



COMP90049 Knowledge Technologies

Exam Preparation 2018

Rao Kotagiri

School of Computing and Information Systems
The Melbourne School of Engineering

Some of slides are derived from Prof Vipin Kumar and modified, <http://www-users.cs.umn.edu/~kumar/>

What we have covered so far

1. Exact String Search [Click for more options](#)
2. Approximate String Search [Click for more options](#)
3. Information Retrieval, updated [Click for more options](#) (Older version is [here](#) [Click for more options](#))
4. Web-scale Information Retrieval, updated [Click for more options](#) (Older version is [here](#) [Click for more options](#))

What we have covered so far

- Probability [Click for more options](#)
- Data Mining Fundamentals [Click for more options](#)
- Classification: Naive Bayes/Evaluation [Click for more options](#)
- Decision Trees [Click for more options](#)
- Random Forests [Click for more options](#)
- Geometric Methods [Click for more options](#)
- Clustering [Click for more options \(extra slides below\)](#)
- Association Rules [Click for more options](#)

Topics to prepare

Part 1:

- Need to know how to search for strings using regular expressions.
- Need to know how to compare for similarity of two strings.
- Information retrieval system
- How ranking is done
- What factors can improve precision and recall

Topics to prepare

Part 2:

- Understand various aspects of Data Mining and Machine learning
- Classification –NB; DT; RF; etc. You need to know how to compute everything NB. Need to understand Precision, Recall, Confusion Matrix, etc.
- Association Rule Mining.