UNSTRUCTURED DATA ANALYTICS (STQD6114)
SEMESTER II: ACADEMIC YEAR 2024/2025

Name and Course Code: Unstructured Data Analytics, STQD 6114

Lecturer: Dr. Nor Hamizah Miswan

E-mail: norhamizah@ukm.edu.my

Course Evaluation:

• Project 1	15%
• Project 2	20%
• Final Project	35%
• Test 1	10%
• Test 2	20%

Short synopsis:
The aim of this course is to introduce students to basic and current methods used to compile, summarize and analyze unstructured and semi-structured data. Unstructured data includes texts, images and audios/videos. Focus are given to algorithms and techniques for mining, exploring and analyzing unstructured data using suitable packages. Students are also exposed to sources for unstructured data. Related applications of unstructured data such as sentiment analysis, document clustering and information extraction are also discussed.

Reference Books:

1. Miner, G. 2012. *Practical Text Mining and Statistical Analysis for Non-structured Text Data Application*. Academic Press.

2. Ingersoll, G.S., Morton, T.S. & Farris, A.L. 2013. *Taming Text*. Manning Publications.

3. Russell, M.A. 2011. *Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites*. O'Reilly.

4. Sarkar, D. 2016. *Text Analytics with Python: A Practical Real-World Approach to Gaining Actionable Insights from your Data*. Bangalore: Apress.

5. Struhl, S. 2015. *Practical Text Analytics: Interpreting Text and Unstructured Data for Business Intelligence (Marketing Science)*. London: Kogan Page.

HPK: Students should be

1. Able to differentiate structured and unstructured data including the sources of the data.
2. Able to perform techniques for unstructured data mining.
3. Able to apply suitable techniques and algorithms on real data.

Course Contents

Lecture Week	Contents	Remarks
Week 3 13/4 – 9.00 am	Discussions on Structured and unstructured Data	
Week 3 13/4 – 2.00 pm	Unstructured Text Mining	
Week 6 4/5 – 9.00 am	Unstructured Text Exploration	
Week 6 4/5 – 2.00 pm	Text Data Analysis: Topic Modelling	
Week 8 17/5 – 18/5	**MID SEMESTER BREAK**	
Week 11 8/6 – 9.00 am	Text Data Analysis: Text Cluster Analysis	Test 1 Project 1
Week 11 8/6 – 2.00 pm	Text Data Analysis: Sentiment Analysis	
Week 13 22/6 – 9.00 am	Image Data Analysis	
Week 13 22/6 – 2.00 pm	Audio Data Analysis	Project 2
Week 14 29/6 – 9.00 am	Audio Data Analysis	Test 2
Week 14 29/6 – 2.00 pm	Project presentation	Final Project