tranches Documentation

Release 0.1.0

Rene-Jean Corneille

Sep 24, 2018

Contents

Non-optimised implementation of neural networks for educational purposes.

CHAPTER 1

User's Guide

1.1 Requirements

tranches is build on top of the following libraries:

- *Numpy* (http://www.numpy.org)
- *matplotlib* (http://http://matplotlib.org/)

1.2 Issues

Should you encounter any issue with the library you can raise them here: https://github.com/RonsenbergVI/tranches/issues

1.3 Installing tranches

1.3.1 Installation with pip (recommended)

The library is available on pypi and to install the last available version, run this command:

```
$ pip3 install tranches
```

To test the installation:

```
>>> import tranches
>>> tranches.__version__
```

this should display the version installed on your system.

1.3.2 Installation from GitHub

tranches releases are also available on github (https://github.com/ronsenbergVI/tranches).

You first need to clone (or fork if you want to modify it) and

```
$ git clone https://github.com/ronsenbergVI/tranches.git
$ cd tranches
$ python setup.py build
$ python setup.py install
```

CHAPTER 2

Machine Learning Basics

Deep learning is a vibrant subfield of machine learning.. The field has attracted a lot of attention over the past few years. It is a consequence of the continuous progress made in data collection, processing and storage and of the advances in parallelisation allowing to state of the art models.

Even though deep learning has its specificities, it is based on learning theories and it is worth going over some of the fundamentals. This chapter is a brief introduction to the theory of learning. It is the bare minimum needed to go forward but is in no case complete. Feel free to skip to the next chapter.

2.1 The supervised learning problem

We assume that we have at hand a set of examples $(\mathbf{x}_i, y_i)_{i\in\{0,\dots,N\}}$ where $\mathbf{x}_i$ is a vector of features and y_i are labels that are either draw from a discrete set (classification) or the real line (regression). In other words, we have a set of inputs and their desired output.

The goal of supervised learning is to find a function f that ultimately is able to predict the desired output for new examples by learning from the examples at our disposal. Ideally, to solve this problem, we would like to be able to find the joint probability:

$$F(x, y) = \mathbb{P}(\mathbf{x}, y)$$

The learning problem is usually defined over a parameter set and thus we are looking for a function $f_\theta(x)$ with $x \in \Theta$ and Θ is the parameter set. The learning procedure will define the parameter update rules. The parameters are updated so that the model output $\hat{y_n} = f_\theta(x_n)$ are the closest possible to the expected outputs y_n. We need to define metrics to define how close is the model.

2.2 Loss and Risk Functions

In order to learn the optimal parameters θ^* supervised learning algorithms want to minimise what is referred to as a loss function.

A loss function is defined a map taking as input a pair of elements from the possible of labels to a real number. Ideally we would like to have a function that:

1. is regular enough (continuous and convex is the jackpot bu hardly ever the case for any interesting model);
2. sensitive to magnitude (this allows to distinguish between the types of error).

The type of loss function chosen depends on the supervised task. For classification the most common loss function is the cross entropy defined as:

$$C(y, \hat{y}) = \sum_k \mathbb{P}(y = k) log \mathbb{P}(\hat{y} = k)$$

and for regression it is the mean square root (MSE) defined as:

$$C(y, \hat{y}) = ||y - \hat{y} = k||$$

The loss function computes the "cost" of the learner's error for one example (intuitively a zero cost means a correct prediction). We want to minimise the cost for the whole dataset. For this we define the (empirical) risk function as:

$$J(\theta) = \frac{1}{N} \sum_n C(y_n, \hat{y}_n)$$

2.3 Gradient Descent

Gradient based optimisation methods are quite popular in machine learning and particularly in deep learning. The gradient of a function gives meaningful information about how to update its parameters to reduce its value. Gradient based methods are iterative algorithms that produce a sequence (theta_n) of parameters where each new iteration improves the risk function.

1. defining an initial value x_0
2. an iteration process to get x_{n+1} given x_n
3. a convergence criteria to stop the training when reached

The simplest gradient based optimisation method is the gradient (or steepest) descent algorithm where each new iteration is defined by going in small steps in the opposite direction of the risk function gradient:

1. initialise x_0
2. $x_{n+1} = `x_n - \gamma \nabla f(x_n)$
3. until $||x_{n+1} - `x_n|| < \epsilon$

The interesting feature of this method is that it is guaranteed to reach a local minima for a convex function on a convex set.

The hyperparameter γ is referred to as the learning rate. Choosing a learning rate is no easy task (well nothing is really in deep learning) and is critical for the performance of the training. Choosing a too large learning learning rate prevents the algorithm to target precisely enough regions of local minima and thus oscillates. A large enough can event prevent the algorithm to converge. Too small a convergence makes the algorithm slow as it takes too many steps towards a local minima.

A common method is to start with a large learning rate when the model starts training and is (arguably) far from the optimal solution and as training goes to decay the rate as the model gets closer to the solution we need to take more precaution and not miss it. The popular learning rate schedules used in practice are:

1. exponential decay:

$$\gamma_r = \gamma_0 \exp(-H \frac{r}{R})$$

2. power decay:

power decay:

$$\gamma_r = \gamma_0(1 - \frac{r}{R})^H)$$

2.4 Batch Learning

2.5 Metrics

CHAPTER 3

Neural Networks

<intro to chapter here>

3.1 Fully connected, feed forward networks

3.2 Activation Function

3.3 Backpropagation

3.4 Initialization

3.5 Beyond SGD

3.6 Regularization

CHAPTER 4

The MNIST dataset (yet again)

<intro to chapter here>

4.1 Model selection

4.2 Results

CHAPTER 5

API Reference

If you are looking for information on a specific function, class or method, this part of the documentation is for you.

CHAPTER 6

tranches changelog

We detail here the changes made to the library

6.1 Version 0.1.1

Release day: September 27 2018

- Initial public release.

CHAPTER 7

License

tranches is licensed under the Apache 2.0 Licence. It means that the source code provided in the binaries can be used, modified, or distributed freely for commercial or personal use with conditions only requiring preservation of copyright and license notices.

The full license text can be found below (tranches-license).

7.1 Authors

tranches is written and maintained by Rene-Jean Corneille

7.2 Contributing

Contribution will be welcomed once a first stable release is ready.

7.3 License Definitions

The following section contains the full license texts for tranches and the documentation.

- "AUTHORS" hereby refers to all the authors listed in the authors section.
- The "tranches-license" applies to all the source code shipped as part of tranches (tranches itself as well as the examples and the unittests) as well as documentation.

7.4 tranches License

Apache License

Version 2.0, January 2004

http://www.apache.org/licenses/

TERMS AND CONDITIONS FOR USE, REPRODUCTION, AND DISTRIBUTION

1. Definitions.

   "License" shall mean the terms and conditions for use, reproduction, and distribution as defined by Sections 1 through 9 of this document.

   "Licensor" shall mean the copyright owner or entity authorized by the copyright owner that is granting the License.

   "Legal Entity" shall mean the union of the acting entity and all other entities that control, are controlled by, or are under common control with that entity. For the purposes of this definition, "control" means (i) the power, direct or indirect, to cause the direction or management of such entity, whether by contract or otherwise, or (ii) ownership of fifty percent (50%) or more of the outstanding shares, or (iii) beneficial ownership of such entity.

   "You" (or "Your") shall mean an individual or Legal Entity exercising permissions granted by this License.

   "Source" form shall mean the preferred form for making modifications, including but not limited to software source code, documentation source, and configuration files.

   "Object" form shall mean any form resulting from mechanical transformation or translation of a Source form, including but not limited to compiled object code, generated documentation, and conversions to other media types.

   "Work" shall mean the work of authorship, whether in Source or Object form, made available under the License, as indicated by a copyright notice that is included in or attached to the work (an example is provided in the Appendix below).

   "Derivative Works" shall mean any work, whether in Source or Object form, that is based on (or derived from) the Work and for which the editorial revisions, annotations, elaborations, or other modifications represent, as a whole, an original work of authorship. For the purposes of this License, Derivative Works shall not include works that remain separable from, or merely link (or bind by name) to the interfaces of, the Work and Derivative Works thereof.

   "Contribution" shall mean any work of authorship, including the original version of the Work and any modifications or additions to that Work or Derivative Works thereof, that is intentionally submitted to Licensor for inclusion in the Work by the copyright owner or by an individual or Legal Entity authorized to submit on behalf of the copyright owner. For the purposes of this definition, "submitted" means any form of electronic, verbal, or written communication sent to the Licensor or its representatives, including but not limited to communication on electronic mailing lists, source code control systems, and issue tracking systems that are managed by, or on behalf of, the Licensor for the purpose of discussing and improving the Work, but excluding communication that is conspicuously marked or otherwise designated in writing by the copyright owner as "Not a Contribution."

   "Contributor" shall mean Licensor and any individual or Legal Entity on behalf of whom a Contribution has been received by Licensor and subsequently incorporated within the Work.

2. Grant of Copyright License. Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable copyright license to reproduce, prepare Derivative Works of, publicly display, publicly perform, sublicense, and distribute the Work and such Derivative Works in Source or Object form.

3. Grant of Patent License. Subject to the terms and conditions of this License, each Contributor hereby grants to You a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable (except as stated in this section) patent license to make, have made, use, offer to sell, sell, import,

and otherwise transfer the Work, where such license applies only to those patent claims licensable by such Contributor that are necessarily infringed by their Contribution(s) alone or by combination of their Contribution(s) with the Work to which such Contribution(s) was submitted. If You institute patent litigation against any entity (including a cross-claim or counterclaim in a lawsuit) alleging that the Work or a Contribution incorporated within the Work constitutes direct or contributory patent infringement, then any patent licenses granted to You under this License for that Work shall terminate as of the date such litigation is filed.

4. Redistribution. You may reproduce and distribute copies of the Work or Derivative Works thereof in any medium, with or without modifications, and in Source or Object form, provided that You meet the following conditions:

   (a) You must give any other recipients of the Work or Derivative Works a copy of this License; and

   (b) You must cause any modified files to carry prominent notices stating that You changed the files; and

   (c) You must retain, in the Source form of any Derivative Works that You distribute, all copyright, patent, trademark, and attribution notices from the Source form of the Work, excluding those notices that do not pertain to any part of the Derivative Works; and

   (d) If the Work includes a "NOTICE" text file as part of its distribution, then any Derivative Works that You distribute must include a readable copy of the attribution notices contained within such NOTICE file, excluding those notices that do not pertain to any part of the Derivative Works, in at least one of the following places: within a NOTICE text file distributed as part of the Derivative Works; within the Source form or documentation, if provided along with the Derivative Works; or, within a display generated by the Derivative Works, if and wherever such third-party notices normally appear. The contents of the NOTICE file are for informational purposes only and do not modify the License. You may add Your own attribution notices within Derivative Works that You distribute, alongside or as an addendum to the NOTICE text from the Work, provided that such additional attribution notices cannot be construed as modifying the License.

   You may add Your own copyright statement to Your modifications and may provide additional or different license terms and conditions for use, reproduction, or distribution of Your modifications, or for any such Derivative Works as a whole, provided Your use, reproduction, and distribution of the Work otherwise complies with the conditions stated in this License.

5. Submission of Contributions. Unless You explicitly state otherwise, any Contribution intentionally submitted for inclusion in the Work by You to the Licensor shall be under the terms and conditions of this License, without any additional terms or conditions. Notwithstanding the above, nothing herein shall supersede or modify the terms of any separate license agreement you may have executed with Licensor regarding such Contributions.

6. Trademarks. This License does not grant permission to use the trade names, trademarks, service marks, or product names of the Licensor, except as required for reasonable and customary use in describing the origin of the Work and reproducing the content of the NOTICE file.

7. Disclaimer of Warranty. Unless required by applicable law or agreed to in writing, Licensor provides the Work (and each Contributor provides its Contributions) on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied, including, without limitation, any warranties or conditions of TITLE, NON-INFRINGEMENT, MERCHANTABILITY, or FITNESS FOR A PARTICULAR PURPOSE. You are solely responsible for determining the appropriateness of using or redistributing the Work and assume any risks associated with Your exercise of permissions under this License.

8. Limitation of Liability. In no event and under no legal theory, whether in tort (including negligence), contract, or otherwise, unless required by applicable law (such as deliberate and grossly negligent acts) or agreed to in writing, shall any Contributor be liable to You for damages, including any

direct, indirect, special, incidental, or consequential damages of any character arising as a result of this License or out of the use or inability to use the Work (including but not limited to damages for loss of goodwill, work stoppage, computer failure or malfunction, or any and all other commercial damages or losses), even if such Contributor has been advised of the possibility of such damages.

9. Accepting Warranty or Additional Liability. While redistributing the Work or Derivative Works thereof, You may choose to offer, and charge a fee for, acceptance of support, warranty, indemnity, or other liability obligations and/or rights consistent with this License. However, in accepting such obligations, You may act only on Your own behalf and on Your sole responsibility, not on behalf of any other Contributor, and only if You agree to indemnify, defend, and hold each Contributor harmless for any liability incurred by, or claims asserted against, such Contributor by reason of your accepting any such warranty or additional liability.

END OF TERMS AND CONDITIONS

APPENDIX: How to apply the Apache License to your work.

To apply the Apache License to your work, attach the following boilerplate notice, with the fields enclosed by brackets "[]" replaced with your own identifying information. (Don't include the brackets!) The text should be enclosed in the appropriate comment syntax for the file format. We also recommend that a file or class name and description of purpose be included on the same "printed page" as the copyright notice for easier identification within third-party archives.

Python Module Index

t

Index

T