Request for Information (RFI) SN08-42: Deep Learning

Machine learning and artificial intelligence techniques that can efficiently process and "understand" massive data streams would have far-reaching military implications. Potential applications include: anomaly detection, object recognition, language understanding, task assistants, information retrieval, pattern recognition, robotic task learning and automatic metadata extraction from video streams, sensor data, and multi-media objects. All of these applications currently require considerable human intervention in data selection and labeling, supervision and maintenance of training regimes, parameter tweaking and programmed representation of data. Systems fail when asked to perform tasks for which they have not been specifically designed and trained. Even for related tasks, where some progress in transfer learning has been made, trained systems often display brittleness. For example, they may be able to recognize faces in images, but not other objects.

All of this begs the following questions:

- How can computer systems be made more robust, adaptive, and multi-functional?
- Can we devise architectures and algorithms that do more than respond to specific problems for which they have been tuned?
- How can machines acquire knowledge from huge amounts of data by creating invariant representations of the input data and applying them to multiple applications?
- How can this be done in a scalable way?
- How can previously learned representations jumpstart new learning tasks so less training is needed? Can this facilitate knowledge representation and transfer in key applications?
- How can machines extract symbolic information from perceptual inputs and enable logical reasoning over the internal symbolic representations created?
- Can machines improve situational awareness in a robust way by incorporating sensory and symbolic information?

Answering these questions will require huge advances in the state of the art. The Defense Advanced Research Projects Agency (DARPA) is soliciting new ideas in machine learning methods, especially those supporting unsupervised and reinforcement learning, temporal pattern learning and hierarchical learning. Recently, there has been progress in deeply layered systems that may incorporate feedback for top-down expectation and/or for temporal information processing.

Enormous amounts of sensor data and surveillance video are collected every day that are never distilled into usable form. In the near future, data streams that are orders of magnitude larger will become common. Information processing is failing to keep pace with information collection. Is it possible to make the rising floodwaters of information manageable and usable for intelligence extraction?

Machine learning, as it has evolved over the last twenty years, is mostly shallow in its architecture. There were two-level neural nets in the 1980's, and today there are two-level support vector machines. The internal data representations created by such shallow, broad systems are necessarily simple so that complexity is captured by combining many of them to, for example, form an object recognition system. Training these systems requires massive amounts of expensive labeled data. Problem-specific kernels are another symptom of these architectures. By contrast, visual representations in cortex range from simple in the retina to complex as the processing proceeds through many levels, and require little labeled data. Perhaps this deep architecture is the reason humans learn simple concepts before they are able to understand more complex ones and do so with much less supervision than current learning systems. Complex representations are built up from copious unlabeled inputs before the need or ability to process precious labeled data arises.

DARPA is also interested in new algorithms for learning from unlabeled data in an unsupervised manner to extract emergent symbolic representations from sensory input which may be spatiotemporal in nature. Methods may include supervised learning using labeled data tuned for specific military applications after representations emerge from less supervised techniques using massive amounts of unlabeled sensory data.

WORKSHOP

A DARPA-sponsored workshop has been planned for mid-September 2008 in Northern Virginia, for the purpose of reviewing on-going research in machine learning and artificial intelligence. Information presented at the Deep Learning workshop will contribute to the formulation of future areas of DARPA research with the objective of creating prototype systems that can use a multi-purpose cognitive architecture requiring minimal human input or parameter tweaking. The systems must be capable of distilling knowledge from massive amounts of spatiotemporal sensory input, detecting anomalous conditions, and applying these capabilities in multiple application domains. Of particular interest are hierarchical systems whose internal representations become increasingly complex and specific in ascending the hierarchy much as mammalian visual systems do to create symbolic representations at the higher levels which can capture the gist of a situation.

The DARPA Information Processing Techniques Office (IPTO) invites participation from all those engaged in related research activities. Authors may be invited to present related work and on-going research activities associated with machine learning and intelligence including:

- Cognitive architectures based on multiple layers of increasingly refined representation that can be applied to multiple application domains;
- Learning complex, hierarchical representations from unlabeled data;
- Unsupervised learning of invariant representations;
- Parameter-free learning;
- Situational awareness from unlabeled sensor data;
- Anomaly detection in distributed sensor systems; and
- Attentional mechanisms for focused learning.

Space for the workshop is limited and attendance will be by invitation only. **Invitations will be based on white papers submitted per the instructions below no later than 1200 Noon ET, 22 July 2008.** These white papers should briefly summarize approaches, and not exceed 5 pages, including figures. Accepted authors will be notified via email by 15 August 2008 and will be invited to provide a 15-minute, unclassified, non-proprietary presentation, with five minutes for questions. If there are more interesting presentations submitted than time allows, a poster session may be added. No expenses will be paid to the participants.

WHITE PAPER SUBMISSION INSTRUCTIONS AND FORMAT

DARPA will employ an electronic upload system for responses to this RFI. To respond to the RFI, interested parties must complete an online cover sheet for each white paper response, which will include the information outlined below. The coversheet submission site is https://www.csc-ballston.com/rfi/rfiindex.asp?RFIid=08-42. Upon completion of the online cover sheet, a confirmation screen will appear, along with instructions on uploading the white paper.

Since candidate authors may encounter heavy traffic on the web server, they SHOULD NOT wait until the day submissions are due to fill out a coversheet and submit the white paper!

DARPA will acknowledge receipt of submissions via email within 3 business days of the deadline.

White papers should adhere to the following formatting and outline instructions:

1. Format specifications include 12 point font, single spaced, single-sided, 8.5 by 11 inches paper, with 1-inch margins in either Microsoft Word or Adobe PDF format and zipped with either Winzip or PKZip.
2. Cover Page (1 page)
    a. Title
    b. Organization
    c. Responder's technical and administrative points of contact (names, addresses, phones and fax numbers, and email addresses)
    d. Indication of willingness to attend the Workshop
3. Technical Ideas (up to 5 pages)
    a. Executive summary
    b. A discussion of the capability/challenge addressed (from your perspective)
    c. Technical response. Your discussion should address the following: What is your proposed innovative technology/concept? How does it address the specific capability/challenge in Deep Learning? What is the current capability versus the desired capability? What extensions or advances are needed to achieve the vision of Deep Learning?
    d. Brief summary of any relevant experience in Deep Learning implementations.
4. An optional list of citations, on a separate page, including URLs, if available.

Respondents are encouraged to be as succinct as possible while at the same time providing actionable insight.

ELIGIBILITY

DARPA appreciates responses from all capable and qualified sources including, but not limited to, universities, university-affiliated research centers, federally-funded research centers, private or public companies and Government research laboratories.

DISCLAIMERS AND IMPORTANT NOTES

This is an RFI issued solely for information and new program planning purposes; the RFI, workshop do not constitute a formal solicitation for proposals. In accordance with FAR 15.201(e), responses to this notice are not offers and cannot be accepted by the Government to form a binding contract. Submission of a white paper, and/or attendance at the workshop, is voluntary and is not required to propose to subsequent Broad Agency Announcements (if any) or research solicitations (if any) on this topic. DARPA will NOT provide reimbursement for costs incurred in responding to this RFI or participating in the RFI workshop. NO PROPRIETARY OR CLASSIFIED INFORMATION SHALL BE INCLUDED IN THE RFI RESPONSE. Respondents are advised that DARPA is under no obligation to acknowledge receipt of the information received, or provide feedback to respondents with respect to any information submitted under this RFI.

Submissions may be reviewed by: the Government (DARPA and partners); Federally Funded R&D Centers (such as MIT Lincoln Laboratory); and Systems Engineering and Technical Assistance (SETA) contractors (such as Schafer Corporation, Science and Technology Associates, CACI International, and System Analysis, Inc.).

POINT OF CONTACT

Dr. Joshua Alspector, IPTO Program Manager, DARPA, Email SN08-42@darpa.mil. ANY INQUIRIES ON THIS RFI AND/OR WORKSHOP MUST BE SUBMITTED TO SN08-42@darpa.mil. NO TELEPHONE INQUIRIES WILL BE ACCEPTED.