

## FÖRSLAG TILL PROGRAMNÄMND INFÖR ÅR

NÄMND/NÄMNDER:

Förslagsställare (Namn, funktion, Inst/Enhet)

Matthew Cooper, Universitetslektor i Informationsvisualisering, ITN

### FÖRSLAGET GÄLLER:

a) EXISTERANDE KURS (Ange kurskod och kursnamn)

b) NY KURS (Ange kursnamn, årskurs, önskad läsperiod, schemablocksplacering. Bifoga utkast till kursplan.)

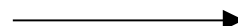
Advanced Visual Data Analysis. Grade A. Starting 2016, VT2 (due to dependency on existing course TNM048 held in VT1).

c) ÄNDRING I EXISTERANDE PROFIL/INRIKTNING (Ange Program och Profil/Inriktning. Bifoga beskrivning över vad förslaget går ut på.)

d) NY PROFIL/INRIKTNING (Ange Program och Profilnamn. Bifoga utkast till Profilbeskrivning.)

e) ÖVRIGT (Bifoga beskrivning över vad förslaget går ut på.)

### PROGRAMNÄMNDENS BESKED:



## FÖRSLAGET I DETALJ:

The proposed course is designed to build on the students' studies in the course on Information visualization (TNM048) currently held in VT1. The existing course touches on such topics as data modelling and data mining but does not cover these aspects in detail, focussing primarily on the visual representation of data for visual exploration. The proposed course will build on the visual exploration aspects from the existing course but will, instead, focus on mathematical methods for data analysis as preprocessing steps which will lead to better analysis of complex, abstract data such as are commonplace in the modern world of business, government, science, healthcare. and industry.

The proposed course is anticipated to include 8 lectures, 4 lab sessions (1 ECTS credit) and a substantial project work (5 ECTS credits). Themes for the lectures in the course will include:

Introduction.

Scoring functions and similarity.

Models, patterns and rules.

Data modelling - Descriptive modelling.

Data modelling - Predictive modelling.

Temporal data

Uncertainty identification and representation.

Retrieval by content.

Data Organization for massive data sets.

The labs will be organized around the exploration of sample data sets using standard methods to experience numerous alternative analysis methods and become aware of the characteristics of each. The objective of the labs is not to code the algorithms but to use them with respect to different data and so learn the benefits of each over a range of types of data.

The examination of the course will be based around a project work, carried out in small groups, to examine problems created for the various 'challenge' contests which have been run at the IEEE VIS conference over the last 10 years. The data sets provided for these contests are diverse and the problems have clear objectives in the aspects of the data which the participants are expected to be able to identify through a combination of analysis and (interactive) visual representation.

Aim: After completing the course, the student should be able to:

- Analyse complex data sets and identify relevant features which might be extracted.
- Select and apply advanced algorithmic methods for analysis of large complex data sets
- Extract relevant information from such data sets for display using standard visualization methods.

Prerequisites: TNM048 - Information Visualization

Organization: The course is composed of lectures, laboratory assignments and a project assignment

Course text: The anticipated course text is "Principles of Data Mining" by David Hand, Heikki Mannila and Padhraic Smyth. This book, while an older text, is very well written as an introduction and has the advantage of being freely available via google books.

The need for this course derives from the growing need for skills in data analysis methods. The students taking the information visualization course experience a little of this and many are quite enthusiastic to explore this area further but there is currently no opportunity within the MT programme for them to do so. There is also a clear and growing need for skills in this area from all aspects of business and industry as well as in the sciences.