Abstract:
    In order to answer the open questions of modern cosmology and galaxy
    evolution theory, robust algorithms for calculating photometric
    redshifts (photo-z) for very large samples of galaxies are needed.
    Correct estimation of the various photo-z algorithms' performance
    requires attention to both the performance metrics and the data used
    for the estimation. In this work, we use the supervised machine
    learning algorithm MLPQNA (Multi-Layer Perceptron with Quasi-Newton
    Algorithm) to calculate photometric redshifts for the galaxies in the
    COSMOS2015 catalogue and the unsupervised Self-Organizing Maps (SOM)
    to determine the reliability of the resulting estimates. We find that
    for $z\_spec\_<1.2$, MLPQNA photo-z predictions are on the same level of
    quality as spectral energy distribution fitting photo-z. We show that
    the SOM successfully detects unreliable zspec that cause biases in the
    estimation of the photo-z algorithms' performance. Additionally, we
    use SOM to select the objects with reliable photo-z predictions. Our
    cleaning procedures allow us to extract the subset of objects for
    which the quality of the final photo-z catalogues is improved by a
    factor of 2, compared to the overall statistics.

Description:
    We present here a catalogue of photometric redshifts obtained with a
    supervised Machine Learning algorithm called Multi Layer Perceptron
    with Quasi Newton Algorithm software (MLPQNA, Brescia et al.,
    2013ApJ...772..140B, 2014A&A...568A.126B, Cat. J/A+A/568/A126) for
    more than 200000 galaxies from the COSMOS2015 catalogue (Laigle et
    al., 2016ApJS..224...24L, Cat. J/ApJS/224/24). Following the
    limitations imposed by the training sample, the photo-z are reported
    for the sources with presumed true redshifts <1.2. ML photo-z are
    obtained using 10-band IR, visual and UV photometry. For the test
    sample of galaxies ML photo-z have std of residuals ~0.048 and
    percentage of catastrophic outliers ~1.64. In addition to this we
    provide reliability indicators for the photo-z obtained with
    Self-Organizing Maps. These indicators allow to detect anomalous
    spectral redshifts (in the train and test samples; the nature of these
    anomalous spec-z can be either physical (e.g. AGNs) or instrumental
    (e.g. misclassification of a spectral line)) and unreliable photo-z
    (in the whole dataset). Using these indicators it is possible to
    select highly reliable photo-z samples. The detailed description of
    the methodology for calculating and using the reliability indicators
    can be found in the paper.

    The catalogue contains information for 214398 galaxies selected from
    the COSMOS2015 dataset (Laigle et al., 2016ApJS..224...24L, Cat.
    J/ApJS/224/24). The catalogue reports basic information about these
    galaxies according to the COSMOS2015: their sky coordinates (DEJ2000
    and RAJ2000), their identifier within the COSMOS2015 (Seq) and SED

fitting photo-z (photoZ_SED). Additionally, the catalogue contains ML
photo-z (photoZ_ML), residual between ML and SED photo-z, a flag,
reporting whether the given galaxy was included in the train, test or
run datasets during the training of the ML model, and reliability
metrics for ML photo-z, SED photo-z and spec-z. The in-cell outlier
coefficients (photoZ_ML_outlCoeff, photoZ_SED_outlCoeff,
specZ_outlCoeff) have the meaning of the number of sigmas by which the
redshift of a given galaxy differs from the mean redshift of all
galaxies belonging to the same SOM cell as this galaxy (see paper for
the details on these indicators). Occupation of the cell
(trainMapOccupation) reports how many galaxies from the train set
belong to the cell of the given galaxy; the higher this number, the
higher is the reliability of the photo-z prediction. For a highly
reliable dataset it is recommended to discard galaxies with
trainMapOccupation<5.

File Summary:

 FileName   Lrecl Records   Explanations

ReadMe   80   .   This file
mlphotoz.dat   194   214398   COSMOS2015 machine learning photometric
redshifts

                                with reliability indicators derived with
SOM

See also:
      J/ApJS/224/24 : The COSMOS2015 catalog (Laigle+, 2016)

Byte-by-byte Description of file: mlphotoz.dat

   Bytes Format Units   Label   Explanations

   1- 18 F18.14 deg   RAdeg   [149.41/150.79] Right ascension (J2000)
  20- 37 F18.16 deg   DEdeg   [1.61/2.82] Declination (J2000)
  39- 44 I6   ---   Seq   Object ID in the original COSMOS2015
                                     catalog, Laigle et al., 2016,
                                     Cat. J/ApJS/224/24)
  46- 50 A5   ---   dataset   [Run Test Train] A flag indicating
whether
                                     the object was included in the train,
test
                                     or run samples during MLPQNA training
  52- 71 F20.18 ---   zphMl   [0.02/1.47] Photometric redshift
obtained
                                     with MLPQNA (photoZ_ML)
  73- 95 E23.17 ---   zphMlCoeff ?=-99.99 In-cell outlier coefficient
for
                                     ML photo-z (photoZ_ML_outlCoeff) (1)
  97-116 F20.18 ---   zphSED   [0.0/4.72] SED fitting photometric
redshift
                                     derived from the COSMOS2015
(photoZ_SED)
 118-140 E23.17 ---   zphSEDCoeff ?=-99.99 In-cell outlier coefficient
for SED

```
                                            fitting photo-z (photoZ_SED_outlCoeff)
(1)
 142-164 E23.17 ---      resML-SED    [-1.11/0.75] Residuals between ML and
SED
                                            fitting photo-z calculated as
                                            resid=(z_SED-z_ML)/(1+z_SED)
(residML_SED)
 166-188 E23.17 ---      zspCoeff     ?=-99.99 In-cell outlier coefficient
for
                                            spec-z (specZ_outlCoeff) (1)
 190-194  F5.1  ---      tMO          Occupation of the SOM cell, to which
this
                                            object belongs, by the train dataset
                                            (trainMapOccupation)
--------------------------------------------------------------------------
-----
Note (1): objects are considered to be outliers if |*Coeff|>3.
--------------------------------------------------------------------------
-----

History:
    From Oleksandra Razim, shr.razim(at)gmail.com
```

```
==========================================================================
=====
```
(End)   Patricia Vannier [CDS]   26-
Oct-2021