CSC 589: Introduction to Machine Learning Instructor: Dr. Nathalie Japkowicz

e-mail: ja pokogwio@ta Proejricta Execute Help

https://powcoder.com

Objectives of Preliminaries

Some Information

- Instructor: Dr. Nathalie Japkowicz
- - much!) https://powcoder.com
- **E-mail:** japkowic@american.edu (best way to contact me!)

 Add WeChat powcoder me!)
- Office Hours:
 - Thursdays, 2:30pm-4pm
 - By arrangement, on Skype

Machine Learning: A Case Study

- Malfunctioning gearboxes have been the cause for CH-46 US Navy helicopters to crash.
- Although gearbox malfunctions can be diagnosed by a mechanic prior to melfic prior take of the what if a malfunction occurs while in-flight, when it is impossible for a human to detect?



• Machine Learning was shown to be useful in this domain and thus to have the potential of saving human lives!

How did it Work?

Consider the following common situation:

- You are in your car, speeding away, when you suddenly hear a "funny" noise.
- To prevent an accide 15:1/10 we down and either stop the car or bring it.

■ The in-flight helicopter gearbox fault monitoring system was designed following the same idea. The difference, however, is that many gearbox malfunction cannot be heard by humans

So, Where's the Learning?

•Imagine that, instead of driving your good old battered car, you were asked to drive this truck:



- Would you know a Aften My Chait promoter normal" one?
- •Well, probably not, since you've never driven a truck before!
- While you drove your car during all these years, you effectively <u>learned</u> what your car sounds like and this is why you were able to identify that "funny" noise.

What did the Computer Learn?

- Obviously, a computer cannot hear and can certainly not distinguish between a normal and an abnormal sound.
- Sounds, however, can be <u>represented</u> as wave patterns such as this <u>represented</u> Project Exam Help

https://powcoder.com

Add WeChat powcoder which in fact is a series of real numbers indicating intensity.

- And computers can deal with strings of numbers!
- For example, a computer can easily be programmed to distinguish between strings of numbers that contain a "3" in them and those that don't.

What did the Computer Learn? (Cont'd)

- In the helicopter gearbox monitoring problem, the assumption is that functioning and malfunctioning gearboxes emit different sounds. Thus, the strings of numbers that represent there sounds have different characteristics.
- The exact characteristics of these different categories, however, are unknown and/or are too difficult to describe.
 Therefore, they cannot be programmed, but rather, they
- Therefore, they cannot be programmed, but rather, they need to be learned by the computer.
- There are many ways in which a computer can learn how to distinguish between two patterns (e.g., decision trees, neural networks, bayesian networks, etc.) and that is the topic of this course!

What else can Machine Learning do?

- Medical Diagnostic (e.g., breast cancer detection)
- •Credit Card Fraud Detection Project Exam Help
- Sonar Detection (e.g., submarines versus shrimps (!))
- Speech Recognition (e.g., Telephone automated systems)
- Autonomous Vehieles (Weführtrpmzardder missions or to assist disabled people)
- Personalized Web Assistants (e.g., an automated assistant can assemble personally customized newspaper articles)
- And many more applications...

Text Books and Reading Material

- Peter Flach, Machine Learning: The art and science of algorithms that make sense of data. Cambridge University Fresh, 2012. roject Exam Help
- Nathalie Japkowicz/pod MohakoShah, Evaluating Learning Algorithms: A Classification Perspective, Cambridge Unwersity Phersp2010.der
- Research papers saved in the directory entitled literature on Blackboard
- On Blackboard, you will also find a list of <u>non-required</u> books that you may find useful.

Objectives of the Course:

- To present a broad introduction of the principles and paradigms Aunderlying magained earning including discussions and hands-on evaluations of some of the major approaches currently being investigated.
- To introduce the stweether the relation of the relation of
- To initiate the students to formulating a research problem and carrying this research through.

Format of the Course:

- The course is lecture based.
- Each student will write 6 research paper critiques as part of a group of 2 or 3 students.
- Each student is expected to present in class, one of the 13 papers provided in the literature packet. (Each student will present a different paper). The student presentations will take place during the weeks of **November 26** and **December 3**.
- On the <u>last day of classes</u>, each student will present a poster to the entire class (and the whole department). This poster will be based on the research they will have carried out for their final project.
- There will be an in-class midterm exam on <u>November 1</u>, but no final exam.
- There will be two assignments and a final project.

Contents of the Course I:

- The course will teach machine learning algorithms, theoretical issues and contemporary problems in Assignment Project Exam Help machine learning.
- Machine learning salgor thins covered:

Decision Trees hat powcoder Artificial Neural Networks

Bayesian Learning

Instance-Based Learning

Support Vector Machines

Ensemble-Learning Algorithms

Rule Learning/Associative Rule Mining

Unsupervised Learning/Clustering

Contents of the Course II:

The roots of Machine learning (Philosophy, AI, Computational Learning Theory, Statistics)

The dratical cissultant in the light of the

https://powcoder.com

Add WeChat powcoder

Data Exploration

Data Preparation

Practical issues considered:

The Class Imbalance Problem

Contents of the Course III:

Contemporary themes considered:

We will be specifically looking at the area of big data analysis, and, possibly, at some of the following topics:

Assignment Project Exam Help

Deep Learning
Graph Mining https://powcoder.com
Mining Social Networks
Data Streams Mining WeChat powcoder
Unstructured or Semi-Structured Data Mining
Data Mining with Heterogeneous Sources
Spatio-Temporal Data Mining
Issues of Trust and Provenance in Data Mining
Privacy in Data Mining
Cybersecurity

Course Requirements:

Percent of the Final Grade

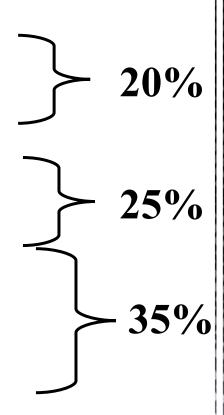
Written commentaries

\rightarrow 12%

- Oral presentation of great the project Exam Help
- 2 Assignments (little programming packages will be provided)

Add WeChat powcoder

- Midterm Exam
- Final Project: Project Proposal
 - Project Report
 - Poster Presentation



More on Assignments and Final Exam:

- Assignment Assignment Project Exam Help
 - Handed out on: Thursday September 27, 2018
 Due on: Thursday Oct 18, 2018
- Assignment 2: Add WeChat powcoder
 - Handed out on: Thursday Oct 18, 2018
 - Due on: Thursday November 15, 2018

Midterm Exam:

In class on Thursday November 1, 2018

More on the Research Project (Also see the Project Description on Blackboard)

- Research Project including a literature review and the design and implementation of a novel learning scheme or the comparison of several existing schemes.
- Projects Propositions of the Projects Proposition of the Projects Projec
- https://powcoder.com
 Project Report are due on December 6, 2018
- Project Presentations Well Chatepplace der December 6, 2018 in the form of a poster presentation.
- Suggestions for project topics are listed on Blackboard, but you are welcome (and that's even better) to propose your own idea.

Start thinking about the project early!!!!!