



Andreas Madsen

Independent Researcher

MSc Eng. – Machine Learning

I'm interested in developing interpretable machine learning models and methodology & mathematics for understanding existing models in general.

I believe that without interpretability, machine learning in some areas is socially irresponsible. Unfortunately, I don't think there is enough research in this area, as most research revolves around beating the state-of-the-art. I want to change that, to do good.

I've published 1) At [ICLR 2020](#), where I received a spotlight award. 2) In the [SEDL workshop at NeurIPS 2019](#). 3) In the [Distill.pub](#) journal. I've been interviewed several times about my publications and work.

In the past, I have made critical open-source contributions to JavaScript, such as implementing [clinic.js](#), that has become a de-facto profiling tool. I have [helped developed Node.js](#) core modules and infrastructure.

I've written a [blog post about my life as an Independent Researcher](#) that went quit viral. I'm looking for a PhD or Research Software Engineer position. Please [reach out](#) if you like my work!

[Twitter](#) [https://twitter.com/andreas_madsen] • [Scholar](#) [<https://scholar.google.com/citations?hl=en&user=XOzwAXYAAAAJ>] • [GitHub](#) [<https://github.com/AndreasMadsen>] • [LinkedIn](#) [<https://linkedin.com/in/andreasmadl>]

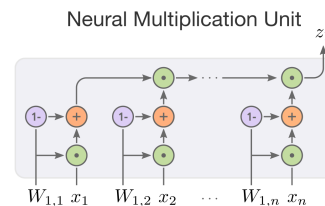
PORTFOLIO

[Open Source](#) | [Textual Heatmap](#)

MAR 2020

A python library for creating the interactive textual heatmap visualization, as I demonstrated in my [Distill paper](#). This library works with Jupyter and Google Colab making it easy for researchers to apply in their interpretability research.

[code](#) [<https://github.com/AndreasMadsen/python-textualheatmap>]



Proposes two new arithmetic units (addition and multiplication), that improves the state-of-the-art by 3x to 20x, over existing units such as the "Neural Arithmetic Logic Unit" (NALU). The improvements were achieved by rigorous theoretical analysis. The new units allow for more interpretable models and potentially perfect extrapolation.

This received a spotlight award at ICLR, as it was among the 5% best-reviewed publications.

Mentions: [Spotlight award by ICLR](https://openreview.net/forum?id=H1gNOeHKPS¬elid=g-OmkaOVen) [https://openreview.net/forum?id=H1gNOeHKPS¬elid=g-OmkaOVen] • [Best papers award by neptune.ai](https://neptune.ai/blog/iclr-2020-deep-learning) [https://neptune.ai/blog/iclr-2020-deep-learning] • [Interview by TWIML AI](https://twimlai.com/twiml-talk-382) [https://twimlai.com/twiml-talk-382]

[paper](https://openreview.net/forum?id=H1gNOeHKPS) [https://openreview.net/forum?id=H1gNOeHKPS] • [peer-review](https://openreview.net/forum?id=H1gNOeHKPS) [https://openreview.net/forum?id=H1gNOeHKPS] • [spotlight talk](https://iclr.cc/virtual_2020/poster_H1gNOeHKPS.html) [https://iclr.cc/virtual_2020/poster_H1gNOeHKPS.html] • [code](https://github.com/AndreasMadsen/stable-nalu) [https://github.com/AndreasMadsen/stable-nalu]

Open Source | lrcurve - pip package

NOV 2019

Creates a learning-curve plot for Jupyter/Colab notebooks that is updated in real-time. This was first developed for a workshop at NodeConfEU, I later made it into its own pip package.

[code](https://github.com/AndreasMadsen/python-lrcurve/) [https://github.com/AndreasMadsen/python-lrcurve/]

Product | AI smartwatch badge for NodeConf EU

NOV 2019



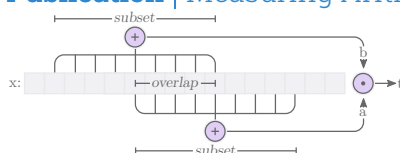
Developed the hand-gesture recognition machine learning model for the IoT smartwatch badge, given out at NodeConfEU 2019. The model ran on a system-on-chip using TensorFlow Lite for Microprocessors. This was done in collaboration with the TensorFlow team.

Mentions: [Article by hackster.io](https://hackster.io/news/27a4899dc3bd) [https://hackster.io/news/27a4899dc3bd] • [Article by Geeky Gadgets](https://geeky-gadgets.com/hackable-smartwatch-11-11-2019/) [https://geeky-gadgets.com/hackable-smartwatch-11-11-2019/] • [Review by Gary Explains](https://youtu.be/NNa_xxLCJ3k) [https://youtu.be/NNa_xxLCJ3k]

[blogpost](https://www.nearform.com/blog/banglejs-excitement-at-nodeconf-eu/) [https://www.nearform.com/blog/banglejs-excitement-at-nodeconf-eu/] • [talk by Gordon](https://youtu.be/6BAqGNOyEhw) [https://youtu.be/6BAqGNOyEhw]

Publication | Measuring Arithmetic Extrapolation Performance – SEDL|NeurIPS 2019

OCT 2019

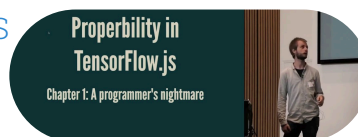


Proposes a new evaluation-criteria, with special confidence intervals, for extrapolation tasks. It uses these criteria in a reproduction study of the "Neural Arithmetic Logic Unit" (NALU), and shows that in some cases its performance is drastically worse than previously assumed.

[paper](https://arxiv.org/abs/1910.01888) [https://arxiv.org/abs/1910.01888] • [code](https://github.com/AndreasMadsen/stable-nalu) [https://github.com/AndreasMadsen/stable-nalu] • [poster](#) [download/2019-sedl-poster.pdf]

Talk | Probability in TensorFlow.js – CopenhagenJS

AUG 2019

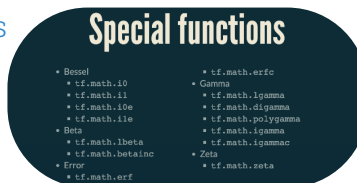


Talked about my TensorFlow.js implementation of the special functions, and especially how to survive a really difficult programming project, with lots of unknowns.

[slides](https://andreasmdsen.github.io/talk-properbility-in-tfjs-ch1/) [https://andreasmdsen.github.io/talk-properbility-in-tfjs-ch1/] • [video](https://youtu.be/VP8Sa7z7R5Y) [https://youtu.be/VP8Sa7z7R5Y]

Open Source | TensorFlow.js Special Functions

JUL 2019

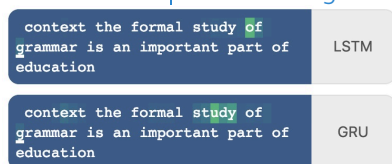


Implementation for TensorFlow.js of the special functions used in probability, calculus, differential equations, and more. Such as the beta, gamma, zeta, and bessel functions.

[code](https://github.com/AndreasMadsen/tfjs-special) [https://github.com/AndreasMadsen/tfjs-special]

Publication | Visualizing memorization in RNNs – Distill.pub

MAR 2019

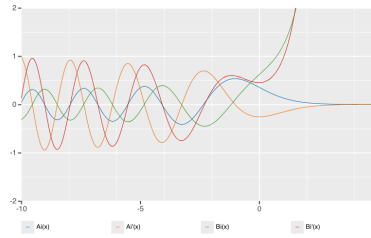


Proposes a visualization method for qualitatively comparing different RNN architectures' ability to memorize and understand what parts of an input-sentence make a prediction, which is great for interpretability.

Distill is a peer-reviewed journal, chaired by Chris Olah from OpenAI, and other famous researchers.

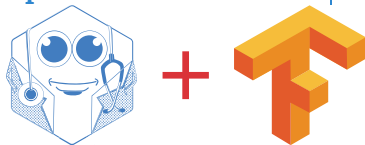
Mentions: [Interview by PracticalAI](https://changelog.com/practicalai/46) [https://changelog.com/practicalai/46] • [Video by Two Minute Papers](https://youtu.be/iKrrKyeSRew) [https://youtu.be/iKrrKyeSRew]

[paper](https://distill.pub/2019/memorization-in-rnns/) [https://distill.pub/2019/memorization-in-rnns/] • [code](https://github.com/distillpub/post--memorization-in-rnns) [https://github.com/distillpub/post--memorization-in-rnns]



By compiling the [cephes](#) library to WebAssembly, this module allows JavaScript developers to use mathematical special functions.

[blogpost](#) [https://nearform.com/blog/webassembly-cephes/] • [code](#) [https://github.com/nearform/node-cephes]



TensorFlow.js Implementation of Hidden Markov Model, that is now used filter background noise from V8 runtime in Node.js from general CPU usage signal, leaving just the main application CPU usage.

[blogpost](#) [https://clinicjs.org/blog/clinic-doctor-just-got-more-advanced-with-tensorflow-js/] • [code](#) [https://github.com/nearform/node-hidden-markov-model-tf]

Implemented the collection runtime and analysis backend of Clinic.js Bubbleprof. The currently most advanced tool for profiling and debugging asynchronous delays in Node.js.

Mentions: [Technology Radar award Vol. 22 by ThoughtWorks](#) [https://thoughtworks.com/radar/languages-and-frameworks?blipid=202005064]

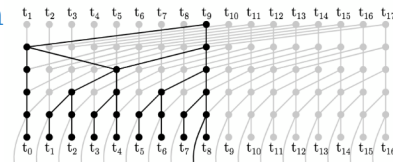
[blogpost](#) [https://clinicjs.org/blog/introducing-bubbleprof/] • [code](#) [https://github.com/nearform/node-clinic-bubbleprof] • [talk by Matteo & David](#) [https://youtu.be/DOlpzLJ1oN8]

Implemented the collection runtime, analysis backend, and frontend of Clinic.js Doctor. Clinic.js Doctor collects runtime usage data from the application runtime and uses machine learning and advanced non-parametric statistics to classify data into a recommendation for what tool to use next.

I was later involved in hiring and managing the team that now maintains it.

Mentions: [Article by IBM Developer](#) [https://developer.ibm.com/technologies/node-js/tutorials/learn-nodejs-debugging-and-profiling-node-applications/] • [Mention at TensorFlow Dev Summit keynote](#) [https://youtu.be/x35pOvZBJk8?t=775]

[blogpost](#) [https://nearform.com/blog/introducing-node-clinic-a-performance-toolkit-for-node-js-developers/] • [code](#) [https://github.com/nearform/node-clinic-doctor] • [talk by Matteo & David](#) [https://youtu.be/DOlpzLJ1oN8]



A semi-supervised neural machine translation model for small bilingual datasets. The model used the ByteNet model (Kalchbrenner, et. al.) together with a beam-search marginalization approach for semi-supervised learning.

[code](https://github.com/AndreasMadsen/master-thesis) [https://github.com/AndreasMadsen/master-thesis] • [thesis](https://github.com/AndreasMadsen/master-thesis/blob/master/thesis.pdf) [https://github.com/AndreasMadsen/master-thesis/blob/master/thesis.pdf] • [slides](https://github.com/AndreasMadsen/master-thesis/blob/master/presentation.pdf) [https://github.com/AndreasMadsen/master-thesis/blob/master/presentation.pdf]



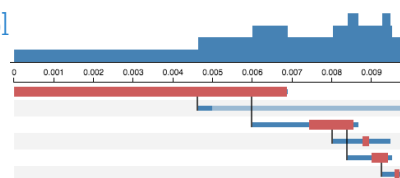
I was a critical part of getting the `async_hooks` module implemented in the Node.js core runtime. This module allows users to monitor all asynchronous operations happening in the application.

[documentation](https://nodejs.org/dist/latest-v8.x/docs/api/async_hooks.html) [https://nodejs.org/dist/latest-v8.x/docs/api/async_hooks.html] • [main pull request](https://github.com/nodejs/node/pull/12892) [https://github.com/nodejs/node/pull/12892]

$$\begin{aligned} \min_{\mathbf{p}} \quad & \mathbf{p}^T \mathbf{p} - 2\mathbf{z}^T \mathbf{p} \\ \text{s.t.} \quad & \mathbf{1}^T \mathbf{p} = 1 \\ & \mathbf{p} \geq 0 \end{aligned}$$

Implemented the sparsemax operator in the TensorFlow core, as part of a course project. This involved Python, C++, and CUDA.

[white paper](https://github.com/AndreasMadsen/course-02456-sparsemax/blob/master/latex/report/report.pdf) [https://github.com/AndreasMadsen/course-02456-sparsemax/blob/master/latex/report/report.pdf] • [main pull request](https://github.com/tensorflow/tensorflow/pull/6387) [https://github.com/tensorflow/tensorflow/pull/6387]



Implemented interactive profiling software for monitoring all asynchronous operations in a node.js application. This used, at the time, an internal version of `async_hooks`, and the tool was instrumental in debugging the `async_hooks` implementation in Node.js.

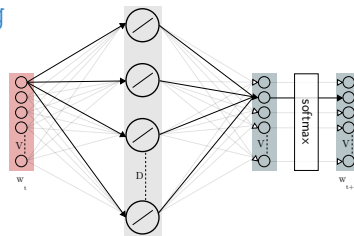
[demo](https://dprof.js.org/) [https://dprof.js.org/] • [code](https://github.com/AndreasMadsen/dprof) [https://github.com/AndreasMadsen/dprof]



After having introduced statistics into the Node.JS open source project for their benchmarking suite, I was invited to speak at NodeConf EU in Ireland.

The challenge was to communicate both how a Welch's t-test works to people that often dislike mathematics, and provide the psychological background of why statistics is necessary.

[slides](https://andreasmdsen.github.io/talk-benchmarking-nodejs-with-statistics/) [https://andreasmdsen.github.io/talk-benchmarking-nodejs-with-statistics/] • [video](https://www.youtube.be/IPH4C9cHJLg) [https://www.youtube.be/IPH4C9cHJLg]



A comparison of paragraph2vec (a word2vec variant) and an LSTM encoder-decoder (Sutskever et al.), for generating semantic vectors that are precise enough to cluster documents according to the story.

The thesis also proposes a quasi-linear-time clustering algorithm, useful for dated documents such as new articles.

[code](https://github.com/AndreasMadsen/bachelor-code) [https://github.com/AndreasMadsen/bachelor-code] • [thesis](https://github.com/AndreasMadsen/bachelor-report/blob/master/Thesis.pdf) [https://github.com/AndreasMadsen/bachelor-report/blob/master/Thesis.pdf]
