

Using AI to design AI: Game-changing neural network design with NeuNetS

By Ruchir Puri,
CTO and Chief Architect, IBM Watson AI

AI is changing the way businesses work. However, it's important to remember that every business has unique challenges to solve, and the range of use-cases for AI is constantly expanding. While pre-built AI models exist for certain scenarios, to try and truly differentiate your business, your teams will need to think about developing custom AI models of their own.

Artificial neural networks are arguably the most powerful tool currently available to data scientists and businesses. When we think about major AI breakthroughs in the past few years, neural networks have played a critical role in those advancements.

Neural networks, though, are also the most complex and difficult type of AI models to design and train. Additionally, over time the model's performance will drift from the data with which it was originally trained. Only a small proportion of data scientists have the skills and experience needed to create a neural network from scratch, and the demand far exceeds the supply. As a result, getting a new neural network even to the proof-of-concept stage requires a level of investment that most enterprises struggle to afford.

At IBM, we've been working on a project that aims to change this paradigm completely. The result is an AI technology called Neural Network Synthesis – NeuNetS – that automatically configures itself to the needs of the user and the use case. This will help businesses reduce the complexity and skills required to build AI models, making their data science teams more productive and enabling them to scale AI across their workflows. NeuNetS is a beta feature of OpenScale that automatically creates customized neural network models using the latest training data from the business.

Building an “AI for AI”

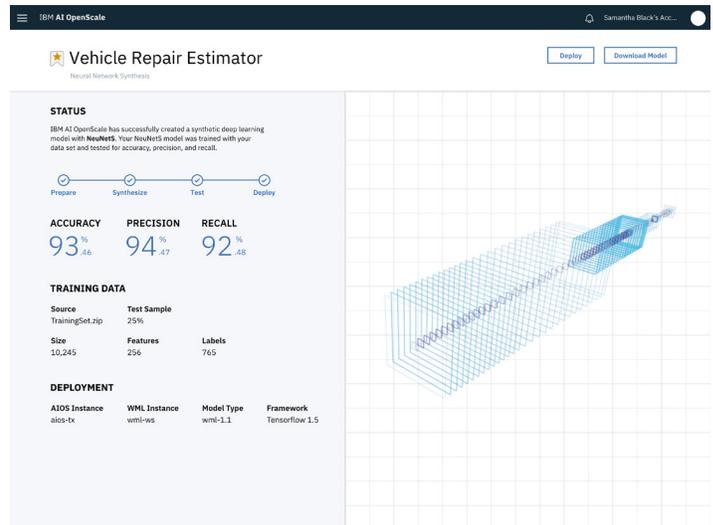
How does it work? Instead of using AI to watch grandmasters play millions of games of chess, we set NeuNetS to learn the art of designing neural networks. Based on this learning, NeuNetS will be able to synthesize new neural networks from scratch, using an iterative, two-stage process.

The first stage is coarse-grained synthesis, which determines the scaffolding of the network: How many layers there should be, how they are connected, how they should handle bypasses, feedback, different architectural components like convolution layers, long short-term memory (LSTM), and so on.

Next comes fine-grained synthesis, where NeuNetS takes a deeper dive into each layer and optimizes the individual neurons and connection—for example, what kind of convolution filter should be applied, and which neurons and edges should be optimized.

One of the critical breakthroughs that have enabled this capability is a very high-fidelity approach to performance estimation, which allows us to bypass real-time training and analysis and design neural networks automatically in a matter of hours—compared to the weeks or months that it might take a data scientist to train and optimize the AI model.

The technology is still evolving, and at this stage, NeuNetS can help businesses build neural networks for specific tasks in a fraction of the time it takes today, and with accuracy similar to that of an expert-designed AI model. The data science team can then further fine tune the model, leading to greater productivity and cost-efficiency, and bringing enterprises closer to addressing the critical AI skills gap. As part of IBM Watson OpenScale, NeuNetS is key solution for enterprises, augmenting human expertise with powerful, AI-driven optimization capabilities. It is another vital step in IBM’s journey to democratize AI and help enterprises of all sizes move confidently into the AI era.



© Copyright IBM Corporation 2018

IBM Corporation
Route 100
Somers, NY 10589

Produced in the United States of
America
October 2018

IBM, the IBM logo, ibm.com, and Watson are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at <http://www.ibm.com/legal/us/en/copytrade.shtml>

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The information in this document is provided “as is” without any warranty, express or implied, including without any warranties of merchantability, fitness for a particular purpose and any warranty or condition of non-infringement.

IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

Statement of Good Security Practices: IT system security involves protecting systems and information through prevention, detection and response to improper access from within and outside your enterprise. Improper access can result in information being altered destroyed or misappropriated or can result in damage to or misuse of your systems, including to attack others. No IT system or product should be considered completely secure and no single product or security measure can be completely effective in preventing improper access. IBM systems and products are designed to be part of a comprehensive security approach, which will necessarily involve additional operational procedures, and may require other systems, products or services to be most effective.

IBM does not warrant that systems and product are immune from the malicious or illegal conduct of an party.

