This Is My Title
Along With More Clarification

by

Anthony J La Barca

Submitted in Partial Fulfillment of the

Requirements for the Degree

Bachelor of Science in Physics and Astronomy

Supervised by Segev BenZvi, Ph.D.

Department of Physics and Astronomy
School of Arts and Sciences

University of Rochester

Rochester, New York

2023

To whomever

Acknowledgments

Lorem Ipsum

Abstract

Lorem Ipsum

Contents

Acknowledgments **iii**

Abstract **iv**

List of Tables **vii**

List of Figures **vii**

1 Introduction **1**
1.1 Citation styles 1
1.2 Figures 2
1.3 Tables 3

2 Machine Learning Techniques **4**
2.1 Convolutional Neural Networks 4
2.1.1 CNNs on Spectroscopic Data 5
2.2 Previous and Current Attempts at Machine Learning 6
2.3 Introduction of Transformers 6

3 Validity of Transformers for Spectral Classification **7**
3.1 Creation of artificial spectra 7
3.1.1 Features of `GenData` class 7
3.1.2 Creation of spectra 8

Bibliography **9**

A Detailed Diagrams of Various Neural Networks **10**
A.1 Fully Connected Neural Networks 10
A.2 Convolutional Neural Networks 10
A.3 Transformers 10

B Title of Appendix B **11**

List of Tables

1.1 This is a table 3

List of Figures

1.1 This is a black box 2

2.1 CNN Diagnostics 5
2.2 Max value of the output vector for the CNN. 6

3.1 Examples of different continuum options (from left to right: linear, linear increasing, linear decreasing, exponential increasing, and exponential decreasing). 8

1. Introduction

1.1 Citation styles

These are the different citation styles for author-year.

The standard \cite command produces the following output: Clarke 1990.

The \textcite command produces the following output: Clarke (1990).

The \parencite command produces the following output: (Clarke 1990).

The \footcite command produces a footnote citation[1].

[1] Clarke 1990.

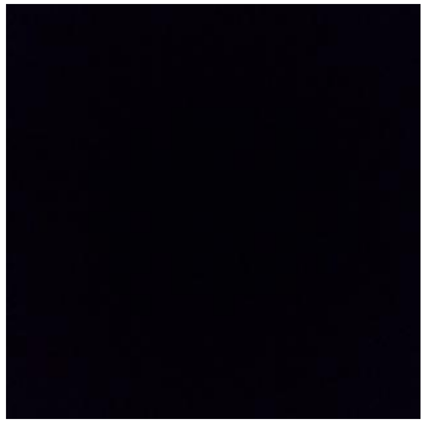


Figure 1.1: This is a black box

1.2 Figures

Always prefer to float figures to the top of the page using the [t] option in the figure environment.

1.3 Tables

Use vertical lines sparingly in tables. They're unnecessary bloat. Write the code for tables in a separate tex file, and include it in when required. Also preferrably use sans serif font for tables (because of their information density) using `sffamily` in table definition.

Spam	Ni	Swallow	Shrubbery
A	1	2	3
E	3	4	5
C	6	9	3
M	4	1	1

Table 1.1: This is a table

2. Machine Learning Techniques

To aid in the classification of supernovae, machine learning techniques provide flexibility and scalability that have been previously unattainable by manual methods. The simplest form of machine learning algorithms are the feed-forward fully connected neural networks (FCNs). These networks apply a series of linear transformations to the input data, followed by a bias and a non-linear activation function. The output of the network is another vector tailored for whatever the purpose of the FCN is (i.e. for classification, is a vector of probabilities that the input data belongs to, see Appendix app:FCN).

This architecture, while simplistic, is very powerful. Cybenko (1989) showed that for any boolean function (i.e. 2 class classification), a FCN with a single hidden layer could approximate any function arbitrarily well; this is known as the universal approximation theorem. This theorem is the basis for the success of FCNs in many fields, but also the reason why they are not the best choice for all problems. The main issue with FCNs is that they are not invariant to translations, which is essential for many problems such as image recognition, time series analysis, and natural language processing. Therefore, alternative options have been proposed.

2.1 Convolutional Neural Networks

Convolutional neural networks (CNNs) were originally developed as a solution to positional differences in the input data (Fukushima 1979). This architecture was shown to be effective in image recognition tasks by LeCun, Huang, and Bottou (2004), and was popularized after the success of AlexNet in the ImageNet challenge Krizhevsky, Sutskever, and Hinton (2012), and have since been applied for use in many other vision tasks, as well as other fields such as NLP and audio processing. CNNs are composed of a series of convolutional layers, usually followed by a pooling layer, and then a fully connected layer (Appendix app:CNN).

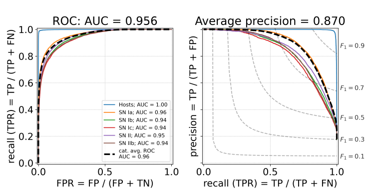
The convolutional layers are composed of a series of convolutional filters (usually referred to as feature maps) which are applied across the entire input space. The application of the same filter across the entire input space is what gives the CNN its invariance to translations.

2.1.1 CNNs on Spectroscopic Data

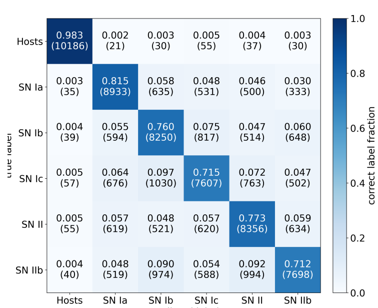
Due to their affinity for disregarding spacial invariance, CNNs theoretically should translate to spectral classification quite well. In fact, CNNs have been applied t o spectroscopic data in the past, even on the DESI dataset. Parks et al. (2018) used a CNN to detect strong emission lines in DESI spectra with great success. * Talk about 1D CNN currently running, and preprocessing used *

This work provides a baseline for the use of CNNs on supernovae classification, but leaves more to be desired. Therefore, an alternate architecture was proposed by * Cite Eddies work * which augments the preprocessing of the spectra with a conversion to a 2D image. This 2D image was then fed to a more traditional vision CNN architecture (Appendix app:CNN).

This architecture was shown to train quickly (less than 1 hour on a single GPU), but was not able to achieve astonishingly high accuracy. Figure fig:cnn$_q ualshowstheROCcurveandconfusionmatrixfortheCNNtrainedonsyntheticspectra$.



((a)) ROC curve



((b)) Confusion Matrix

Figure 2.1: CNN Diagnostics

The maximum value of the output vector was used to determine the predicted class, which may not be the best choice for this problem. Figure fig:cnn$_m axshowsthedistributionofthemaxi$

From

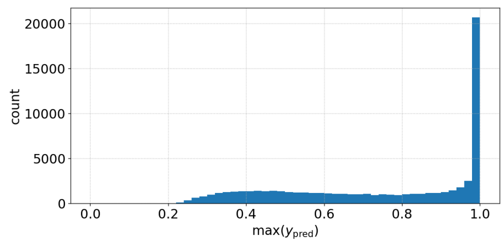


Figure 2.2: Max value of the output vector for the CNN.

2.2 Previous and Current Attempts at Machine Learning

2.3 Introduction of Transformers

3. Validity of Transformers for Spectral Classification

3.1 Creation of artificial spectra

Creating large amounts of high signal-to-noise signals was essential to verifying the application of a Vision Transformer. These signals needed to include the basic features of supernova spectrum, such as the presence of a continuum, absorption lines, and emission lines. The locations of these spectral features in particular were not important, as long as each class of supernovae had consistent features. The `GenData` class was created to generate these signals.

3.1.1 Features of `GenData` class

The `GenData` class must provide a means of generating a large number (order of 10^5) of random spectra. These spectra must exhibit a consistent continuum, variable noise, and a set of spectral features that are unique to arbitrary classes. In order to accomplish this, the `GenData` class must first identify a domain in which to place spectral features, noise, and the continuum. The continuum is a function across the domain that remains consistent for all samples, and examples can be found in Fig. fig:continuumoptions. A set number of spectral features are randomly placed within the domain. The number of spectral features is determined by the user. Then, each class (or type) of spectra is assigned a random combination of these spectral features. This is to ensure that each class has a unique set of spectral features, while maintaining a consistent location of features. Once the different classes are specified, the creation of the spectra can begin.

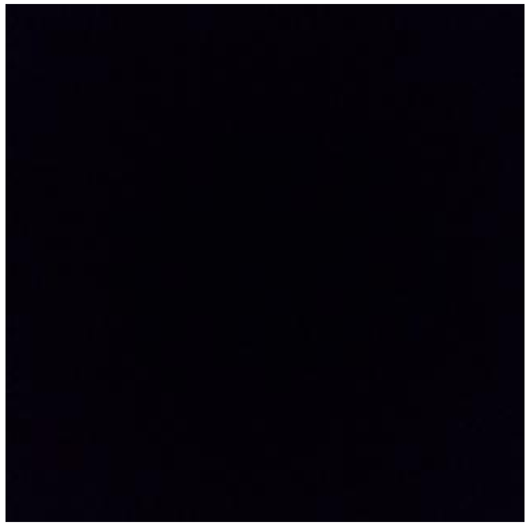


Figure 3.1: Examples of different continuum options (from left to right: linear, linear increasing, linear decreasing, exponential increasing, and exponential decreasing).

3.1.2 Creation of spectra

Once the spectral features present in each class are determined, the spectra can be constructed from the ground up. First, the continuum is created

Bibliography

Clarke, Arthur Charles (1990). *Rendezvous with Rama*. Spectra.

Cybenko, G. (1989). "Approximation by superpositions of a sigmoidal function". In: *Mathematics of Control, Signals and Systems* 2.4, pp. 303–314. DOI: 10.1007/BF02551274. URL: https://doi.org/10.1007/BF02551274.

Fukushima, Kunihiko (1979). "Self-Organization of a Neural Network Which Gives Position-Invariant Response". In: *Proceedings of the 6th International Joint Conference on Artificial Intelligence - Volume 1*. IJCAI'79. Tokyo, Japan: Morgan Kaufmann Publishers Inc., pp. 291–293. ISBN: 0934613478. DOI: 10.5555/1624861.1624928.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton (2012). "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Ed. by F. Pereira et al. Vol. 25. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

LeCun, Y., Fu Jie Huang, and L. Bottou (2004). "Learning methods for generic object recognition with invariance to pose and lighting". In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* Vol. 2, II–104 Vol.2. DOI: 10.1109/CVPR.2004.1315150.

Parks, David et al. (Jan. 2018). "Deep learning of quasar spectra to discover and characterize damped Ly systems". In: *Monthly Notices of the Royal Astronomical Society* 476.1, pp. 1151–1168. ISSN: 0035-8711. DOI: 10.1093/mnras/sty196. eprint: https://academic.oup.com/mnras/article-pdf/476/1/1151/24261051/sty196.pdf. URL: https://doi.org/10.1093/mnras/sty196.

A. Detailed Diagrams of Various Neural Networks

A.1 Fully Connected Neural Networks

A.2 Convolutional Neural Networks

A.3 Transformers

B. Title of Appendix B