Machine Learning and Light Relief:

A Review of *Truth from Trash*

Jean Hayes Michie

As explained in his preface, the author of *Truth from Trash* presents machine learning in the form of "edutainment" for the lay reader. The book might also supply points of interest, although not always dependable instruction, to social scientists, philosophers, and psychologists. Thornton describes his book as a research memorandum "in keeping with the technicolour spirit of our times" and also owns to "importing various devices from the pop-science genre" (Preface, pp. 1–2).

His pop-science offerings include "light relief through a concoction of dialogues, anecdotes, and other forms of non-scientific material" (Preface, p. 2) of which the more historical chapters make the best reading. Here, departures from strict accuracy are offset by the liveliness of Thornton's accounts of Kepler's work (chapter 3) and Turing's part in breaking the German wartime Enigma code (chapter 6). He is less successful with Hume's demonstration of the fallibility of the inductive process. To spice up a long-dead controversy, he concocts an imaginary courtroom dialogue between a prosecutor anxious to outlaw induction and a defense arguing for its retention. This dialogue adds little. The old philosophical concern was that theories inductively inferred from sampled facts cannot be guaranteed true for every single future fact that might be sampled from the same source. This concern ceased to worry experimental scientists when modern statistics, now reinforced by machine learning aids, showed that calculable degrees of certainty can be attached to each such inference. In experimental science, this settled the matter, but apparently not yet in pop science: Thornton's prosecutor and defense both seem oblivious to the twentieth century development of statistical theory. Light relief apart, the book can be considered in four sections. Thornton classifies learning algorithms into two classes: (1) nonrelational data and (2) relational data. Section 1 covers algorithms for nonrelational data. Section 2 provides brief descriptions of the more complex task of classifying relational data. Section 3 sets out brief notes on the relevant science of today as a general background. The nub of the book comes in section 4. Here, the book's unexpected title is explained, leading to the author's speculations on the association of relational learning with creativity and consciousness itself. Against this background, he claims that consciousness is a continuum and cannot be ruled out as a property of any organism or learning mechanism or other artifact.

At heart, *Truth from Trash* is a philosophical chat show, designed to prod the reader with broad questions about machine learning. To the questions of a would-be practitioner, such as "How does it work?" or "How do I use it?" it gives answers that are sometimes perfunctory and sometimes erratic. However, these questions were in any case separately addressed by Thornton himself in his 1992 teaching text *Techniques in Computational Learning: An Introduction (*Springer-Verlag, 1992).

■ *Truth from Trash: How Learning Makes Sense,* Chris Thornton, Cambridge, Massachusetts, The MIT Press, 2000, 216 pp., ISBN 0-262-20127-5, $37.50 (cloth).

Given that the author's prime intention is to amuse and provoke, I am tempted to downplay the new book's errors, inconsistencies, and omissions. However, in one particular respect, I have a fundamental criticism of Thornton's attitude toward learning and the scientific process. *Truth from Trash* stresses several times that all learning can be seen as a form of prediction. The exclusive focus on prediction glosses over the fact that some machine learning procedures yield bare predictions, yet others (usually referred to as symbolic learning algorithms) produce new predictive theories. Some of these algorithms take forms that can readily be comprehended by the human mind and transcribed into the scientific literature.

Recent creative discoveries by relational learning programs of new structure-activity relations for drug design have in recent years been reported (see, for example, King et al. [1996]) from cooperating groups located at the Imperial Cancer Research Foundation, London, and at the Universities of York, Oxford, and Aberystwyth. Although perhaps not meeting Thornton's "edutainment" criterion, it is nonetheless surprising that breakthroughs of this magnitude find no place in the book.

The key point is that learning has a second aspect, namely, to generate from seeming trash not only truth but also understanding. Without understanding, there can hardly be said to be a cumulative scientific process.

Anyone who keeps this costly omission in mind might yet profit from reading *Truth from Trash*. Some readers, well attuned to the technicolor spirit of our times, might even enjoy Thornton's imaginary conversations, imaginary people, and imaginary humorous happenings. Those who find the technical sketchiness and whimsical adornments irritating might still be encouraged to check other books, search the web, and question colleagues. As an experimental psychologist, I have too often used inductive learning tools without having a completely clear picture of what makes them tick. I must now acknowledge a debt to this book for goading me into finding out considerably more.

Reference

King, R. D.; Muggleton, S. H.; Srinivasan, A.; and Sternberg, M. J. E. 1996. Structure-Activity Relationships Derived by Machine Learning: The Use of Atoms and Their Bond Connectivities to Predict Mutagenicity by Inductive Logic Programming. Proceedings of the National Academy of Sciences, Volume 93, 438–442. Washington, D.C.: National Academy of Sciences.

Jean Hayes Michie is currently trustee and chairperson of the Human-Computer Learning Foundation.