Title of tutorial:
Practical Introduction to Predictive Analytics

Aim and Scope:
Industry as well as academia have long seen the benefits of being able to analyse data. In today's ever-expanding sea of data, traditional techniques often reach their limits quickly in particular when analysing data sets with complex patterns, large number of attributes and thousands or maybe even millions of records.

Predictive analytics is the area within data science where an attribute value is being predicted by fitting a model using historic data as input. The general aim of this tutorial is to build a robust model that can be applied to unseen data.

This hands on practical session focuses on general concepts, techniques and algorithms during the various life cycle stages of predictive analytics including data understanding, data preparation, data modelling, evaluation and model deployment. A suitable transport related dataset will be used during the tutorial.

A graphical user interface software called RapidMiner will be used throughout the tutorial.

Outline of topics and respective expected duration (in minutes):
The following is the indicative outline of the tutorial:

- Introduction to Predictive Analytics (25 minutes)
- Terminology (15 minutes)
- RapidMiner as a data mining suite (20 minutes)
- Data Exploration (40 minutes)
- Data Pre-processing (40 minutes)
- Data Modelling (40 minutes)
- Evaluation (35 minutes)
- Deployment (20 minutes)

Intended length
Half day – 4 hours

Intended audience and assumed background knowledge
The tutorial focuses on the introductory part of predictive analytics but contains fundamental concepts that are often ignored when using such techniques leading to incorrect results or

interpretation. Therefore, this tutorial is relevant to all researchers and practitioners who work directly or indirectly with data.

If any, personal computer and software requirements for attendees
The software used is RapidMiner which has a trial version and also a free educational license. It is possible to attend the workshop without participating on a practical level also.

Organizer

Dr. Markus Hofmann | Senior Lecturer in Informatics
Institute of Technology Blanchardstown (ITB)
Blanchardstown Road North
Dublin 15
Ireland

 markus.hofmann@itb.ie www.itb.ie