Report Date: 06/10/2022
To:

- ematson@purdue.edu
- ahsmith@purdue.edu
- lhiday@purdue.edu
- lee3450@purdue.edu
- wang4070@purdue.edu

From: TN

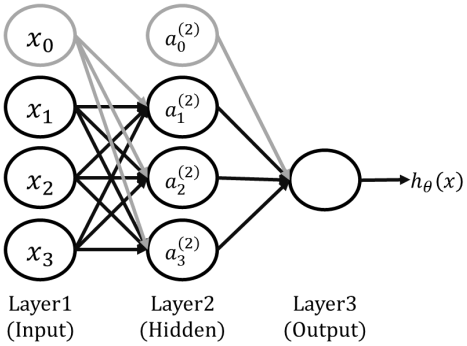
- Eunyoung Bang (yeong35@kangwon.ac.kr)
- Yeongmin Seo (dudals1003@cu.ac.kr)
- Jeongyoun Seo (201810773@sangmyung.kr)
- Raymond Zeng (zeng172@purdue.edu)
- Aminata Niang(aminata.niang@telecom-sudparis.eu)

Summary

- The training Deep Learning code to DataLoader.
- Pytorch version 1.11 was chosen for our project.

What TN completed this week

- Pytorch version is chosen to 1.11.
    - Everyone downloaded Pytorch (1.11.) successfully.
- The training Deep Learning code was written until DataLoader.
    - Examples of how to write train code are [1], [2].
- Machine Learning study and Deep Learning [3], [4].
    - Machine Learning is the science of getting computers to learn, without being explicitly programmed.
    - Supervised learning: When someone who knows the right answers, called a supervisor, points out mistakes during the learning process. This means teach the computer how to do something, then let it use it's new found knowledge to do it.
        - Regression problems: Predict continuous valued output, no real discrete delineation.
        - Classification problems: Predicting a discrete class label output for an example. (making decisions by 0 and 1)
    - Unsupervised learning: The process of learning without training labels. (I.e. clustering, grouping)
    - Reinforcement learning: The process of learning in an environment through feedback from an AI's behavior.
    - Gradient descent is a method of differentiating the cost function to obtain the slope, obtaining the pole, and thus obtaining the minimum value of the cost function. If the cost function is non-convex, it can not obtain the minimum value.
    - The structure and internal equation of the Neural Network inspired by the Neural Network of living things are as follows.



$$\rightarrow a_1^{(2)} = g\left(\theta_{10}^{(1)} \cdot x_0 + \theta_{11}^{(1)} \cdot x_1 + \theta_{12}^{(1)} \cdot x_2 + \theta_{13}^{(1)} \cdot x_3\right)$$

$$\rightarrow a_2^{(2)} = g\left(\theta_{20}^{(1)} \cdot x_0 + \theta_{21}^{(1)} \cdot x_1 + \theta_{22}^{(1)} \cdot x_2 + \theta_{23}^{(1)} \cdot x_3\right)$$

$$\rightarrow a_3^{(2)} = g\left(\theta_{30}^{(1)} \cdot x_0 + \theta_{31}^{(1)} \cdot x_1 + \theta_{32}^{(1)} \cdot x_2 + \theta_{33}^{(1)} \cdot x_3\right)$$

$$\rightarrow h_\theta(x) = g\left(\theta_{10}^{(2)} \cdot a_0^{(2)} + \theta_{11}^{(2)} \cdot a_1^{(2)} + \theta_{12}^{(2)} \cdot a_2^{(2)} + \theta_{13}^{(2)} \cdot a_3^{(2)}\right)$$

$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad z^{(2)} = \begin{bmatrix} z_1^{(2)} \\ z_2^{(2)} \\ z_3^{(2)} \end{bmatrix}$$

$$z^{(2)} = \theta^{(1)} \cdot x$$

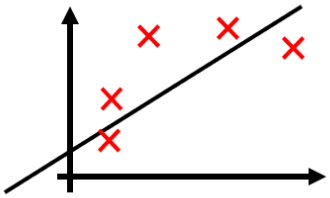
$$a^{(2)} = g\left(z^{(2)}\right)$$

$$z^{(3)} = \theta^{(2)} \cdot a^{(2)}$$

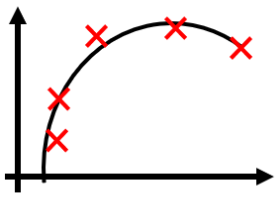
$$h_\theta(x) = a^{(3)} = g\left(z^{(3)}\right)$$

$$a^{(j)} = g\left(z^{(j)}\right)$$

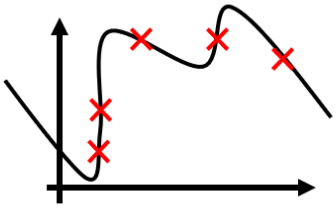
- A state in which the high bias is not properly trained due to too few parameters is called 'underfit'. Conversely, a state in which there are too many parameters and the accuracy is high only for training data and the accuracy is low for new data is called 'Overfit'.



'Underfit'
High bias
Ex) $h_\theta\left(x^i\right) = \theta_0 + \theta_1 \cdot x$



'Just right'
Ex) $h_\theta\left(x^i\right) = \theta_0 + \theta_1 \cdot x + \theta_2 \cdot x^2$



'Overfit'
High variance
Ex) $h_\theta\left(x^i\right) = \theta_0 + \theta_1 \cdot x + \theta_2 \cdot x^2 \vdots \theta_5 \cdot x^5$

- Convolutional Neural Networks (CNNs).
  - It is mainly used for image processing and object detection. CNN's are widely used to identify satellite images, process medical images, forecast time series, and detect anomalies.
  - CNN's have multiple layers that process and extract features from data
- Rectified Linear Unit (ReLU)
  - CNN's have a ReLU layer to perform operations on elements. The output is a rectified feature map.
- Pooling Layer
  - The rectified feature map next feeds into a pooling layer. Pooling is a down-sampling operation that reduces the dimensions of the feature map.
  - The pooling layer then converts the resulting two-dimensional arrays from the pooled feature map into a single, long, continuous, linear vector by flattening it.
- Fully Connected Layer

- A fully connected layer forms when the flattened matrix from the pooling layer is fed as an input, which classifies and identifies the images.

Things to do by next week

- The training code for deep learning will be finished.

Problems or challenges:

- Discuss the method of collecting the flying UAV sound; the dataset was too easy for ML algorithms which make 100% accuracy.
    - There are two different ways to collect the new dataset, could not decide which method will be more helpful for our ML and deep learning algorithms.
    - The UAV will be driven to a circle, and it will be still 10 seconds for each audio file.
    - The UAV still flies back and forth, however the length of the audio file will be shorter between 3 to 5 seconds, to avoid including the obvious audio feature when the UAV turns the direction.

References

[1] Eunyoung Bang, "MusicTranformer-Pytorch." github.com. https://github.com/yeong35/MusicTransformer-Pytorch/blob/master/dataset/e_piano.py (accessed June. 8, 2022)

[2] Ketan Doshi, "Audio Deep Learning Made Simple: Sound Classification, Step-by-Step." medium.com. https://towardsdatascience.com/audio-deep-learning-made-simple-sound-classification-step-by-step-cebc936bbe5 (accessed June. 8, 2022)

[3] holehouse, "Stanford Machine Learning", holehouse.org. http://www.holehouse.org/mlclass/ (accessed June. 02, 2022)

[4] "Initiez-vous au Deep Learning." openclassrooms.com/fr/. https://openclassrooms.com/fr/courses/5801891-initiez-vous-au-deep-learning/5814656-decouvrez-les-cellules-a-memoire-interne-les-lstm (accessed June. 09, 2022)