From: Misha Gromov <[redacted]>
To: "jeffrey E." <jeevacation@gmail.com>
Subject: Re: Fwd:
Date: Wed, 11 Oct 2017 19:47:21 +0000

Like Bach's comments:)

On Wed, 11 Oct 2017 20:01:46 +0200, jeffrey E. wrote:

> ---------- Forwarded message ----------
> From: **Joscha Bach** <[redacted]>
> Date: Wed, Oct 11, 2017 at 7:55 PM
> Subject: Re:
> To: Jeffrey Epstein <jeevacation@gmail.com>
>
> After skimming their paper, the idea seemed unexciting to me at first: basically, if we have enough feature dimensions we can almost always find a linear separation. This is also related to how Support Vector Machines work: they project the data into an extremely high-dimensional space, find a separating hyperplane with linear regression, and then project that plane back into the original space as the separator. A similar idea is behind Echo State networks, which use a randomly wired recurrent neural network and then only train the output layer with a single linear regression.
> The authors take an existing trained neural network, and whenever it makes a mistake, they train a linear classifier on the network state and data, i.e. they try to find out when the network goes wrong. Instead of improving the network (which is also likely to make it worse in other cases), they add an additional layer to it. For engineering, this makes a lot of sense, because large neural networks are cheap to use and deploy but expensive to train.
>
> On a more philosophical level, it is tempting to ask if that might be a general learning principle for brains: when you don't perform well, add more control structure on top. It probably makes sense whenever you are confident that training the existing structure won't improve it that much, but unless training the weights in an existing network, it also adds quite a few milliseconds to the processing time. There is probably an optimal tradeoff for this. The other thing is that the new layer is a linear classifier only (at least in this paper), and it is creating a local override on the system's results, instead of integrating with it, somewhat similar to how reasoning might override our subconscious behavior. One of the drawbacks is that this won't allow us to use the new layer for simulating/understanding the structure of the domain modeled by the rest of the network.
>
> – Joscha
>
> > On Oct 10, 2017, at 09:43, jeffrey E. <jeevacation@gmail.com> wrote:
> >
> > https://www.sciencedaily.com/releases/2017/08/170821102725.htm
> >
> > --
> > please note
> > The information contained in this communication is

> confidential, may be attorney-client privileged, may
> constitute inside information, and is intended only for
> the use of the addressee. It is the property of
> JEE
> Unauthorized use, disclosure or copying of this
> communication or any part thereof is strictly prohibited
> and may be unlawful. If you have received this
> communication in error, please notify us immediately by
> return e-mail or by e-mail to jeevacation@gmail.com, and
> destroy this communication and all copies thereof,
> including all attachments. copyright -all rights reserved

--

please note
The information contained in this communication is
confidential, may be attorney-client privileged, may
constitute inside information, and is intended only for
the use of the addressee. It is the property of
JEE
Unauthorized use, disclosure or copying of this
communication or any part thereof is strictly prohibited
and may be unlawful. If you have received this
communication in error, please notify us immediately by
return e-mail or by e-mail to jeevacation@gmail.com, and
destroy this communication and all copies thereof,
including all attachments. copyright -all rights reserved