

School of Computing and Information Systems
The University of Melbourne
COMP90049 Knowledge Technologies (Semester 2, 2019)
Workshop exercises: Week 7

1. What is link analysis? Briefly describe the PageRank algorithm and the rationale behind it.
2. In the PageRank algorithm:
 - (a) What is the mechanism for the “random walk”?
 - (b) In terms of user behaviour, what is the significance of “teleporting”?
3. [EXTENSION] Review the questions about Video Snippets in Google Results and discuss more possible answers to the four main questions:
 - (a) What knowledge is necessary to provide video snippet recommendation?
 - (b) Where could Google obtain this knowledge from? What data sources could it use?
 - (c) What challenges occur when indexing videos?
 - (d) How could link analysis (like anchor texts) be useful for annotating video snippets?

_____ & _____
4. What is data mining/machine learning? What makes this a knowledge task?
5. Revise the definitions of instances and attributes (or features). For the following problems, identify what the instances and attributes might consist of:
 - (i) Building a system that guesses what the weather (temperature, precipitation, etc.) will be like tomorrow
 - (ii) Predicting products that a customer would be interested in buying, based on other purchases that customer has previously made
 - (iii) Automatically identifying the author of a given piece of literature
6. What are the main differences between supervised and unsupervised machine learning?
7. Based on the problems in question 4, identify the “concept” we might attempt to “learn” for each problem, and conjecture whether a typical strategy is likely to use supervised or unsupervised Machine Learning.