Project Proposal Title: Finding Anomalous Galaxies in DESI Data
Members: John DiNovi, Caleb Painter, Santi Calderon, Adam Boesky

Background and Motivation: With a number of unprecedentedly large Astronomical surveys such as the Legacy Survey of Space and Time (LSST) going online in the upcoming years, rapid object classification will be crucial. Given its easy-to-use and rapid nature, artificial intelligence and machine learning are promising candidates for the task of classification. Our goal will be to use AI to classify galaxies based on images and photometric data from past surveys.

Data: The data we will use comes from observations taken by the Dark Energy Spectroscopic Instrument (DESI). This dataset originally includes ~440,000 unique galaxies, which have been reduced down to around 17,000. These galaxies are labeled with their 'type' from the Galaxy Zoo project, falling into categories like 'spiral', 'merging', 'edge-on', and more. Each image is 256x256 pixels in r, g, and z bands. We also download data from the DESI legacy survey, which has photometric properties for each galaxy, as well as estimates on the mass. D*ata Clean-Up:* A preliminary analysis of the data was performed and found that it is relatively clean and well-organized. The main data processing would be joining the two tables of galaxy type and galaxy properties.

Scope: The focus of our project will be to use properties like photometry and mass to predict galaxy type. We will also explore how other methods such as principal component analysis and k-nearest neighbor regression can be used to group galaxies based on their images, and how their efficacy compares to other methods like CNNs. This project will be a good way for us to learn about the role of AI in the world of data science, and the performance of AI relative to other computational methods.