ELTE
EÖTVÖS LORÁND
UNIVERSITY

Third Semester Report

Student Information

Name	Killian Maxwell
Department	Doctoral School of Physics
Email	maxwelki@student.elte.hu
Neptun ID	FZ72X9
Supervisors	Andras Kovacs, PhD; Istvan Csabai, PhD
Thesis Title	Mapping the Cosmic Web: From Simulations to the Deepest Galaxy Surveys

Overview

As with the previous semesters, a primary goal in my research continues to be attempting to develop new tools for comprehending the dark energy component of the current cosmological model. This semester I took some significant steps forward in those efforts, particularly in the utilization of AI. I continued practicing my coding skills with small side-projects of my own, and I delved even deeper into the literature on cosmology and deep learning models currently popular within the field. I especially focused on the category of graph convolutional networks, as well as works that have previously utilized the network I have been working with (DeepSphere). Additionally, I enrolled in a class–Radio Astronomy; a subject that helped bolster my understanding of the physical and observational science behind much of the data and theories I often work with. I maintained my participation in regular journal review and exploratory seminar groups as well. Furthermore, I continued to participate in Euclid Consortium events when possible, while managing the consortium's LinkedIn page. Ultimately, my primary objective of developing tools for use in analysis and further understanding of dark energy grew much closer to actualization this semester. I managed to successfully train the Graph Convolutional Neural Network (GCNN) DeepSphere to estimate Omega-matter and Sigma-8 parameters of a given patch of the sky, with tomographic density slices and ISW information, with decent accuracy. This was achieved by preparing the data carefully and fine-tuning the neural network (NN) structure and hyperparameters to my specific problem. I shall provide further information on this work–as well as all the other efforts mentioned above–in the Details section below.

Details

Coding Projects

As I did the previous semester, I continued this semester to take on small "continued learning" coding projects of my own to better my coding skills and instincts. Many of the projects I chose to take on were focused on utilizing python libraries specifically designed to implement machine learning (Tensorflow, PyTorch, etc.). These projects included everything from preparing data in an optimal way for AI to learn patterns, to altering layer structures and hyperparameters to improve the learning process, to constructing proper loss functions that best suit the learning task. These projects provided great insight into my continued work with AI, and were key in guiding me in many of the crucial steps I took this semester.

Literature Review

I continued to deepen my understanding of mainstream research in machine learning and cosmology by delving further into literature. I was able to finish "Cosmology," by Daniel Baumann, however, I regularly returned to it for reference and reminder. Additionally, I happened upon a free, web-based textbook on Deep Learning provided by a few collaborating MIT professors referenced in one of the journal articles I read. I have been utilizing the font of information provided therein to further bolster my understanding of the foundations of deep learning. Beyond the textbook, I have continued to regularly read papers covering content from machine learning models (specifically convolutional neural networks (CNNs)), to exploration of the ISW effect, and to various cross-correlations already being made by various teams around the world. This semester, I gave specific focus to articles by teams that also utilized DeepSphere. I found that others have explored different variations in the layer and hyperparameter structure of DeepSphere. I ended up utilizing their findings to build upon my own explorations. As mentioned in my previous semester report, my supervisor and I concluded that seeking a way to optimally construct an AI-driven analysis of correlations between density maps and ISW maps was a gap that could be filled in much of the known literature. Thus, a great deal of my literature review has been geared towards that end. Thanks to much of the results from other teams working with AI (and specifically DeepSphere), significant strides were made in this direction.

University Coursework

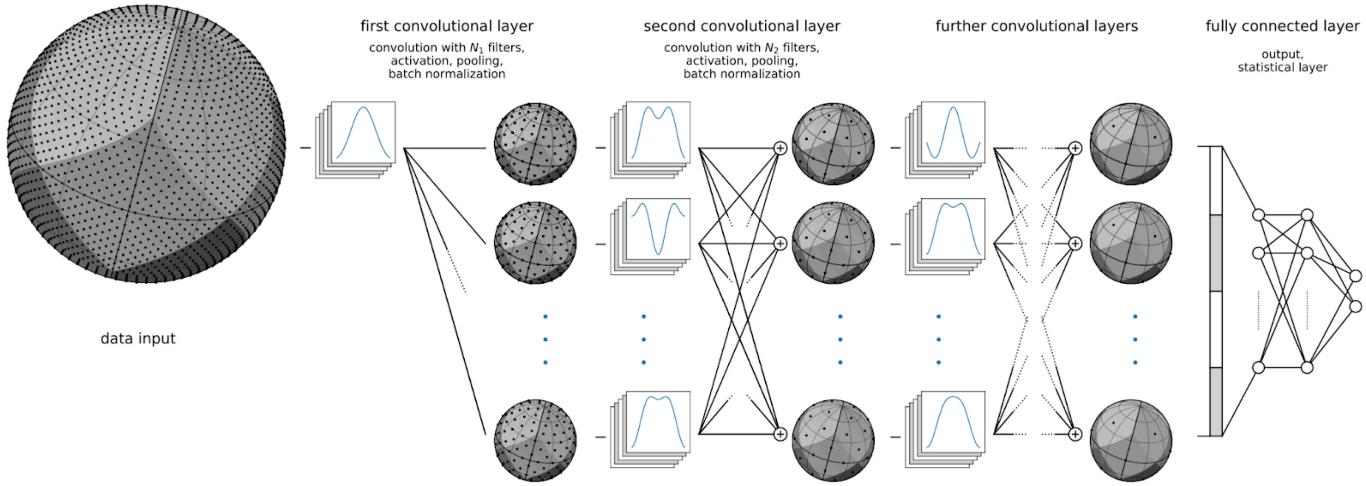
At the university, I enrolled in one course: “Radio Astronomy I.” This class focused on tools of observation and measurement for astronomical objects/features that emit radio signals. It covered much of the structure of radio telescopes and satellites, and how they are able to make the observations that they make. It also expanded on the practical and mathematical theories that drive the derivation of meaning from radio observations. This class was of particular significance to my ongoing efforts. Some of the data I utilize from simulations, as well as many of the theories at the foundation of my explorations, are derived from radio observations. A significant amount of cosmological research can be done using radio-loud tracers. While I am not focused specifically on such tracers, much of the theories I utilize in understanding the data I work with are grounded in the results from research that does utilize them. Therefore, having the understanding gained from this class has helped significantly in the development of my research this semester.

Seminars and Groups

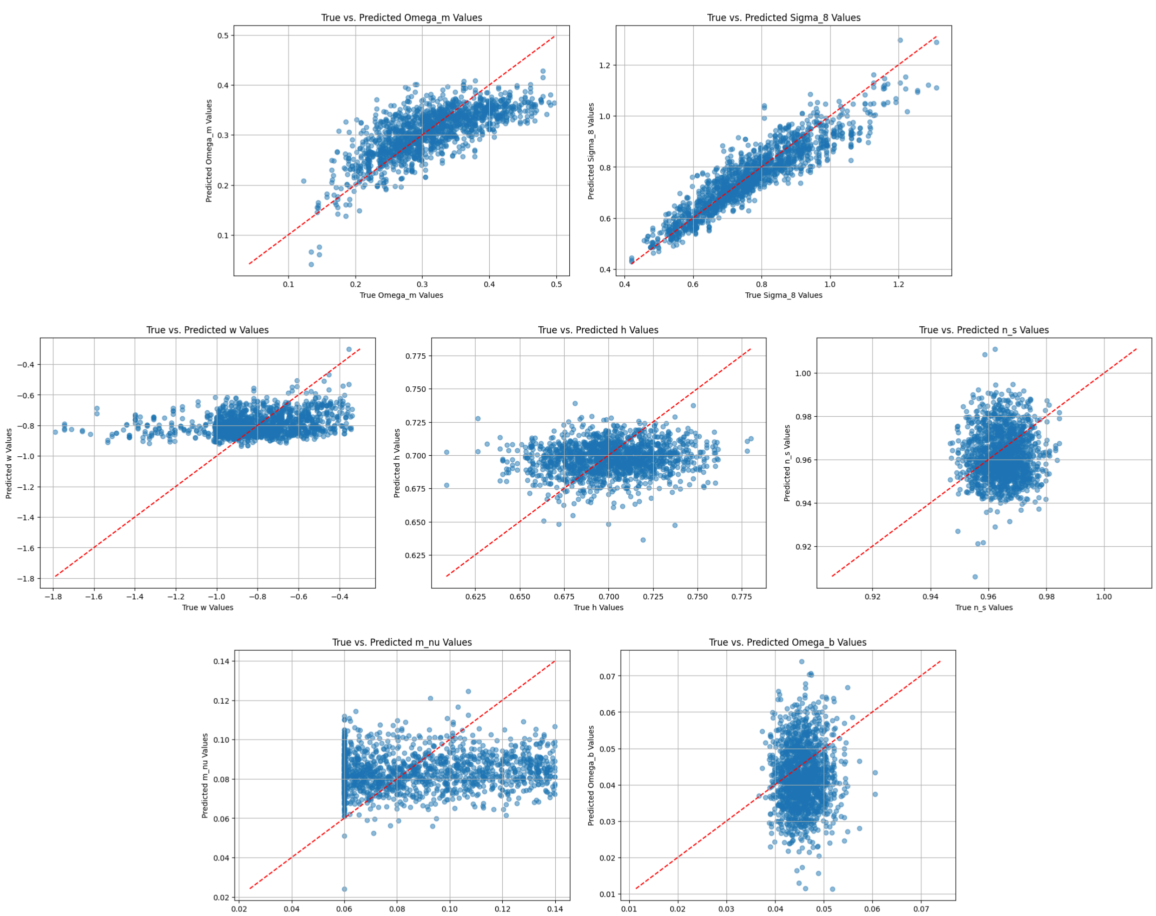
My time participating in seminars and journal review groups remained limited this semester. I became even more engrossed in my research during my time outside of class. I occasionally kept up with the CosmoVerse group, but, more often, I ended up needing to watch the group’s recorded seminars on my own time, outside of the regular weekly meetings. On the flip side, I presented more articles of interest at the regular meetings with my team. In fact, many of the discussions held at these meetings this semester were critical in helping my progress with my research. The frequency of these meetings also increased, and at each we continued to discuss topics for further investigation, present interesting journal articles, and plan out the many progressive next steps we would take in our team’s ongoing projects. It was, in fact, thanks to my fellow PhD student within this group that the ISW maps I utilized in my work this semester were made available to me.

Research Progress

This semester, I made multiple significant steps towards crafting useful AI tools in understanding dark energy. First, I focused my attention on tuning DeepSphere to be able to better predict Omega-matter (Ω_m) and Sigma-8 (σ_8) (this included testing a different layer structure utilized by another team that worked with DeepSphere). Second, I structured the training data to allow DeepSphere to correlate across smaller tomographic redshift bins (rather than a single composite density map across a large difference in redshift). Third, after structuring the training data to be tomographically binned, I also included ISW maps as part of the input. Before delving into the details of these progressions, I’ll first give a brief review of the (original) DeepSphere structure, and where my progress had left off previously.



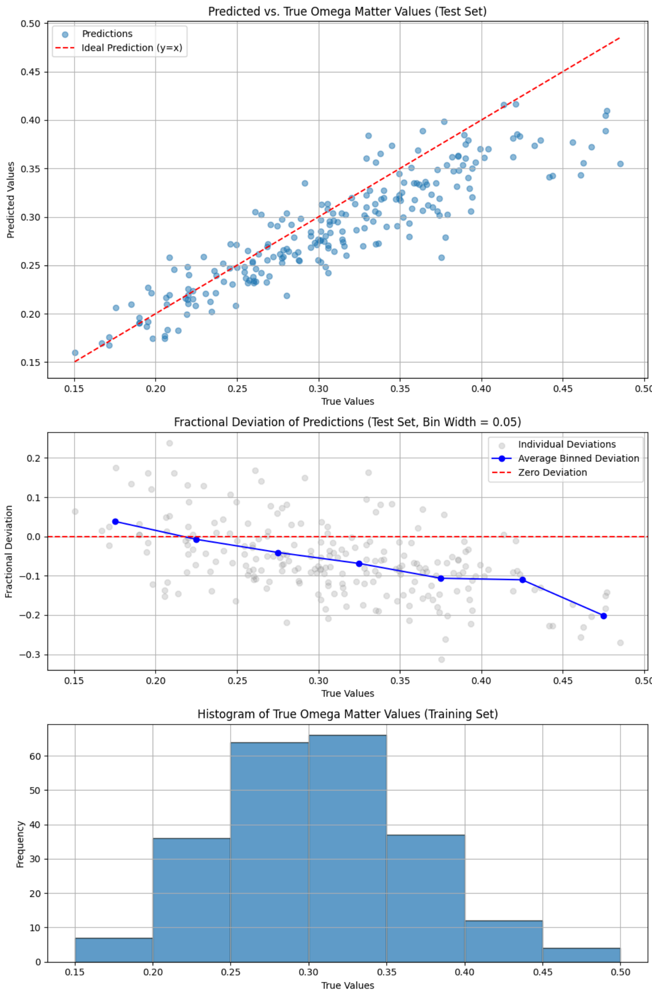
The structure of DeepSphere, as provided in the article by Nathanaël Perraudin, et. al. (arXiv: 1810.12186v2), is shown above. From left to right, the graphic depicts the two or more successive graph convolutions, followed by a system of fully connected (dense) layers. At each layer the data is encoded, via convolution and pooling, and the most significant features of the input are passed on to the next layer. By the time the information reaches the fully connected layer, it has been significantly reduced in size and, in the case of my model's setup, flattened into a single large array of data. Previously, I had DeepSphere structured to output a prediction for each of the seven cosmological parameters implemented in the Gower Street simulations (a suite of full-sky, N-body simulations provided by N. Jeffrey, et. al. (arXiv: 2403.02314v1)) based on a single input map (below, you can see the previous results of these efforts).

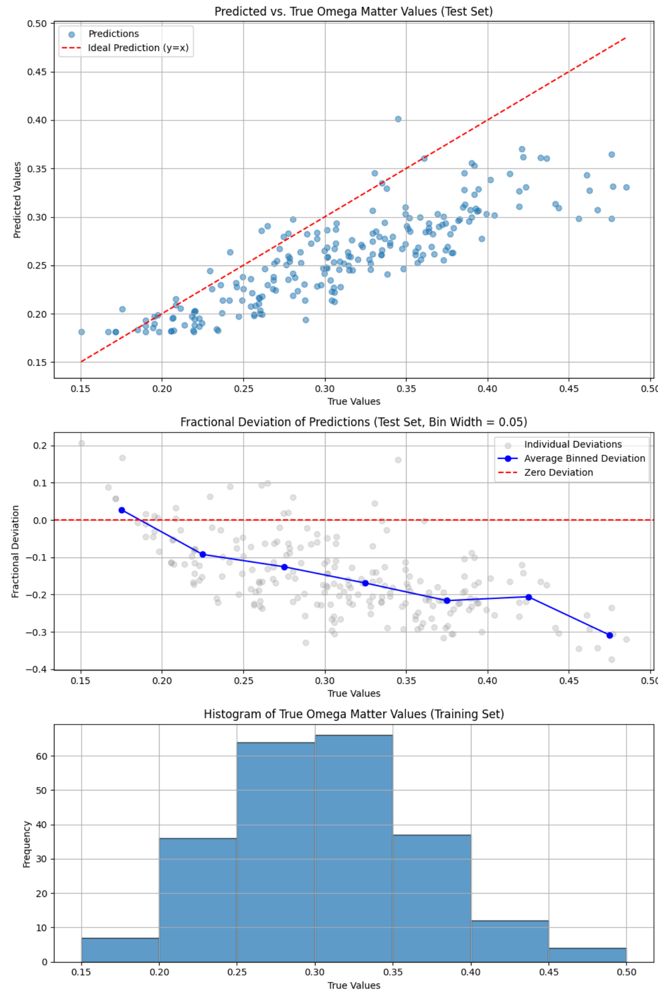


However, after seeing that predictions for Ω_m and σ_8 yielded the best outcomes, I augmented DeepSphere to only make predictions for those two. This allowed me to focus on tailoring the structure and hyperparameters of DeepSphere to these specific predictions. This was also where I was able to test an alternative structure suggested in an article by A. Thomsen, et. al. (arXiv: 2511.04681v1). Below is a table from their article that shows the layout of the structure they used.

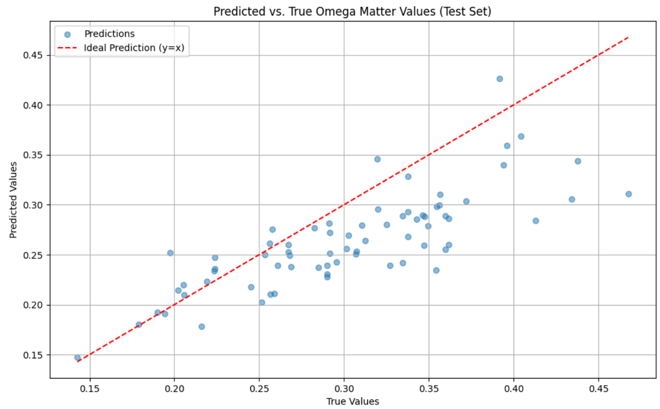
Layer Type	Output Shape	# Parameters
Input	$(N_b, 458\,752, 8)$	0
Smoothing	$(N_b, 458\,752, 8)$	0
Pseudo convolution	$(N_b, 114\,688, 32)$	1 056
Pseudo convolution	$(N_b, 28\,672, 64)$	8 256
Pseudo convolution	$(N_b, 7\,168, 128)$	32 896
Chebyshev convolution	$(N_b, 7\,168, 256)$	163 840
Layer-normalization	$(N_b, 7\,168, 256)$	512
Pseudo convolution	$(N_b, 1\,792, 256)$	262 400
Chebyshev convolution	$(N_b, 1\,792, 256)$	327 680
Layer-normalization	$(N_b, 1\,792, 256)$	512
Pseudo convolution	$(N_b, 448, 256)$	262 400
Residual layer	$(N_b, 448, 256)$	656 896
	$\vdots$	
Residual layer	$(N_b, 448, 256)$	656 896
Flatten	$(N_b, 114\,688)$	0
Layer-normalization	$(N_b, 114\,688)$	229 376
Fully connected	$(N_b, 20)$	2 293 780

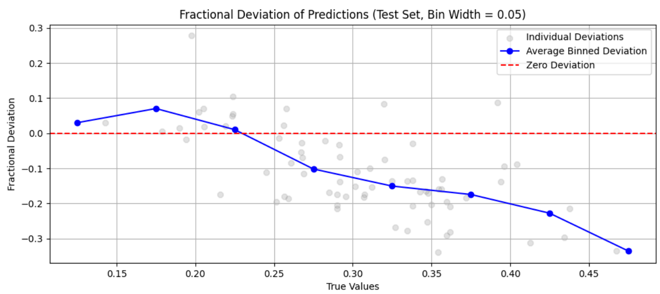
I drew inspiration from this structure, and augmented the original DeepSphere by adding a pseudo-convolution before the main convolutional layers, as well as adding layer-normalizations and an additional fully-connected layer at the end. Furthermore, my previous results came from using sections of only a single composite density map across redshifts z = 0 ~ 1 as input; however, I wanted to allow the NN to consider correlations across different redshifts, hopefully giving it access to information about timewise evolution of the cosmos. So I prepared the data such that the NN would receive an input of five separate redshift bins (0 ~ 0.2, 0.2 ~ 0.4, 0.4 ~ 0.6, 0.6 ~ 0.8, and 0.8 ~ 1) as observable channels from a small patch of the sky. Below is a depiction of the results for predicting Ω_m using this multi-z input for both the model I augmented and a mimicry of the structure used by A. Thomsen, et. al.

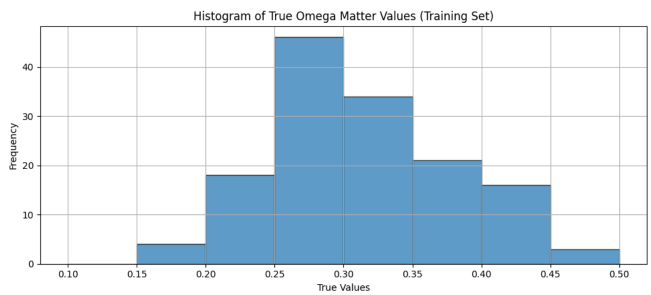




The graph on the left is the results from the original DeepSphere with my included augmentations, and the graph on the right is the results from the A. Thomsen, et. al. structure. The top two panels of each graphic show the 'Predicted vs True' Ω_m values and the fractional deviation from truth in each case. The bottom panels show histograms of the training data that was used (in both cases, the training data was drawn from the same overall data pool, but the random shuffling I implemented would sometimes create slight variations in the distribution of the training data). The results show that the A. Thomsen, et. al. structure yielded a slightly greater bias towards lower predictions. So I decided to stick with the original structure, keeping the few augmentations I'd included. Keeping this as the main focus, the final bit of progress I achieved this last semester was the inclusion of ISW maps in the input data. Below is a graphic depicting the prediction results after training of the NN on the multi-z plus ISW input.







As can be seen, the training data had a different distribution from the previous cases, primarily due to the random shuffle. However, despite the appearance of fewer data points, the amount of data points used was the same; the difference in this case was in the batching method implemented in order to optimize for GPU memory. Aside from all of that, the results display a returned bias towards lower predictions. Moving on from this point, I intend to first alter the shuffling process so as to keep an identical distribution of training data between runs; and second, I will work to fine tune the hyperparameters and layer structure of this NN to optimize it for this inclusion of ISW maps. Some further goals beyond this point–once the network has been optimized–will be to utilize the weights of learned filters and saliency maps or pixel weight analyses in order to analyze which parts of each z-bin and the ISW maps play crucial roles in making predictions about cosmological parameters. Graphics shown here only display the results from training towards Ω_m, but similar training and results have been acquired for σ_8; and the goal is to progress toward being able to make predictions about the other five cosmological parameters included within the Gower Street simulation suite. With the current implementation of tomographically binned data and ISW maps, I hope to find some progress towards accurate predictions of parameters such as w or h, which often require data including timewise evolution to be derived. Ultimately, the goal is that these progressions lead to a utilization of the predictions made by AI for the constraint of cosmological parameter contour maps.

Conclusion

Progress on my research this third semester has included the continual expansion of my own skills and knowledge, deeper exploration of GCNNs, a new augmentation to DeepSphere as presented from other works, an improvement to the preparation of data geared towards my specific task, and the implementation of new information in the form of ISW maps as input data. For the next semester, my focus will be to fine-tune the current structure I have for DeepSphere, and continue to seek connections between my results and possible constraints for dark energy and cosmological parameters.