



PETMATCH



Denise Blady, Matthew Robinson, Theresa Thoraldson

PROBLEM

- Humans often choose the wrong pet (costs ~\$2 billion USD)
 - Post-COVID animal surrenders are overburdening shelters
 - Unwanted pets are often abandoned or euthanized (355K in 2023) => negative mental health impacts
- Lack of tools to choose the best pet to suit their lives
 - Breed characteristics not taken into consideration
 - Implementation of user-preferences may not allow for open-mindedness



SOLUTION

- Provide users personalized, curated pet recommendations
 - Learns from users' liked/disliked pets, pet similarities, and generalized breed characteristics
- Provide on-demand recommendations via Tinder-like UI
 - Using favorite mobile device, find your perfect pet quickly
 - On-demand adoption info for chosen pet(s)
- Success == Increased adoption rates and lower return rates (Hit rate)

120 Billion dollar market for pet ownership!



[CONTENT-BASED FILTERING] PET-BASED SIMILARITY V2

Data

- 46710 cats, 15 features
- 57000 dogs, 20 features

Final Beta Models

- Cosine Similarity - Cat
- Cosine Similarity - Dog

Metrics

- Offline - Cosine Similarity score (1 is best, 0 is worst)

Version 2 Results

	Offline Metric Score
Cat	.922
Dog	.918

Outcomes

- Both cat and dog models improved with more data
- Addresses cold-start issue of collaborative filtering models

[COLLABORATIVE-BASED FILTERING] USER-BASED SIMILARITY V2

Data

- Cats - 433 rankings
- Dogs - 360 rankings

Final Beta Models

- SVD++ - Cat
- SVD++ - Dog

Metrics

- Offline- RMSE score (lower=better)

Validation

- 10-fold Cross Validation (used as proxy to holdout test set)

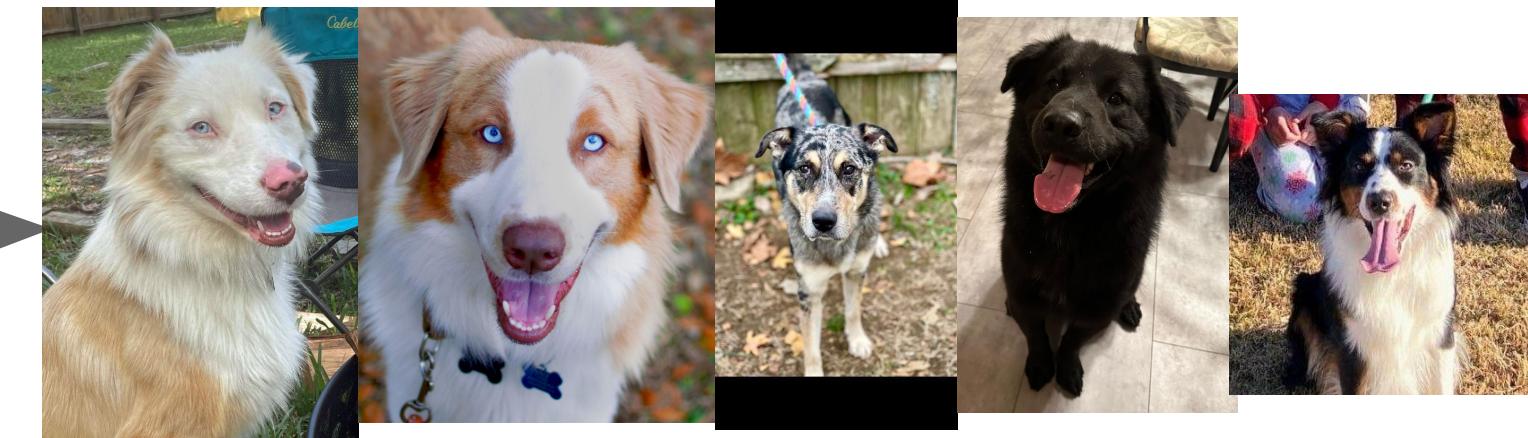
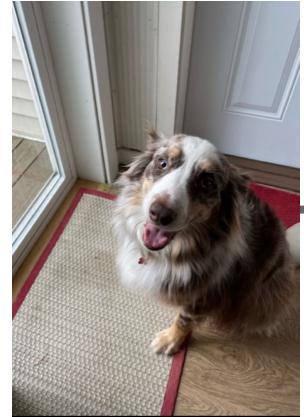
Version 2 Results

	Offline Metric Score
Cat	.342
Dog	.390

Outcomes

- Given target value range (0-1), RMSE needs some improvement
- Hyperparameter tuning was not enough, need more data

DEMO V2 - COSINE SIMILARITY EXAMPLE RESULTS



DEMO V2 - COLLABORATIVE FILTERING EXAMPLE RESULTS

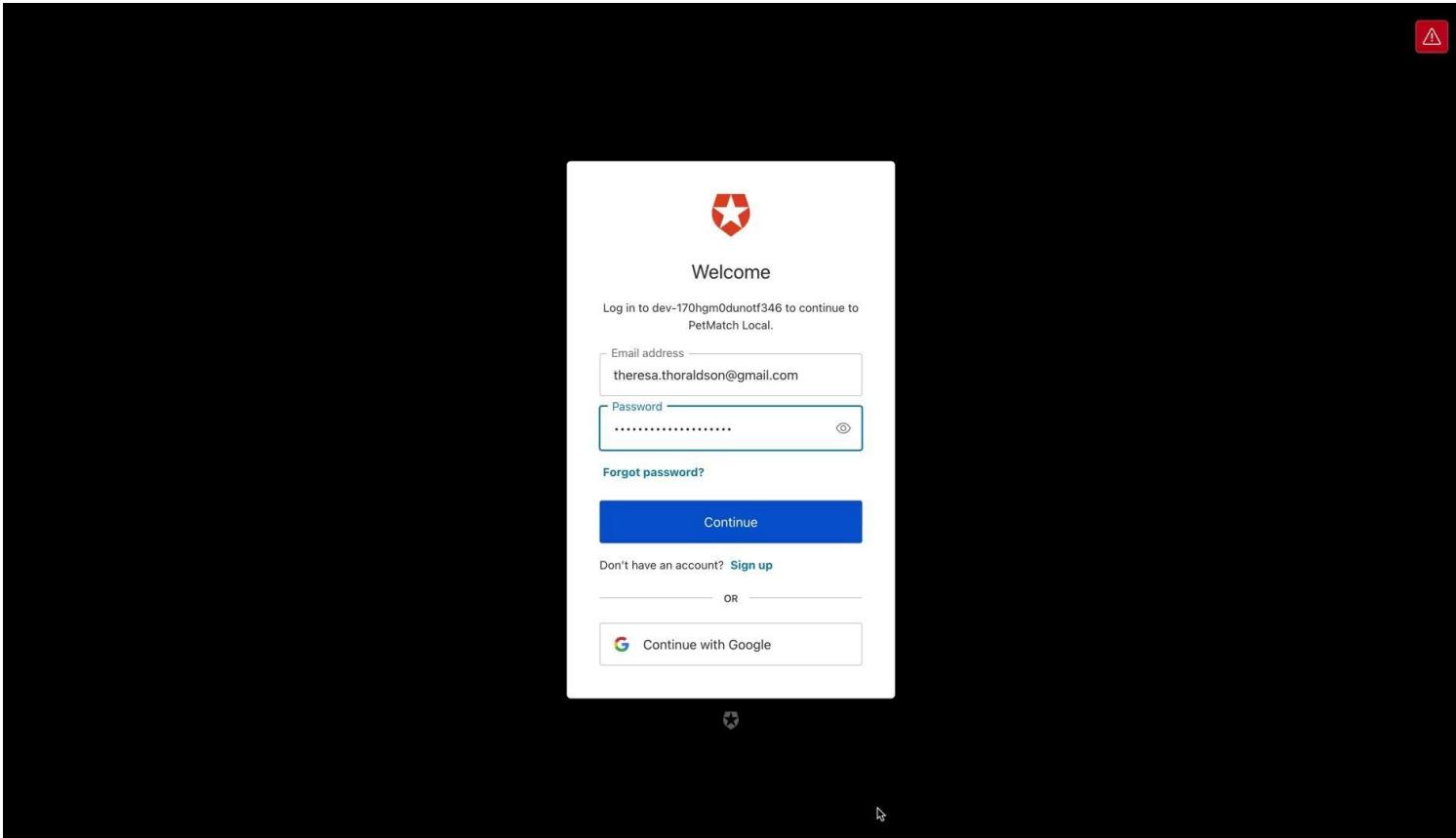
“Denise”
,”dog”



“Denise”
,”cat”



DEMO



[Video Link](#)

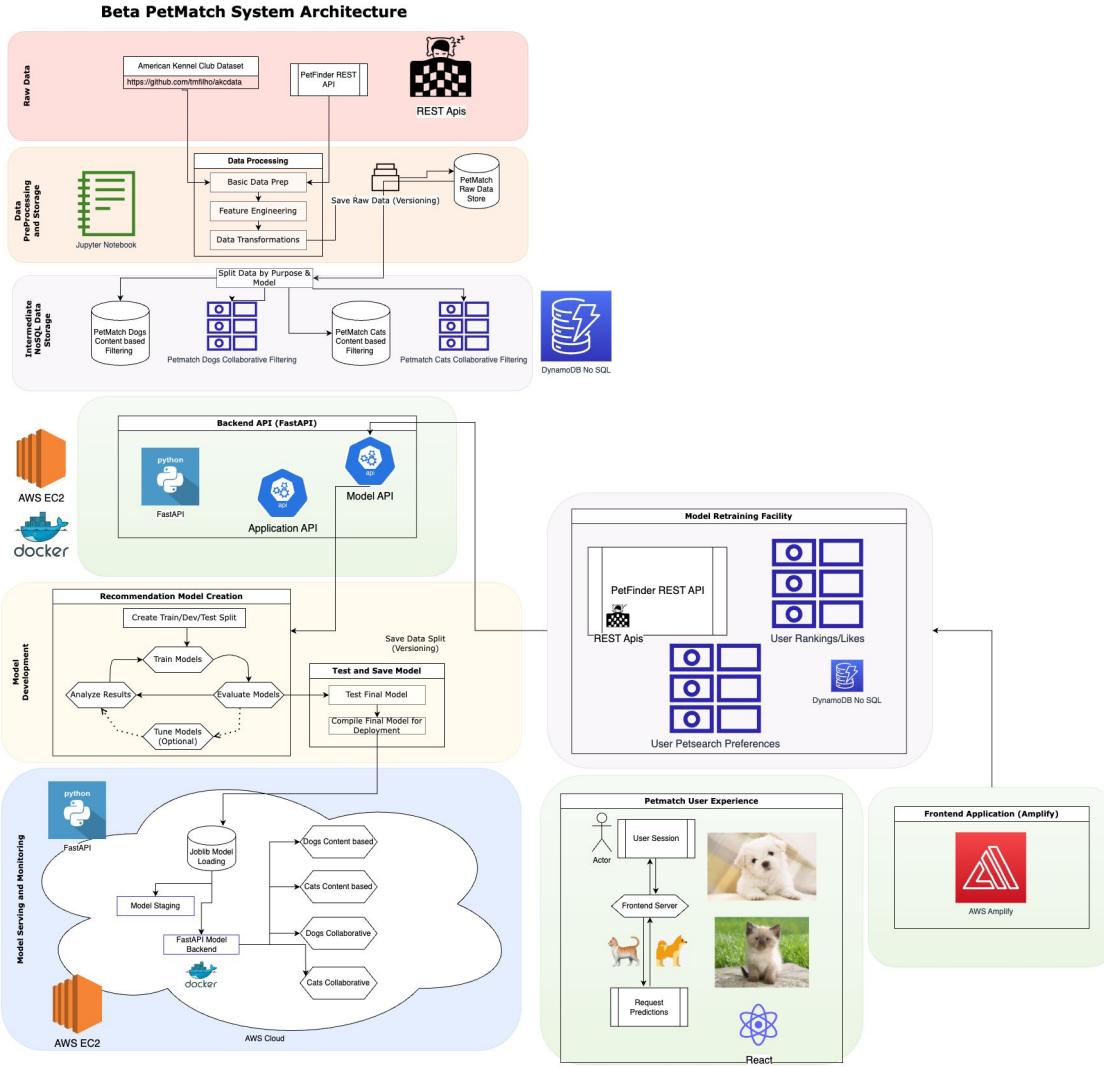
SYSTEM ARCHITECTURE

● Overview

- Two-model recommendation system
- AWS EC2 model serving
- Basic model retrain pipeline

● Key Decisions

- Content-based filtering to solve ‘cold-start’ problem
- Collaborative filtering for long term recommendations
- Animal-type specific models
- Improved models via breed data
- User preferences for initial served pets
- Non-relational database



LESSONS LEARNED & CONCLUSION

- Pet-based models (**solves cold-start issue**) and user-based models (**encourages user open-mindedness**) were an effective combo
- Hyperparameter tuning not enough to get optimal results for user-based models, need more data
- Pet-based similarity models can take up a LOT of ram
- Cosine Similarity better at differentiating pets, compared to similar models
- Collecting data from scratch is hard work
- Full stack solutions are hard but more customization

Personalized pet recommendations will fully unlock a 120 Billion dollar market!

FUTURE WORK

- System Architecture/UI
 - Leverage SageMaker for model lifecycle
 - Implement a ‘match’ page where users view pets they’ve liked
- Metrics
 - Implement Hit Rate online model metric
- Models
 - Expand cosine similarity to more examples without running out of RAM
 - Incorporate distance from pet to user
- Data
 - Fix data pull to get full animal description
 - Get more data for Collaborative Filtering Models
 - Implement time-sensitive recommendations

QUESTIONS?



Beta Waitlist Form



Github link: <https://github.com/tthoraldson/PetMatch>

Become a beta user: <https://forms.gle/dKBGbQ58DyxTX1p19>

APPENDIX

MARKET SIZE

Pets Industry (2021) Breakdown:

Pet Food & Treats

\$50.0 billion

Supplies, Live Animals & OTC Medicine

\$29.8 billion

Vet Care & Product Sales

\$34.3 billion

Other Services

\$9.5 billion

(Other Services include boarding, grooming, insurance, training, pet sitting and walking and all services outside of veterinary care)

COST FOR UNWANTED PETS

- Government-supported animal shelters cost Americans almost \$2 billion every year.
- About 625K euthanizations as of 2019 and average euthanization cost of \$50 is ~\$31 million dollars

Total cost of ~2 billion dollars. (Src: thezebra.com)

CAT CONTENT-BASED MODEL RESULTS AND FINDINGS VI

- Metric: Average Similarity score of all examples for top 5 results
 - Linear Kernel has a maximum value of 0-15, where 15 is best
 - Cosine Similarity and Laplacian Kernel have range 0-1, where 1 is best
- Outcomes
 - Cosine Similarity is better at differentiating one cat from another

Model Name	Score
Linear Kernel	13.7663
Cosine Similarity	0.917753
Laplacian Kernel	0.999386



DOG CONTENT-BASED MODEL RESULTS AND FINDINGS V1

- Metric: Average Similarity score of all examples for top 5 results
 - Linear Kernel has a maximum value of 0-15, where 20 is best
 - Cosine Similarity and Laplacian Kernel have range 0-1, where 1 is best
- Outcomes
 - Cosine Similarity is better at differentiating one dog from another

Model Name	Score
Linear Kernel	18.2681
Cosine Similarity	0.913404
Laplacian Kernel	0.999454



[COLLABORATIVE-BASED FILTERING] USER-BASED SIMILARITY VI

Data

- Cats - 433 rankings
- Dogs - 360 rankings

Initial Models

- Models: SVD, NMF, PMF
- Animal type: Cat, Dog

Metrics

- Offline- RMSE score (lower=better)

Validation

- 10-fold Cross Validation

Version 2 Results

	SVD	NMF	PMF
Cat	.342	.376	.483
Dog	.391	.402	.407

Outcomes

- Hyperparameter tuning can make a big difference
- Use SVD models as base model
- Given target value range (0-1), RMSE needs some improvement

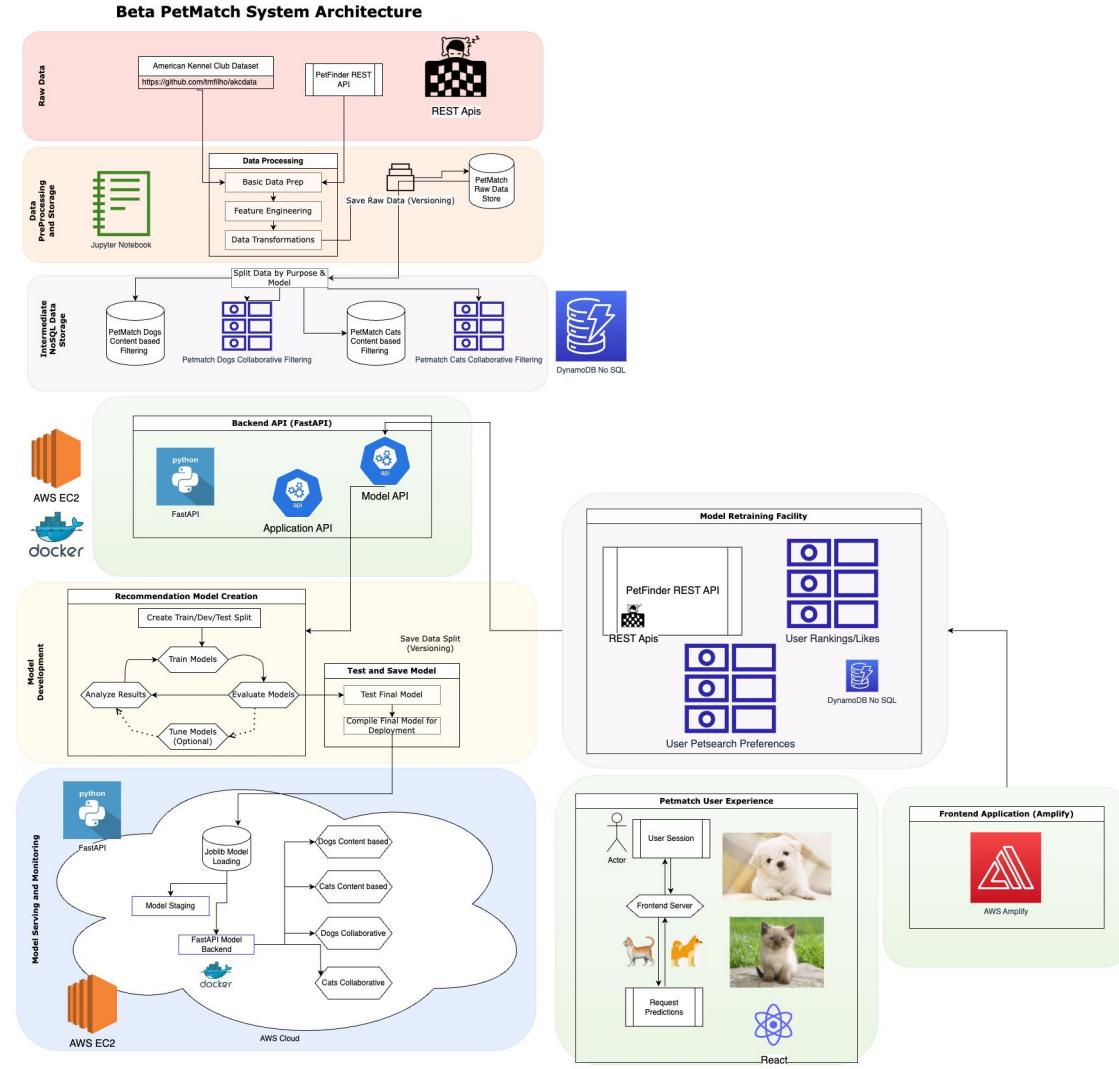
SYSTEM ARCHITECTURE

● Overview

- Non-relational database
- Mobile Responsive FE
- WSGI API in Python
FastAPI

● Key Decisions

- Containers supported rapid development and CI/CD
- Separation of concerns
- Auth0 supporting Authentication
- DynamoDB supports scaling of interactions on the app
- Models are lightweight
- FastAPI supports parallel processing, data validation, both sync/async paradigms, business logic, and model operations



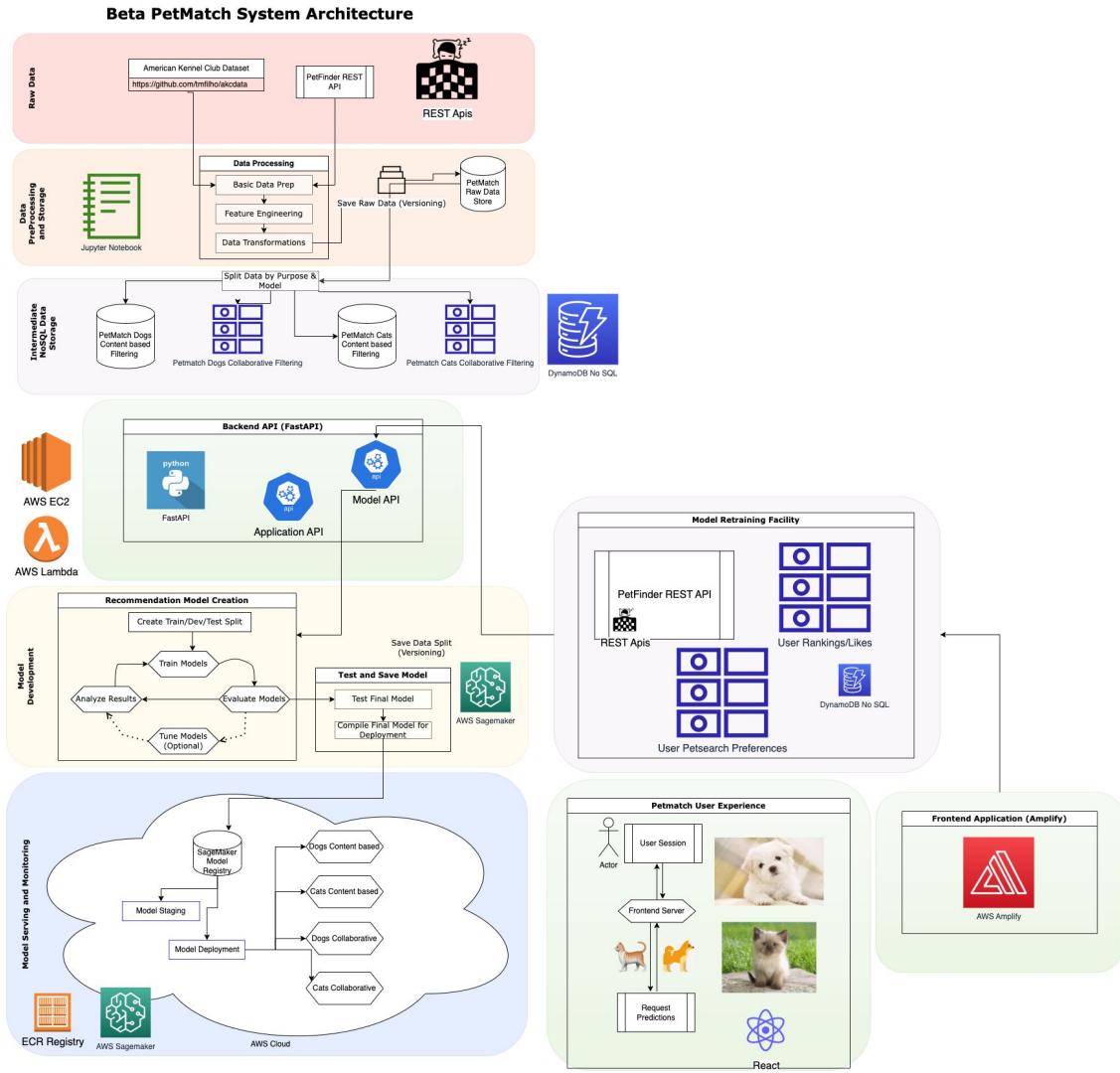
IMPROVED MLE APP STACK

● Overview

- Non-relational database
- Mobile Responsive FE
- Serverless/On-demand compute
- Scalable with some microservices
- Feature flexibility

● Key Production Changes

- Auth0 supporting Authentication and user data
- DynamoDB supports scaling of interactions on the app
- Transitioning to a model serving platform will allow inference, training, and MLops to be responsively scalable as well
- We can expand data sources and data sizes for pets, users, and model features
- FastAPI supports parallel processing, backgrounding, data validation, both sync/async paradigms, business logic, and model operations



EXTRA LESSONS LEARNED

- If you have two identical recommendation sets, Laplacian Kernel model assigns a much higher score than Cosine Similarity model => Latter is more discriminative
- Cosine Similarity is better at differentiating one pet from another
- User rankings data requires more pre-processing than initially expected
- Content-based Filtering doesn't seem to be included in mainstream modeling libraries, required code from scratch
- Sagemaker doesn't like .pkl files for models
- DynamoDB tables break dataframe mindset => must query at scale
- Better to implement core UI functionality in Streamlit before moving to more complex full React app



FUTURE WORK- EXTRAS

- Future Metrics once collab filtering more mature
- Personalization score (higher is better)
- MAP@K (relevancy of reccs) and MAR@K (recall ground truth 1s)
- Petfinder data pulls are rate limiting
- Leverage Serverless/On-demand compute

