AGNet: Weighing Black Holes Using Deep Learning

Sneh Pandya, Devanshi Pratap
Under Instruction of: Joshua Yao-Yu Lin, Xin Liu

Introduction

Goal
To develop a convolutional neural network that can predict mass of supermassive black holes using time series spectra and redshift data.

Supermassive black holes (SMBH) are ubiquitously found at the centers of most galaxies. Measuring SMBH mass is important for understanding their origin and evolution. We train Deep Learning (DL) algorithms that learn from the Sloan Digital Sky Survey (SDSS) Stripe 82 and DR7 data for a sample of ~10,000 quasars to map out the nonlinear encoding between black hole mass and multi-color optical light curves (Sun, 2020). We later incorporate data augmentation, redshift and absolute magnitude implementation, and optimized DL architectures in order to improve the performance of AGNet.

Motivation
Traditional methods of weighing SMBH require spectral data which are expensive to gather, as well as tedious. Our results have direct implications for efficient applications with future observations from the Vera C. Rubin Observatory (LSST).

Data and Methodology

Data Matching
We adopt multi-color photometric light curves from SDSS Stripe 82 as our spectral data. Our baseline sample consists of ~10,000 quasars in the Stripe 82 survey. We assume the viral black hole mass estimates from the SDSS DR7 catalog as the ground truth, and match the two according to their equatorial coordinates. Spectra measure both how fast a network learned most from redshift, and simulating new light curves did not improve results substantially.

Data Augmentation
To eliminate small sample size as a factor in neural network performance, we use data augmentation to simulate 10x new light curves: Simulated light curves share redshift and ID information with initial objects, and we simulate new magnitudes for 5 bands (u,g,r,i,z) using error information within 1σ of original magnitude. It was found later that our network learned most from redshift, and simulating new light curves did not improve results substantially.

Network Implementation
We first reshape our light curves into 224 x 224 numpy images to feed AGNet. We use deep convoluted neural networks (CNNs) to predict black hole mass directly from quasar light curves, employing Pytorch, a deep learning Python library. We use the standard 18/34/50 layers of deep residual network architecture as our baseline, and further modify the last layer by adding a fully connected layer so that it outputs the number of parameters we desire. We later concatenate absolute magnitudes to the last year as it greatly helped AGNet performance. The skip connection helps neural networks with skip connections have a better ability to approach the minimum of highly non-convex loss functions with a smoother loss surface. We use a pre-trained ResNet and add an output neuron at the last layer for outputting the value we wish to obtain.

Find our code AGNet at: https://github.com/devanshipratap/DeepLearningAGN/ (QR code above)

Figure 1: Matched DR7 and Stripe 82 Data Sets
Figure 2: Original and Simulated Light Curve
Figure 3: AGNet Implementation Flowchart

Statistical Analysis and Testing
We split our data set of ~100,000 quasars into a 80/20 training and testing set. We do statistical analysis on ground truth mass values to validate any concerns about error. A degree of error is apparent, however most errors are small (<0.25) and are confined to our lower mass quasars (Figure 4). We quantify our results according to many statistical quantities, such as Pearson-Spearman coefficient, RMSE, and R² score. Our best AGNet model achieves RMSE = 0.389 and R² = 0.372.

Discussion and Next Steps
We are satisfied with both our redshift and mass predictions with AGNet. The success in redshift predictions was initially shown in Pasquet-Itam et al. Pasquet, 2018, however this is a pioneer effort in applying a DL architecture to retrieve quasar masses. In the future, exploring modified network architectures and sampling data from other surveys would be advisable. Support for this project was provided by the National Center for Supercomputing Applications REU Inclusion Program and the Department of Physics, Department of Astronomy at the University of Illinois at Urbana-Champaign. We also acknowledge the NCSA HAL cluster, without which most of our computation could not be possible.

References