

Chieh Lin



NTHU, CS senior bachelor

Programming Languages:

C/C++/Python/Java/JS

Primary ability:

Machine learning,

Frontend development

Learning:

Security,

Database,

Backend development

Github:

[hubert0527](https://github.com/hubert0527)

English ability:

TOEIC 870 points

E-mail:

hubert052702@gmail.com

Website (by myself):

hubert0527.github.io

Online-resume:

goo.gl/Bb7kne

➤ Application

Full-year internship related to machine learning or data science. Can work 3 to 4 days a week during semester.

➤ Introducing myself

I am a student with high self-demand and courage to face challenges. One of the proofs is that I almost completed CS graduate threshold in two years and now attending graduate school courses, co-working with graduate students. I'm now mostly focusing on machine learning, data science and security.

In my free time, I'm interested in frontend/backend development and occasionally attends Kaggle competitions. I have learned Python, JS, some web frameworks, Chrome extension, iOS app, Android ... etc. I think these knowledge and concepts can provide me different new thoughts and ideas.

Quoted from data science professor, "Know your data." I think processing real data, knowing how the knowledge we learned is being used in reality and what problems to be faced are the most important topics for us.

In this internship, I hope I can apply knowledge I have learned so far into real works, solve problems and contributes to our team!

➤ Career Planning

Short Term :

I planned to go to work directly after graduated from university. And consider go back for master degree again after few years working.

Long Term :

Find an interesting area to explore and a great team to work with!

➤ Related Projects

News Popularity (Large-Scale Machine Learning)

Links: [Github](#)、[Kaggle](#)

Competition of graduate school course, predicting popularity of news based on HTML source.

I response for data preprocessing, feature section and extraction, model design, tuning, validation ... etc. And won 2nd place in the competition.

I can prove that I have correct understanding and great implementation skill in this project. I also Implemented several algorithm by myself (Boosting, Stacking, CV, Ensembling). In addition, proves generalizability in private score (public score: 0.572, private score: 0.579).

Image Caption (Deep Learning)

Links: [Github](#)、[Kaggle](#)

Competition of graduate school course, solving image captioning problem using MS COCO image set.

Utilize word-embedding, CNN, RNN and implement attention, beam search, curriculum learning. I also designed my own algorithm during the competition, boosted accuracy of beam search and improved 10% score.

Score 0.58 under CIEDr-D policy, which can become top-40 if scoring 0.4 in 2015 MS-COCO competition. And finally won 5-th place in the competition.

Data Crawling (MLB game information crawling)

Links: [Github](#)

This project is about predicting MLB game result. I was assigned to do data parsing and partially help model training.

I eventually parsed 103 years, 170k games, 652 features, 389MB, cleaned MLB game data. And implemented feature generating API based on team's requirement. Largely reduces time pressure of model training.

Utilize python request and selenium library to conduct data parsing. Dealing with all possible exceptions, re-parsing and data cleaning.

-- Thanks for reading my resume ! --

-- See more projects on Github : [hubert0527](#) --