Teachable Machine Classifier Project Summaries

Recycling Detective

Description

For this project, I created a simple image-classification model using Google Teachable Machine. The project Recycling Detective was designed to explore how a simple machine learning classifier could support children in identifying and sorting common objects while also providing an accessible introduction to how machine learning systems learn from data.

The model included three classes: “Recycling (Water Bottle)”, “Trash (Used tissue)”, and “Blank (Nothing)”. The intended context of use was a recycling station in classroom, cafeteria, or school hallway. In this setting, a child could hold an object in front of a camera and receive immediate visual feedback indicating how the model classified the object.

My initial plan was to create only two classes: one for used water bottles and one for used tissue. During the development process, however, I realized that classifier would continue assigning an input to one of these categories even when no object was visible. I therefore added a third class, “blank (nothing)” to represent the absences of an object.

This modification became particularly significant because the “blank (nothing)” class was the most difficult to classify with consistently high confidence. All training images were collected in front of the same white wall. As a result, the white background appeared not only in “blank (nothing)” class but also in the training images for the bottle and tissue classes. When the camera was later directed only toward the white wall, the model generally predicted “blank (nothing)” with approximately 89% confidence rather than near 100%. I inferred that the repeated presence of the white wall across all three classes may have influenced the features the model associated with each category.

Process

The primary challenge during the project was determining how the composition of the training images affected model performance. During the first training attempt, I positioned the objects relatively far from the camera. Consequently, the objects occupied only a small proportion of each image, while the white wall occupied a comparatively large area.

After training the initial model, the predictions were unstable and did not consistently correspond with the object being presented. At first, I assumed that increasing the number of training images would necessarily improve the model. However, adding more examples did not automatically result in more accurate predictions.

I therefore deleted the original training images and reconstructed the dataset. During the second and third training attempts, I positioned the water bottle and used tissue closer to the camera so that the target object occupied a greater proportion of the image. I also attempted to provide more varied examples of each object. These changes improved the stability of the model’s predictions.

The iterative process led me to recognize an important distinction between the features that I intended the model to learn and the features were actually available in the training data. My intended categories were based on object identity, water bottle, used tissue, or no object. However, each image also included additional visual information, such as the background, lighting conditions, the position of my hand, the distance between the object and the camera, and the orientation of the object. The classifier could potentially use any of these recurring visual patterns when distinguishing among the classes.

I trained each class approximately three times in an effort to improve model performance. Although classification of the bottle and paper became relatively reliable, the “blank (nothing)” class remained less consistent. Rather than treating this as simply a technical failure, I found it useful for understanding the dependence of machine learning systems on the characteristics of their training data.

The main resources I used were the Teachable Machine interface, the in-class demonstration, and the iterative feedback provided directly by the model. Because the tool immediately displayed its prediction confidence, I was able to repeatedly train, test, identify problems, revise the dataset, and retrain the classifier. The design process therefore became an ongoing cycle of experimentation rather than a linear process in which a model was trained only once.

Reflection on Learning

The most important insight I gained from this project was that the quality and diversity of training data are at least as important as the quantity of data. Prior to completing the activity, I understood conceptually that machine learning systems depend on data. However, the relationship became much more concrete when I observed how seemingly minor design decisions, such as distance from the camera or the amount of white background visible, affected the classifier's performance.

This experience connects directly with Tedre, Denning, and Toivonen 2021 discussion of CT 2.0. The authors distinguish machine-learning-based problem solving from traditional programming by emphasizing a workflow that includes collecting data, filtering and cleaning data, labeling data, training a model, and evaluating its performance. They further argue that machine learning systems can be strongly context-dependent because their behavior is shaped by the data from which they are trained. My project reflected this principle on a small scale. Because most of the training images were collected under similar environmental conditions, the resulting classifier may have learned features that were specific to that context rather than only features that characterized the intended object categories.

The project also changed how I interpreted prediction confidence. Initially, I tended to view a high confidence score as evidence that the model had correctly recognized the object. Through repeated testing, I learned that high confidence does not necessarily indicate that a prediction is conceptually correct. Rather, the confidence score reflects the model's classification based on the patterns it has learned from the available training data. This distinction is important because a model can produce a highly confident prediction even when its reasoning is based on unintended or context-specific features.

This observation also aligns with Touretzky, Gardner-McCune, and Seehorn 2023 discussion of machine learning experiences for K–12 learners. The authors argue that students should not stop after successfully training a classifier. Instead, they should test whether the model continues to perform when conditions change. In their Teachable Machine example, students test a gesture classifier using different hands, orientations, and clothing in order to reveal the model's limited ability to generalize. My own experience with the white-wall background demonstrated a similar issue: a model may perform well under familiar training conditions while becoming less reliable when the relationship between object and context changes.

If I were to teach this activity in the future, I would recommend that students deliberately vary the conditions under which training data are collected. For example, they could photograph the same object against multiple backgrounds, under different lighting conditions, at different distances, and from different orientations. I would also encourage students to test their models in conditions that differ from the training environment.

Overall, the project helped me understand that building a machine learning classifier involves more than providing examples and pressing a training button. The model did not independently learn the abstract concepts of “recycling” or “trash.” Instead, it learned statistical patterns from the particular examples I provided. My own decisions therefore influenced what the system was capable of recognizing.

Reference

Tedre, Matti, Peter Denning, and Tapani Toivonen. 2021. “CT 2.0.” Pp. 1–8 in *Proceedings of the 21st Koli Calling International Conference on Computing Education Research*. Joensuu Finland: ACM.

Touretzky, David, Christina Gardner-McCune, and Deborah Seehorn. 2023. “Machine Learning and the Five Big Ideas in AI.” *International Journal of Artificial Intelligence in Education* 33(2):233–66. doi:10.1007/s40593-022-00314-1.