Identifying CaII absorbers using Machine Learning

Edward Du[*]

[*] *Leland High School, San Jose, CA, 95120*

ABSTRACT

The results of this research paper are composed of CaIIλλ3934, 3969 absorption spectra from Sloan Digital Sky Survey(SDSS). In this paper, the quasar spectra from Data Release 14 are used, opposed to the earlier releases such as Data Release 7 and 9. I have identified 1260 different doublets from this data release. Instead of the conventional method of finding these doublets, I employed a new technique meant to hasten the process of discovering these doublets. I used a neural network and principal component analysis (PCA) method in order to find the doublets. Of the over 250000 spectra in Data Release 14, I found that only a small percentage are actually CaII doublets. Although the accuracy of 73.9% is not as high as it could potentially be, my new technique at discovering metal absorption spectra could mean a shift in the methods used to find these absorption lines and similar ones in the near future. The process of discovery for other rare metal absorbers such as FeII, MnII, and NaI (each with a full sample of less than 300 spectra) can be accelerated.

1. INTRODUCTION

1.1 Background

First, there was nothing.

Then, 13.82 billion years ago the universe burst into existence—a huge release of energy that would be called the Big Bang. It happened in total darkness because light didn't even exist. Today, the idea of the Big Bang is accepted. Most astronomers and scientists believe it was just a miniscule and infinitely hot particle expanding at the speed of light, creating the visible world we see today. As it expanded, it cooled down at an extremely fast rate. It created matter in the form of countless trillions of subatomic particles, and finally, hydrogen and helium formed. Galaxies formed, stars were created, and heavier elements were created due to fusion at the center of these stars. As the stars die, these elements are released into space, leaving gas and dust in its wake. Thus, the dust and gases can help us better understand the formation of galaxies.

When gases and other materials get in between the observer and a quasar, analysis of the absorption lines of the spectra from the quasar can help one better understand the type of material. The importance of discovering these systems can be attributed to the importance of the early universe. Being able to discover these quasar absorption lines (QALs) enable us not only to examine the formation of galaxies but the properties of it. Since every element has its own unique absorption spectra, as shown in Figure 1, by figuring out the element that created the spectra, We can infer that somewhere between the object and the Earth, that element is present. Thus, the discovery of the absorption lines can give us the opportunity to study both the physical and chemical properties of regions of the universe that were previously not possible to assess via traditional study methods. According to Bahcall and Spitzer (1969), it is believed that these lines are formed from the halos of galaxies. These beliefs are plausible because galaxies are a center of star formations and deaths, which contains metals.

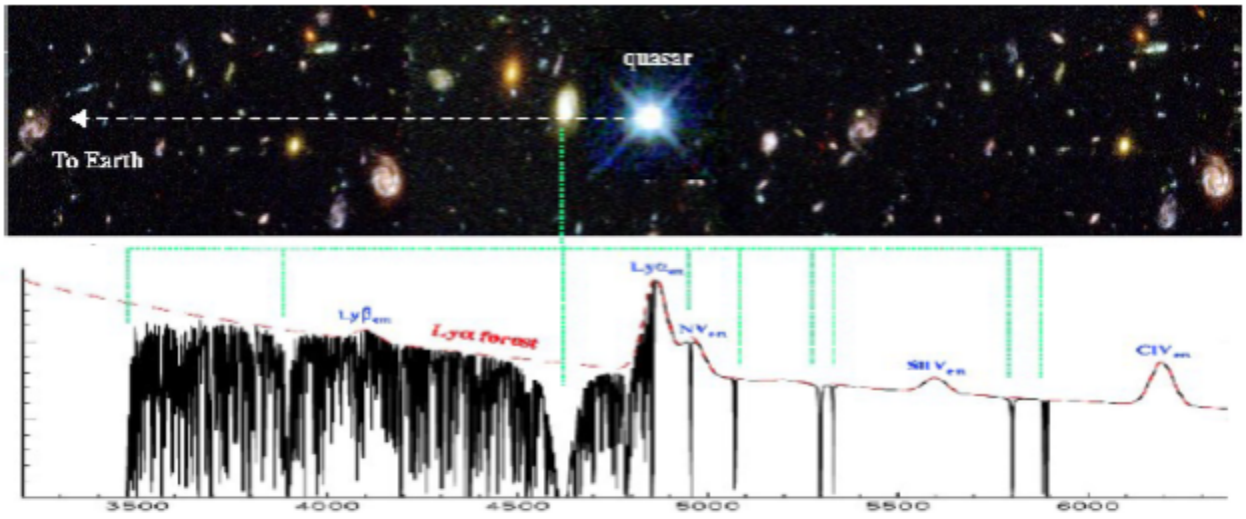

Figure 1. Light from the remote quasar is selectively absorbed by gaseous structure in between. The observed spectrum at the earth is shown at bottom that present many absorption at different wavelength.

Redshift also plays a significant role in the path to the discovery of these systems. The Doppler effect is the main cause of redshift and its counterpart, blueshift. When a police car with its siren on races towards a person, he or she will hear a higher frequency than a person in which the police car is racing away from. This is due to the compression of wavelengths, which leads to an increase in frequency via the formula, $v = \lambda f$. The exact same concept applies to light, but the v is replaced with the constant for the speed of light, c. Thus, the formula would be $c = \lambda f$. Unsurprisingly, an object moving (albeit at high speeds) away from a subject would appear to be shifted towards the red side of the spectrum, while an object moving towards a subject would appear to be shifted towards the violet or blue side of the spectrum. But why is everything in the universe redshifted? In order to understand, one must understand the concept of the expansion of the universe.[1] As the universe expands, everything inside it also moves outwards, becoming further and further apart from any object (unless bound together by gravity). As a result, every object that we observe is moving away from us at an extremely high speed, causing it to become redshifted.

1.2 MOTIVATIONS

Ultimately, the goal of studying absorption lines in the spectra of quasars with high resolution is to learn about the detailed chemical, ionization and kinematic conditions and provide detailed data to be

leveraged against fundamental problems in cosmology and astrophysics. Many of the problems include questions on galactic formation and evolution, cosmic chemical evolution, and cosmic evolution of the UV ionizing background. From the spectrum in Figure 2, we can observe that many different rest wavelengths are from different elements such as CoII, CrII, MgII, FeII, and MnII.

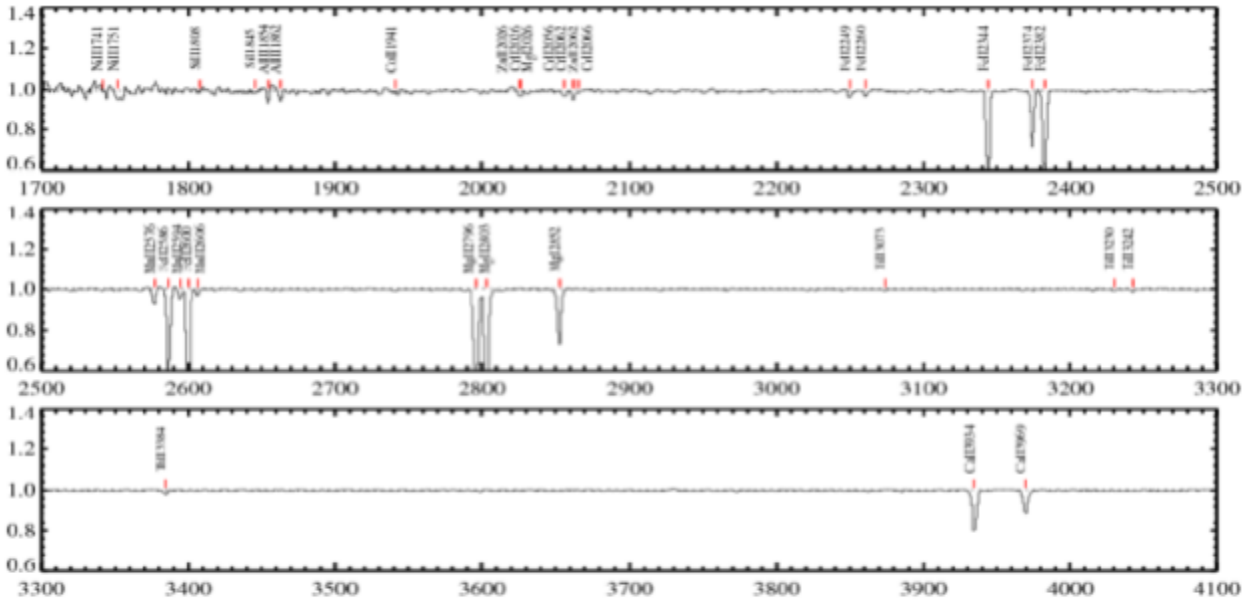

Figure 2. Examples of the Rest Wavelengths of different elements.

In particular, the CaII doublet absorption line is studied in this paper. Compared to other gases that can be studied, it is relatively rare. CaII absorbers have properties that make them unique from other metal absorber types. Being able to discover these allows us to more closely examine the dust makeup in galaxies. Questions such as the importance of CaII absorbers to the evolution of galaxies could be answered more confidently. The CaII H & K doublets have wavelengths of 3934 and 3969, respectively, meaning that the rest frame wavelength of CaII H is $\lambda = 3969.60$ Å, and for CaII K, $\lambda = 3934.78$ Å. These wavelengths are for ionized calcium, Ca^+, with can be photoionized if the energy reaches 11.87 eV. Yet, Calcium only requires 6.11 eV to become photoionized, which hints that Ca^+ may not be the most abundant state of Calcium. CaII is a rare absorber, yet it is still an important measure of properties of dust. Density, self-shielding abilities, and dust content are examples of how CaII can be used to examine interstellar dust in space.

One group, Sardane et al.(2014), discovered CaII 435 doublets in Data Release 7 and 9. However, the 435 discovered CaII doublets may not be enough to experiment with. It would be too difficult to draw statistical results with a small amount of data. Even more, the process of discovering

these metal absorption lines took up to weeks, making the discovery of these less efficient. However, with the huge increase in computing speed and capabilities over the years, it would be useful to utilize that potential. Thus, I turn to the field of machine learning.

Convolutional neural networks (ConvNet) have been widely used in machine learning and signal processing due to their impressive performances. They are of particular interest in this case since the complexity of the prediction of CaII absorber line in the universal radiation. In this project, I have employed a deep neural network, more specifically, a ConvNet, to lessen the manpower and increase the efficiency. Neural networks have already proven to be useful in the real world: computers can recognize handwritten digits and letters, predict the weather, and anticipate heart attacks.

In the ImageNet competition of 2012, Alex Krizhevsky created his convolutional neural network that was able to best state-of-the-art techniques at the time. From then on, his ConvNet has been known as an AlexNet, containing five convolutional layers and three fully connected layers (Krizhevsky et al. 2012).

In this project, I make use of an AlexNet to search for CaII absorbers in SDSS Data Release 14. In Section 2 of this paper, the datasets that I use and the preprocessing process is described. In Section 3, the structure of the neural network is described. I report my results in Section 4, and I go back to discuss ways my technique can be improved and the implications of it in Section 5..

2 DATA AND PREPROCESSING

For my research, I cannot just download a dataset and plug it into my ConvNet, assuming it will work properly. I have to preprocess the data in such a way so that it is "readable" by the machine. In addition to cutting out some absorption spectra, each spectrum is rescaled and interpolated to obtain my final spectrum. When rescaling, I use the emission redshift (Z_{ems}) to determine the wavelength I use.

2.1 SDSS DATA CATALOG

The Sloan Digital Sky Survey has played and will continue to play a pivotal role in the field of astronomy. SDSS provides astronomers with hundreds of thousands of quasar spectra to look at, which can potentially lead to many discoveries. Each dataset, a huge number of new quasar spectra is released to the public, allowing more surveys to be done with the absorption-line spectra. The most recent dataset,

Data Release 14, is the second data release of SDSS-IV.[2] The observations in this dataset were taken up to July 2016, and includes images, three types of spectra (optical, infrared, IFU), and catalog data. The spectra are taken on a 2.5 meter telescope sitting on Apache Point Observatory, with two upgraded spectrographs, one in the blue channel, one in the red channel, that were rebuilt on the original spectrographs. Combined, the two spectrographs can cover wavelengths from 3600 Å to 10400 Å. Each spectrograph also has 1000 fibers with diameters of 2 arcsec, an improvement from the previous 640 fibers with diameters of 3 arcsec. In my search for CaII absorption lines, I only used spectra with a redshift that is less than 4.8.

2.2 PREPROCESSING

One of the catalogs that I used for this project was the DR14 catalogue. First, I had to download all the spectra plate files from the bulk data download webpage: https://www.sdss.org/dr14/data_access/bulk/.
Using the name of the files, the mjd and the plate of the spectra can be determined. For example, from a spectra named "spPlate-3586-55181.fits", the mjd would be 3586 and the plate would be 55181, providing a means of identification for the spectra. Also, I used the catalog that can be found on the website of SDSS:
https://www.sdss.org/dr14/algorithms/qso_catalog/.
The filename is DR14Qv4_4.fits. This is the final version that contains all DR14 quasars. According to the website, compared to the previous version, some wrongly identified objects were removed and missing coordinates were added. In order to obtain my final dataset for testing, I had to filter and process the catalog as shown in Figure 3. When I filter the catalog, I print out the strong MgII information. Obtaining the MgII absorbers is helpful to finding the CaII absorbers because many of the CaII doublets are from MgII systems. After running a catalog check, approximately 50000 quasar spectra are determined to be in DR14. Then, preparing the raw data, I get the final data matrix containing a list of possible doublets. Then, I can input the data into my trained ConvNet to get the results.

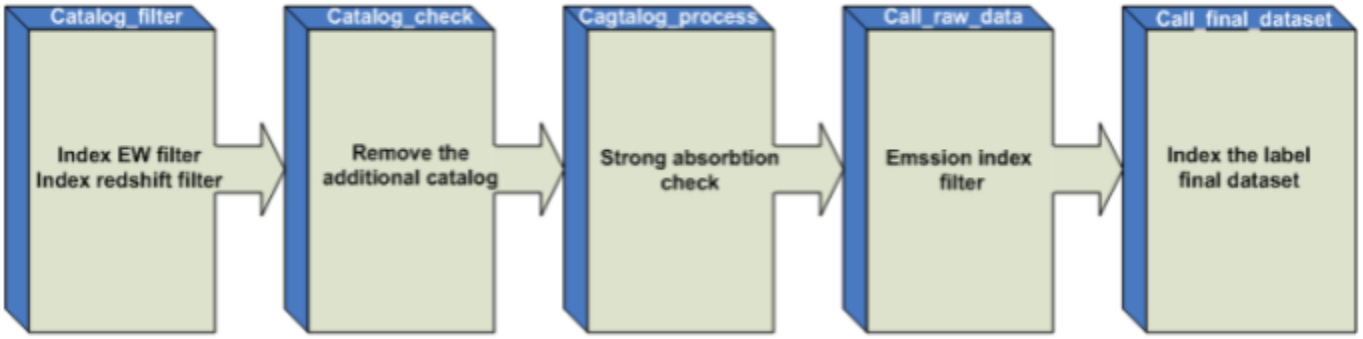

Figure 3. This is the SDSS-DR14 quasar catalog preprocess flow that is suitable for identifying CaII absorbers. The quasars with null coordinates, empty columns are removed.

2.3 ARTIFICIAL SPECTRA FOR TRAINING

In this project, artificial training sets are used to train the neural network. The reason for this is because there are very few discovered CaII absorbers: too few to be able to to train a neural network. While preprocessing, 24 bins are created to separate the spectrum up into sections. For example, bin1, which contains the redshift range from 0.0 to 0.2, has a wave window between 3166 and 9200 Å. These are saved into a folder named Qgen_bins, which contains 24 files, named Qgen_zbin0.hdf5, Qgen_zbin1.hdf5… Qgen_zbin23.hdf5. Then, for each of the bins, the nan values in the spectra are removed, Iterative Principal Component Analysis (IPCA) is performed, and the spectra is projected in eigenvector space. These files are saved into a folder called Qgen_eigen in the format of eigenvector_%s, where %s is from 0 to 23. After this, a PCA fitting is created, saved into a folder called Qgen_fit. Using the files generated from the fitting, the signal-to-noise ratio(SNR) can be measured by taking the average of all the final flux values divided by the final error values. Again, I save these files into a folder called Qgen_SNR. The Figure 4 demonstrates the quasar sampling spectra from Data Release 7. This is the raw data model for the artificial training data set creation. It is plotted from the first stage of processing.

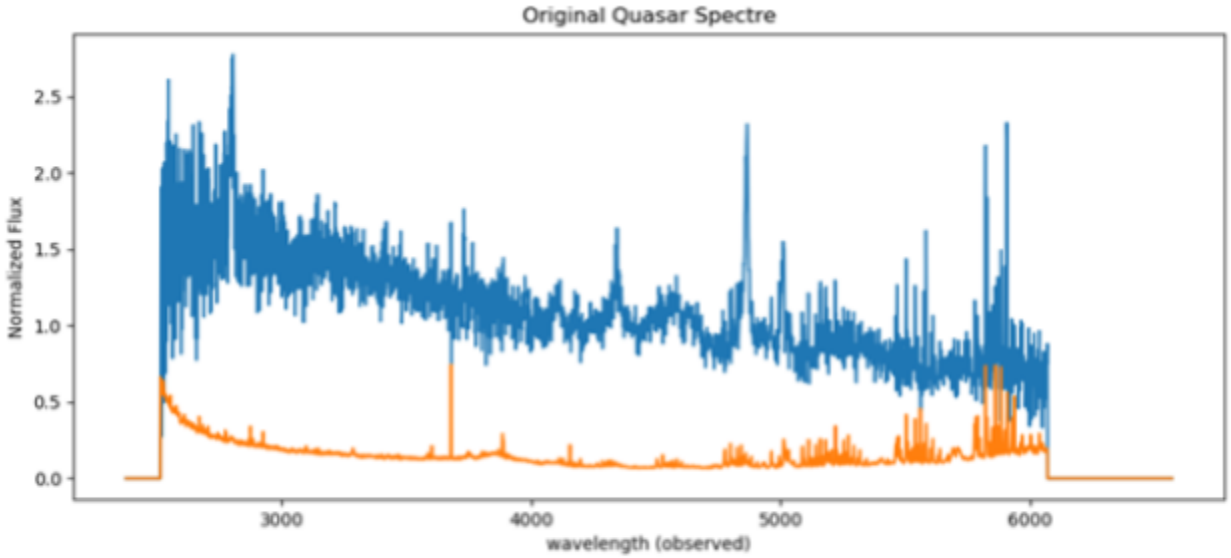

Figure 4. Illustration of the original quasar spectrum. The blue color wave is the spectra data. The orange color wave is the error.

The following step is to creating an artificial training data set. This training data set is to inject absorption lines into the spectra in my subsamples. I import the files from Qgen_fit and Qgen_SNR, and filter out the spectra that are outside smaller than 3800 Å or greater than 9200 Å. As shown in the Figure 5, The CaII doublet absorber data is injected in which is pointed by red arrow.

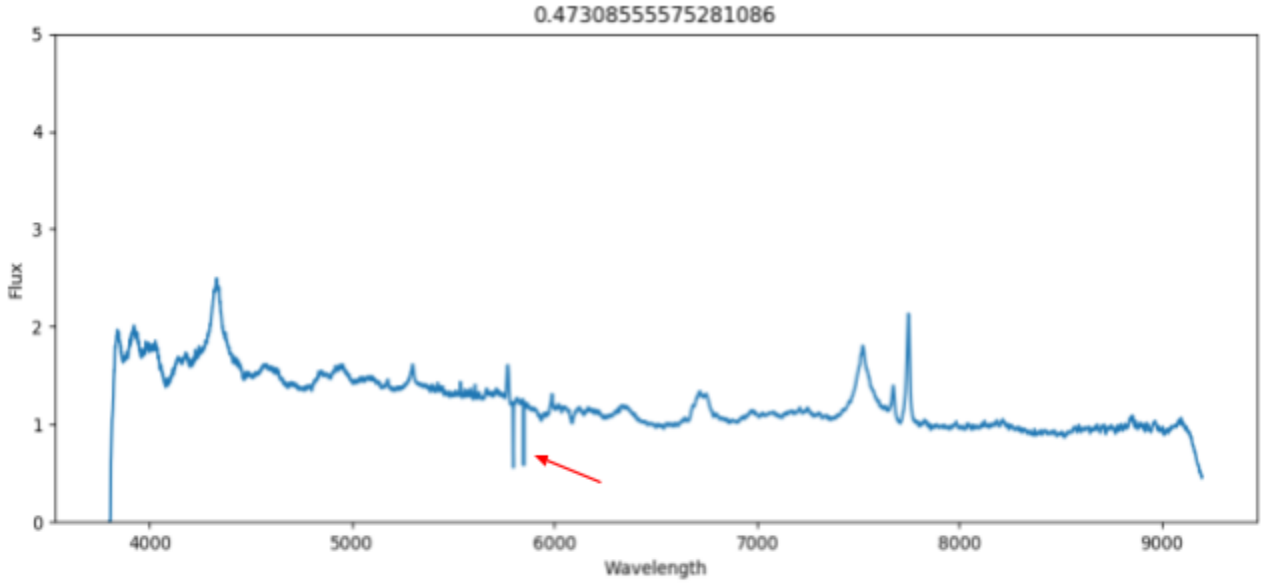

Figure 5. The blue curve is the continuum fitting of the quasar spectrum with PCA method. The quasar continuum is inserted with a CaII doublet absorber, where the dip on the left is H and the dip on the right is K.

After this, I filter out all the nan values, and keep the spectra with a SNR that is greater than 3.0 ($\sigma > 3.0$). Noise is injected into the spectra with the error being greater than 3.0, following the Gaussian distribution. Overflow is then prevented by normalization. After this, I fit the absorption lines into the continuum and derive the full width at half maximum (FWHM). These are saved into a folder called Qgen_training, and in total, 69584 simulated spectra are generated, with a portion of them having a strong CaII profile. I combine the subsamples by running another script, which provides us with the dataset that is actually used in training my ConvNet. Figure 6 shows the final CaII doublet absorber line.

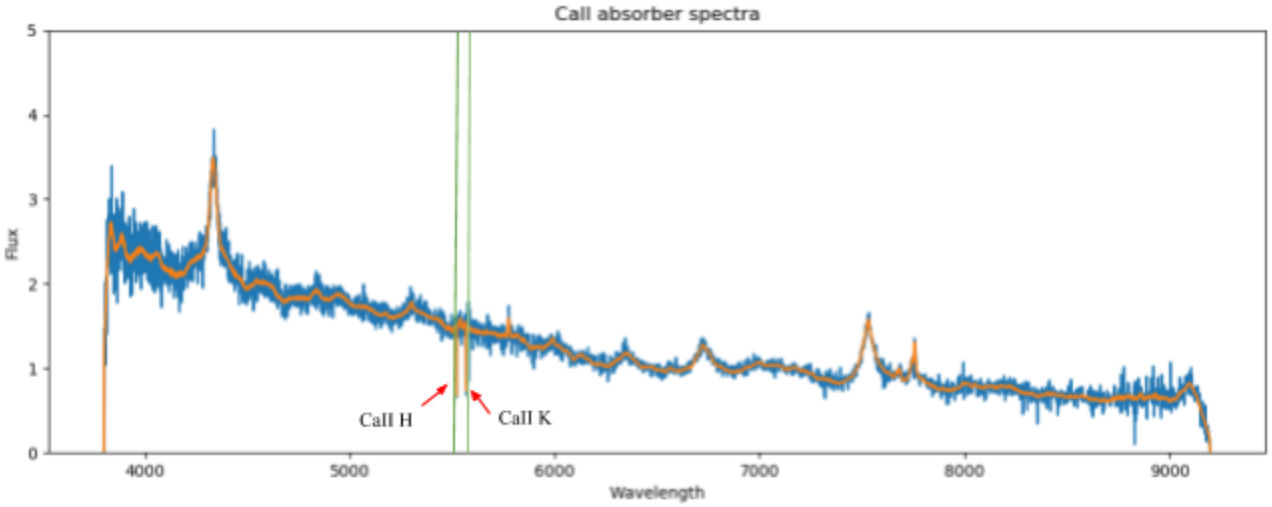

Figure 6. The inserted CaII doublets absorber is marked with straight blue line. Gaussian noises is injected into the quasar continuum with CaII absorber.

3 STRUCTURE OF THE NEURAL NETWORK

In this project, the deep neural network is one of the most important steps. It is what creates the prediction that is ultimately used to generate results. The components of a neural network are an input layer, some hidden layers, and a final output layer. An input layer is layer in which the photo or data is inputted. The input units in this layer are sent to each hidden node in the hidden layer. In this project, the input layer has a feature vector of 3841 units, because quasar spectra has a dimension of 3841.

The next layer in a neural network are the hidden layers. These hidden layers are meant to simulate parts of the human brain—a brain takes in inputs as probabilities and converts them to outputs. In this project, the weighted inputs are fine-tuned through a method called backpropagation.

The last layer is called the output layer that produces the end result. The outputs that they provide are not concrete: they are in the form of a probability.

3.1 THE CONVOLUTIONAL NEURAL NETWORK

In this project, a Convolutional neural network (ConvNet) is used. What makes a ConvNet different from a typical neural network is its used of neurons in 3-dimensions.[3] A ConvNet consists of multiple layers, such as a Convolutional Layer, a Pooling Layer, and a Fully-Connected Layer. Generally, the architecture of a ConvNet is as follows: CONV, RELU, POOL, FC. Some hyperparameters are also needed to be specified in order for the ConvNet to function; some of the parameters are depth, stride, and zero-padding. The depth is the amount of filters needed, the stride is needed in order to specify the distance in which the sliding window jumps, and the zero-padding is the amount of layers of zeroes on the border of the input volume. In doing so, the output volume can be controlled. The output layer can be calculated as:

$$O = (W-K+2P)/S + 1, (K = \text{filter size}, W = \text{input length}, S = \text{stride}, P = \text{padding}).^{4}$$

After the Convolutional Layer, I apply a ReLU (Rectified Linear Units) layer. Nonlinear activation functions such as tanh and sigmoid were used.[5] Both sigmoid and tanh were great because they had the activations bound to a certain range, and near the middle, the slope of the graph is steeper than on the ends. However, a problem arises from this: vanishing gradients. This occurs because at the ends of the function, a change in x corresponds to a miniscule change in y. This can be fixed with the ReLU function, given by $f(x) = \max(0,x)$. It seems linear, yet it is a nonlinear function, with a range of $[0,\infty)$.

After the ReLU layer, I have a pooling layer. The point of a pooling layer is to get rid of the information that is not as useful. In this project, I apply max pooling opposed to the other two types: L2-norm pooling and average pooling. In max pooling, a filter is used and goes around the input layer, outputting the maximum number in each section. By reducing the spatial dimension, the amount of computing power needed is much less. Pooling is also utilized in order to prevent overfitting, which is when a model is overtrained and starts connecting irrelevant parts of an input to an output. A symptom of overfitting is if the program gets a very high accuracy rate on a training set, but a much lower accuracy rate on the test data.

After the pooling layer, a dropout layer is utilized: another method to prevent overfitting. In the dropout layer, random units and their connections are dropped, by setting them to zero (G. Hinton et al. 2014).

In the nodes of the hidden layers, a sliding window of a fixed length slides across the input with static weights. For each given node, a feature map, or activation map, is created from the output activations. A neuron can be written as

$$\Sigma(\text{weight} * \text{input}) + \text{bias}.$$

The feature map can be written as:

$$A^k_{ij} = (\sigma)\ ((W^k * s)_{ij} + b_k)$$

Where A^k is the feature map at node k, W^k are the weights, and b_k are the biases.

In order to identify CaII absorption lines from other metal absorption lines, I employ binary classification. With binary classification, I use a sigmoid function, given by the function:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

To learn about thousands of objects from the half millions of the data, I need the model with a large learning capacity. Around 7 popular deep neural network models can be used as model structure for the training. I adopt an AlexNet training model which is popularized Convolutional Networks in Computer Vision was the AlexNet, developed by Alex Krizhevsky, Ilya Sutskever and Geoff Hinton. The revised model presented in Figure 7.

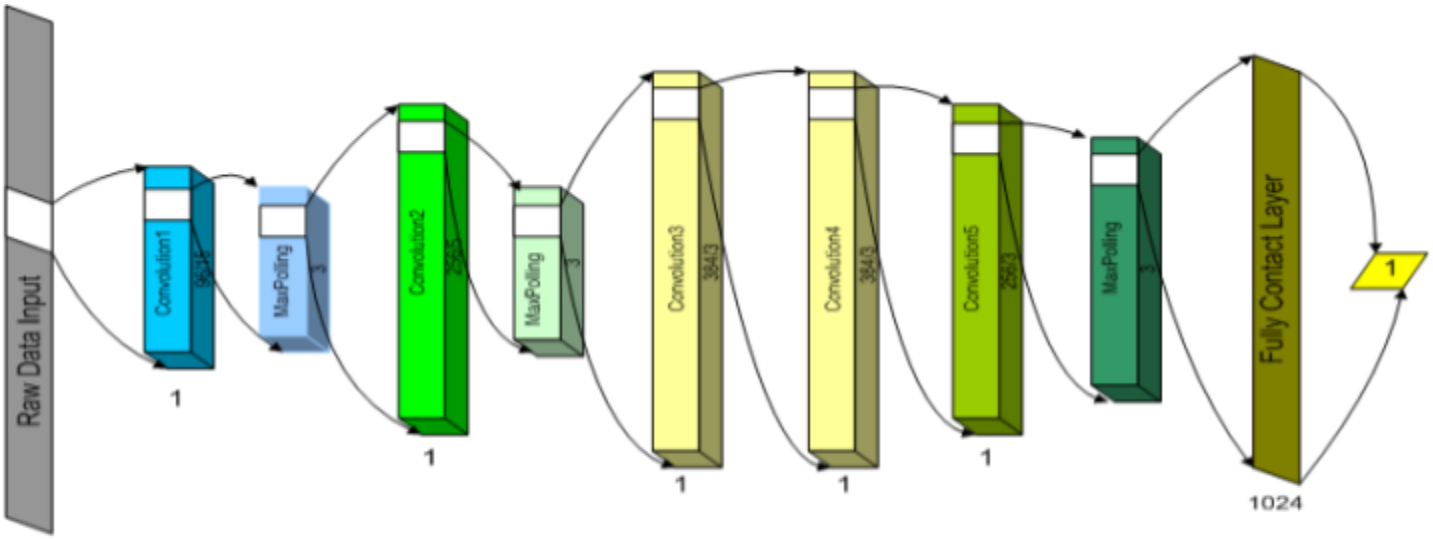

Figure 7. This diagram is the diagram of a Convolutional Neural Network AlexNet model. In my project, I used 5 convolutional layers and 3 pooling layers, with an input layer of dimension 3841 single dimension quasar spectra.

The neural network in this project was trained by an artificial training set and a real spectral dataset, and it has 5 hidden layers, 3 pooling layers, and 5 stages. In my first layer, I have a kernel size of 15, and for the other 4 layers, the kernel size is 5, 3, 3, and 3. I choose to use Glorot normal initialization as described in Bengio and Glorot (2010). It was programmed using Python 2.7 with libraries such as Keras and Tensorflow.

3.2 MODEL IMPLEMENTATION AND PARAMETERS

The model implementation has been done using the Python toolboxes Keras and Theano that allow for a fast and scalable prototyping. They implement the use of a GPU for learning the neural network (with SGD) and provide efficient compiled functions for predicting with the model. Learning a neural network requires to choose a number of parameters that is reported here for research reproducibility. I set the learning rate (step of the gradient descent) to 0.001 with a momentum of 0.9 and I use a Nesterov-type acceleration. The size of the minibatch is set to 50 and the number of epochs(number of times the whole dataset is scanned)is limited to 15. In addition, a callBack function is implemented and the tensorBoard can load the callBack data to get a view on internal states and statistics of the model during training.

3.3 PLATFORM OF CONVOLUTIONAL NEURAL NETWORK

Despite the attractiveness of ConvNets and despite the relative efficiency of AlexNet architecture, it still needs the high performance CPU to handle the high volume data. Luckily, current GPUs, paired with a highly-optimized implementation of 2D convolution, are powerful enough to facilitate the training of large ConvNets without severe overfitting. Training occurred on both a laptop with a Geforce GT X 1060 GPU, and also a PC with a Geforce GTX 1080 Ti GPU. With 69584 artificial spectra and 15 epochs, training takes approximately 2 hours on GTX 1060 and around 1 hour on the 1080 Ti.

3.4 TRAINING DATA SET IN CONVNET

The artificial CaII doublets training set that I used contains 69584 simulated spectra samples. This data set is as the training set. There is another artificial data set without CaII doublets which contains 17396 samples, it is used as validation data to run test. Both data set are obtained from the DR7 catalog.

4. RESULTS AND CONCLUSION

4.1 WHY ARTIFICIAL SPECTRA

As mentioned in the 'Introduction', Sardane et al.(2014) discovered CaII 435 doublets in SDSS Data Release 7 and 9. However, I used my ConvNet to predict 1260 CaII doublets in the SDSS DR14. The amount of data available is extremely small. It is very common to collect larger datasets in order to make essential use of machine learning methods. Here, a problem arises: some elements have less than a thousand of their absorption line spectra discovered, making it difficult to train a neural network based on real data. However, as presented in section 2.3, I can create an artificial spectra, making a bunch of fake CaII absorbers to train the machine to look for similar patterns. The number of artificial spectra is around 60000, which is suitable for training, yet definitely could be improved. In this way, the problem of not having enough data is minimized. Creating artificial spectra is obvious not as good as having real data, but once more and more CaII absorbers are discovered, training sets can evolve from artificial spectra to real spectra.

4.2 CAII DOUBLET RESULTS

Though the ConvNet training on the artificial spectra data set with CaII strong doublet absorption lines and the SDSS DR14 Quasar Catalog data, the prediction model is built, by using this ConvNet model on DR14Q_v4_4.fits (Final SDSS_DR14 quasar catalog - Pâris et al. 2018).

From the ConvNet model prediction, there are the number of 1260 predicted CaII doublets. A part of set is presented in the table below.

MJD	Plate	Fiber	Z_{qso}	Z_{abs}	W_0^{3934}	W_0^{3936}	SL^{3934}	SL^{3936}
55499	4296	616	1.432	0.819	2.161	0.720	6.248	5.013
56567	7145	200	2.801	0.861	1.479	1.267	3.322	3.522
56544	6877	388	2.270	0.812	2.689	2.285	3.188	5.077
56238	6173	459	1.174	1.157	1.884	1.757	3.751	6.809
55477	4216	860	2.155	1.008	2.520	1.796	8.532	7.020

Then, in a histogram, I graphed the number distribution for Z_{abs}.

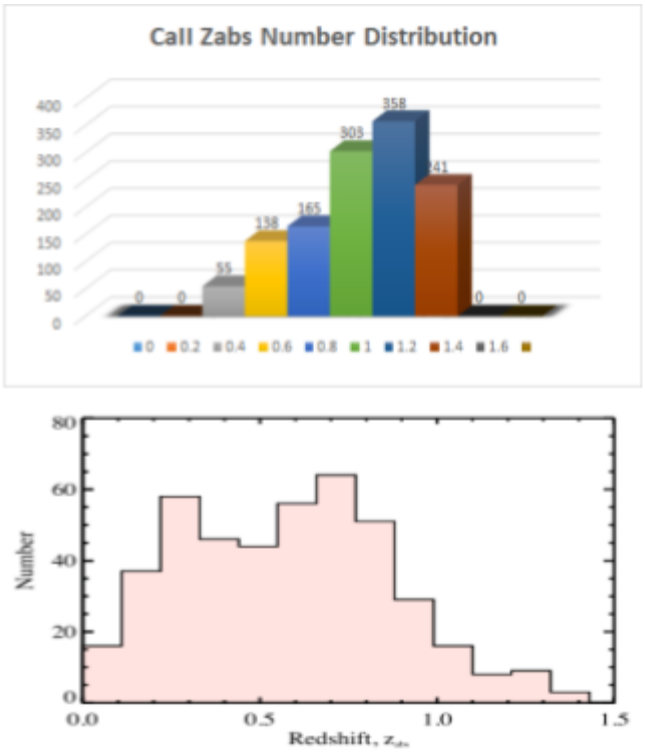

Figure 8. The histogram on top is from my research. Based on my predictions, we identified 1260 CaII doublets within $0 < Z_{abs} <= 1.5$ from the DR14 dataset contains 526356 different quasar spectra, as shown in Pâris et al.(2018). The observed CaII doublets distribution is shown in Figure. Most of the spectra fall in the redshift range of 1.0 - 1.2

The bottom image is from Sardane et al. (2014), where they identified 435 CaII doublets from about 95000 Data Release 7 and Data Release 9 quasar spectra. Most of their spectra have redshifts between 0.5 and 0.8.

When comparing figures, I realized that the figure of Sardane et al. (2014) had a mean of z_{abs} = 0.579, while mine had a mean of 1.047. A source of this discrepancy may be because my machine had an accuracy of only 73.9%, which means that about one-fourth of the spectra I identified may be false. Another source of this discrepancy may be because Data Release 14 contained more quasars with redshifts a relatively high redshift than low redshift.

Figure 9 is an example of CaII doublet which is predicted by this ConvNet training model. The position of CaII absorption lines are marked with the green lines. A pair of Gaussian line profile separated in wavelength by the redshifted doublet separation is fitted to the accepted candidate's profile of W^{3934}_{0} distribution.

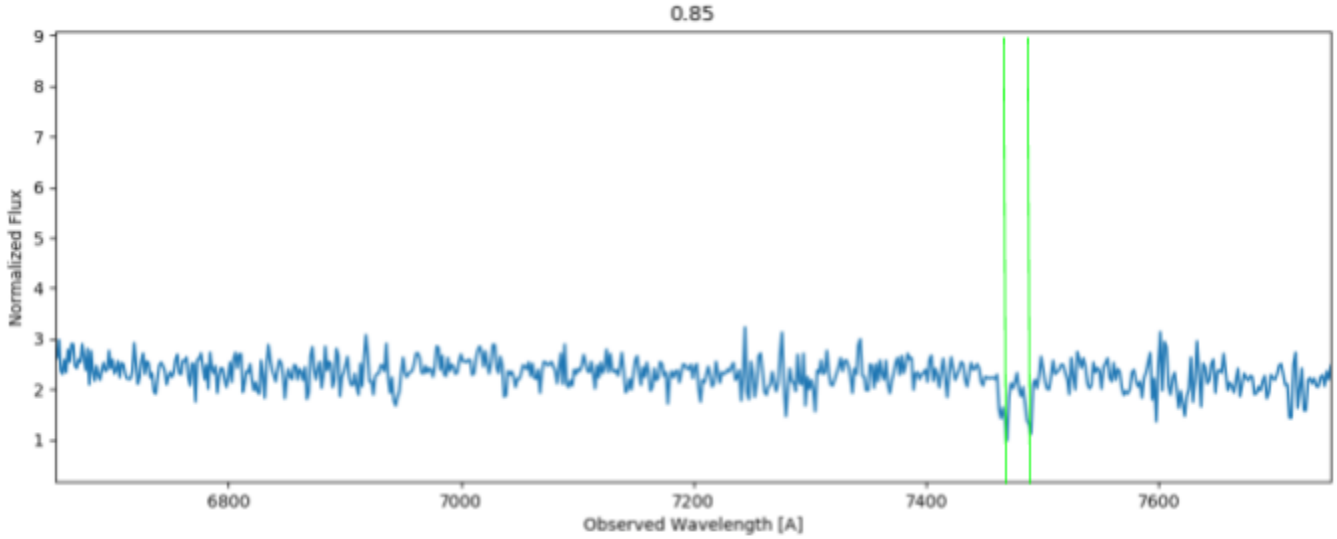

Figure 9. Quasar spectra that contain CaII doublets. The x-axis is the wavelength in Angstrom units, and the y-axis is the normalized flux. The title of the graphs is the name of the spectra, in this format: spSpec-mjd-plate-fiber.fit.

4.3 SUMMARY

In this paper, I developed a ConvNet for the CaII doublet absorption spectra detection in the SDSS DR14 quasar spectra catalog. The CaII absorbers in the SDSS research accomplishment by Gendith M. Sardane, David A. Turnshed and Sandhya M. Rao provided very fundamental information. Many techniques were implemented in this project, such as a continuum fitting, a line-finding algorithm, and a PCA fitting. Thanks to these methods, I am able to create an artificial spectra for my neural network, bolstering the accuracy of it. I use a Gaussian function to inject noise into my artificial CaII doublet absorption line spectra in order to emulate the real world. At the moment, there are only 435 CaII doublets identified in the SDSS DR7 and DR9 quasar spectra. This is not enough to draw conclusions out of, so with the

artificial spectra, the accuracy rate is raised. With SDSS Data Release 14, I discovered 1260 CaII doublets, the most every discovered to date, thanks to machine learning. Below, in Figure 10, is my Tensorflow training accuracy chart with the AlexNet model.

4.4 IMPLEMENTATIONS OF THE NEURAL NETWORK

As mentioned in Section 1 and in Figure 2, there are different absorption line spectra for every element, such as Cr and Mn. Even though the focus of my research paper is on CaII doublets, the functionality of the neural network can spread to these other metal absorption lines. Since there are many rare metal absorption lines out there, some even more rare than CaII, the technique of creating artificial training sets for the spectra can be extremely helpful. With the methodology of this research paper, metal absorption lines will be able to be identified quicker and more efficient than before, allowing more and more discoveries to be made.

4.5 THE TENSORBOARD CALLBACK REPORT

As shown in Figure 10, the training accuracy is around 90% and the validation accuracy is about 75%. This tells that my approach still has plenty of room for improvement. During my training, I noticed the artificial training set is one of most important factors that affect the training and validation accuracy rate. The neural network modeling and related filter size also is one of factors that affected the training and validation accuracy rate. The next stage of my study is to build analysis model to find out to minimize the amount of missed predictions of the neural network. Once the model has been perfected, the methodology used can be transferred to other metal absorbers in the sky, and datasets can be looked at again with this neural network to discover missed spectra.

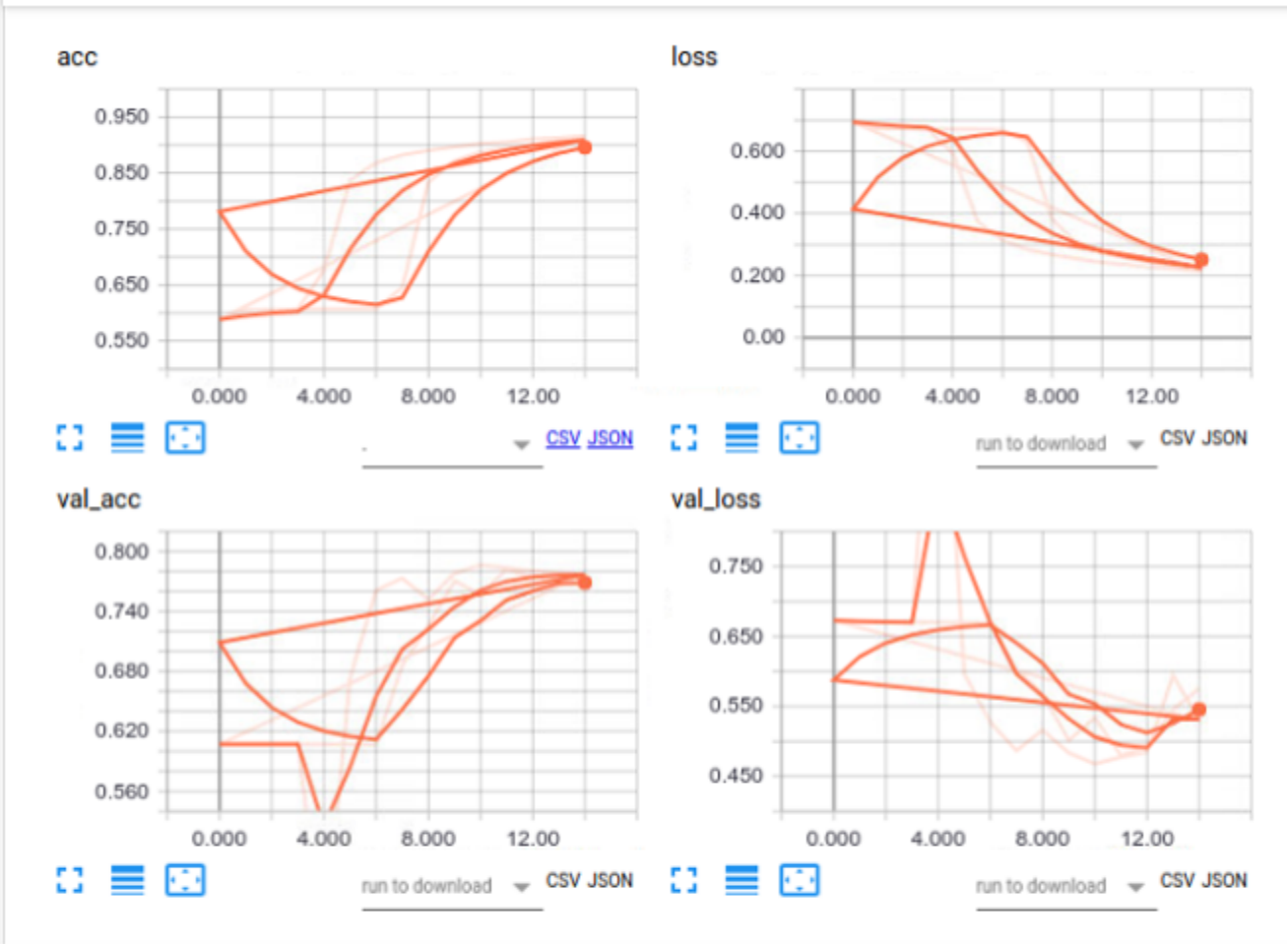

Figure 10. In these graphs, the accuracy and loss of training and testing are given. The top two graphs are from the training set of artificial spectra, while the bottom two graphs are from the testing set that we got from Data Release 14. Over 15 epochs, our training with the artificial training set had a higher accuracy than our testing set.

Author

First Author – Edward Du

REFERENCES

Bahcall, John N., Spitzer, Lyman, Jr. 1969, APJ, L63

Glorot, X., Bengio, Y. 2010, PMLR, V9

Isabelle Pâris, Patrick Petitjean, Éric Aubourg, Adam D. Myers…, Astronomy & Astrophysics manuscript no. DR14Q_v1, January 16, 2018

Krizhevsky, A., Sutskever, I. and Hinton, G. E. 2012, NIPS

Paris et al. (2018)

Sardane, G. M., Turnshek, D. A., & Rao, S. M. 2014, MNRAS, 444, 1747

Zhu, G. T., Menard, B. 2013, ApJ, 770, 130

[1]Loc.gov. (2018). What does it mean when they say the universe is expanding? (Everyday Mysteries: Fun Science Facts from the Library of Congress). [online] Available at: http://www.loc.gov/rr/scitech/mysteries/universe.html [Accessed 3 Nov. 2018].

[2]Sdss.org. (2018). Data Release 14 | SDSS . [online] Available at: https://www.sdss.org/dr14 [Accessed 4 Nov. 2018].

[3]Cs231n.github.io. (2018). CS231n Convolutional Neural Networks for Visual Recognition. [online] Available at: http://cs231n.github.io/convolutional-networks/ [Accessed 2 Nov. 2018].

[4]Deshpanda, A. (2018). A Beginner's Guide To Understanding Convolutional Neural Networks Part 2. [online] Adeshpanda3.github.io. Available at: https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks-Part-2/ [Accessed 26 Oct. 2018].

[5]Sharma V, A, (2018). Understanding Activation Functions in Neural Networks. [online] Medium. Available at: https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0 [Accessed 1 Nov. 2018].