Photometric Redshift Estimation:

Using Machine Learning Algorithms

**Physical relevance:**

**Redshift** measures how much light from a distant object has been "stretched" due to the expansion of the universe, which correlates with its distance from Earth. However, obtaining redshifts from spectroscopy is resource-intensive and time-consuming, particularly for faint, distant objects.

Photometric redshift estimation offers a quicker alternative by using multiband photometry (e.g., colours in specific wavelength bands), allowing astronomers to estimate distances for millions of galaxies efficiently​

The aim of Photometric redshift estimation is to approximate the redshift of astronomical objects, such as galaxies and quasars, based on their observed photometric data.

This method is essential in astrophysics and cosmology for several reasons some of them are:

**1. Studying Large-Scale Structure of the Universe**

* Understanding the distribution of galaxies across different redshifts (distances) helps map the universe's large-scale structure, including galaxy clusters, superclusters, and voids.
* This large-scale structure is fundamental for testing cosmological models, such as those describing dark matter and dark energy, as it reveals the distribution and movement of galaxies over cosmic time​

**2. Cosmic Evolution and Galaxy Formation Studies**

* Photometric redshift estimation allows researchers to study galaxy evolution across cosmic time, tracing how galaxies form, grow, and cluster.
* By examining galaxies at different redshifts (distances), astronomers can observe how properties like star formation rates, galaxy morphology, and chemical composition have changed since the early universe​

**Algorithm Used:**

Our current model uses Radam forest algorithm along with GMM (Gaussian Mixture Model) to find probability density function.

The Random Forest algorithm is an ensemble method that builds multiple decision trees for tasks like classification and regression. A decision tree is a model that makes decisions by splitting data into branches based on feature values, resulting in a straightforward, tree-like structure where each "leaf" represents a prediction.

But Using only Random Forest Algorithm may give misleading outputs as the measurement of photometric values may have inherent Uncertainty. To avoid this generating PDF (Probability Density Function) using GMM offer insight into the possible redshift values along with its probability or uncertainty.

**Novelty:**

**Enhanced Uncertainty Quantification**:

Using both RF and GMMs to estimate probability density functions (PDFs) offers a more sophisticated approach to handling uncertainty in photometric redshift predictions. Unlike traditional RFs, which provide point estimates or simple variances, the combination with GMMs enables multi-modal uncertainty representation, showing not just a range but the likelihood of different values. This is particularly novel for complex datasets like galaxy surveys, where uncertainties are inherently high due to observational noise.