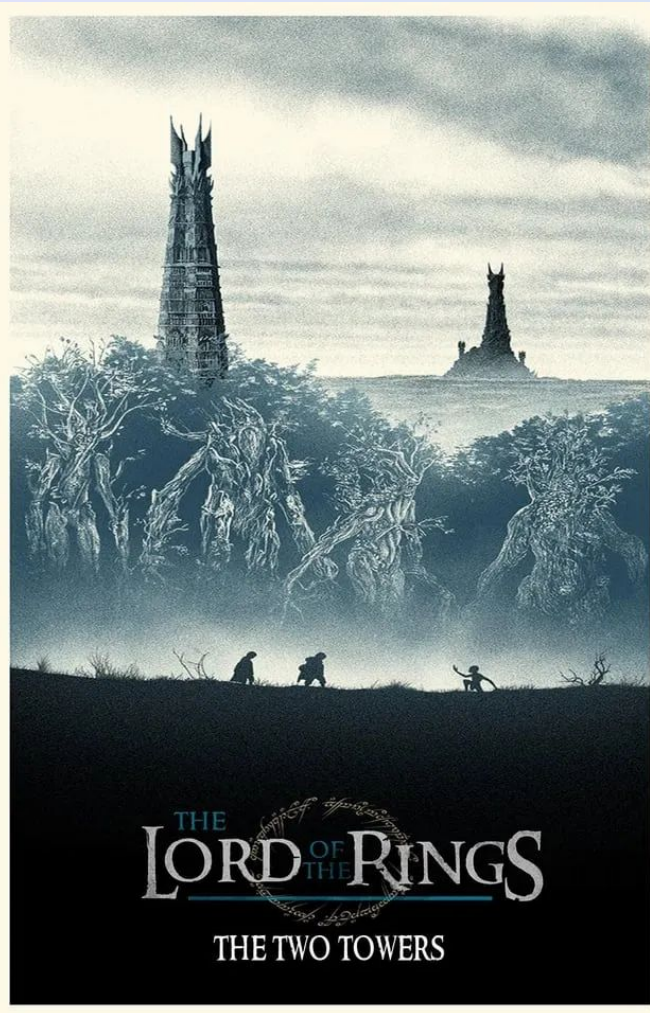
Uses query  
vector to find  
top 5 docs



Has doc embeddings



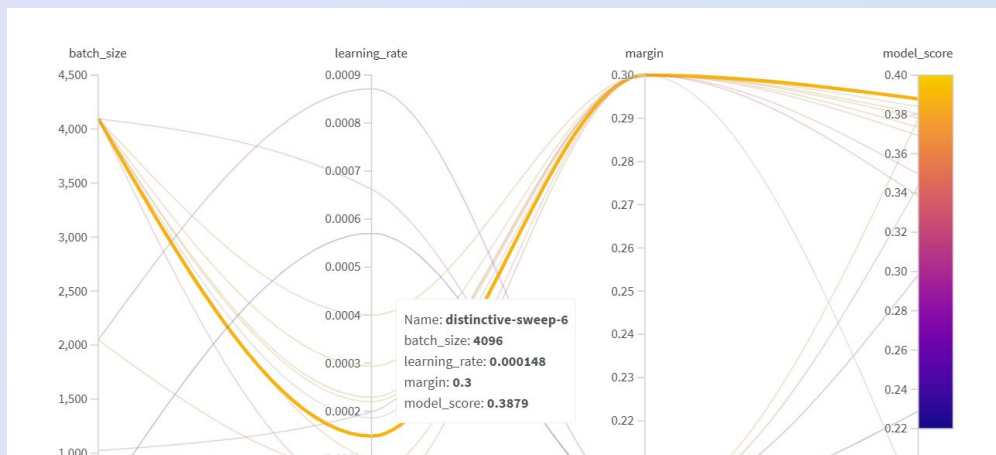
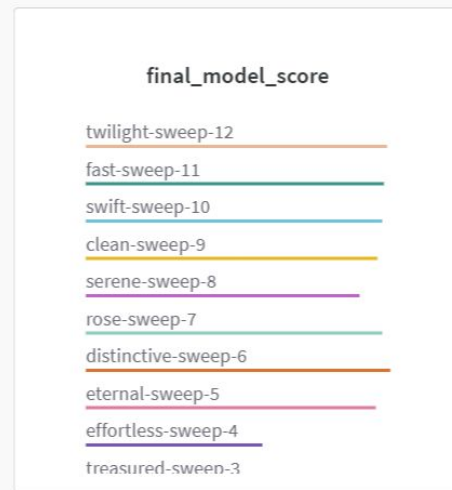
It's ...  
DEMO  
TIME!





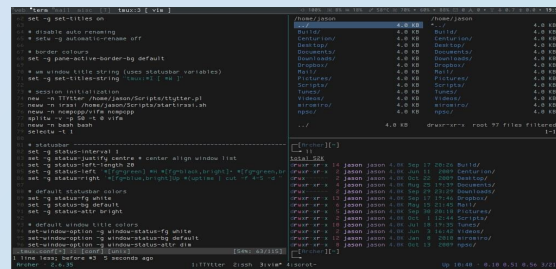
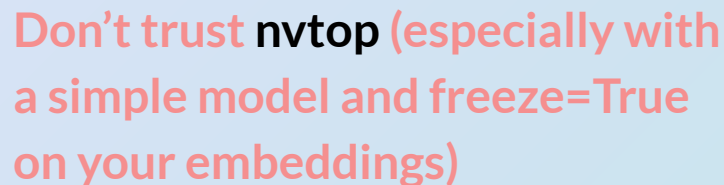
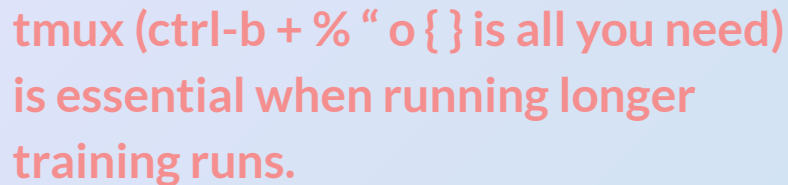
# Wunderbar (wandb) Sweeps

Distinctive-sweep  
6 won the award  
for best final  
model score 🏆 ...



...to give us the  
one set of  
params to rule  
them all 🏆

# You never have enough Memory



# Learnings

---

M series CPUs are very very fast  
especially at smaller batch sizes



Sweeps over smaller amounts of  
data in Wandb give lots of info on  
potential hyperparameters



# More Learnings

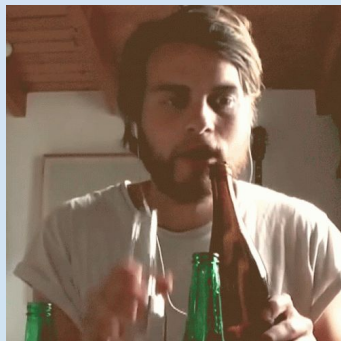


EmbeddingBag was worth learning  
and a useful tool for the future

Evaluating your results with real test  
data is vital to improving the model

Test test test test test test etc.

Debug debug debug



Dataloader  
probably isn't  
the bottle neck

# What went well

---

- Sensational Team (even the README is clear and high quality - Thanks Helen)
- Whiteboard to start and make everyone on the same page
- Being patient and believing everything would come together helped.
- “Pairing” with everyone on the team allowed people to have a core understanding and branch off as needed.



# Improvements

---

- Build scripts, frameworks and set of utils for data, wandb, servers, etc.
- Small samples of the data at the start gets you testing things faster, usually it's enough to get a direction
- Improving at seeing through the data and AI learnings and getting results
- Don't fixate on performance, it'll probably be okay as you leverage more complex models, larger data and longer training runs
- `collate_fn` might be a bottleneck so try creating the full tensor on GPU right away.

# To finish off on... 😂



Enter your search query:



What is **page** rank?

Search

Top matching documents found:

## Rank 1

CDC Fact Sheets. Use the links below to visit the STD fact sheet web pages in English and Spanish. Print versions of these fact sheets are available on the web pages. If you'd like to include this content on your web site with automatic updates, visit the Sexually Transmitted Diseases (STDs) Syndicated Content page.



Similarity Score: 0.7776