SYSTEMS DESIGN ENGINEERING

Information, Patterns, Pattern Analysis, Synthesis and Discovery

SYDE 770 - *Topic 6* **Fall**

The objective of this course is to study the fundamental concepts that information measures and statistical patterns can be derived or discovered from data. The study includes first a critical literature survey related to this central theme including some of the pioneering work of Professor Andrew K.C. Wong. The second part of the study will focus in formulating a methodology that will reveal the inherent multiple statistical patterns at the event level and/or probabilistic patterns at the pattern subgroup level explicitly to render greater understanding of the data. The third part involves the identification of a few sets of appropriate data, preferably from the protein families for demonstration purposes. While the survey part applies to both tabulated relational data and multiple sequence data, the algorithmic system proposed on the second part and third part of the study is focussed on sequence data. The target of the study is to come up with a methodology to derive from the sequence data their implicit structural and relational relationships and organize them explicitly in the form of attributed hypergraphs and/or random hypergraphs. While the course will cover the fundamental theories and methodologies, Professor Wong will direct the problem formulation. In addition to the algorithms developed in the study, Professor Wong will provide certain pattern discovery software wherever appropriate to expedite the development pace. The end target of this study is to produce a journal paper quality report that includes the survey, the problem formulation and data identification, hopefully with experimental results that demonstrate the feasibility of the proposed system. The course performance will be evaluated based on the content and the impact of the report.

READING COURSE

INSTRUCTOR: Andrew K.C. Wong
Email: akcwong@uwaterloo.ca