**ONE YEAR PLAN**

**UPDATE 20201210**

FOCUS: D4TBook + Kaggle master + ChoiceNet library.

**WRITE A BOOK:**

DEEP LEARNING [**Data Science**] FOR TRANSPORTATION: A PRATICAL GUIDE **(This year)**

+) Introduction

+) Transport data

* Travel demand data
* Network data

+) Data collection and preprocessing

* From API
* From OSMap
* Web scraping

+) Visualisation + Spatial analysis.

+) Discrete choice model

+) Decision trees => Random forest => XGBoost

+) Deep learning models

+) Combining deep learning with discrete choice

**WRITE A PYTHON LIBRARY:**

INTERGRATE DEEP LEARNING WITH DISCRETE CHOICE

**PROJECTS:**

**TASKS:**

* Data collection and preparation: 1 month (Jan)
* Data visualisation and publishing on web (Feb)
* Model and Prediction ( March)
* Apply jobs: From March

Project

[**1. STOCK TRANSFORMER** 2](#_Toc60301420)

[**2. HOUSING PREDICTION** 3](#_Toc60301421)

[**3. LEFT – CENTRE – RIGHT** 4](#_Toc60301422)

## **1. STOCK TRANSFORMER**

Web scraping:

* CafeF
* Batdongsan

Cafef news vs Vnexpress news vs …?

## **2. HOUSING PREDICTION**

Scraping Melbourne & sydney housing data

Prediction using different method (RF, GD boost, DL, CNN, …)

## **3. LEFT – CENTRE – RIGHT**

Show a twitter sentiment (positive/negative) and biases (left/right)

Show what is the attention of left vs right wring newspaper?

Source of left vs. right:

<https://guides.lib.umich.edu/c.php?g=637508&p=4462444>

<https://www.allsides.com/unbiased-balanced-news>

<https://blog.thefactual.com/what-are-the-most-credible-liberal-news-sources>

**Towards the Left**

<https://www.nytimes.com/>

<https://www.washingtonpost.com/>

#### Far Left

<http://www.huffingtonpost.com/>

<https://www.motherjones.com/>

#### Towards the Right

<http://www.foxnews.com/>

#### Far Right

<http://www.theblaze.com/>

<http://www.drudgereport.com/>

