Making Machine Learning a Maker

1st Author
1st author's affiliation
City, State/Province
Country
Author's (or affiliation's) website
1st author's E-mail address

ABSTRACT

Education is changing. New methodologies such as project-based learning (PBL) and making activities have opened their ways into schools. With these new learning experiences new types of data is being generated: artifacts, with video and pictures of them, project documentation, diagrams, etc. What can this new data tell us about the new learning experiences? How can we interpret the data? This paper is an exploration of what can be done with maker and project based learning picture data using out-of-the-shelve machine learning (ML). One discovery is that artificial intelligence (AI) can provide educators interesting insights about the learning experiences represented in the pictures, which might serve as a later reflexion and understanding of the different phases of 'making' and PBL. Another possible application of this exploration is that it can also be done by students and function as an ML project that provides a window into understanding their own learning. The code of this paper is available for any educator interested, please email the author.

Keywords

Making, Project-Based Learning, Machine Learning, Artificial Intelligence, Google Vision API

1. INTRODUCTION

This is an exploratory analysis of 7640 images taken through the different learning units of Portfolio School (PS) during 2016-2018 using Google's Vision API. Two immediate goals of this exercise are i) to understand what relevant information can ML provide educators and ii) what needs to be in place for this analysis to be used as an ML activity for learners.

Portfolio School is a project-based learning school. Teachers include "making" activities as part of the daily learning through projects, there is also half day dedicated to computational thinking and digital fabrication, as well as half day per week dedicated to personal projects. The pictures were taken from September 2016 through November 2018 by PS's staff, mainly teachers. The pictures portrait children's collaborations, activities, diagrams, drawings, games, artifacts, projects in progress, etc and were taking daily. PS staff took the pictures mainly as a way to document the daily activities for the weekly newsletter or because they saw something worth taking a picture. So the pictures represent what's happening daily at the school, what PS's staff considers worthy of taking a picture and teachers' perspective.

I'm using Google's Vision Label Detection API to categorize the content of the images and get a sense of what were the most common labels in the pictures. I'm also exploring how to answer questions such as how reliable that label classification is given the complexity and context of the pictures, how granular is the classification, what levels of abstraction are present in the labels ('communication' vs 'finger'), what (if any) can the relationship between labels tell us about learning experiences. The images where divided by units so I'm also interested in learning if ML can point to some interesting differences between units. There pictures belong to six units: 2016-2017: Learning is Delicious, Color, Domestication. 2017-2018: Mars, Water. 2018: Identity.

What I learned so far: given that the label detection seems to work well enough with these type of pictures it might be possible to characterize distinct learning experiences via AI. In my analysis, I found at least 2 different learning experience signatures from contrasting 2 of the units. One experience is more traditional: kids sitting at a table, looking at a teacher or notebook, etc. Mainly taking information in. The labels corresponding to this experience were 'class', 'workshop', 'communication', 'reading', which were highly represented in the Mars Unit. Another experience was exemplified in the Learning is Delicious unit, were labels such as 'font', 'text', 'writing', which on inspection generally indicate children externalizing their learning through diagrams, drawings, and writing pieces. Another interesting result is that correlation between labels can tell us information about the experiences. For example: overall, the labels communication and electronic device were correlated. Reviewing the pictures I found that usually when kids are interacting with technology they are doing so together rather than in solitude. This type of correlation can help educators to understand context that accompanies learners activities.

2. METHODOLOGY

2.1 Data Cleaning and Labelling

I wanted this exploration to include as much data as possible so the only files I removed were duplicated pictures. This simple process can be an starting point for students to design a function that remove duplicates and the beginning of a conversation about which data can be useful.

Labels were extracted via the Google Vision API's label annotation feature. This was chosen over open-source alternatives for operational simplicity, and because in future iterations I would like to use other modalities of the API, such as text, speech, & video. Each image may receive 0 or more labels. Although the API returns scores associated per each label, I did not use the score.

From the results of the API I removed some labels that were highly frequent but that may not bring learning insights, such labels were: 'flooring', 'floor', 'child', 'furniture', 'shoulder', 'hand', 'arm', 'whiteboard', 'student', 'leg', 'desk', 'toddler', 'angle' ,'room', 'table', 'human behavior', and 'classroom'. Even though I removed 'hand' 'arm' and 'finger', I'm interested in looking at them later to see if they are a proxy for hands-on-learning experiences.

2.3 Reliability

What do the labels really mean? To better understand what's the meaning of the labels I inspected 192 pictures from each of the 30 top labels, 4860 pictures in total. This investigation was done per unit (32 pictures from each of the 6 units) to ensure that the labels were consistent across units/time. Even though this process needs much further investigation, I got a few insights:

1. For 28 out of 30 top labels the label matches the object it tries to identify. The two labels that had a problematic matching result were ‘product’ and ‘human behavior’. The labels 'product' labels a really wide range of pictures: kids working on a project, kids with a lot of objects, kids looking an object, etc. Also 'product' was not strongly correlated with any other label making its interpretation harder. Another label with an even more amplified generalization problem was 'human behavior'.
2. In low-abstraction level labels like 'font', 'personal computer' or 'plant' I found a range of 0-3 (0-1.5%) pictures over the 192 that did not match the label.
3. In labels that match a group of objects like 'class', 'school' or 'technology' I found a range of 0-8 (0-4.1%) pictures over the 192 that did not match the label.
4. In high-abstraction level labels like 'fun', 'learning' or 'education' I found a range of 5-10 (2.6-5.2%) pictures over the 192 that did not match the label.
5. For some labels the name of the label can be rename to better understand the learning experience that they represents. By example: the 'material' label is closely related to activities where crafts were present, but there is not label for craft activities.

It's clear that more research needs to be done on this topic to understand better what the different levels of labels abstraction represent and what's the best way to use them.

Figure 1. Images that with the label ‘text’

3. Exploratory Data Analysis

3.1 Top Labels

Exploring if the top total labels can help us visualize the general learning experience at PS, I plotted the top labels and compared them per unit. The top labels are: ‘fun’, ‘play’, ‘learning' and ‘design’ (labels ‘product’ and ‘girl’ are hard to interpret). Right now I don’t have a way to compare this result with the result of another school but given that the design of PS is: no grades, multi-age, project-based the frequency of those labels seems aligned.

Figure 1 helps to understand better where the top labels frequency came from and make sure that one single unit is not having a lot of one label I plotted the top labels of each unit to visualize the differences. In the future, I’ll like to compare more the different units and do it through time. For now I’ll just describe the process and insights of a simple comparison between 2 units: Learning is Delicious (LD) and Mars Unit (MU). I tried to pick units that were different to make the contrast more salient. These two units happened in different school years (LD first unit of 2016, MU first unit of 2017), their length was different (LD- 6 weeks, MU- 14 weeks), number of children in the pictures was different (LD- 10, MU- 19) and their final project was very different (LD- an electric ice-cream machine, MU- a movie).

We can see that there are many similarities in the labels: fun, girl, learning, play and product. But there are labels that appear exclusively per unit. Labels such as ‘class’, ‘workshop’, ‘communication’, ‘reading’ are exclusive of the Mars Unit. When taking a look at those pictures I found similarities between them, they represent a similar experience: one where children are gathered with other children or by themselves in a table working or looking at a book, notebook or whiteboard. In the Learning is Delicious unit the labels that were different from the Mars Unit but are similar between them are 'font', 'text', and 'writing'. After inspecting these labels I found that they represent drawing, diagrams, or writing pieces, and children writing. These experiences represent a specific type of learning that is mostly guided by externalizing information. These differences might tell us that those two different experiences -the one represented by the labels: class, workshop, communication and reading vs another experience for text, font, and writing- were more valued or might had happened more frequently on each unit.

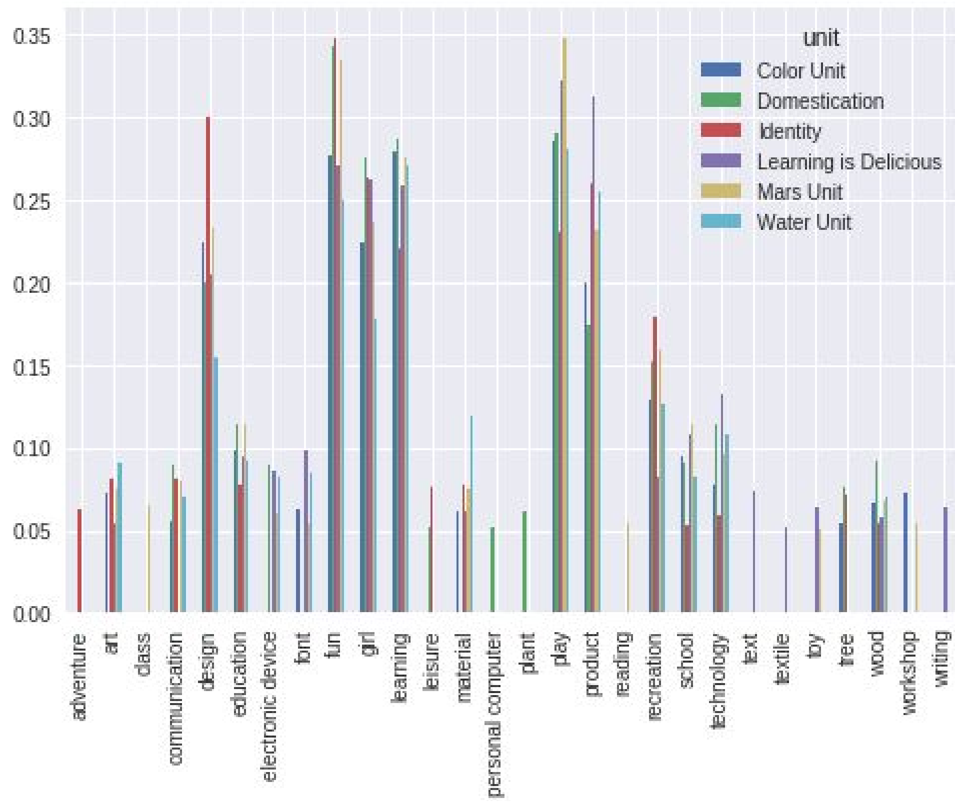

Figure 2. Top 28 labels per unit

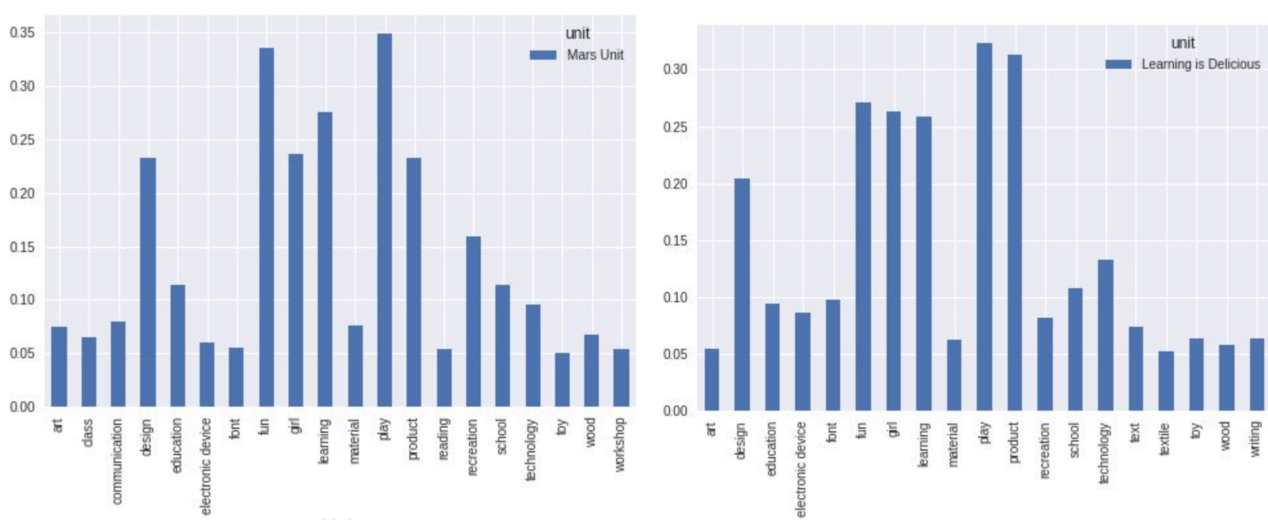

Figure 3. On the right top 19 labels for the Learning is Delicious unit. On the left top 19 labels for the Mars unit

3.2 Labels' Correlation

Google's vision API returns several labels per image. I was interested in learning more about what are the relations between labels: Is 'fun' related with 'learning'? Is 'art' related with 'electronic devices'? Are these relations mediated by unit? For this exploration I correlate all the labels. To understand visually what does correlation mean I use a system that can provide me with the pictures that have a specific label or labels but not other ones. For example: there was a strong correlation between 'plant' and 'recreation' when I use this system I was able to see that the label ‘plant’ was giving the context of the picture being taken in an outdoor area where learners normally have recess. This process gave the opportunity to look for pictures for where the correlation was low and see what activities were happening there. For example: there was a very low correlation between ‘plant’ and ‘class’ but there were pictures with both labels where learners where having a learning activity with plants inside the school. Correlating labels help me to find less obvious relationships between labels, it was also a way to see clusters of labels. Figure 4 represents the correlations between the labels, as visualized by a forced-directed graph layout algorithm. Each label corresponds to a node, and an edge is created between two nodes if their correlation exceeds 0.2.

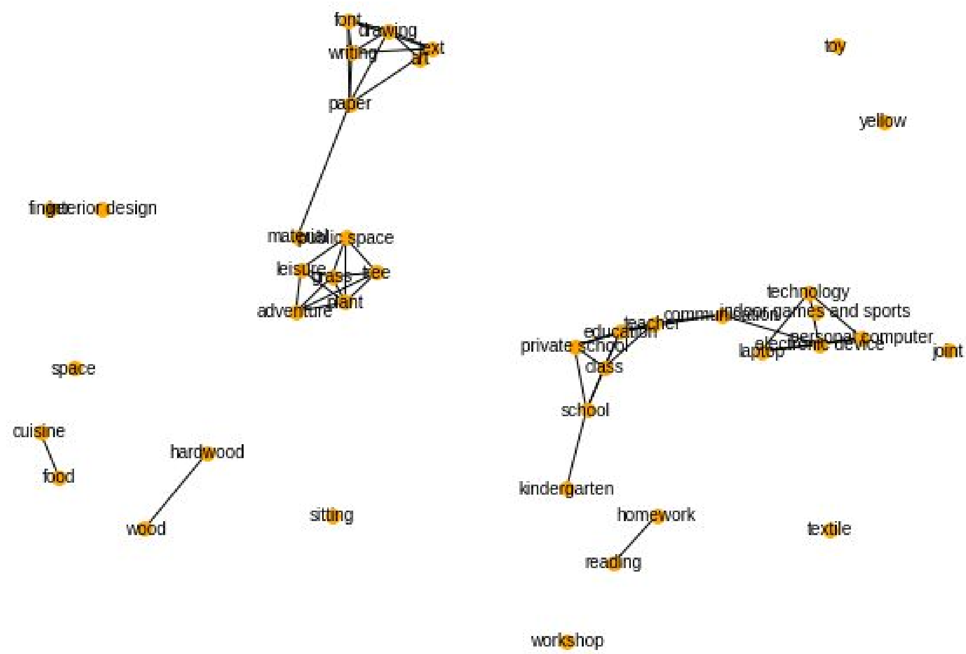

Figure 4. Correlation between top 40 labels

3.3 Results

Even though this is the tip of the iceberg and more analysis is needed, this explorations pointed at some interesting patterns to develop further:

1. A comparison between units can be possible using the relationship between labels. This relationship can be used to see the learning experiences that occurred in the different units and find more specific patterns around that experience. For example: was technology used in those learning experiences, what materials were present the most? were children mostly creating diagrams, working by themselves or building artifacts?
2. Some labels don't represent anything specific enough to be useful, e.g. ‘human behavior’, so it's important to understand what are the label referring to before making assumptions. Another problem are labels such as ‘girl’, they might be pointing to a problem of classification with the API because its counterpart 'boy' never appears. It can be that the label 'girl' has higher probability than 'boy' because it's more specific. But further analysis should be done about it. Learners discussion about labels such as ‘girl’ can provide a real example of ML bias. That activity can giving an example of how classification works, what are their limits and how can it be improved.
3. These pictures were taken by the school staff, so their analysis doesn't represent an objective take on the different units but mostly teachers’ perspective. For this analysis to be less subjective a device that can take pictures every 5-10 minutes without human intervention will be better. But for educators using ML to analyze their own pictures can still provide many insights.
4. Using Google’s Vision API was pretty simple but still the exploration gave interesting results. For learners that have some experience with python it can be a fun project for ML. I used a web platform called colab which connects with google drive, google vision api and can be easily shared. This might be a good approach for educators and their learners.

4. FUTURE DIRECTIONS

1. We have three main learning phases at the school: Exploration, Formalization and Implementation. Adding the time stamps to the pictures and visualizing what labels arise during each learning phase could better differentiate learning experiences that are not time-independent.

2. Adding time stamps can also help to see the evolution through time of children's collaboration and group formation. Pose and gesture analysis can help understand better those changes.
3. Comparing these pictures taken by teachers with pictures taken by a raspberry-pi every 5 minutes and pictures taken by learners can help undercover teachers points-of-view and differences between observers.
4. Human coding some of the pictures that represent specific experience and use supervised learning to create other models of defining labels like 'learning', 'classroom', 'fun', etc

5. BIOS

Nancy Otero is the Founding Director of Research and Learning Design at Portfolio School. She has a masters in Learning, Design and Technology from Stanford University and collaborated with Transformative Learning Technology Lab for four years. She co-founded FAB!, a non profit in Mexico that works with public schools.

5. REFERENCES

[1] Worsley, M. & Blikstein, P. Int J Artif Intell Educ (2018) 28: 385. https://doi.org/10.1007/s40593-017-0160-1

[2] Blikstein, P. & Worsley, M. (2013). Multimodal Learning Analytics and Assessment of Open-Ended Artifacts. Bill and Melinda Gates Foundation, Learning Analytics Working Group .