Weeks 10, 11

Large Scale Machine Learning,
Application Example, Photo OCR

Large Scale Machine Learning

- **Consider a huge dataset, eg $m=10^9$**
- **Gradient descent would be slow...**
- **Solutions:**
  - Train on subset, e.g. m=100,1000
  - SGD: randomly order, train on individual examples

Learning rate in Large Scale ML

- **Same problems faced with small-scale**
- **Iteratively reduce rate:**

$$\alpha(i) = \frac{k_1}{i + k_2}$$

- **Check convergence, reduce rate**

$$|J(i) - J(i-1)| \tilde{=} 0$$
$$\alpha(i) \rightarrow k\alpha(i); k < 1$$

Map Reduce & Parallelization

- **Split operations in ind. operations**
- **Execute ind. on parallel nodes/CPUs**
- **Combine results on central node/CPU**
- **For example, train subsets in parallel, combine results**

ML Pipeline

- **Split total ML algorithm into steps**
- **Assign different group to each step**
- **Optimize each step to determine bottlenecks**

ML Pipeline (Con’t)

Example:

ML Pipeline (Con't)

Example:

1. Text detection

2. Character segmentation

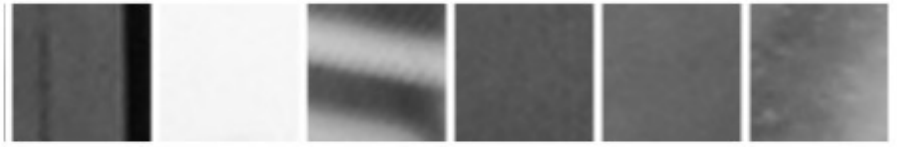
3. Character classification

ML Pipeline (Con’t)

Example:

Text detection

Positive examples $(y = 1)$



Negative examples $(y = 0)$

Ceiling Analysis

- **Motivation:**
  - Determine which areas to improve
- **Procedure:**
  - Perform basic test
  - Tune one component to be ‘perfect’
  - Check performance
  - Repeat

Ceiling Analysis (Con’t)

- **Example:**
  - Consider a classifier which sees an image possibly containing:
    - Text
    - Face
  - If face is present, estimate gender

Baseline Accuracy: ***68%***

Perfect Text Detection: ***69%***

Perfect Face Detection: ***78%***

Perfect Gender Detection: ***100%***

https://medium.com/@rossbulat/ceiling-analysis-in-deep-learning-and-software-development-8bc41e59364a