

TULIP Lab, Deakin University
<https://www.tulip.org.au>

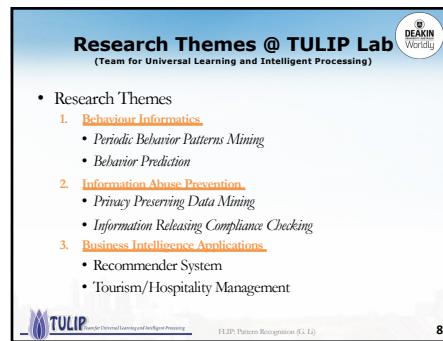
- Prof. Gang Li
 - IEEE Technical Committee
 - IEEE Task Force on EDM (Chair 2020-2023)
 - Data Mining and Big Data Analytics (Vice Chair 2017)
 - Enterprise Architecture and Engineering
 - Enterprise Information Systems



7

Research Themes @ TULIP Lab
 (Team for Universal Learning and Intelligent Processing)

- Research Themes
 1. **Behaviour Informatics**
 - Periodic Behavior Patterns Mining
 - Behavior Prediction
 2. **Information Abuse Prevention**
 - Privacy Preserving Data Mining
 - Information Releasing Compliance Checking
 3. **Business Intelligence Applications**
 - Recommender System
 - Tourism/Hospitality Management



8

Deakin University, Australia

- Melbourne named world's most liveable city for a record seven years in a row



Rank	City	Score
1	Melbourne, Australia	9.3
2	Sydney, Australia	9.4
3	Vancouver, Canada	9.3
4	Toronto, Canada	9.2
5	Singapore, Singapore	9.0
6	Auckland, New Zealand	8.9
7	Brisbane, Australia	8.9
8	Helsinki, Finland	8.8
9	Vienna, Austria	8.6
10	Frankfurt, Germany	8.6
11	Sydney, Australia	8.6

9

Deakin University, Australia



10

Deakin University, Australia

We live 15 miles from here....



11

Deakin University, Australia

We live 125 miles from here....



12

History of Deakin

- Established in the 1970s as ‘one of the new generation of Australian universities’
 - A pioneer in distance education
 - Named after **Alfred Deakin**, three-time Australian Prime Minister
 - A comprehensive University
 - Excellent teaching and research
 - New ways of developing and delivering courses
 - Operating revenue of more than A\$830 million

 FLIP: Pattern Recognition (G 1.2)

13

Our rankings

A growing reputation

 FLIP: Pattern Recognition (G 1.2)



 FLIP: Pattern Recognition (G 1.2)

14

Our rankings

Table 1: Major Rankings of Deakin University

Benchmark	2014-2015	2015-2016	2016-2017	2017-2018
ARWU	World 301-400	201-300	201-300	201-300
	National 12-19	9-14	11-14	10-15
CWTS Leiden	World 284	325	314	226
	National 11	17	17	11
QS#	World 360	324	355	293
	National 19	17	19	17
THE	World 301-350	301-350	251-300	301-350
	National 13-15	18-19	12-18	17-21
THE<50	World 45	50	48	50
	National 6	8	8	8

<https://www.australianuniversities.com.au/rankings/>

 FLIP: Pattern Recognition (G 1.2)

15

What?



From Data to Knowledge
Selected P.R. Applications

 FLIP: Pattern Recognition (G 1.2)

16

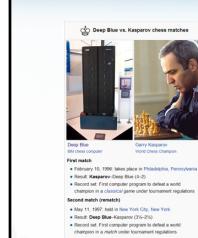
Artificial Intelligence



 FLIP: Pattern Recognition (G 1.2)

17

Intelligent Applications



Deep Blue vs. Kasparov chess matches

First match:

- Deep Blue 1996: IBM place in Philadelphia, Pennsylvania
- Result: Kasparov - Deep Blue (3-2)
- Reason: Kasparov claimed that a small change in a classical game rule tournament regulations

Second match (1997):

- IBM place in New York, New York
- Result: Deep Blue - Kasparov (3½-2½)
- Reason: Kasparov claimed that a small change in a match under tournament regulations

 FLIP: Pattern Recognition (G 1.2)

18

Intelligent Applications

Watson (computer)
From Wikipedia, the free encyclopedia
"IBM Watson" redirects here. For the television, see [Thomas J. Watson](#).
Watson is a supercomputer intelligent computer system capable of answering questions posed in natural language.¹ Developed in IBM's DeepQA project by a research team led by principal investigator David Ferrucci, it is based on technologies developed for the [Lotus Notes](#) system. The computer system was specifically developed to answer questions on the quiz show [Jeopardy!](#).² In 2011, Watson competed on [Jeopardy!](#) against former champions Brad Rutter and Ken Jennings.³ Watson received the first prize of \$1 million.⁴

Stanley (vehicle)
From Wikipedia, the free encyclopedia
Stanley is an autonomous car created by Stanford University's Stanford Racing Team in cooperation with the Volkswagen Group. It was the first robot to cross the finish line in the 2005 DARPA Grand Challenge.¹ It earned the Stanford Racing Team the \$2 million dollar prize.

Watson, Ken Jennings, and Brad Rutter, all in their Jeopardy! exhibition match.
Stanley parked after the 2005 DARPA Grand Challenge.

TULIP: Center for Universal Learning and Intelligent Processing

19

19

Data, Information and Knowledge

- Data** are raw facts and figures that on their own have no meaning
 - these can be any alphanumeric characters
 - i.e. text, numbers, symbols
- For example:
 - Yes, Yes, No, Yes, No, Yes, No, Yes
 - 42, 63, 96, 74, 56, 86
 - 111192, 111234
- None of the above data sets have any meaning until they are given a **CONTEXT** and **PROCESSED** into a usable form

TULIP: Center for Universal Learning and Intelligent Processing

20

20

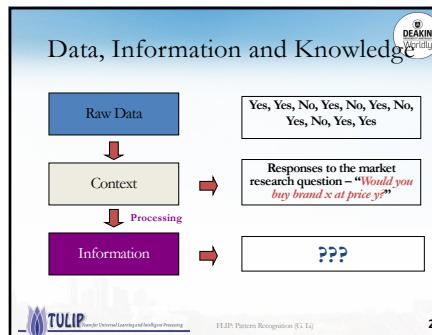
Data, Information and Knowledge

- Information** is data that has been processed within a context to give it meaning, or data that has been processed into a form that gives it meaning
 - Accurate, relevant, and timely *information* is key to good decision making

TULIP: Center for Universal Learning and Intelligent Processing

21

21



22

Data, Information and Knowledge

- Knowledge** is the appropriate collection of information, such that its intent is to be useful
 - Or put it simply, *useful information*, for example
 - Based on previous information, a Marketing Manager could use this information to decide whether or not to raise or lower price y

consciousness

information (understanding patterns)

wisdom (understanding principles)

knowledge (understanding patterns)

understanding

Data

Analysis

Experience

TULIP: Center for Universal Learning and Intelligent Processing

23

23

Data: The New Currency

- Data is everywhere:
 - Facebook, blogs, emails, cars, sensors, ship, genome sequence, your Fitbit, and even your body.
 - Everyone talks about data these days, and people think that data is the new currency!

Data: the new currency?

TULIP: Center for Universal Learning and Intelligent Processing

24

Data Science

- Data science
 - the Fourth Paradigm of Science
 - *The intellectual and practical activity encompassing the systematic study of facts and statistics collected together for reference or analysis.* (1942-2012)

25

CSIR-4PI (India)

26

Confluence of Multiple Disciplines

27

Confluence of Multiple Disciplines

- Art & Science of **Refining Data to Knowledge**
 - **Pattern Recognition**, Machine Learning, Data Mining, Multivariate Statistics, Statistical Inference
 - Exploring Data → Visualization
 - Discover Patterns in Data → Unsupervised Learning
 - Learning Rules (from Training Data) that allow future predictions → Supervised Learning
 - Different Origins
 - P.R. has origins in *Engineering*
 - M.L./D.M. has origins in *Computer Science*
 - Statistics Learning has origins in *Statistics*.

28

P.R. Applications
(Case 1: Face Age Recognition)

Xin Geng, Zhi-hua Zhou, Yu Zhang, Gang Li, Benghua Guo, *Learning from facial aging patterns for automatic age estimation*, Proceedings of the 14th annual ACM international conference on Multimedia, 2006

29

P.R. Applications
(Case 2: Speech-based Emotion Detection)

Jia Rong, Gang Li, Yi-Ping Phoebe Chen, *Acoustic Feature Selection for Automatic Emotion Recognition from Speech*, Information Processing and Management, 46(3), 315-329, 2010

30

P.R. Applications
(Case 3: K-Complex Detection)

Sleep Stage Identification

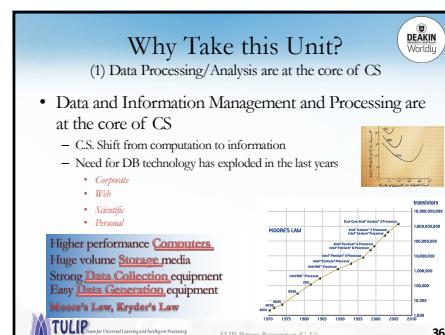
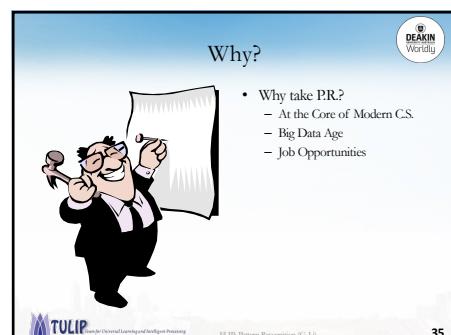
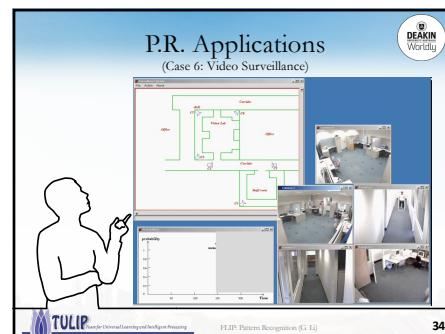
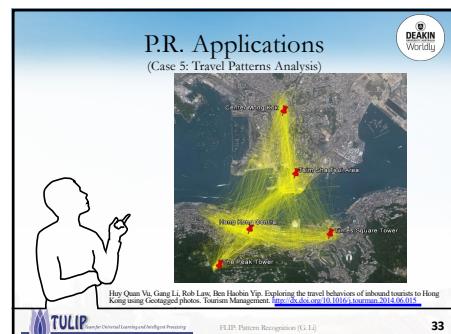
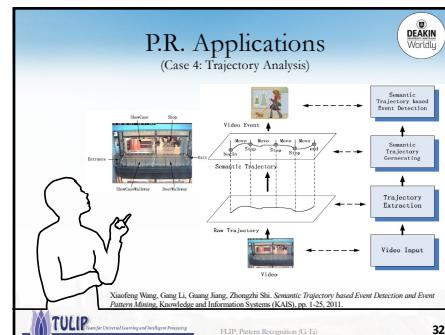
- **Rapid Eye Movement (REM) Sleep**
 - A normal stage of sleep characterized by the rapid movement of the eyes, easily discernible in PSG
 - 20%-25% of total sleep time in human adults, memorable dream
- **Non-REM Sleep**
 - N1 slow eye movement, hypnic jerks at beginning, believe fully awake
 - N2: non eye movement, rare dreams, easily awaken
 - N3: deep sleep, sometimes dreamless but disconnected less vivid
- General sleep order in cycles of REM and NREM: N1
 \rightarrow N2 \rightarrow N3 \rightarrow N2 \rightarrow REM
 - Early in the night, a greater amount of deep sleep (N3)
 - Later in the night, REM sleep increases, and before natural awakening

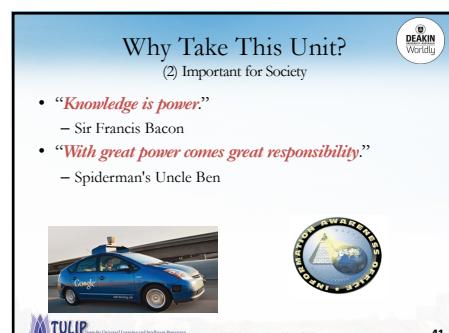
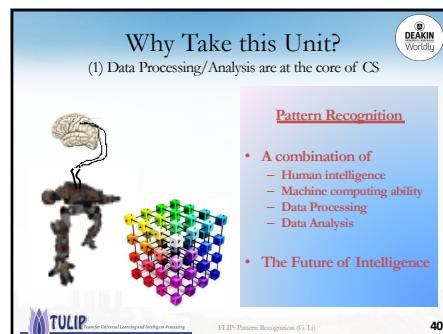
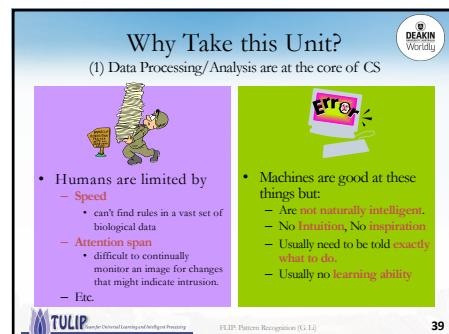
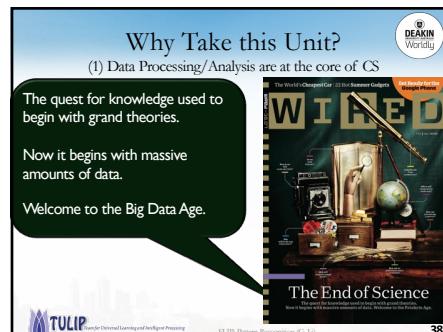
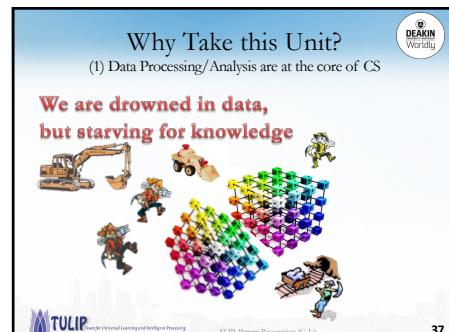
Vu Quoc Li; Gang Sakharukova, Nadezhda Belakova, Gleb Litvinov, Shaojun et al. K-complex Detection using a Hybrid-Synergistic Machine Learning method. IEEE Transactions on SMC(C), P1999: 1-13.

TULIP: Tutorials for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G.1.2)

31





Why Take this Unit?

- (3) Looks good on my CV
- Top Academic Conference Cup

The 11th International Cybersecurity Data Mining Competition (COMAC 2020) logo features a golden seal above a traditional Asian temple tower at sunset. Below the seal, the text reads: "THE 11TH INTERNATIONAL CYBERSECURITY DATA MINING COMPETITION (COMAC 2020)". The competition is described as a research competition organized by the City University of Hong Kong and the University of Deakin, with support from the Chinese University of Hong Kong and the University of Macau. It aims to promote international cooperation in research and innovation in the field of cybersecurity.

43

Why Take this Unit?

- (3) Looks good on my CV
- Top Industry Data Competition

The Tianchi Big Data Competition page features a dark background with a network diagram. The text "Tianchi Big Data Competition" is prominently displayed, along with the subtext "Big data and distributed computing resources, cutting-edge solutions for real-world applications". Below this, there's a section for the "LUCAI-PRICAI 2020 3D AI Challenge: Image-based 3D Shape Retrieval".

44

What future work might look like?

- Use data from the crowd to derive new insights!
- This is also known as “collective intelligence”

A horizontal bar chart titled "What future work might look like?" comparing different treatment types. The y-axis lists treatments: Behavioral therapy, Dietary Supplements (Vitamins, etc.), Calcium with Vitamin D, Educational therapy, Speech/Language Therapy, Omega-3 supplement, Vitamin B-12, Gluten and Casein Free Diet, Antidepressant medications, Play Therapy, Cod Liver Oil, and Music Therapy. The x-axis shows the number of users (n) for each treatment: Behavioral therapy (23), Dietary Supplements (Vitamins, etc.) (16), Calcium with Vitamin D (6), Educational therapy (5), Speech/Language Therapy (4), Omega-3 supplement (7), Vitamin B-12 (12), Gluten and Casein Free Diet (13), Antidepressant medications (14), Play Therapy (11), Cod Liver Oil (4), and Music Therapy (10).

46

What future work might look like?

- Use data from the crowd to derive new insights!
 - Repurposing the data: reusing the data to derive new knowledge.

A scatter plot titled "What future work might look like?" showing the relationship between "Popularity" (x-axis, 0.0 to 0.9) and "Patient Reported Effectiveness" (y-axis, -0.2 to 1.2). The plot area is divided into four quadrants: Top-Left (High Effectiveness, Low Popularity), Top-Right (High Effectiveness, High Popularity), Bottom-Left (Low Effectiveness, Low Popularity), and Bottom-Right (Low Effectiveness, High Popularity). Various treatments are plotted as points, including Acupuncture, Chiropractic, Cognitive Behavioral Therapy, Exercise, Mindfulness, Meditation, Physical Exercise, and Yoga.

47

What future work might look like?

- Use data to connect demand and supply.
 - UCar and UBer

Two side-by-side images illustrating the use of data to connect demand and supply. On the left is a screenshot of the Uber app interface with the text "YOUR RIDE, ON DEMAND". On the right is a screenshot of the Ucar app interface with the text "专人专车 随叫随到". Both images show a woman in a car.

48

What future work might look like?

- Use data to understand customer behaviours.
 - Taobao uses data to build **recommender system**!

A screenshot of a Taobao product listing for Microsoft Surface Pro 4 (i5 256G RAM 8GB). The page shows the product image, price (\$998.00), and a list of recommended products. Recommended items include Microsoft Surface Pro 4 (i5 256G RAM 8GB), and Microsoft Surface Laptop.

49

What future work might look like?

- Summary:
 - Use data to extract new values, insights and hypothesis.
 - Reuse existing data to derive new knowledge.
 - Use data to understand customers' behavior.
 - Use data to facilitate the demand market to suppliers.
 - Use data to build recommender system.
 - Use data to build predictive systems.
 -
 - DATA = NEW CURRENCY?

 FLIP: Pattern Recognition (G 1.2) 50

50

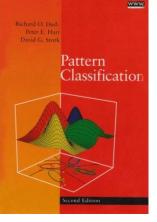
How?

- Textbook
- Session Plan
- Resources

  FLIP: Pattern Recognition (G 1.2) 51

51

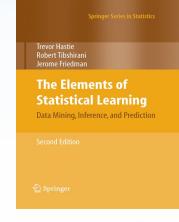
Textbook

- Duda, Hart and Stork, **Pattern Classification**, Second Edition, Wiley, 2001 

 FLIP: Pattern Recognition (G 1.2) 52

52

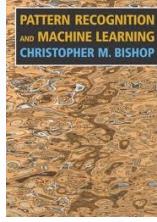
Highly Recommended References

- Trevor Hastie, Robert Tibshirani, Jerome Friedman. *Statistical Learning: Data Mining, Inference, and Prediction*. Second Edition. February 2009, Springer 

 FLIP: Pattern Recognition (G 1.2) 53

53

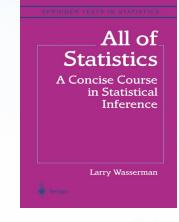
Highly Recommended References

- Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Second Edition. 2004, Springer 

 FLIP: Pattern Recognition (G 1.2) 54

54

Highly Recommended References

- Larry Wasserman, *All of Statistics: A Concise Course in Statistical Inference*, 2004, Springer 

 FLIP: Pattern Recognition (G 1.2) 55

55

Pattern Recognition Topics

Tentative Session Plan

1. Introduction
2. Review of L.A.
3. Review of L.A. and PCA and LDA
4. Review of P.T. and PAC Learning
5. Bayesian Decision Theory
6. Parameter Estimation
7. Parametric Model (1)
8. Parametric Model (2)
9. Parzen Window
10. KNN and Distance
11. Discriminant Function (1)
12. Discriminant Function (2)



JUST DO IT.

TULIP Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G 1.1) 56

56

Unit Materials

- Unit Website:
 - <https://github.com/tulip-lab/pattern-classification>
 - All materials
- Practical References
 - <https://github.com/tulip-lab/flip01>
 - <https://github.com/tulip-lab/flip02>



TULIP Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G 1.1) 58

58

Software Tools

<u>Open source or Free software</u>	<u>Licensed software</u>
<ul style="list-style-type: none"> • Python • R • Julia • GitHub 	<ul style="list-style-type: none"> • Matlab <ul style="list-style-type: none"> – With <i>Statistic Toolbox</i>



TULIP Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G 1.1) 59

59

Unit Assessment

- Assessments
 - **Group Work**
 - Assignment 1:
 - 2 Presentations for each group
 - » 1 Presentation from Seminar Topic
 - » 1 Presentation from Group project Work
 - Assignment 2:
 - Hands on project
 - » Code
 - » Report
 - » Presentation Recordings

TULIP Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G 1.1) 60

60

Plagiarism Notice

- “the practice of taking someone else's work or ideas and passing them off as one's own”.
 - Oxford Dictionary
- It is a serious academic offence, and you can be expelled from the University!



TULIP Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G 1.1) 62

62

Plagiarism Notice

Plagiarism and collusion

Plagiarism is the act of using other people's words, ideas, research findings or information without acknowledgement, that is, without referencing the source. Plagiarism is regarded as a very serious offence in Western academic institutions and Deakin University has procedures and penalties to deal with instances of plagiarism.

In order not to plagiarise, all material from all sources must be correctly referenced. It is necessary to reference direct quotes, paraphrases and summaries of sources; statistics, diagrams, images, experiment results and laboratory data - anything taken from sources.

- How can plagiarism occur?
- Why is it important not to plagiarise?
- Why is plagiarism an issue of concern?
- What is collusion?

<http://www.deakin.edu.au/students/study-support/referencing/plagiarism>

TULIP Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G 1.1) 63

63



Advice 1: Don't be too ambitious with your course load.

 TULIP
Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G.1.2)

64



Advice 2: Don't wait until the last minute to get help

 TULIP
Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G.1.2)

65

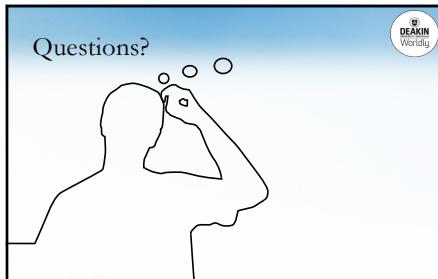


Advice 3: leave yourself time for mistakes.

 TULIP
Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G.1.2)

66



 TULIP
Tool for Universal Learning and Intelligent Processing

FLIP: Pattern Recognition (G.1.2)

67