ECML/PKDD-01 Workshop

The Predictive Toxicology Challenge 2000-2001

Edited by
Christoph Helma, Ross King, Stefan Kramer
and Ashwin Srinivasan

Freiburg, September 2001

Schedule

9.15-10.15: Introduction and Summary

R. King (University of Wales), D. Bristol (US NIEHS) and S. Kramer (University of Freiburg): *Summary of the Predictive Toxicology Challenge (PTC) 2000–2001*

C. Helma (University of Freiburg): *Toxicological Background of the PTC*

10.45-12.45: Contributions

H. Blockeel, K. Driessens, N. Jacobs, R. Kosala, S. Raeymaekers, J. Ramon, J. Struyf, W. Van Laer, S. Verbaeten (Katholieke Universiteit Leuven): *First-Order Models for the Predictive Toxicology Challenge 2000-2001*

H. Ohwada, M. Koyama, Y. Hoken (Science University of Tokyo): *ILP-based Rule Induction for Predicting Carcinogenicity*

K.S. Ng, J.W. Lloyd, A.W. Slater (Australian National University): *Predictive Toxicology Using a Decision Tree Learner*

B. Pfahringer (University of Waikato): *(The Futility of) Trying to Predict Carcinogenicity of Chemical Compounds*

T. Okada (Kwansei Gakuin University): *Characteristic Substructures and Properties in the Chemical Carcinogenicity Studied by the Cascade Model*

V.G. Blinova, D.A. Dobrynin, S.O. Kuznetsov, E.S. Pankratova (All-Russia Institute for Scientific and Technical Information): *Toxicology Analysis by Means of Simple JSM Method*

J.F. Boulicaut, B. Cremilleux (INSA Lyon): *Delta-Strong Classification Rules for Characterizing Chemical Carcinogens*

T.H. Reijmers and M. Engels (Janssen Research Foundation): *The application of several variable selection methods in the Predictive Toxicology Challenge 2000-2001*

14.30-15.40: Evaluation

D. Bristol (US NIEHS): *Carcinogenesis Modeling : Results Reflect Representation*

Y.-T. Woo (US EPA): *Predictive Toxicology Challenge (PTC) 2000-2001: A Toxicologist's View and Evaluation*

16.15-17.30: General Discussion

Table of Contents

Introduction

Prevention of environmentally-induced cancers is a health issue of unquestionable importance. Almost every sphere of human activity in an industrialized society faces potential chemical hazards of some form. It is estimated that nearly 100,000 chemicals are in use in large amounts every day. A further 500-1000 are added every year. Only a small fraction of these chemicals have been evaluated for toxic effects like carcinogenicity. The US National Toxicology Program (NTP) `http://ntp-server.niehs.nih.gov/` contributes to this enterprise by conducting standardized chemical bioassays – exposure of rodents (mice and rats) to a range of chemicals – to help identify substances that may have carcinogenic effects on humans. However, obtaining empirical evidence from such bioassays is expensive and usually too slow to cope with the number of chemicals that can result in adverse effects on human exposure. This has resulted in an urgent need for carcinogenicity models based on chemical structures and properties alone. It is envisaged that such models would:

- generate reliable toxicity predictions for chemicals;
- enable low cost identification of hazardous chemicals; and
- refine and reduce the reliance on the use of large numbers of laboratory animals

The outcome of the bioassays conducted by the NTP has resulted in a large (by toxicological standards) database of compounds classified as carcinogens or otherwise. Predicting the outcome of these tests using chemical structure (and related information) presents a formidable test for techniques concerned with knowledge discovery from databases.

The Predictive Toxicology Challenge (PTC) 2000-2001

The Challenge was to obtain models that predict the outcome of biological tests for the carcinogenicity of chemicals using information related to chemical structure only. As rats and mice differ substantially in their response to chemical carcinogens we required separate models for each species. These models have been assessed in regard to predictive accuracy. A subset of “optimal” models has

been reviewed by toxicological experts, to identify models that are particularly relevant to toxicology.

Detailed information about the Predictive Toxicology Challenge 2000-2001, datasets and submissions can be obtained from the PTC homepage `http://www.informatik.uni-freiburg.de/~ml/ptc/`.

Chairs

C. Helma University Freiburg

R.D. King University of Wales

S. Kramer University Freiburg

A. Srinivasan Oxford University

Organizing Committee

Romualdo Benigni Italian National Institute of Health

Doug Bristol NIEHS

Diane Cook University of Texas

Giuseppina Gini Politecnico di Milano

John Lloyd Australian National University

Bernhard Pfahringer University of Waikato

Ann Richard US EPA

Herbert Rosenkranz University of Pittsburgh

Yin-tak Woo US EPA