

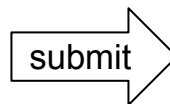
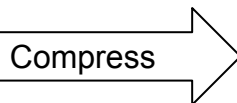
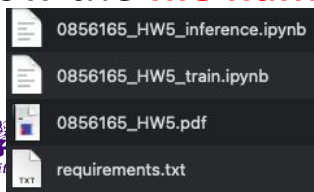
Introduction to Machine Learning

Homework 5 announcement

TA: 楊証琨, Jimmy
Ph.D. student at National Taiwan University
d08922002@csie.ntu.edu.tw

Homework 5

- **Deadline: Dec. 20, Tue. at 23:59.**
 1. Code assignment (100%): Implement the deep neural network by any deep learning frameworks, e.g., Pytorch, TensorFlow and Keras, and then train DNN model on the provided dataset (The pre-trained models are allowed to use.)
 - Submit your **1) code (.py/.ipynb)** **2) reports (.pdf)** and **3) requirements.txt** (if you submit .py file) on [E3](#)
- [Sample Code](#) and [HW5 questions](#)
- Please follow the **file naming rules below**, and **put the model weight link in your report!**



[E3](#)

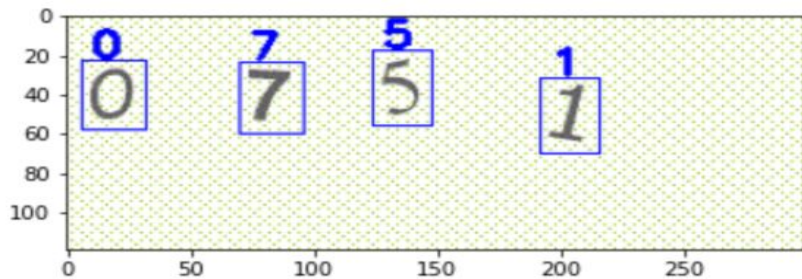


Digits captcha

- Train a model to predict all the digits in the image
 - Task 1: Single character in the image
 - Task 2: Two characters in the image (order matters)
 - Task 3: Four characters in the image (order matters)

Showing the digits on the Captcha:

```
40 plt.imshow(output_img)  
plt.show()  
  
predicted_number
```



```
40 '0751'
```



Join the competition [[HERE](#)]



Community Prediction Competition

Captcha Hacker

ML HW5

8 days to go

[Overview](#)

[Data](#)

[Code](#)

[Discussion](#)

[Leaderboard](#)

[Rules](#)

[Join Competition](#)



Overview

[Description](#)

HW5 - captcha recognition

[Evaluation](#)



國立交通大學
National Chiao Tung University

Download the dataset

[Overview](#)[Data](#)[Code](#)[Discussion](#)[Leaderboard](#)[Rules](#)[Team](#)[Submissions](#)[Submit Predictions](#)[...](#)

task1/WYGjnmKoKBpDYA
SP.png

0

task1/WWfY9kDFGzrbLo
c8.png

0

task1/VN1kHNVPUGaaQ1
HH.png

0

task1/MRYIrbeYm7c2mB
x2.png

0

Summary

- 📁 17.5k files
- 📊 4 columns

[Download All](#)

```
>_ kaggle competitions download -c captcha-hacker
```



Kaggle Team Name

- You MUST set the team name as your Student_ID**



Your Team

Everyone that competes in a Competition does so as a team - even if you're competing by yourself. [Learn more.](#)

General

TEAM NAME

0856165

This name will appear on your team's leaderboard position.



Submit your predictions

- Screenshot your result on public leaderboard and paste it on the report!
- Your score is **based on the private leaderboard**, but we will check whether your results match
- You may submit a maximum of 5 entries per day


[Overview](#) [Data](#) [Code](#) [Discussion](#) [Leaderboard](#) [Rules](#) [Team](#) [Submissions](#) [Submit Predictions](#) [...](#)

Leaderboard

[Raw Data](#) [Refresh](#)

[Public](#) [Private](#)

This leaderboard is calculated with approximately 50% of the test data. The final results will be based on the other 50%, so the final standings may be different.

#	Team	Members	Score	Entries	Last	Code
	sample-code-submission.csv		0.42120			



Reports

- Including
 - Environment details
 - ◆ python version
 - ◆ ... (the more complete and clear, the better)
 - Implementation details
 - ◆ model architecture
 - ◆ hyperparameters
 - ◆ used deep learning framework
 - ◆ ... (the more complete and clear, the better)
- Get **extra penalty** if not provide environment or implementation details



Submission

- Compress the report, all your code for this homework, and requirements.txt(if .py file) and then submit it to E3.
 - Model training (<STUDENT_ID>_HW5_train.ipynb/.py)
 - Generate prediction (<STUDENT_ID>_HW5_inference.ipynb/.py)
 - **Report in pdf with your downloadable model weight link**
(<STUDENT_ID>_HW5.pdf)
 - requirements.txt (if you submit .py file)



Environment setup

- For python file .py (will be checked in our lab's servers)
 - Please provide the requirements.txt
 - The requirements.txt can help us rapidly re-build your environment and then reproduce your result. (see [tutorial](#))
- For jupyter notebook file .ipynb (will be checked in [google colab](#))
 - Please keep the pip install instructions you need in the first cell.

Example `requirements.txt` file

```
matplotlib>=2.2
numpy>=1.15.0, <1.21.0
pandas
pytest==4.0.1
```

```
[1] 1 # 若你要的套件原本 colab 沒有，請保留這些手動安裝指令，讓我們能快速重建你的環境
    2 !pip install transformers
    3 !pip install wandb
```

```
00 1 # 以下套件 colab 原本就有，不用額外裝
    2 import numpy
    3 import torch
    4 import sklearn
    5
    6 # 以下套件 colab 原本沒有，須額外裝（例如cell 1）
    7 import wandb
    8 import transformers
```



We will test inference.py file on our lab's server

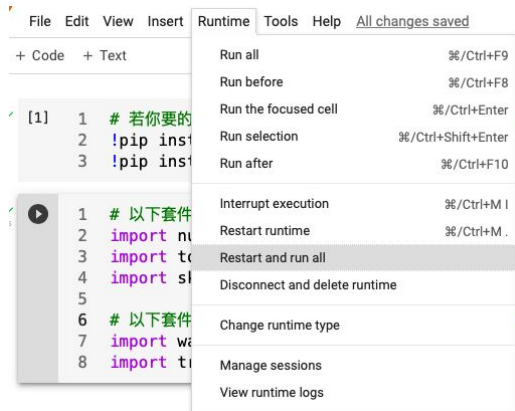
- We will take the following steps to reproduce your result. (**You can check it by yourself before submission**)
 - Build the virtual environment ([tutorial](#)) via

```
$ virtualenv -p <path to python version you provided> myenv
```
 - `$ pip install -r requirements.txt`
 - Download the `<pre-trained weight that you provided>`
 - **Only modify the necessary path** in your `inference code` (testing data, model weight, and so on...)
 - After the modification, the code should be run successfully (**Otherwise, no points will be given**)
- Check and test the generated csv file



We will test inference.ipynb file on [google colab](#)

- We will take the following steps to reproduce your result. (**You can check it by yourself before submission**)
 - Download the <pre-trained weight that you provided>
 - **Only modify the necessary path** in your **inference code** (colab mount, testing data, model weight, and so on...)
 - Restart and run all, the code should be run successfully (**Otherwise, no points will be given**)
 - Check and test the generated csv file



Reference

- [Convolutional Neural Networks Tutorial in PyTorch](#)
- [Building a Convolutional Neural Network \(CNN\) in Keras](#)



Grading policy

- Plagiarism
 - **No points** will be given if we found any plagiarism
- Late policy
 - We will deduct a late penalty of 20 points per additional late day
 - For example, If you get 90 points of this HW but delay for **two days**, your will get only $90 - (20 \times 2) = 50$ points!



FAQ

- Can we use the pre-trained weights, e.g., ImageNet?
 - Pre-trained weights are available. Please specify it in your reports
- Why my model does not learn anything?
 - See the [reference](#) to debug your deep learning model
- How can I get better performance
 - Data augmentation
 - Hyperparameter searching for model structure and optimizer (learning rate)
 - Find some techniques from SoTA paper in the [Paper with codes](#)



Notice

- Submit your homework on [E3-system](#) !
- Check your email regularly, we will mail you if there are any updates or problems of the homework
- If you have any questions or comments for the homework, please mail TAs and cc Prof. Lin
 - Prof. Lin, lin@cs.nctu.edu.tw
 - TA Jimmy, d08922002@csie.ntu.edu.tw
 - TA 政儒, ace52751208@gmail.com
 - TA 季嘉, jijiawu.cs@gmail.com
 - TA 睿哲, benchiang.cs07@nctu.edu.tw



Have fun!

