

SWS3023

WEB MINING

INTRODUCTION

ABOUT ME



LEK Hsiang Hui

Senior Lecturer

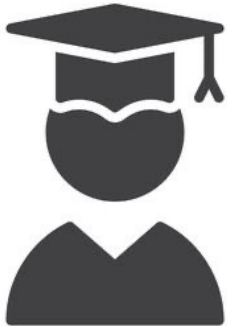
Department of Information Systems and Analytics

Teaches Courses on:

- Analytics
- Software Engineering

Undergraduates, Graduates, and Executives

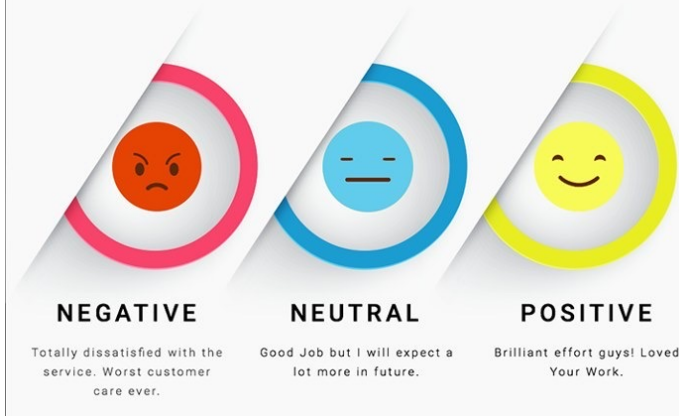
ABOUT ME



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

ABOUT ME

SENTIMENT ANALYSIS



Scott, Sydney "Just point and shoot" [More like this](#)



+

Picture quality is brilliant. It's small, light and does everything you'd want - and more!

-

Power button is rather difficult to access.

Was this review helpful? [Yes](#) [No](#)

1 of 1 people found this review helpful



Picture Quality

brilliant



Size



small

Weight

light



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



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

It's just a day to find out more about the services that we offer. We'll be glad to make your 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

✉ qiuyu.87@gmail.com

OBJECTIVES AND OVERVIEW

Aim:

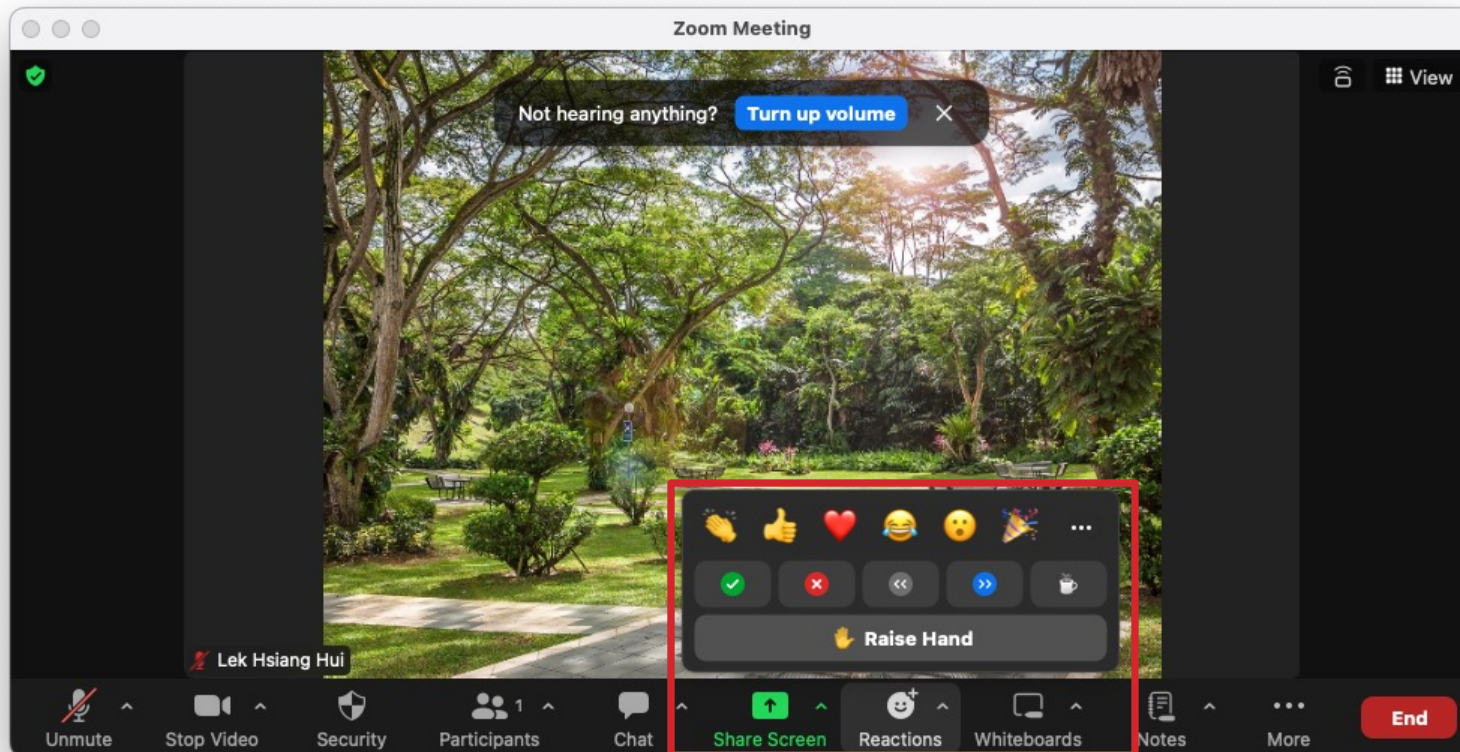
- Provide a good understanding of predictive analytics and its applications
- Provide a systematic approach to mine web content

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
and Web Mining (12 May)**

Predictive Analytics I

Predictive Analytics II

Mining Web Content I

Mining Web Content II

Mining Web Content III

Recommender Systems



Lectures include
interactive in-class
activities

TEACHING/CONSULTATION SCHEDULE*

July 2024				
Mon	Tues	Wed	Thurs	Fri
1	2 CRISP-DM & Predictive Analytics I Project Group Formation (10am-12pm) Lab 1 (1-3pm)	3 Predictive Analytics II (10am-12pm) Lab 2 (1-3pm) Consultation (3-6pm)	4 Mining Web Content I (10am-12pm) Lab 3 (1-3pm) Consultation (3-6pm)	5 Mining Web Content II (10am-12pm) Lab 4 (1-3pm) Consultation (3-6pm)
8 Mining Web Content III (10am-12pm) Lab 5 (1-3pm) Consultation (3-6pm)	9 Recommender System (10am-12pm) Lab 6 (1-3pm) Consultation (3-6pm)	10 Consultation (10-6pm)	11 Consultation (10-6pm)	12 Consultation (10-6pm)
15 Consultation (10-6pm)	16 Consultation (10-6pm)	17 Consultation (10-6pm)	18 Consultation (10-6pm)	19 Poster Submission
22	25	26	27	28

TEACHING/CONSULTATION SCHEDULE*

July 2024				
Mon	Tues	Wed	Thurs	Fri
1	2 CRISP-DM & Predictive Analytics I Project Group Formation (10am-12pm) Lab 1 (1-3pm)	3 Predictive Analytics II (10am-12pm) Lab 2 (1-3pm) Consultation (3-6pm)	4 Mining Web Content I (10am-12pm) Lab 3 (1-3pm) Consultation (3-6pm)	5 Mining Web Content II (10am-12pm) Lab 4 (1-3pm) Consultation (3-6pm)
8 Mining Web Content III (10am-12pm) Lab 5 (1-3pm) Consultation (3-6pm)	9 Recommender System (10am-12pm) Lab 6 (1-3pm) Consultation (3-6pm)	10 Consultation (10-6pm)	11 Consultation (10-6pm)	12 Consultation (10-6pm)
15 Consultation (10-6pm)	16 Consultation (10-6pm)	17 Consultation (10-6pm)	18 Consultation	19 Roster
22	25	26	<div> 3 additional lectures that provide the classic techniques to perform predictive analytics and/or handle web data (in red) </div>	

TEACHING/CONSULTATION SCHEDULE*

July 2024

Mon	Tues	Wed	Thurs	Fri
1	2 CRISP-DM & Predictive Analytics I Project Group Formation (10am-12pm) Lab 1 (1-3pm)	3 Predictive Analytics II (10am-12pm) Lab 2 (1-3pm) Consultation (3-6pm)	4 Mining Web Content I (10am-12pm) Lab 3 (1-3pm) Consultation (3-6pm)	5 Mining Web Content II (10am-12pm) Lab 4 (1-3pm) Consultation (3-6pm)
8 Mining Web Content III (10am-12pm) Lab 5 (1-3pm) Consultation (3-6pm)	9 Recommender System (10am-12pm) Lab 6 (1-3pm) Consultation (3-6pm)	10 Consultation (10-6pm)	11 Consultation (10-6pm)	12 Consultation (10-6pm)
15 Consultation (10-6pm)	16 Consultation (10-6pm)	17 Consultation (10-6pm)	18 Consultation (10-6pm)	19 Poster Submission
22	3 additional lectures that teach you how to mine pretty much any website (in red)			27
				28

TEACHING/CONSULTATION SCHEDULE*

July 2024				
Mon	Tues	Wed	Thurs	Fri
1	2 CRISP-DM & Predictive Analytics I Project Group Formation (10am-12pm) Lab 1 (1-3pm)	3 Predictive Analytics II (10am-12pm) Lab 2 (1-3pm) Consultation (3-6pm)	4 Mining Web Content I (10am-12pm) Lab 3 (1-3pm) Consultation (3-6pm)	5 Mining Web Content II (10am-12pm) Lab 4 (1-3pm) Consultation (3-6pm)
8 Mining Web Content III (10am-12pm) Lab 5 (1-3pm) Consultation (3-6pm)	9 Recommender System (10am-12pm) Lab 6 (1-3pm) Consultation (3-6pm)	10 Consultation (10-6pm)	11 Consultation (10-6pm)	12 Consultation (10-6pm)
15 Consultation (10-6pm)	16 Consultation (10-6pm)	17 Consultation (10-6pm)	18 Consultation (10-6pm)	19 Poster Submission
22	25	<div> Lab sessions equip you with hands on skills how to handle data with Python (e.g. Basics of Python, Pandas, etc) </div>		28

TEACHING/CONSULTATION SCHEDULE*

July 2024				
Mon	Tues	Wed	Thurs	Fri
1	2 CRISP-DM & Predictive Analytics I Project Group Formation (10am-12pm) Lab 1 (1-3pm)	3 Predictive Analytics II (10am-12pm) Lab 2 (1-3pm) Consultation (3-6pm)	4 Mining Web Content I (10am-12pm) Lab 3 (1-3pm) Consultation (3-6pm)	5 Mining Web Content II (10am-12pm) Lab 4 (1-3pm) Consultation (3-6pm)
8 Mining Web Content III (10am-12pm) Lab 5 (1-3pm) Consultation (3-6pm)	9 Recommender System (10am-12pm) Lab 6 (1-3pm) Consultation (3-6pm)	10 Consultation (10-6pm)	11 Consultation (10-6pm)	12 Consultation (10-6pm)
15 Consultation (10-6pm)	16 Consultation (10-6pm)	17 Consultation (10-6pm)	18 Consultation	19 Poster Session
22	25	26		

Lab sessions to learn how to perform modeling with Python

TEACHING/CONSULTATION SCHEDULE*

July 2024				
Mon	Tues	Wed	Thurs	Fri
1	2 CRISP-DM & Predictive Analytics I Project Group Formation (10am-12pm) Lab 1 (1-3pm)	3 Predictive Analytics II (10am-12pm) Lab 2 (1-3pm) Consultation (3-6pm)	4 Mining Web Content I (10am-12pm) Lab 3 (1-3pm) Consultation (3-6pm)	5 Mining Web Content II (10am-12pm) Lab 4 (1-3pm) Consultation (3-6pm)
8 Mining Web Content III (10am-12pm) Lab 5 (1-3pm) Consultation (3-6pm)	9 Recommender System (10am-12pm) Lab 6 (1-3pm) Consultation (3-6pm)	10 Consultation (10-6pm)	11 Consultation (10-6pm)	12 Consultation (10-6pm)
15 Consultation (10-6pm)	16 Consultation (10-6pm)	17 Consultation (10-6pm)	18 Consultation (10-6pm)	19 Poster Submission
22	25	26	<div>Lab sessions to practice web scraping</div>	
				28

TEACHING/CONSULTATION SCHEDULE*

July 2024				
Mon	Tues	Wed	Thurs	Fri
1	2 CRISP-DM & Predictive Analytics I Project Group Formation (10am-12pm) Lab 1 (1-3pm)	3 Predictive Analytics II (10am-12pm) Lab 2 (1-3pm) Consultation (3-6pm)	4 Mining Web Content I (10am-12pm) Lab 3 (1-3pm) Consultation (3-6pm)	5 Mining Web Content II (10am-12pm) Lab 4 (1-3pm) Consultation (3-6pm)
8 Mining Web Content III (10am-12pm) Lab 5 (1-3pm) Consultation (3-6pm)	9 Recommender System (10am-12pm) Lab 6 (1-3pm) Consultation (3-6pm)	10 Consultation (10-6pm)	11 Consultation (10-6pm)	12 Consultation (10-6pm)
15 Consultation (10-6pm)	16 Consultation (10-6pm)	17	18	19 Poster Submission
22	25	<div>Lab session on Recommender Systems</div>		28

GROUP PROJECT

Work in groups of 4-5 (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 consultations with the lecturer, ad-hoc consultations with the lecturer/TA

4 planned consultations 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)

ASSESSMENT*

Mainly project

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