Deep Learning

Exercise 4: Multi-Output Networks and Batch Processing

Instructor: Manuel Günther
Email: guenther@ifi.uzh.ch
Office: AND 2.54

Friday, March 19 2020

Outline

Outline

Multi-Output Regression

Task 1: Batch Processing

- Signature: `def batch(X, T, B):`
- Shuffle data each epoch
- Generate batches of a fixed size
- Hint: `yield` is your friend

Task 2: Multi-Target Network

- `def forward(X, W1, W2):`
- Hidden unit output $\mathbf{H}$
- Network output $\mathbf{Y}$

Task 3: Gradient Descent Step

- `def descent(X, T, W1, W2, eta):`
- Forward pass: $\mathbf{Y} = f(\mathbf{X})$
- Compute loss: $\mathcal{J}^{L_2} = \|\mathbf{Y} - \mathbf{T}\|^2$
- Compute gradients: $\nabla_{\mathbf{W}^{(1)}}, \nabla_{\mathbf{W}^{(2)}}$
- Perform weight update:

$$\mathbf{W}^{(1)} -= \eta \nabla_{\mathbf{W}^{(1)}}$$

$$\mathbf{W}^{(2)} -= \eta \nabla_{\mathbf{W}^{(2)}}$$

Multi-Output Regression

Task 4: Data Set Loading

- Download dataset from UCI
- Read CSV file `student-mat.csv`
- Convert binary into $\{-1., 1.\}$
- Convert integral values to `float`
- Ignore categorical values
- Input $\mathbf{X}$: columns 1–8,13–30
- Targets $\mathbf{T}$: columns 31–33

Task 5: Learning

- Load data $\mathbf{X}, \mathbf{T}$
- Initialize weights $\mathbf{W}^{(1)}, \mathbf{W}^{(2)}$
- Select parameters η and B
- Perform gradient descent for a given number of epochs
- Plot loss over epochs

Dataset URL

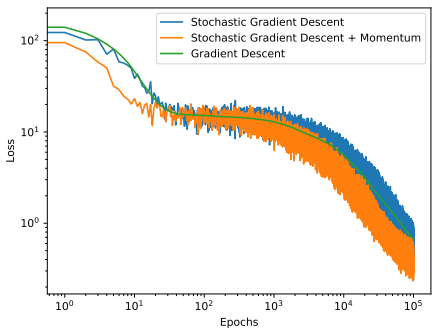
https://archive.ics.uci.edu/ml/datasets/Student+Performance

Multi-Output Regression

Task 6: Optional

- Add momentum term to `descent`

Losses for Training Methods



Task 7: Evaluation

- Select all students with:
  1. male, female (index 2)
  2. paid classes (index 18)
  3. romantic relationship (index 23)
  4. daily alcohol (index 27)
- Compute estimated average grades from network output
- Is there a difference between courses: `"mat"` or `"por"`

Multi-Output Regression

Example Outcomes for "mat"

Variable	Gender		Paid Classes	
	male	female	yes	no
Grade 1	10.75	10.93	11.05	8.32
Grade 2	10.23	10.78	10.87	7.82
Grade 3	9.77	10.50	10.61	6.81