SWS3023
WEB MINING

INTRODUCTION

ABOUT ME

LEK Hsiang Hui

Senior Lecturer

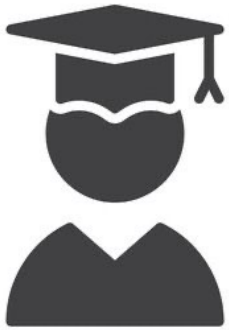
Department of Information Systems and Analytics

Teaches Courses on:

- Analytics
- Software Engineering

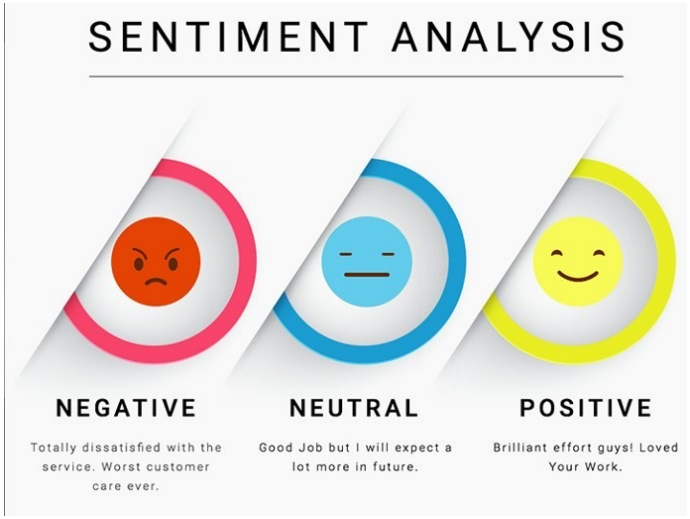
Undergraduates and Executives

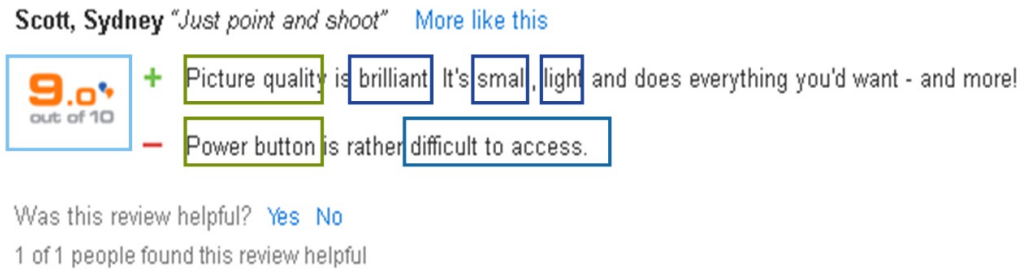
ABOUT ME



Did PhD in the area Natural Language Processing (Sentiment Analysis)

ABOUT ME





Picture Quality — brilliant

Weight — light

Size — small

Power Button — difficult to access

about us

Based in Singapore's silicon valley, Blk 73 Launchpad, we are a big data analytics company, revolutionizing global and local brand analytics ranking and profiling. With billions of online conversations, we have the data and technologies to determine any brands' relative brand position globally. We have the ability to automatically conduct extensive user profiling for brands wanting to know more about their customers' preferences and desires. Through workshops and trainings, we help brands make sense of data and get them started with digital marketing.

what we offer

CUSTOMER INSIGHTS

Want to know what netizens are talking about your brand, competitors or a brand? How can you improve your services or business offerings?

INDUSTRY KNOWLEDGE

What is the latest trend happening - for your brand, your industry or target audience? What can we learn from others in your industry to help you in your marketing strategy?

INFLUENCER PROFILING

Who are your influencers? Can you turn them into leads? Or use them to bring in leads?

Data made simple

STANDICES

how we work + how we can help you

REPORTS

With our technology and the immense data that are available on the web, we can provide customer insights, industry knowledge and influencer profiling.

TRAININGS

We conduct trainings and workshops to help brands solve problems using their data and to provide knowledge to propel your marketing strategies. Let us help you understand analytics and marketing tactics.

TALK TO US

day to find out more about the services that We'll be glad to life simple just over coffee and tea. You'd just need to let us do the thinking!

email: enquiry@standices.com

Also Co-founded a Big Data Analytics Company

CONTACT

https://www.facebook.com/hsianghui

@hsianghui

https://www.linkedin.com/in/hsianghui/

hsianghui@nus.edu.sg

TEACHING ASSISTANT

Mr TAN Qiu Yu

e0261614@u.nus.edu

OBJECTIVES AND OVERVIEW

Aim:

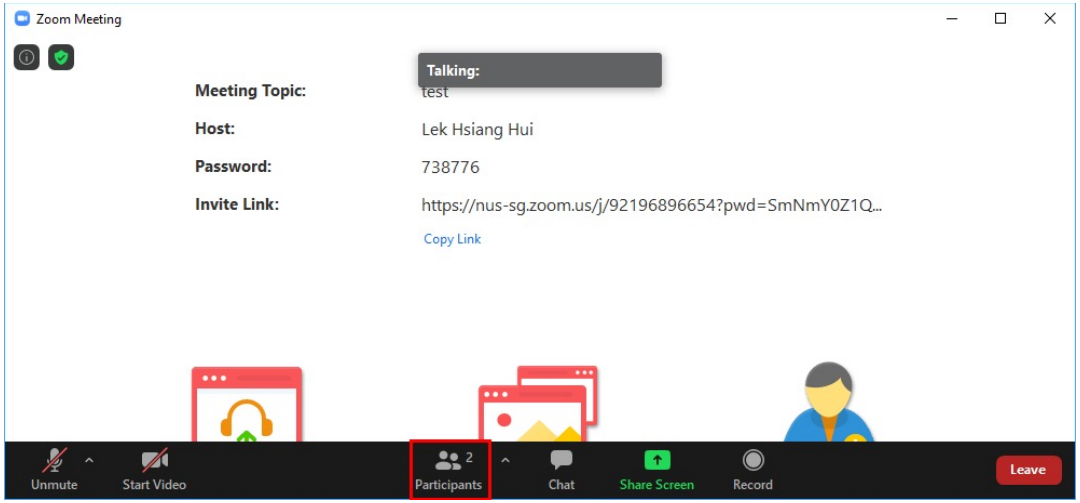
- Provide a good understanding of predictive analytics and its applications
- Provide a systematic approach to mine web content

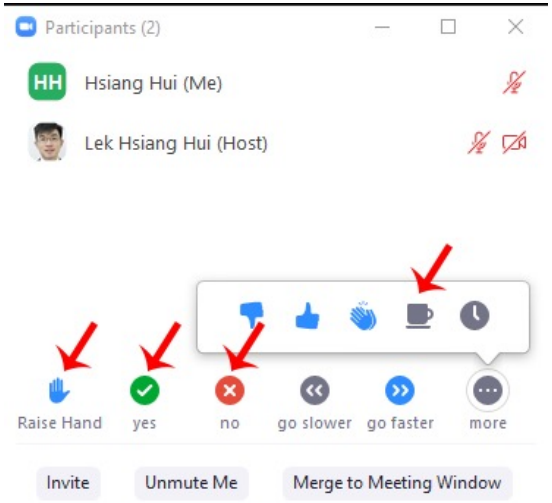
Objectives

- Understand the conceptual foundations of predictive analytics
- Be able to programmatically mine web content
- Apply predictive modeling on web data

SOME HOUSE RULES (ZOOM)

Please respond in Zoom





TOPICS

Introduction to Analytics (5 June)

Introduction to Web Mining (5 June)

Mining Web Content I (5 June)

Predictive Analytics I

Predictive Analytics II

Mining Web Content II

Mining Web Content III

Recommender Systems

Lectures include interactive in class activities

TEACHING/CONSULTATION SCHEDULE

July 2021					
Mon	Tues	Wed	Thurs	Fri	Sat
12 **Predictive Analytics I** (10-12pm) **1st consultation** (1-6pm)	13 **Predictive Analytics II** (10-12pm) **Lab 1 / 1st consultation** (1-6pm)	14 **Mining Web Content II** (10-12pm) **Lab 2 / Ad hoc consultation** (1-6pm)	15 **Mining Web Content III** (10-12pm) **Lab 3 / 2nd consultation** (1-6pm)	16 **Recommender System** (10am-12pm) **Lab 4 / 2nd consultation** (1-6pm)	17 **Ad hoc consultation** (10-6pm)
19 **Lab 5 / 3rd consultation** (10am-6pm)	20 **Lab 6 / 3rd consultation** (10am-6pm)	21 **Lab 7 / Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	22 **4th consultation** **Ad hoc help from TA** (10am-6pm)	23 **4th Consultation** **Ad hoc help from TA** (10am-6pm)	24 **Ad hoc consultation** (10-6pm)
26 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	27 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	28 **Project Showcase** (12-6pm)	29	26	26

TEACHING/CONSULTATION SCHEDULE

July 2021					
Mon	Tues	Wed	Thurs	Fri	Sat
12 **Predictive Analytics I** (10-12pm) **1st consultation** (1-6pm)	13 **Predictive Analytics II** (10-12pm) **Lab 1 / 1st consultation** (1-6pm)	14 **Mining Web Content II** (10-12pm) **Lab 2 / Ad hoc consultation** (1-6pm)	15 **Mining Web Content III** (10-12pm) **Lab 3 / 2nd consultation** (1-6pm)	16 **Recommender System** (10am-12pm) **Lab 4 / 2nd consultation** (1-6pm)	17 **Ad hoc consultation** (10-6pm)
19 **Lab 5 / 3rd consultation** (10am-6pm)	20 **Lab 6 / 3rd consultation** (10am-6pm)	21 **Lab 7 / Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	22 **4th consultation** **Ad hoc help from TA**	23 **4th Consultation** **Ad hoc help from TA**	24 **Ad hoc consultation** (10-6pm)
26 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	27 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	28 **Project Showcase** (12-6pm)			

3 additional lectures that provide the classic techniques to perform predictive analytics and/or handle web data (in red)

TEACHING/CONSULTATION SCHEDULE

July 2021					
Mon	Tues	Wed	Thurs	Fri	Sat
12 **Predictive Analytics I** (10-12pm) **1st consultation** (1-6pm)	13 **Predictive Analytics II** (10-12pm) **Lab 1 / 1st consultation** (1-6pm)	14 **Mining Web Content II** (10-12pm) **Lab 2 / Ad hoc consultation** (1-6pm)	15 **Mining Web Content III** (10-12pm) **Lab 3 / 2nd consultation** (1-6pm)	16 **Recommender System** (10am-12pm) **Lab 4 / 2nd consultation** (1-6pm)	17 **Ad hoc consultation** (10-6pm)
19 **Lab 5 / 3rd consultation** (10am-6pm)	20 **Lab 6 / 3rd consultation** (10am-6pm)	21 **Lab 7 / Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	22 **4th consultation** **Ad hoc help fro**	23 **4th Consultation** **Ad hoc help from TA**	24 **Ad hoc consultation** (10-6pm)
26 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	27 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	28 **Project Showcase** (12-6pm)			

2 additional lectures that teach you how to mine pretty much any website (in red)

TEACHING/CONSULTATION SCHEDULE

July 2021					
Mon	Tues	Wed	Thurs	Fri	Sat
12 **Predictive Analytics I** (10-12pm) **1st consultation** (1-6pm)	13 **Predictive Analytics II** (10-12pm) **Lab 1 / 1st consultation** (1-6pm)	14 **Mining Web Content II** (10-12pm) **Lab 2 / Ad hoc consultation** (1-6pm)	15 **Mining Web Content III** (10-12pm) **Lab 3 / 2nd consultation** (1-6pm)	16 **Recommender System** (10am-12pm) **Lab 4 / 2nd consultation** (1-6pm)	17 **Ad hoc consultation** (10-6pm)
19 **Lab 5 / 3rd consultation** (10am-6pm)	20 **Lab 6 / 3rd consultation** (10am-6pm)	21 **Lab 7 / Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	22 **4th consultation** **Ad hoc help fr**	23 **4th Consultation** **Ad hoc help from TA**	24 **Ad hoc consultation** (10-6pm)
26 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	27 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	28 **Project Showcase** (12-6pm)			

Lab sessions equip you with hands on skills how to handle data with Python (e.g. Basics of Python, Pandas, etc)

TEACHING/CONSULTATION SCHEDULE

July 2021					
Mon	Tues	Wed	Thurs	Fri	Sat
12 **Predictive Analytics I** (10-12pm) **1st consultation** (1-6pm)	13 **Predictive Analytics II** (10-12pm) **Lab 1 / 1st consultation** (1-6pm)	14 **Mining Web Content II** (10-12pm) **Lab 2 / Ad hoc consultation** (1-6pm)	15 **Mining Web Content III** (10-12pm) **Lab 3 / 2nd consultation** (1-6pm)	16 **Recommender System** (10am-12pm) **Lab 4 / 2nd consultation** (1-6pm)	17 **Ad hoc consultation** (10-6pm)
19 **Lab 5 / 3rd consultation** (10am-6pm)	20 **Lab 6 / 3rd consultation** (10am-6pm)	21 **Lab 7 / Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	22 **4th consultation** **Ad hoc help from TA**	23 **4th Consultation** **Ad hoc help from TA**	24 **Ad hoc consultation** (10-6pm)
26 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	27 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	28 **Project Showcase** (12-6pm)			

Lab sessions to learn how to perform modeling with Python

TEACHING/CONSULTATION SCHEDULE

July 2021					
Mon	Tues	Wed	Thurs	Fri	Sat
12 **Predictive Analytics I** (10-12pm) **1st consultation** (1-6pm)	13 **Predictive Analytics II** (10-12pm) **Lab 1 / 1st consultation** (1-6pm)	14 **Mining Web Content II** (10-12pm) **Lab 2 / Ad hoc consultation** (1-6pm)	15 **Mining Web Content III** (10-12pm) **Lab 3 / 2nd consultation** (1-6pm)	16 **Recommender System** (10am-12pm) **Lab 4 / 2nd consultation** (1-6pm)	17 **Ad hoc consultation** (10-6pm)
19 **Lab 5 / 3rd consultation** (10am-6pm)	20 **Lab 6 / 3rd consultation** (10am-6pm)	21 **Lab 7 / Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	22 **4th consultation** **Ad hoc help from TA**	23 **4th Consultation** **Ad hoc help from TA**	24 **Ad hoc consultation** (10-6pm)
26 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	27 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	28 **Project Showcase** (12-6pm)			

Lab sessions to practise web scraping

GROUP PROJECT

Work in groups of 4 (tentatively)

Self-propose an interesting analytics problem that requires some form of predictive analytics using web data

- Formulate a strategy to mine the web data
- Use appropriate analytics techniques to process the data
- Guidance will be provided along the way
- Planned consultation with the lecturer, Ad-hoc consultation with the lecturer/TA

4 planned consultation with lecturer

- 1st consultation – Problem formation and project scoping
- 2nd consultation – Formulate strategies to mine sites (web scraping)
- 3rd consultation – Project fine-tuning (analysis)
- 4th consultation – Project fine-tuning (presentation)

TEACHING/CONSULTATION SCHEDULE

July 2021					
Mon	Tues	Wed	Thurs	Fri	Sat
12 **Predictive Analytics I** (10-12pm) **1st consultation** (1-6pm)	13 **Predictive Analytics II** (10-12pm) **Lab 1 / 1st consultation** (1-6pm)	14 **Mining Web Content II** (10-12pm) **Lab 2 / Ad hoc consultation** (1-6pm)	15 **Mining Web Content III** (10-12pm) **Lab 3 / 2nd consultation** (1-6pm)	16 **Recommender System** (10am-12pm) **Lab 4 / 2nd consultation** (1-6pm)	17 **Ad hoc consultation** (10-6pm)
19 **Lab 5 / 3rd consultation** (10am-6pm)	20 **Lab 6 / 3rd consultation** (10am-6pm)	21 **Lab 7 / Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	22 **4th consultation** **Ad hoc help from TA** (10am-6pm)	23 **4th Consultation** **Ad hoc help from TA** (10am-6pm)	24 **Ad hoc consultation** (10-6pm)
26 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	27 **Ad hoc consultation** **Ad hoc help from TA** (10am-6pm)	28 **Project Showcase** (12-6pm)	29	26	26

ASSESSMENT

Mainly project

- Participation (in labs/consultation, individual): 10%
- Analysis and Results: 60%
- Project presentation: 30%