NUI Galway
OÉ Gaillimh

Semester IIExaminations 2016/ 2017

Course Instance Code(s)	1CSD, 1SPE, 2SPE, 3SPE
Exam(s)	MSc in Computer Science (Data Analytics), 1st,2nd,3rd Structured PhD
Module Code(s)	CT5107
Module(s)	Advanced Topics in Machine Learning and Information Retrieval.
Paper No.	1
Repeat Paper	No
External Examiner(s)	Professor Liam Maguire
Internal Examiner(s)	Dr. Michael Schukat *Dr. Colm O'Riordan *Dr. Enda Barrett

Instructions: Answer 2 questions from each section (4 in total). All questions will be marked equally. Use a separate answer book for each section.

Duration	2 hours
No. of Pages	3
Discipline(s)	Information Technology
Course Co-ordinator(s)	Dr. Conor Hayes

Requirements:

Release in Exam Venue	Yes	☐	No	☐
MCQ	Yes	☐	No	☐
Handout	None			
Statistical/ Log Tables	None			
Cambridge Tables	None			
Graph Paper	None			
Log Graph Paper	None			
Other Materials	None			
Graphic material in colour	Yes	☐	No	☐

PTO

Section A: Machine Learning

1.

a) Markov Decision Processes (MDPs) are mathematical frameworks for handling uncertainty. Explain the components of a typical MDP formalism. In your answer you should include a sketch of a simple two state, two action MDP. You should also include a brief discussion of the Markov property and its importance to MDPs.

[7]

b) In Reinforcement Learning explain the difference between model based and model free learning methods. Describe a scenario/problem where the application of one is more suitable than the other.

[5]

c) Parallel reinforcement learning has been proposed as a method for dealing with large state and action spaces. Explain what parallel learning is. Describe the approach proposed by R.M. Kretchmar for combining experiences with multiple agents learning in parallel which ensures that as agents continue to share information their experiences are not duplicates of the same trials.

[7]

d) Explain the purpose of the Kullback-Liebler divergence. In your answer state typically what the distributions P and Q represent. Why is the Kullback-Liebler divergence not considered a true metric?

[6]

2.

a) Explain briefly why linear classifiers when presented with unknown data can produce different classifications. In your answer you should explain why this has led to the development of Linear Support Vector machines.

[5]

b) In the context of Linear Support Vector Machines, explain with a diagram what the support vectors are and what the maximum margin is. Is there more than one solution for the maximum margin hyperplane? What are the consequences of this for finding one?

[8]

c) In a Linear Support Vector Machine, the decision function can be expressed in either the primal or dual form. Provide equations for both forms of the decision function, explaining all terms in the equations. How do the two forms relate to each other?

[8]

d) Describe the “kernel trick” as used in non-linear Support Vector Machines.

[4]

3.

a) Based on your experience, what would you consider to be the most significant two challenges in implementing a neural network? Explain your answer.

[5]

b) Considering an example of a fully-connected feed-forward neural network with 2 input nodes, 3 hidden nodes and one output node, draw a diagram showing all nodes and weights. Write down equations for how the inputs are propagated forward to produce the output.

[6]

c) Is the cost function used in training feed-forward neural networks convex? What are the implications of this? Explain how a training curve may be used to monitor training progress.

[5]

d) Explain what an auto-encoding neural network is. Building on this, describe in detail the principle of pre-training a deep neural network using stacked auto-encoders. As part of your answer, identify which problem(s) encountered in classic shallow neural networks areaddressed by this approach.

[9]

Section B: Information Retrieval

4.

a) In many modern information systems, the following properties hold: distributed collections, many information providers and information seekers, costs and benefits associated with the retrieval (e.g. quality, relevance, and cost). In these domains, there is often a conflict of interests leading to an incentive for a participant to behave unfairly in order to exploit other participants. Choosing a suitable instantiation, discuss an approach to represent and reason about these conflicts. [10]

b) Evolutionary computation represents a form of search loosely inspired by the principles of Darwinian evolution. Discuss briefly the main components of an evolutionary computation approach. [5]

c) Term weighting schemes are fundamental to the performance of an IR system. Outline an approach using evolutionary computation to evolve weighting schemes for information retrieval. Discuss the advantages and limitations of this approach. [10]

5. Traditional Information Retrieval systems adopt a 'bag of words' model, where the context of the words is ignored. More recently, graph models have been adopted to provide a richer expression of documents and queries. With reference to graph models, write notes of any two of the following:

i) Graph models that may be adopted; i.e. given a piece of text what approaches could be used to build a graph
ii) Local graph properties that may be used to capture properties of documents that may be useful in designing better weighting schemes.
iii) Global graph properties that capture features at a document level and potential applications of these properties.
iv) Pre-processing approaches that may be applied to documents and how they may affect graph properties.
v) Weighting schemes based on graph properties and how to combine different features.
vi) Axiomatic approaches to analysing graph based weighting schemes
vii) Applications in information retrieval other than term weighting schemes
viii) Approaches to visualising and its use in graph based information retrieval.

[25]