Electronic Transactions on Artificial Intelligence
Vol. 5 (2001), Section B: pp. *37-45*
http://www.ep.liu.se/ej/etai/2001/002/

Refereed Article

Boris Ryabko:

The nonprobabilistic approach to learning the best prediction

Author's affiliation: Siberian State University of Telecommunication and Computer Science, Novosibirsk, Russia

Abstract: *The problem of predicting a sequence* $x_1, x_2, ...$ *where each* x_i *belongs to a finite alphabet* A *is considered. Each letter* x_{t+1} *is predicted using information on the word* $x_1, x_2, ..., x_t$ *only. We use the game theoretical interpretation which can be traced to Laplace where there exists a gambler who tries to estimate probabilities for the letter* x_{t+1} *in order to maximize his capital . The optimal method of prediction is described for the case when the sequence* $x_1, x_2, ...$ *is generated by a stationary and ergodic source. It turns out that the optimal method is based only on estimations of conditional probabilities. In particular, it means that if we work in the framework of the ergodic and stationary source model, we cannot consider pattern recognition and other complex and interesting tools, because even the optimal method does not need them. That is why we suggest a so-called nonprobabilistic approach which is not based on the stationary and ergodic source model and show that complex algorithms of prediction can be considered in the framework of this approach.*

The new approach is to consider a set of all infinite sequences (over a given finite alphabet) and estimate the size of sets of predictable sequences with the help of the Hausdorff dimension. This approach enables us first, to show that there exist large sets of well predictable sequences which have zero measure for each stationary and ergodic measure. (In fact, it means that such sets are invisible in the framework of the ergodic and stationary source model and shows the necessity of the new approach.) Second, it is shown that there exist quite large sets of such sequences that can be predicted well by complex algorithms which use not only estimations of conditional probabilities.

Publication and review history

This work was first presented at the Machine Intelligence 18 Workshop, held on September 19-21, 2001 at the University of York, UK. Attendance at this workshop was based on invitations. Speakers at the workshop were invited to submit their articles for consideration in the forthcoming workshop volume. The subsequent publication and review history for the present article was as follows.

1. First versions *published* by Linköping University Electronic Press on in August - October, 2001 and permanently available at
   http://www.ep.liu.se/ea/cis/2001/016/
2. First version *accepted* after due refereeing and according to scientific journal standards by the Electronic Transactions on Artificial Intelligence (ETAI) on 21.9.2001. Authors were invited to revise their articles, if they so desired, based on comments by the reviewers.
3. Revised version *accepted* on 4.12.2001.
4. Revised version *published* by Linköping University Electronic Press on 28.2.2002 and permanently available at the same URL as mentioned above.
5. Revised version appears in Electronic Transactions on Artificial Intellligence, Volume 5 (2001), section B, pages 37–45. The section and the annual volume are made permanently available at
   http://www.ep.liu.se/ej/etai/
6. Paper editions of the ETAI issue and the ETAI volume containing this artice *republished* by the Royal Swedish Academy of Sciences.

The review policy and the quality requirements for acceptance are documented at http://www.etaij.org/info/

Article maintenance

The Electronic Transactions on Artificial Intelligence maintains an Article Interaction Page for this article at
http://www.etaij.org/ra/mi/022/
Besides the publication and review history, it contains links for contacting the author(s), as well as to amendments and other related information for the article. It is intended to keep these links up-to-date in the future.

Copyright conditions