Comparative Testing and Evaluation of Statistical and Logical Learning Algorithms for Large-Scale Applications in Classification, Prediction and Control

Fact Sheet

Objective

STATLOG has completed an evaluation of the performance of machine learning, neural and statistical algorithms on large-scale, complex commercial and industrial problems. The overall aim has been to give an objective assessment of the potential for classification algorithms in solving significant commercial and industrial problems, and to widen the foundation for commercial exploitation of these and related algorithms both old and new.

This result describes the learning causal network algorithms and software currently being developed as part of the casual structures from inductive learning project (CASTLE), a software package which allows the user to learn the polytree's structure from raw data, to propagate knowledge throughout a polytree either interactively or in
batch mode, to simulate data from a given casual network and to create and edit casual networks. CASTLE has been created to test and evaluate Bayesian learning algorithms. The user can edit a polytree, i.e., draw nodes, link them by arrows, give names to the nodes and cases, and define the conditional or marginal probabilities in each node. This option can be combined with a simulation process to offer a way of testing the performance of the algorithms implemented. Included in CASTLE is the possibility of propagating knowledge throughout a polytree. Using this module the learned net can be consulted to reason about the interpretation of specific input data. The interpretation process involves instantiating a set of variables corresponding to the input data, calculating its impact on the probabilities of a set of variables designated as hypotheses, and finally selecting the most likely combination of these hypotheses. There is now a batch version of CASTLE. This version allows the user to execute the learning algorithms in batch mode. The user can provide the program with a new type of file containing a set of samples of observed values of any variable but the last one (the one thought of as classifier). The program propagates the observed knowledge throughout the net and outputs a file containing the posterior probability of the cases of the classifier given the observed values of the rest of the variables.

The project involves the evaluation of the performance of machine learning algorithms on large scale, complex commercial and industrial problems. The overall aim is to give an objective assessment of the potential for machine learning algorithms in solving significant commercial and industrial problems, and to widen the foundation for commercial exploitation of these and related algorithms both old and new. The 3 main approaches to decision problems are machine learning algorithms using decision trees, Bayesian methods of classical statistics and discrimination or regression methods generally. Partly due to the limited field of application of these methods, newer methods have emerged in response to new problems: relational learning algorithms deal with complex data in the form of rules; neural net algorithms mirror the behaviour of neural networks in the brain; while genetic algorithms solve problems by following an evolutionary path.

The main results expected are:
evaluation and comparison of the main artificial intelligence/machine learning algorithms, with a full specification of their merits and demerits and their range of application;
establishment of an objective set of criteria for the evaluation and comparison of algorithms;
establishment of an interactive environment for comparative testing of classification algorithms;
establishments of a set of measures for datasets, by which the performance of algorithms can be predicted;
a draft manuscript for a handbook of machine learning and statistical classification algorithms.
procedures, giving practical guidance for large scale complex classification problems in commerce and industry;
an objective assessment of several novel techniques for controlling a simulated spacecraft model.

Progress has been made in the following areas:
an improved version of a neural network algorithm (back propagation) to incorporate the differing costs of wrong decisions;
proposals for incorporating decision costs into the learning and testing phases of machine learning algorithms;
an objective assessment of the performance of machine learning, neural net techniques and traditional forecasting techniques in the prediction of economic datasets.

Historically, the three main approaches to decision problems have been (i) machine learning algorithms using decision trees; (ii) Bayesian methods of classical statistics and (iii) discrimination or regression methods generally. Partly due to the limited field of application of these methods, more recent methods have emerged in response to new problems: relational learning algorithms deal with complex data in the form of rules; neural net algorithms are linked to the fascination of mankind with understanding and emulating the human brain; while genetic algorithms solve problems by following an evolutionary path. The fact that the various methods may sometimes be applied to the same dataset with contradictory results is partly due to their treatment of the data but is more to do with the different emphasis put on the classification/prediction/optimisation aspects of the problem.

By testing around 23 algorithms from this list on about 22 large-scale and commercially important problems, this project has determined to what extent the various algorithms meet the needs of industry and has provided improved software designed to extend the commercial exploitation of advanced data analysis, including machine learning type algorithms.

The objectives of the project have been to:

- provide critical performance measurements, and criteria for measurement on available classification algorithms which will improve confidence in full exploitation
- indicate the nature and scope for next-stage development which particular algorithms require to meet commercial performance expectations
- indicate the most promising avenues of development for commercially immature approaches.

Programme(s)
### Coordinator

**DAIMLER-BENZ AG**  
Address  
Plieningerstraße 150  
70567 Stuttgart  
Germany

### Participants (10)

**Brainware Gesellschaft für Artificial Intelligence Systementwicklung und -beratung mbH**  
Germany  
Address  
Gustav-meyer-allee 25  
13355 Berlin

**DEUTSCHE AEROSPACE AG**  
Germany  
Address  
Otto-hahn-straße 28-30  
81611 München

**FRAUENHOFER-INSTITUT FÜR INFORMATIONS UND DATENVERARBEITUNG**  
Germany  
Address  
Kurstraße 33  
10117 Berlin

**INSTITUT FÜR BIOPHYSIK UND KYBERNETIK DER UNIVERSITÄT LUBECK**  
Germany  
Address  
Ilshahl 5  
24536 Neumünster

**ISOFT**  
France  
Address
28 Rue Georges Clemenceau
91400 Orsay

TECHNISCHEN UNIVERSITÄT DRESDEN
Germany
Address
Am Hulsenbusch 54
44803 Bochum

Turing Institute Ltd
United Kingdom
Address
George House 36 North
Hanover Street
G1 2AD Glasgow

UNIVERSIDAD DO PORTO
Portugal
Address
Rua Dr. Roberto Frias
4200 Porto

UNIVERSITAT DE GRANADA
Spain
Address
Cuesta Del Hospicio
18071 Granada

UNIVERSITY OF STRATHCLYDE
United Kingdom
Address
16 Richmond Street
G1 IXQ Glasgow

Last update: 17 June 1994
Record number: 8791