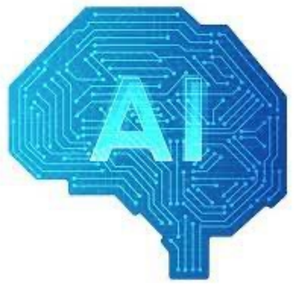
DLHub: A Tutorial for the First Time Undergraduate Users to Learn Deep Learning

Jeehyun Oh, Yonghyun Kim, Madhukar Shrestha
Faculty Advisor: Dr. Junghwan (John) Rhee

1 PROBLEM AND MOTIVATION

Deep learning is getting much attention as a widely used machine learning technique because of its strength and usefulness for research, industry, and society. However, this technique requires a stiff learning curve for first-time users including a setup of a computing environment, applying deep learning to his/her problem, and writing code for it.

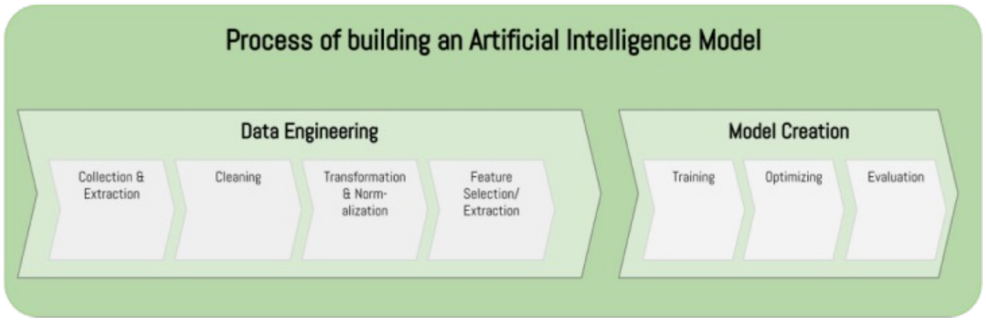


2 BACKGROUND

Artificial intelligence (AI) imitates human intelligence by generating and applyi ng algorithms embedded in machines. There are two big techniques to imple ment artificial intelligence, which are supervised and unsupervised learning te chniques. Supervised learning uses training data that include the desired outp ut, so it may learn what output should be presented by given input data. Unsu pervised learning uses training data without the desired output, so it may pred ict the output itself, so needs a lot of time.

Preprocessing data is an important step, as raw data can be inconsistent or in complete in its formatting. Effectively preprocessing raw data can increase its accuracy, which can increase the quality of projects and improve its reliability. Data preprocessing is an important step in the data mining process. It refers t o the cleaning, transforming, and integrating of data in order to make it ready for analysis. The goal of data preprocessing is to improve the quality of the da ta and to make it more suitable for the specific data mining task.

To learn deep learning in an easy procedure and make beginners of AI more i nterested, we created a website called Deep Learning Hub (DLHub), which is a website with a deep learning tutorial customized for first-time users. It repre sents steps to install programs that are needed for AI and provides a test cas e to make their own custom object detection model.



3 IMPLEMENTATION

Setting up an environment in Windows and Linux (Ubuntu):

- python, jupyter, and Anaconda for the base environment.
- CUDA, cuDNN, and PyTorch to use a GPU in local computer.
- LabelImg to preprocess image data.
- Provide a page for troubleshooting.

Preparing custom image data:

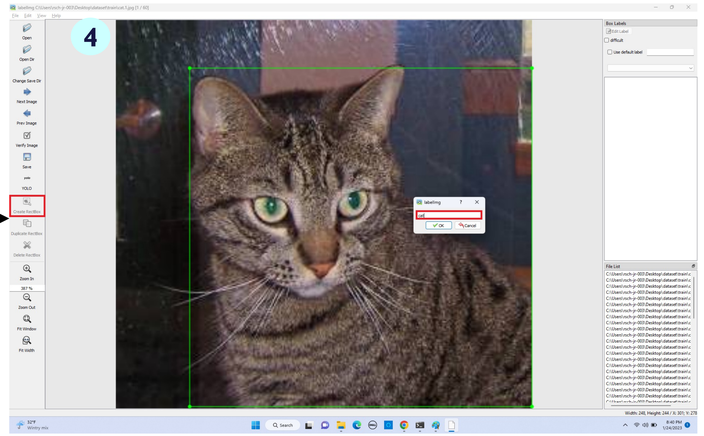
- Preprocess image data (Labeling).
- PASCAL, YOLO, and CreateML formats are available

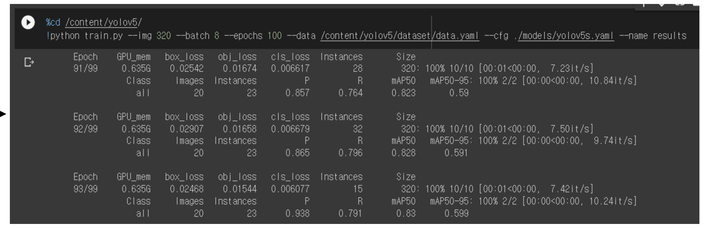
Data recognition using deep learning:

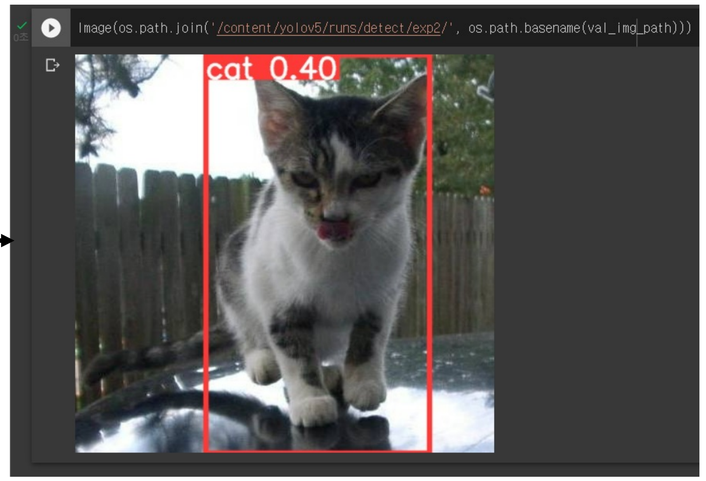
- Follow the steps to build an example model.
- Build a custom Image detection model with YOLOv5.
- Validate images and evaluate the model.

4 TOOLS

- YOLOv5 is a pretrained object detection model, it represents open-source research into future vision AI methods.
- LabelImg is a graphical image annotation tool. It can export an image as XML files in PASCAL, YOLO, and CreateML formats.
- Google Collab is a Python development environment that runs in the browser using Google Cloud.
- Anaconda offers an easy way to perform Python/R data science and machine learning on a single machine and it also includes other useful tools.
- Mkdocs is a fast and simple static site generator that is geared towards building project documentation. The documentation of this project was written in the Markdown language and the website was generated using this tool.







5 RESULT

DLHub, a tutorial for deep learning, was successfully built using Mkdocs and Github repository. We prepared the step-by-step documentation to install and try deep learning. The troubleshooting page is also made to take feedbacks and any issue of errors from users. We wish DLHub can help students to learn and try deep learning in a more friendly way.