Project Title: Chihuahua vs Muffin Classification Hackathon
Team Name: The Muffins
Team Members: Lavanya Bansal, Jonas Banks, Riya Yadav
Date: 11/20/25

---

1. Methodology

1. We began by setting up our environments according to the starter note book.

2. We then used multiple different methodologies to label our data in the 3LC dashboard.
    a. Lavanya filtered by predicted labels. She deleted ambiguous data, and also filtered by high confidence and accepting and labeling the data the model was highly confident and correct on.
    b. Jonas looked at the model embeddings and filtered by predicted labels and low loss to find shapes in the embeddings. He accepted the predictions in areas where the predictions were clustered closely together or created “lines” in the embeddings.
    c. Riya filtered the labels by their confidence level.She altered the labels and even deleted them if the confidence level was below 0.8.The 3D chart and the lasso tool was really helpful to her in finding the unlabeled and less confident labels and selecting them all at once made it an easier task.

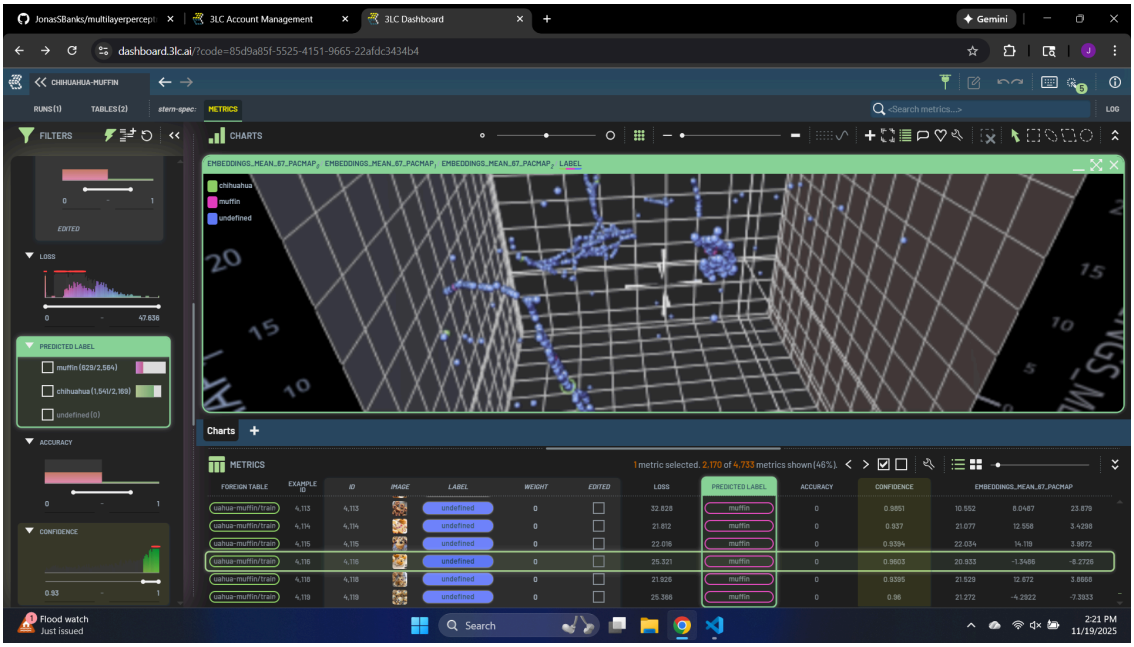
3. We then reran our models adding more labels in batches each time.

4. We finally tuned our hyper parameters to achieve higher accuracy.

Tools Used:
Our team utilized the train.py file and built in shortcuts in the 3LC dashboard such as the accept prediction in this row short cut and selecting multiple images with shift and right click.

2. Data-Centric Improvements Using 3LC

- Embedding Analysis:
    - We used the 3LC embedding visualization to identify the unlabeled images with high confidence in clustered areas, and then analyzed if these clusters were distinctly muffins or chihuahuas.



(embeddings filtered by confidence)

- Outlier/Ambiguity Detection:

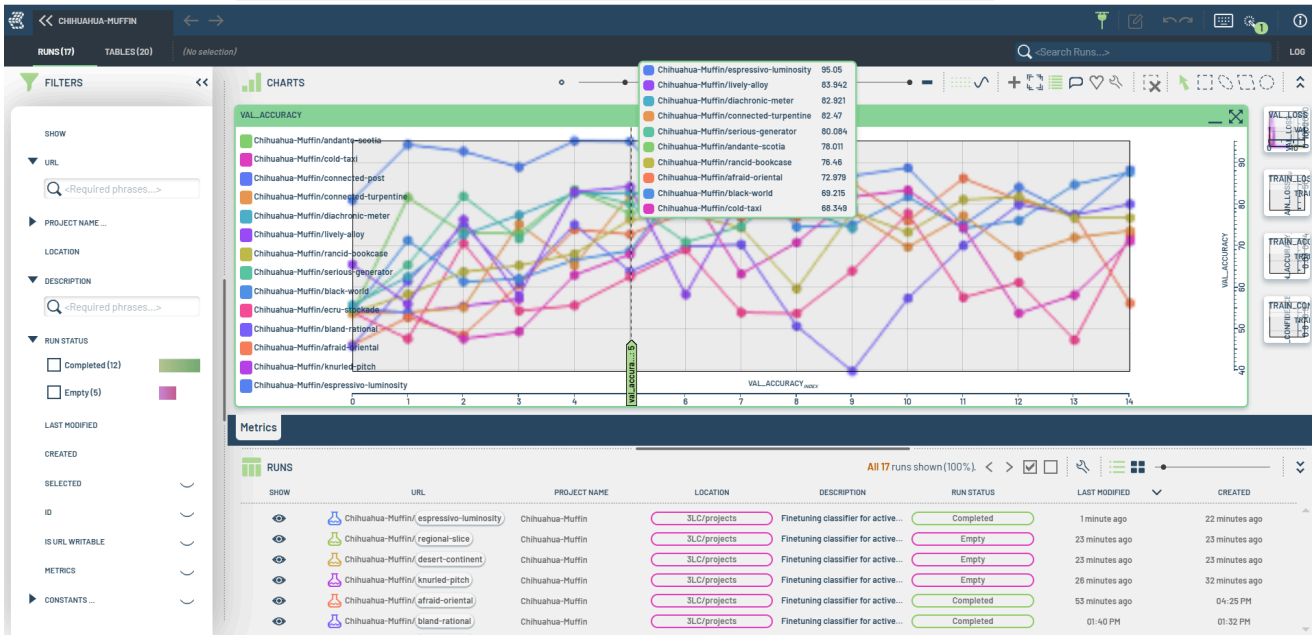
  Some images in the data set were ambiguous. There were samples with both muffins and chihuahuas in the image along with samples that did not contain either a chihuahua or a muffin. These samples typically had higher loss though we did not specifically search for these samples.

3. Iteration Details (v1 → v2 → v3 …)

Iteration	Description of Changes	Labeled Samples Added/Removed	Accuracy (Test Set)
v0	Baseline (200 images)	+0	85.135%
v1	Labeled ~250 samples via embedding selection	+250	84.478%
v2	Labeled ~2500 samples via embedding selection	+2500	86.242%
v3	Fine tuning of hyper parameters	+0	89.33%

- Final Accuracy:

  Our final model (v3) achieved an accuracy of 95.05% on the private test set.



5. Challenges & Learnings

➢ Challenges:

- The primary challenge was the similarity between the muffin images that often resembled dog eyes/noses.
- Setting up the 3LC database path was tricky at first, ensuring all team members were at the same page.
- In the pool of unlabelled samples,there were limited labelled samples to assign

➢ Learnings:

- We learned how to analyze data based on the loss,confidence level,accuracy and how it can affect the overall result
- Cleaning ambiguous data provided us with a greater accuracy boost.
- We learned how to use 3LC to improve datasets,and the behaviour of the model
- Building AI systems with limited labeled data

6. Future Improvements
    - We would have liked to spend more time setting up our git repository and making sure we could all work on the same 3LC tables together instead labeling the data independently.
    - We would have liked to tune the hyper parameters for more intensive training instead of compromising for the sake of time.