# Homework 1 - CMSC-25400

Patrick Collins

January 22, 2014

## 1 Proofs

1. Derive an algorithm that can learn any monotone conjunction over $0, 1^n$ with a mistake bound of $n$.

**Algorithm 1** Addition for Monotone Conjunctions

$f \leftarrow \bigwedge_{n=1}^{n} x_n$
$i \leftarrow 1$
**while** True **do**
  predict $\hat{y}^t = f(x^t)$
  **if** $(\hat{y}^t == 1)$ and $(y^t == 0)$ **then**
    remove $\{x | x \in x^t \text{ if } x == 1\}$ from $f$.
  **end if**
  $t \leftarrow t + 1$
**end while**

Proof of mistake bound:

Note that: (a) this algorithm will only ever produce false negatives, and (b) it removes at least one term from the concept for every false negative. Therefore, because there are only $n$ terms total, it can make a maximum of $n$ mistakes.

2. Prove that, for any concept class $C$ and online algorithm $A$ with finite mistake bound $M$, there is a conservative algorithm $A'$ with mistake bound $M$ on $C$.

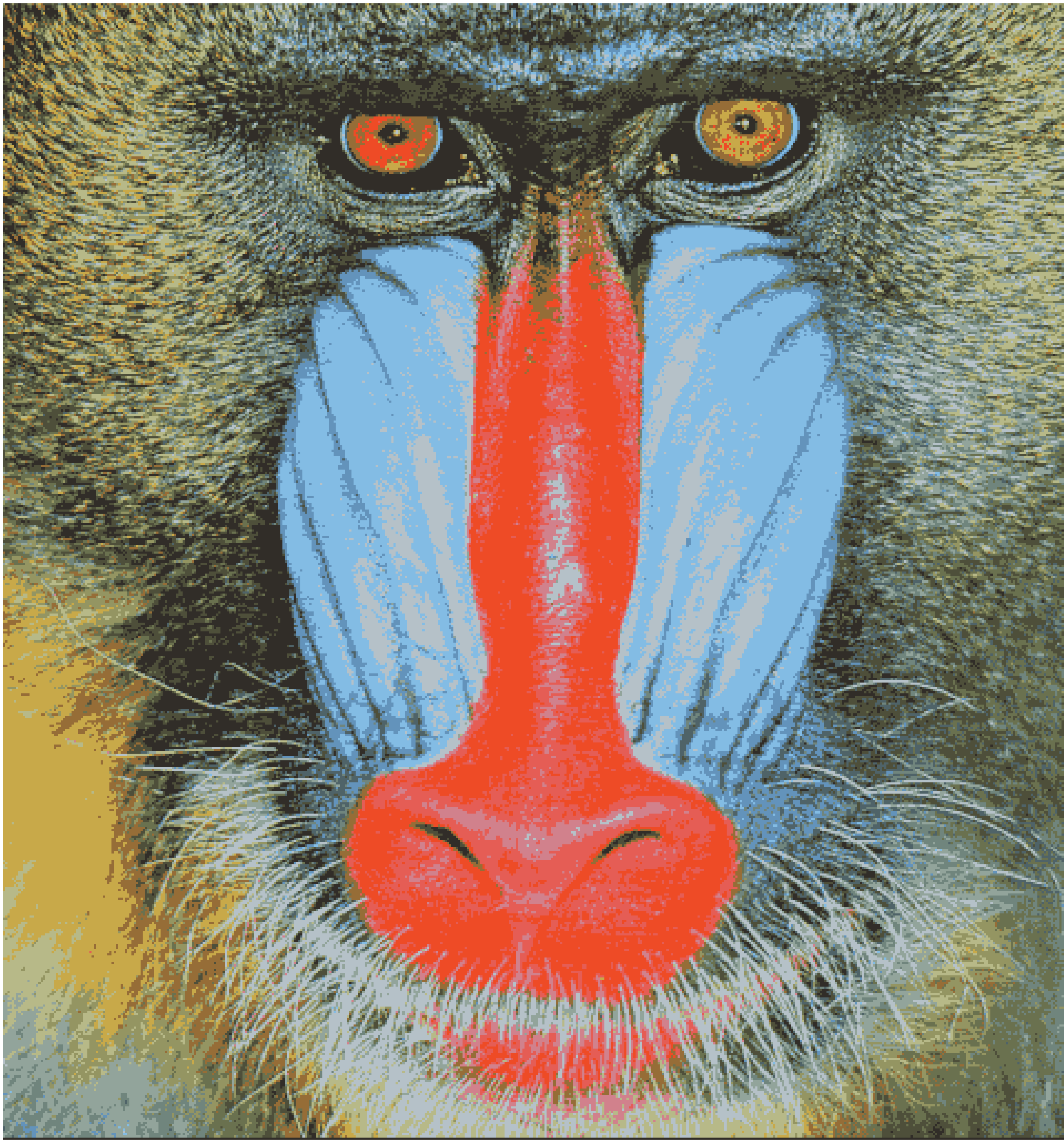
Suppose that there exists some algorithm $A$ that can select a true hypothesis from concept class $C$ with a mistake bound of $M$. Consider an arbirary set of training inputs, $x^1, x^2, \ldots, x^n$, and let $x^i$ be the training input where $A$ develops a true hypothesis. Consider $x^j$, the last input for which $A$ makes a mistake. It must be the case that $j \leq i$, since, if $A$ already had a true concept, it would no longer make mistakes.

Suppose $j < i$. Then there exists some series of inputs $x^{j+1}, x^{j+2}, \ldots, x^{i-1}$ such that $A$ provides the correct answer for each. By assumption, it is also true that $A$ provides the correct answer for $x^i, \ldots$ But if $A$ provides the correct answer for all inputs after $x^j$, then it must be the case that $A$ has a true concept at $x^j$, therefore $i = j$.

Therefore, there exists an $A'$ such that $A'$ has a mistake bound $M$ on $C$ and $A'$ is conservative.

3.

4. Results from applying KMeans:



The contrast in the resulting, compressed image is sharper, reflecting the fact that the number of colors in the pictures has been halved. It also appears to be somewhat more washed-out than the original.

This compression reduces the total number of colors to 16, meaning that each pixel can be represented with three 4-bit numbers rather than 8-bit ones, resulting in a compression ratio of approximately 50% for large images, because the storage requirement for each pixel has been halved.