An Exploration of Auto Encoders on MNIST

Jonathan E. Mathews*
Department of Electrical and Computer Engineering
South Dakota School of Mines and Technology
Rapid City, SD 57701
`jonathan.mathews@mines.sdsmt.edu`

Abstract

The abstract paragraph should be indented ½ inch (3 picas) on both the left- and right-hand margins. Use 10 point type, with a vertical spacing (leading) of 11 points. The word **Abstract** must be centered, bold, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

1 Introduction

Supervised machine learning has been used to solve complex classification problems but it requires large amounts of preclassified data for training. Unsupervised machine learning attempts to solve this problem by extracting order from data without a pre-determined solution to compare with. Auto encoders are a specific class of unsupervised network that do this by transforming the data into a latent space with desirable properties and then transforming the latent representation back into a form resembling the original input. The loss function usually involves the accuracy of the reconstruction combined with a parameter representing the conformance of the latent space to the desired form.

This paper will explore the properties of three different auto-encoders and their facility for data compression. One of the common goals of an auto-encoder is to find a lower-dimensional representation of the data that, with the appropriate decoder, can be used to reconstruct the original high-dimensional input with minimal reconstruction loss. This is contrasted with a denoising auto-encoder that may have a higher-dimensional latent space than the input, but whose purpose is to reconstruct the original data given a noisy version of the original input. First, this paper will provide a discussion of a linear auto-encoder and its cognates in the machine learning world. Next, the structure will be extended to use a variety of different non-linear activation functions with the goal of improving reconstruction performance in the presence of a non-linear manifold. Finally, the variational auto-encoder will be explored and it's capability to extract the data manifold will be contrasted with the other auto-encoder methods.

2 The significance of manifold learning

The crux of solving problems in engineering and science is often data visualization. The real world is full of noise and phenomena that distort and obscure the desired trends. Finding an appropriate transformation that moves the data from the realm of seeming chaos to the realm of interpretable results is therefore key. In the thermal sciences, this is often done by non-dimensionalizing the data to extend applicability to untested cases. In control theory, the LaPlace and Fourier transforms can be used to show the frequency-dependence of the system. Even something so simple as using a logarithmic plot instead of a linear plot is an example of data transformation for visualization.

*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

The fundamental concept behind manifold learning is that the physical processes that give rise to the data are relatively few in number and typically exist in a continuous space. A simple example would be using a camera to record the trajectory of a soccer-ball after it was kicked. The problem could be somewhat accurately simulated using a small number of parameters such as the impulse imparted to the ball by the kick, the point of contact between foot and ball, the presence of wind or friction of the grass, the shape of the ball, and the mass of the ball. This is only six parameters that could vary continuously in the space of all experiments. If the measurement system is a 720x1080 pixel rgb camera at 30fps, and each experiment takes 5 seconds, then the raw measurement data contains 349,920,000 individual pixel values between 0 and 255. Obviously, there is a huge amount of unnecessary data collected by the system. Manifold learning says that if we represent each experiment as a single point in 349,920,000-dimensional space, all of the experiments should lie close to some 6-dimensional manifold or surface where traversing the manifold is equivalent to traversing the original six simulation parameters. It concerns itself with finding an appropriate transform that can map the results of a new experiment to a low-dimensional space where each dimension has a direct physical corollary.

This concept coincides with the function of an auto-encoder. The encoding portion of an auto-encoder takes a high-dimensional input and reduces it to a low dimensional latent-space. This would correspond to taking the video of the soccer ball and extracting from it the kicking parameters. The decoding portion takes that latent-space and maps it back to the high-dimensional space. In the soccer example, if the latent-space truly represents the physical variables in the system, then then the decoder could also be thought of as a simulator that could reproduce an accurate video of a soccer-ball being kicked based on the desired physical parameters. An auto-encoder can therefore be thought of as a deep network that attempts to learn the relationship between data and the processes that generated the data. It guesses the appropriate encoding and decoding relationships and then tries to extract the latent parameters and simulate the result. It then compares the simulation with the original input and iterates until it finds an acceptable solution. In this way, the complex relationships between process and data can be determined and a useful visualization developed.

This paper will explore auto-encoders using the MNIST dataset which consists of 28x28-pixel images of hand-drawn numeric digits from 0 to 9. The physical process used to create those images was a person moving a pen across paper. We would therefore expect the manifold extracted to have some correspondence to the various strokes of a pen with some additional parameters to account for the variation within each stroke. There are perhaps 10 to 20 distinct strokes that are used to make each character along with a few parameters each to represent degree of curvature or width of each stroke, etc. The hypothesis is therefore that an appropriate latent-space will have a dimensionality of $10^1 < N < 10^2$.

3 Linear auto-encoder

A linear auto-encoder resolves the input onto a low-dimensional plane

4 Non-linear auto-encoders

5 Variational auto-encoder

The variational auto-encoder

5.1 Style

Papers to be submitted to NIPS 2017 must be prepared according to the instructions presented here. Papers may only be up to eight pages long, including figures. This does not include acknowledgments and cited references which are allowed on subsequent pages. Papers that exceed these limits will not be reviewed, or in any other way considered for presentation at the conference.

The margins in 2017 are the same as since 2007, which allow for $\sim 15\%$ more words in the paper compared to earlier years.

Authors are required to use the NIPS LaTeX style files obtainable at the NIPS website as indicated below. Please make sure you use the current files and not previous versions. Tweaking the style files may be grounds for rejection.

5.2 Retrieval of style files

The style files for NIPS and other conference information are available on the World Wide Web at

`http://www.nips.cc/`

The file `nips_2017.pdf` contains these instructions and illustrates the various formatting requirements your NIPS paper must satisfy.

The only supported style file for NIPS 2017 is `nips_2017.sty`, rewritten for LaTeX 2_{ε}. **Previous style files for LaTeX 2.09, Microsoft Word, and RTF are no longer supported!**

The new LaTeX style file contains two optional arguments: `final`, which creates a camera-ready copy, and `nonatbib`, which will not load the `natbib` package for you in case of package clash.

At submission time, please omit the `final` option. This will anonymize your submission and add line numbers to aid review. Please do *not* refer to these line numbers in your paper as they will be removed during generation of camera-ready copies.

The file `nips_2017.tex` may be used as a "shell" for writing your paper. All you have to do is replace the author, title, abstract, and text of the paper with your own.

The formatting instructions contained in these style files are summarized in Sections 6, 7, and 8 below.

6 General formatting instructions

The text must be confined within a rectangle 5.5 inches (33 picas) wide and 9 inches (54 picas) long. The left margin is 1.5 inch (9 picas). Use 10 point type with a vertical spacing (leading) of 11 points. Times New Roman is the preferred typeface throughout, and will be selected for you by default. Paragraphs are separated by ½ line space (5.5 points), with no indentation.

The paper title should be 17 point, initial caps/lower case, bold, centered between two horizontal rules. The top rule should be 4 points thick and the bottom rule should be 1 point thick. Allow ¼ inch space above and below the title to rules. All pages should start at 1 inch (6 picas) from the top of the page.

For the final version, authors' names are set in boldface, and each name is centered above the corresponding address. The lead author's name is to be listed first (left-most), and the co-authors' names (if different address) are set to follow. If there is only one co-author, list both author and co-author side by side.

Please pay special attention to the instructions in Section 8 regarding figures, tables, acknowledgments, and references.

7 Headings: first level

All headings should be lower case (except for first word and proper nouns), flush left, and bold.

First-level headings should be in 12-point type.

7.1 Headings: second level

Second-level headings should be in 10-point type.

7.1.1 Headings: third level

Third-level headings should be in 10-point type.

Paragraphs There is also a `\paragraph` command available, which sets the heading in bold, flush left, and inline with the text, with the heading followed by 1 em of space.

8 Citations, figures, tables, references

These instructions apply to everyone.

8.1 Citations within the text

The `natbib` package will be loaded for you by default. Citations may be author/year or numeric, as long as you maintain internal consistency. As to the format of the references themselves, any style is acceptable as long as it is used consistently.

The documentation for `natbib` may be found at

`http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf`

Of note is the command `\citet`, which produces citations appropriate for use in inline text. For example,

```
\citet{hasselmo} investigated\dots
```

produces

> Hasselmo, et al. (1995) investigated...

If you wish to load the `natbib` package with options, you may add the following before loading the `nips_2017` package:

```
\PassOptionsToPackage{options}{natbib}
```

If `natbib` clashes with another package you load, you can add the optional argument `nonatbib` when loading the style file:

```
\usepackage[nonatbib]{nips_2017}
```

As submission is double blind, refer to your own published work in the third person. That is, use "In the previous work of Jones et al. [4]," not "In our previous work [4]." If you cite your other papers that are not widely available (e.g., a journal paper under review), use anonymous author names in the citation, e.g., an author of the form "A. Anonymous."

8.2 Footnotes

Footnotes should be used sparingly. If you do require a footnote, indicate footnotes with a number[2] in the text. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a horizontal rule of 2 inches (12 picas).

Note that footnotes are properly typeset *after* punctuation marks.[3]

8.3 Figures

All artwork must be neat, clean, and legible. Lines should be dark enough for purposes of reproduction. The figure number and caption always appear after the figure. Place one line space before the figure caption and one line space after the figure. The figure caption should be lower case (except for first word and proper nouns); figures are numbered consecutively.

You may use color figures. However, it is best for the figure captions and the paper body to be legible if the paper is printed in either black/white or in color.

[2] Sample of the first footnote.
[3] As in this example.

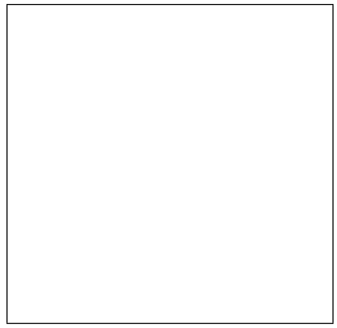

Figure 1: Sample figure caption.

Table 1: Sample table title

Part		
Name	Description	Size (μm)
Dendrite	Input terminal	$\sim$100
Axon	Output terminal	$\sim$10
Soma	Cell body	up to 10^6

8.4 Tables

All tables must be centered, neat, clean and legible. The table number and title always appear before the table. See Table 1.

Place one line space before the table title, one line space after the table title, and one line space after the table. The table title must be lower case (except for first word and proper nouns); tables are numbered consecutively.

Note that publication-quality tables *do not contain vertical rules.* We strongly suggest the use of the `booktabs` package, which allows for typesetting high-quality, professional tables:

`https://www.ctan.org/pkg/booktabs`

This package was used to typeset Table 1.

9 Final instructions

Do not change any aspects of the formatting parameters in the style files. In particular, do not modify the width or length of the rectangle the text should fit into, and do not change font sizes (except perhaps in the **References** section; see below). Please note that pages should be numbered.

10 Preparing PDF files

Please prepare submission files with paper size "US Letter," and not, for example, "A4."

Fonts were the main cause of problems in the past years. Your PDF file must only contain Type 1 or Embedded TrueType fonts. Here are a few instructions to achieve this.

- You should directly generate PDF files using `pdflatex`.
- You can check which fonts a PDF files uses. In Acrobat Reader, select the menu Files>Document Properties>Fonts and select Show All Fonts. You can also use the program `pdffonts` which comes with `xpdf` and is available out-of-the-box on most Linux machines.
- The IEEE has recommendations for generating PDF files whose fonts are also acceptable for NIPS. Please see `http://www.emfield.org/icuwb2010/downloads/IEEE-PDF-SpecV32.pdf`

- `xfig` "patterned" shapes are implemented with bitmap fonts. Use "solid" shapes instead.
- The `\bbold` package almost always uses bitmap fonts. You should use the equivalent AMS Fonts:

  ```
  \usepackage{amsfonts}
  ```

  followed by, e.g., `\mathbb{R}`, `\mathbb{N}`, or `\mathbb{C}` for $\mathbb{R}$, $\mathbb{N}$ or $\mathbb{C}$. You can also use the following workaround for reals, natural and complex:

  ```
  \newcommand{\RR}{I\!\!R} %real numbers
  \newcommand{\Nat}{I\!\!N} %natural numbers
  \newcommand{\CC}{I\!\!\!\!C} %complex numbers
  ```

  Note that `amsfonts` is automatically loaded by the `amssymb` package.

If your file contains type 3 fonts or non embedded TrueType fonts, we will ask you to fix it.

10.1 Margins in LaTeX

Most of the margin problems come from figures positioned by hand using `\special` or other commands. We suggest using the command `\includegraphics` from the `graphicx` package. Always specify the figure width as a multiple of the line width as in the example below:

```
\usepackage[pdftex]{graphicx} ...
\includegraphics[width=0.8\linewidth]{myfile.pdf}
```

See Section 4.4 in the graphics bundle documentation (`http://mirrors.ctan.org/macros/latex/required/graphics/grfguide.pdf`)

A number of width problems arise when LaTeX cannot properly hyphenate a line. Please give LaTeX hyphenation hints using the `\-` command when necessary.

Acknowledgments

Use unnumbered third level headings for the acknowledgments. All acknowledgments go at the end of the paper. Do not include acknowledgments in the anonymized submission, only in the final paper.

References

References follow the acknowledgments. Use unnumbered first-level heading for the references. Any choice of citation style is acceptable as long as you are consistent. It is permissible to reduce the font size to `small` (9 point) when listing the references. **Remember that you can go over 8 pages as long as the subsequent ones contain *only* cited references.**

[1] Alexander, J.A. & Mozer, M.C. (1995) Template-based algorithms for connectionist rule extraction. In G. Tesauro, D.S. Touretzky and T.K. Leen (eds.), *Advances in Neural Information Processing Systems 7*, pp. 609–616. Cambridge, MA: MIT Press.

[2] Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural SImulation System.* New York: TELOS/Springer–Verlag.

[3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.