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 ~100,000 quasars from the Dark Energy Survey Supernova fields.

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 Rubin Observatory (LSST).

Data Matching
Our baseline sample consists of ~10,000 quasars in the Stripe 82 survey. We assume that 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.

Data Augmentation
To eliminate small sample size as a factor in neural network performance, we use data augmentation to simulate 10x new light curves using random seeds. 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.

Network Implementation
We first reshape our light curves into 224 x 224 numpy images to feed our network. We use deep convolutional 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. The skip connection helps the neural network to be trained without the vanishing gradient problem. It’s been shown that 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).

Discussion and Next Steps
It is evident from our results that our network performs well in predicting redshift, which was initially shown in Pasquet-Itam & Pasquet, 2016. In mass estimations (Figure 5) the network shows signs of learning but estimations tend to consolidate around the mean. In the future, modified network architectures would be advisable. Sampling datasets outside of SDSS Stripe 82 and DR7, such as the Dark Energy Survey (DES) and Catalina Real-Time Transient Survey (CRTS) and eventually LSST may also improve network results.

Support for this project was provided by the National Center for Supercomputing Applications SPIN 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