DSAA Final Project Report

Team Members:
Sajal Asati - 20171183
Preet Thakkar - 20171163
Arnav Juneja - 20171068

Concepts Used:

To find similarity between the images we used a simple version of normalized 2D cross correlation, and the output the best matching slide for a given frame.

Steps followed:

1. Initially, from the given system path of frames and slides folder we read the slide images and store it in a numpy array (to avoid reading it multiple times for computations).
2. The image is converted to grayscale - because we saw that mostly the slides we need to match contain textual content. (Converting it into grayscale improved efficiency from 90.1% to 92.5%)
3. For each image, we find its mean and subtract it from the whole image.
4. We also stored the L2 norm of the image matrix in a python list.
5. Now for each frame given in the folder, we read it in grayscale form and find its single correlation with all the slides in memory. Wherever we get the highest correlation value, we output that slide as the matching slide.

For each cell of the image, we did:
X_new = (X - mean)
where X_new is the new cell value, and mean is the mean of all values of the image.

Now upon with this preprocessing, we just pointwise multiply both the matrices and divide by the product of their L2 norms for normalization. Now higher the correlation values, more closely the images are related.

Results:

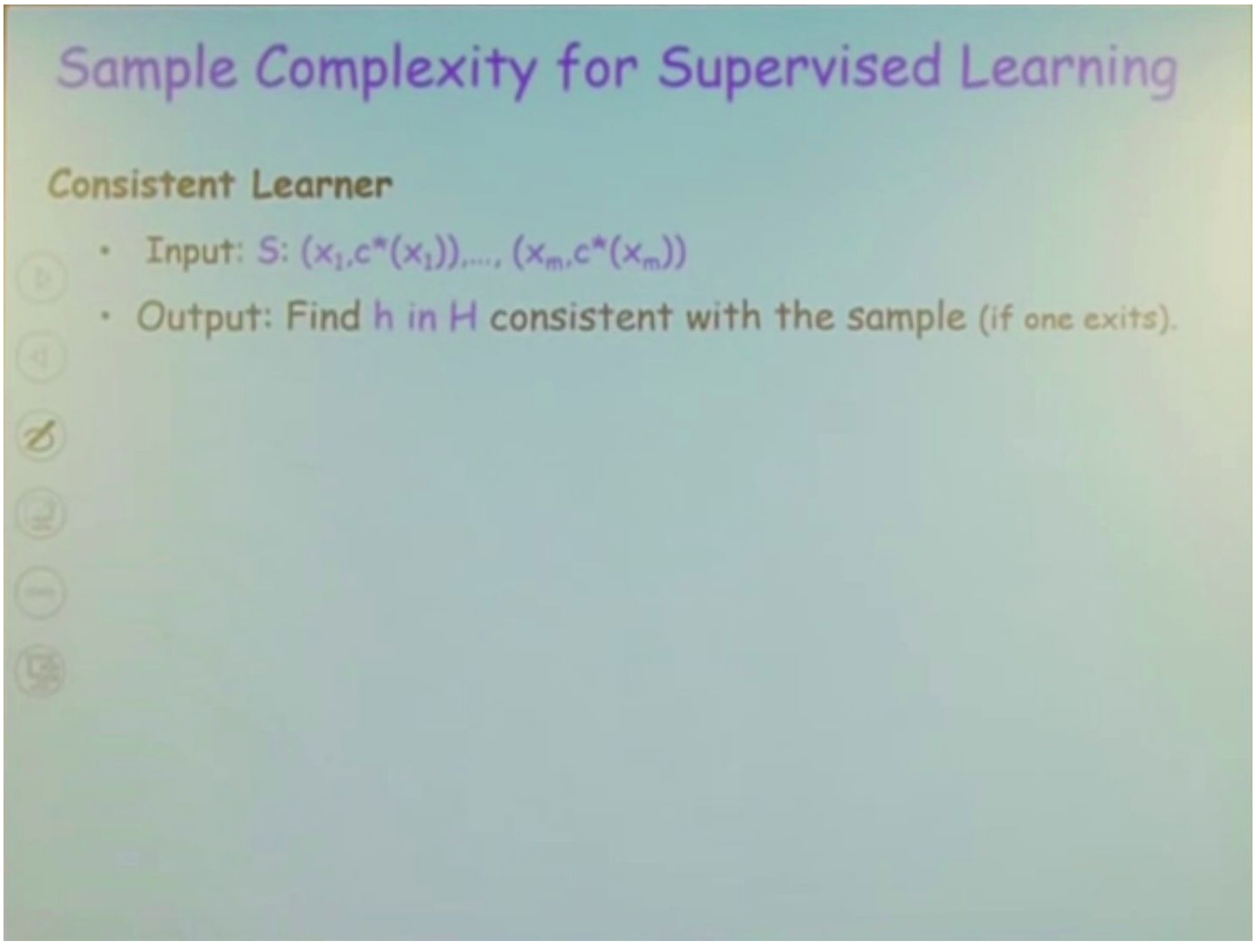
1. Total images = 835
2. Correct Predictions = 773
3. Accuracy= **92.5748503** %

Observations:

- Incomplete slides can match to multiple slides which have the same part of that content but the remaining part is different.

See the next pages

Given Sample frame



Matched with:

Sample Complexity for Supervised Learning

Consistent Learner

- Input: S: $(x_1,c^*(x_1)),\ldots,(x_m,c^*(x_m))$
- Output: Find h in H consistent with the sample (if one exits).

Theorem

$$m \geq \frac{1}{\varepsilon}\left[\ln(|H|) + \ln\left(\frac{1}{\delta}\right)\right]$$

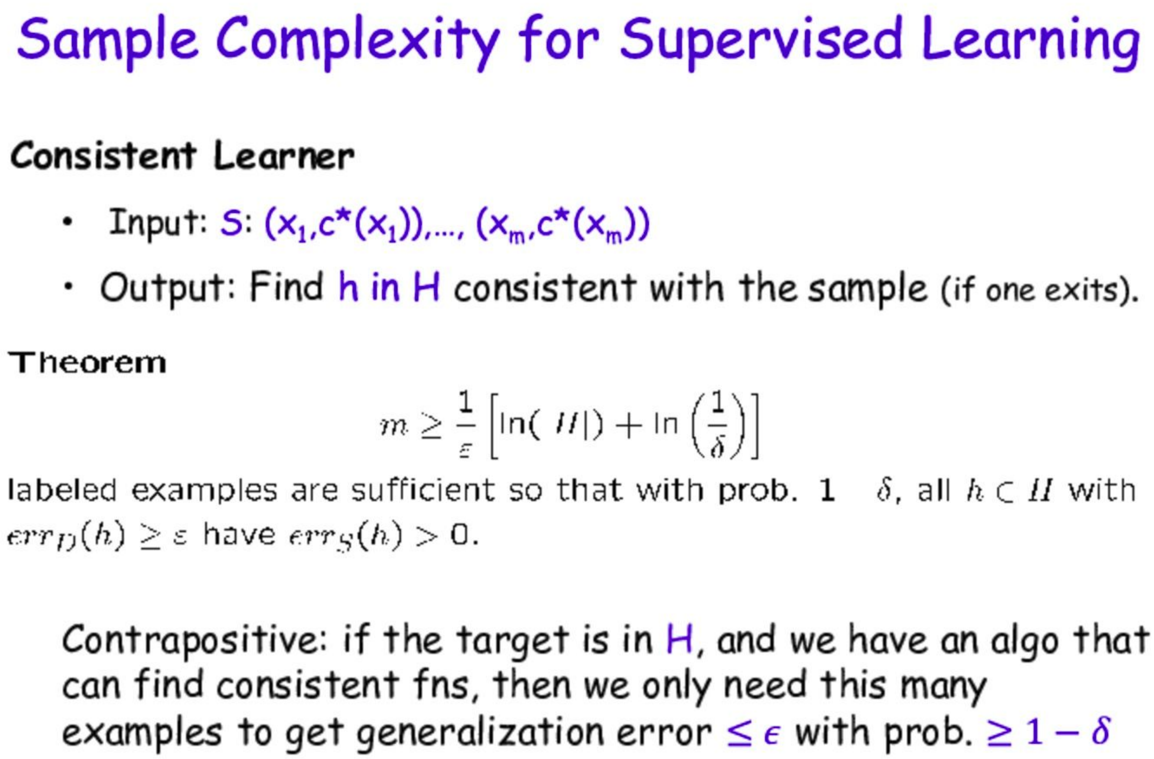
labeled examples are sufficient so that with prob. $1 - \delta$, all $h \in H$ with $err_D(h) \geq \varepsilon$ have $err_S(h) > 0$.

Example: H is the class of conjunctions over $X = \{0,1\}^n$. $|H| = 3^n$

E.g., $h = x_1 \overline{x_3} x_5$ or $h = x_1 \overline{x_2} x_4 x_9$

Then $m \geq \frac{1}{\epsilon}\left[n \ln 3 + \ln\left(\frac{1}{\delta}\right)\right]$ suffice

Correct Slide:



- Many slides are tilted which are sometimes not matched properly with the correct slide sometimes.
- Presence of exactly same slides in different ppt images.
- To avoid the effect of color and to account for matching text in the slide with the frame (which is mostly the case) we converted the slides into grayscale.