BAIS:6100 Text Analytics

Python Basics for Text Processing Part 2

Kang-Pyo Lee

Course Schedule (Subject to Change)

Week	Date	Topics	Due
1	Jan 28	Introduction to Text Analytics Introduction to Python, Jupyter Notebook, and UI Interactive Data Analytics Service (IDAS)	
2	Feb 4	Module 1. Python Basics for Text Processing, Part 1 : Strings, Collections, Built-in Functions, Flow Control, and User-Defined Functions	
3	Feb 11	Module 2. Python Basics for Text Processing, Part 2 : Files, Dataframes, and Pattern Matching Using Regular Expressions	HW 1
4	Feb 18	Module 3. Basic Natural Language Processing (NLP) Techniques : Tokenization, Part-of-Speech Tagging, Stemming, Lemmatization, N-grams, Noun Phrase Extraction, Language Detection and Translation, and Gender Prediction Module 4. Keyword Analysis and Visualization	HW 2
5	Feb 25	Test 1	HW 3 (Feb 23)
6	Mar 4	Modules 5 & 6. Text Data Acquisition Using Twitter APIs and Web Scraping Group Project Announcement	
7	Mar 11	Module 7. Document-Term Representation Module 8. Text Classification	Hw 4
8	Mar 18	Module 9. Text Clustering and Topic Modeling	Project Proposal
9	Mar 25	Module 10. Text Similarity Module 11. Keyword Network Analysis	
10	Apr 1	Test 2	HW 5 (Mar 30)
11	Apr 8	Group Project Presentations and Course Wrap-Up	Project Deliverables

Homework 2, which corresponds to Module 2, is due by 6:00 PM on Thu, Feb 18 via ICON Assignments

13 questions, 7 points in total

You will receive a 20% deduction for each day that it is late, including the first/same day

Homework

Homework 2 is expected to take a student with a good understanding of the topics 2-3 hours to complete it

There will be a question at the end of the notebook asking you to share how much time it actually took you to complete the homework

Remember...

- **Pay attention to every little detail in homework assignments**
- **Do not share code under any circumstance**
- **Validate your submission on ICON**

Python Workshop Opportunities

Tippie Python Workshop Series (in Collaboration with IIAI)

Workshop #1: Introduction to Data Programming in Python

- WHO: Anyone interested in learning Python as a tool for data analytics

- WHEN: Fri, Feb 19 from 1 pm to 5 pm

- WHERE: online via Zoom

- REGISTRATION: by Sun, Feb 14 on https://forms.gle/uiurjCrWMq3JWvdD6

Workshop #2: Web Scraping Using Python

- WHO: Anyone interested in using Python for web scraping

- WHEN: Fri, Apr 23 from 1 pm to 5 pm

- WHERE: online via Zoom

- WHAT: Some specific financial applications will be covered as examples. Suggestions on specific applications will be collected during the first workshop.

- REGISTRATION: by Sun, Apr 18 on https://forms.gle/XjZjGgk4Y2xEniay8

An email with workshop details will be sent out to registrants prior to each workshop.

Please contact Dr. Kang Lee at kangpyo-lee@uiowa.edu with questions.