EUROPEAN AI FOR FUNDAMENTAL PHYSICS CONFERENCE
EuCAIFCon 2024

Contribution ID: **108** Type: **Flashtalk with Poster**

Normalizing flows for jointly predicting photometry and photometric redshifts

Tuesday, 30 April 2024 14:25 (3 minutes)

Traditionally, machine-learning methods have mostly focused on making predictions without providing explicit probability distributions. However, the importance of predicting probability distributions lies in its understanding of the model's level of confidence and the range of potential outcomes. Unlike point estimates, which offer a single value, probability distributions offer a range of potential outcomes and their likelihoods. MDNs enable capturing complex and multi-modal distributions. Unfortunately, the modeling process encounters challenges in N-dimensional spaces due to the introduction of covariance parameters.

Normalizing flows (NF) are generative models that learn to transform a simple probability distribution into a more complex one through a series of invertible mappings. Conditional normalizing flows (cNF) are an extension of NF that incorporate conditioning variables, which acting as inputs guide the generation of the probability distribution in a context-specific manner. While the conventional use of NF and cNF primary focus on generating data samples, an emerging trend employs them as a probabilistic framework for regression problems. Therefore, cNF have emerged as a valuable option for addressing probability and covariance estimation challenges in N-dimensional spaces.

Commonly, in any observational field, the data reduction pre-processing and the prediction of physical quantities like the redshift are two independent steps of the data analysis. In the case of astronomy, the information from the observed images is compressed into photometric catalogs, which contain the light captured by the telescope using different photometric filters. Unfortunately, this approach results in the loss of valuable details from the image, which can significantly impact the accuracy of subsequent inferences. Moreover, errors introduced at each stage of the image processing propagate in subsequent estimations, compromising the quality of derived quantities like the photometric redshift.

In this project, we are developing a cNF to predict the photometry and the photometric redshift of a galaxy directly from its image observations. The NF learns the transformations from a simple N-dimensional Gaussian distribution to the joint probability distribution of the photometry and the photometric redshift conditioned on the observed images. By applying the cNF directly to the images, we can bypass the data pre-processing step and reduce the probability of introducing errors. This allows the network to learn the mapping between the images and the photometry and photometric redshift directly without the need for intermediate steps and data compression that introduce noise or biases. Furthermore, the cNF can use the full information content of the images to make predictions.

Moreover, by predicting both the photometry and the photometric redshift simultaneously, the cNF benefits from multi-task learning. This approach involves training a model to perform multiple related tasks simultaneously, allowing it to learn shared representations that benefit all tasks. Our unified model enables the extraction of common data representations, which enhance the overall performance by exploiting the synergies between the photometry and redshift.

The integration of data reduction with the prediction of derived quantities holds significant relevance across all observational fields. Using a cNF architecture enables the entire process to be addressed from a probabilistic approach. Moreover, our end-to-end approach avoids the need for data compression, preserving the integrity of information. This accelerates the transformation of raw observations into scientific measurements, minimizing the risk of errors introduced during manual processing.

Primary author: CABAYOL GARCIA, Laura (IFAE/PIC)

Co-author: ERIKSEN, Martin Boerstad (IFAE-PIC)

Presenter: CABAYOL GARCIA, Laura (IFAE/PIC)

Session Classification: 1.2 Generative models & Simulation of physical systems

Track Classification: Session A