

New theorems help robots to correct errors on-the-fly and learn from each other

August 21 2017

Errors in Artificial Intelligence which would normally take a considerable amount of time to resolve could be corrected immediately with the help of new research from the University of Leicester.

Researchers from the University of Leicester's Department of Mathematics have published a paper in the journal *Neural Networks* outlining mathematical foundations for new algorithms which could allow for Artificial Intelligence to collect error reports and correct them immediately without affecting existing skills - at the same time accumulating corrections which could be used for future versions or updates.

This could essentially provide robots with the ability to correct errors instantaneously, effectively 'learn' from their mistakes without damage to the knowledge already gained, and ultimately spread new knowledge amongst themselves.

Together with Industrial partners from ARM, the algorithms are combined into a system, an AI corrector, capable of improving performance of legacy AIs on-the-fly.

ARM is the world's largest provider of semiconductor IP and is the architecture of choice for more than 90% of the smart electronic products being designed today.

Professor Alexander Gorban from the University of Leicester's

Department of Mathematics said: "Multiple versions of Artificial Intelligence Big Data analytics systems have been deployed to date on millions of computers and gadgets across various platforms. They are functioning in non-uniform networks and interact.

"Industrial technological giants such as Amazon, IBM, Google, Facebook, SoftBank, ARM and many others are involved in the development of these systems. Performance of them increases, but sometimes they make mistakes like false alarms, misdetections, or wrong predictions. The mistakes are unavoidable because inherent uncertainty of Big Data.

"It seems to be very natural that humans can learn from their mistakes immediately and do not repeat them (at least, the best of us). It is a big problem how to equip Artificial Intelligence with this ability.

"It is difficult to correct a large AI system on the fly, more difficult as to shoe a horse at full gallop without stopping.

"We have recently found that a solution to this issue is possible. In this work, we demonstrate that in high dimensions and even for exponentially large samples, linear classifiers in their classical Fisher's form are powerful enough to separate errors from correct responses with high probability and to provide efficient solution to the non-destructive corrector problem."

There is a desperate need in a cheap, fast and local correction procedure that does not damage important skills of the AI systems in the course of correction.

Iterative methods of machine learning are never cheap for Big Data and huge AI systems and therefore the researchers suggest that the corrector should be non-iterative with the reversible correctors needed to

reconfigure and merge local corrections.

Dr Ivan Tyukin from the University of Leicester's Department of Mathematics said: "It is often infeasible just to re-train the systems for several reasons: they are huge and re-training requires significant computational resources or long time or both; it may be impossible retrain the system locally, at the point where mistake occur; and we can fix one thing but break another leading to that important skills could vanish.

"The development of sustainable large intelligent systems for mining of Big Data requires creation of technology and methods for fast non-destructive, non-iterative, and reversible corrections. No such technology existed until now."

The researchers have discovered and proved stochastic separation theorems which provide tools for correction of the large intelligent data analytic systems.

With this approach, instantaneous learning in Artificial Intelligence could be possible, providing AI with the ability to re-learn following a mistake after an error has occurred.

The research has been partially supported by Innovate UK through Knowledge Transfer Partnership grants: KTP009890 between ARM/Apical Ltd and the University of Leicester and KTP010522 between Visual Management Systems Limited and the University of Leicester.

The Knowledge Transfer Partnership (KTP) scheme helps businesses to innovate and grow. It does this by linking them with a university and a graduate to work on a specific project.

Dr Ilya Romanenko, Director of R&D, Computer Vision, Imaging and Vision Group from ARM said: "Having a system like this is indispensable in large-scale deployment of AI services to customers and end-users. Customer-specific usage of AI enabled devices gives rise to customer-specific errors, which is perceived by the end-user as non-acceptable situation. Retraining core AI to deal with these errors is technically challenging and potentially risky. Newly trained AI, taught to avoid specific errors may exhibit unexpected behaviour in another situation. In this scenario, the scale of a problem grows exponentially as the size of AI deployment grows, making it practically impossible."

"New technology enables to remove these obstacles altogether, making AI enabled devices collaborative in error removal process. This new quality allows large AI enabled deployments to become more intelligent as their size grows. In practice it means that error-less AI powered devices become reality. Recently we filed a patent application to secure our priority in this area."

More information: A.N. Gorban et al. Stochastic separation theorems, *Neural Networks* (2017). [DOI: 10.1016/j.neunet.2017.07.014](https://doi.org/10.1016/j.neunet.2017.07.014)

A.N. Gorban et al. One-Trial Correction of Legacy AI Systems and Stochastic Separation Theorems. arXiv. arxiv.org/abs/1610.00494

Provided by University of Leicester

Citation: New theorems help robots to correct errors on-the-fly and learn from each other (2017, August 21) retrieved 23 April 2024 from <https://techxplore.com/news/2017-08-theorems-robots-errors-on-the-fly.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private

study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.