We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter

Bayesian Deep Learning for Dark Energy

Celia Escamilla-Rivera

Abstract

In this chapter, we discuss basic ideas on how to structure and study the Bayesian methods for standard models of dark energy and how to implement them in the architecture of deep learning processes.

Keywords: cosmology, dark energy, Bayesian analyses, machine learning, cosmological parameters

1. Introduction

The dark sector of the universe has been the issue of study for cosmologists who are striving to understand the world around us in its entirety. The composition of the current universe is an age-old inquiry that these researchers have probed into. And while we do have estimates of the likely percentages of baryonic matter, dark matter, and dark energy at 5, 27 and 68%, respectively, researchers have been trying to improve these estimates and optimise the computational expense of the statistical methods employed to analyse cosmological data available.

These thoughts have opened the path of the following chapter, in where we will discuss from the standard dark energy models to explain the cosmic acceleration until the design of a numerical architecture in order to understand the constrains over the cosmological parameters that can describe the current universe and its effects.

2. Dark energy as a solution to the cosmic acceleration

A highlight in observational cosmology is the origin and nature of the cosmic accelerated expansion. The standard cosmological model that is consistent with current cosmological observations is the so-called concordance model or Λ CDM. According to this scenario, the observed accelerating expansion is related to the repulsive gravitational force of a Cosmological Constant Λ with constant energy density ρ and negative pressure p. This proposal has been the backbone of the standard cosmology since the nineties, but simple enough as it is the proposal that has a couple of theoretical problems; two of them are the fine tuning argument and coincidence problem [1, 2]. In order to solve or at least relax these problems, some proposals have led to alternative scenarios that can modify the general relativity (GR) or consider a landscape with a dynamical dark energy. It is in this way that dark energy emerges as a cosmological solution since it can be described as a fluid

parameterised by an equation of state (EoS), which can be written in terms of the redshift, $w(z)$. So far, the properties of this EoS remain under-researched. Just to mention a few, there are a zoo of proposals on dark energy parameterisations discussed in the literature (see, e.g., [3–9]), addressing from parameterisation as Taylor-like series to dynamical $w(z)$ that can provide oscillatory behaviours [10–13].

Nowadays, the techniques to discriminate between models and confront them with Λ CDM are based on the calculations of the constraints on the EoS-free parameter(s) of the models. This methodology has been done using observables that can show the cosmic acceleration such as supernovae type IA (SNeIa), baryon acoustic oscillations (BAO), cosmic microwave background (CMB), weak lensing spectrum, etc. The relevance of using these observations is due to the precision with which dark energy can be probed. Currently, some measurements such as the Pantheon from supernovae [14], BOSS [15], just to cite a few, point out a way to constrain these EoS parameters. These observations allow deviations from the Λ CDM model, which are usually parameterised by the EoS- free parameters [16–20]. In past years, there have been many observations related to the verification of the cosmic acceleration, for example, from Union 2.1[1] to the Joint LightCurve Analysis [21, 22]. But the statistics has been improved due to the density of data this kind of supernovae.

3. On how to model dark energy

One of the first steps to understand the behaviour of the cosmic acceleration remains in that we require an energy density with negative pressure at late times [23]. To achieve this, we need to express the ratio between the pressure and energy density as negative, i.e., $w(z) = P/\rho < 0$. In order to develop the evolution equations for a universe with this kind of *fluid*, we start by introducing in Einstein equations a Friedmann-Lemaitre-Robertson-Walker metric to obtain the Friedmann and Raychaudhuri equations for a spatially flat universe:

$$E(z)^2 = \left(\frac{H(z)}{H_0}\right)^2 = \frac{8\pi G}{3}(\rho_m + \rho_{DE})\left[\Omega_{0m}(1+z)^3 + \Omega_{0(DE)}f(z)\right], \tag{1}$$

and

$$\frac{\ddot{a}}{a} = -\frac{H^2}{2}[\Omega_m + \Omega_{DE}(1+3w)], \tag{2}$$

where $H(z)$ is the Hubble parameter in terms of the redshift z, G is the gravitational constant and the subindex 0 indicates the present-day values for the Hubble parameter and matter densities.

From Eq. (2), it is possible to obtain the energy conservation equation, in that way, the energy density of the non-relativistic matter is $\rho_m(z) = \rho_{0m}(1+z)^3$. And the ρ_m is given by:

$$\rho_m(z) = \rho_{0m}(1+z)^3, \tag{3}$$

and the dark energy density can be modulated as $\rho_{DE}(z) = \rho_{0(DE)}f(z)$, where can be written as:

$$\rho_{DE}(z) = \rho_{0(DE)}\, f(z). \tag{4}$$

[1] http://supernova. lbl.gov/Union/

If we assume that the energy-momentum tensor (on the right side of the Einstein's equations) $T^{\mu\nu}$ is a perfect fluid (without viscosity or stress effects), i.e., $\nabla_\mu T^{\mu\nu} = 0$, the form of $f(z)$ can be restricted to be:

$$f(z) = e^{\left[3\int_0^z \frac{1+w(\tilde{z})}{1+\tilde{z}} d\tilde{z}\right]}. \tag{5}$$

Now, the behaviour of the latter is restricted directly to the form of $w(z)$, which can give a description of the Hubble function (which can be normalised by the constant Hubble H_0), as for example, in the case of quiescence models ($w = const.$) the solution of $f(z)$ is $f(z) = (1+z)^{3(1+w)}$. If we consider the case of the cosmological constant ($w = -1$), then $f = 1$.

Some interesting insights of the above forms for $w(z)$ has been reported in [4, 24] and references therein, where a dark energy density ρ_{DE} with varying and non-varying $w(z)$ is considered.

As an extension, with the later equations we can calculate the dynamical age of the universe using the follow relationship:

$$\Omega_m + \Omega_{DE} = 1 \quad \text{or} \quad \frac{\rho_m}{\rho_{DE}} = \frac{\Omega_m}{\Omega_{DE}}. \tag{6}$$

Integrating, we can obtain:

$$t_0 = \int_0^\infty \frac{dz}{(1+z)H(z)}, \tag{7}$$

$$t_0 = {H_0}^{-1} \int_0^\infty \frac{dz}{(1+z)\sqrt{\left[\Omega_{0m}(1+z)^3 + \Omega_{0(DE)} f(z)\right]}}. \tag{8}$$

From here, we can set a functional form of $f(z)$, in which contribution of the dark energy density to $H(z)$ in Eq. (1) goes to a region of negative values of $w(z)$. The physics behind this behaviour is an impact on the evolution of dark energy using the dynamical age of the universe Eq. (8). When we compare several theoretical models in the light of observations, a model approach is essential. As we mentioned in the "Introduction" section, to obtain a dark energy model with late-time negative pressure, we can think in two scenarios:

- a quiescence model, which can show a wide application in tracker the slow roll condition of scalar fields and demands a constant value of w. As an example, for a flat universe and according to the Planck data [21], the dark energy EoS parameter gives $w = -1.006 \pm 0.045$, which is consistent with the cosmological constant. These data constrain the curvature parameter at $2\,\sigma$ and are found to be very close to 0 with $|\Omega_k| < 0.005$.

- a kinessence model; where when the EoS is a function of redshift z. For this case, several dark energy models with different parameterisations of $w(z)$ have been discussed in the literature [24].

4. Standard dark energy models

One of the most commonly used proposals in the literature are Taylor series-like parameterisations [25–29]:

$$w(z) = \sum_{n=0} w_n x_n(z), \tag{9}$$

where w_n are constants and $x_n(z)$ are functions of the redshift z, or, the scalar factor a. As brief examples, in this section, we present three models that have bidimensional forms in the since that they depend only of two free parameters w_i. A first target is to express the exact form of the Hubble function using a specific expression for w given by Eq. (5). Once integrated, we can normalise this function by a Hubble parameter H_0, and from now on, we called this normalisation function depending of the redshift as $E(z) = H(z)/H_0$. The second target is to test these equations with the current astrophysical data available.

4.1 Lambda cold dark matter-redshift parameterisation (ΛCDM)

This model is given by:

$$E(z)^2 = \Omega_m(1+z)^3 + (1-\Omega_m), \tag{10}$$

where Ω_m represented the matter density (including the non-relativistic and dark matter terms). We consider in $f(z)$ the value of $w = -1$. As it is well known in the literature, this standard model provides a good fit for a large number of observational data surveys without addressing the important theoretical problems mentioned above.

4.2 Linear-redshift parameterisation (LR)

One of the first attempts using Taylor series—at first order—is the EoS given by [30, 31].

$$w(z) = w_0 - w_1 z, \tag{11}$$

from we can recover ΛCDM model if $w(z) = w = -1$ with $w_0 = -1$ and $w_1 = 0$. We notice immediately that due the linear term in z, this proposal diverges at high redshift and consequently yields strong constraints on w_1 in studies involving data at high redshifts, e.g., when we use CMB data [32].

As usual, we can use the later to obtain an expression for the Hubble normalised function as:

$$E(z)^2 = \Omega_m(1+z)^3 + (1-\Omega_m)(1+z)^{3(1+w_0+w_1)} \times e^{-3w_1 z} \tag{12}$$

4.3 Chevallier-Polarski-Linder parameterization (CPL)

Due the consequence of the LP parameterisation divergence, Chevallier, Polarski and Linder proposed a simple parameterisation [33, 34] that in particular can be represented by two w_i parameters that are given by a present value of the EoS w_0 and its overall time evolution w_1. The proposal is given by the expression:

$$w(z) = w_0 + \left(\frac{z}{1+z}\right) w_1, \tag{13}$$

and its evolution is

$$E(z)^2 = \Omega_m(1+z)^3 + (1-\Omega_m)(1+z)^{3(1+w_0+w_1)} \times e^{-\left(\frac{3w_1 z}{1+z}\right)}. \tag{14}$$

As we can notice, the divergence at high redshift relaxes, but still this ansatz has some problems in specific low redshift range of observations.

5. Estimating the cosmological parameters

After we have defined a specific cosmological model, we can then perform their test using astrophysical observations. The methodology can be described by a simple calculation of the usual χ^2 method and then process the MCMC chains computational runs around a certain value [observational(s) point(s)] and obtain the best fit parameter(s) of this process. Parameter estimation is usually done by computing the so-called *likelihood function* for several values of the cosmological parameters. For each data points in the parameter space, the likelihood $\mathscr{L}$ function gives the minimised probability of obtaining the observational data that was obtained if the hypothesis parameters had the given values (or priors). For example, the standard cosmological model Λ CDM is described by six parameters, which include the amount of dark matter and dark energy in the universe as well as its expansion rate H. Using the CMB data (which is the accuracy data that we understand very well so far), a likelihood function can be constructed. The information given by $\mathcal{L}$ can tell which values of these parameters are more likely, i.e. by probing many different values. Therefore, we are able to determine the values of the parameters and their uncertainties via error propagation over the free parameters of the model.

Now, the following question is that what kind of astrophysical surveys[2] can we use to test the cosmological models. In the next sections we described the most used surveys that are employed to analyse the cosmic acceleration. It is important to mention that these surveys spread depending upon their own nature. We have three types of observations classified as: standard candles (e.g., supernovae, in which characteristic function is the luminosity distance), standard rulers (e.g., supernovae, in which characteristic function is the angular/volumen distance), and the standard sirens (e.g., gravitational waves, which can be described by frequencies or chirp masses depending the observation) [35–45]. The set of all of them can describe a precise statistics, but by separate, each of them have intrinsic problems due to their physical definition. For supernovae, the luminosity distance has in their definition an integral of the cosmological model; therefore, when we perform the error propagation, the uncertainty is high. This disadvantage can be compensated by the large population of data points in the sampler. On the other hand, the uncertainty is less for standard rulers in comparison to supernovae. For this case, the definition of angular distance does not include integrals. The price that we pay in order to use this kind of sampler is that the population of data is very small (e.g., from surveys like BOSS or CMASS, we have only seven data points). Moving forward, the observation of gravitational wave standard sirens would be developed into a powerful new cosmological test due that they can play an important role in breaking parameter degeneracies formed by other observations as the ones mentioned. Therefore, gravitational wave standard sirens are of great importance for the future accurate measurement of cosmological parameters. In this part of the chapter, we are going only to develop the use of the first two kinds of observations.

[2] This word in the coloquial language also can be replaced by *likelihood* –do not misunderstood with the function $\mathcal{L}$. Or simple we can called as *samplers*.

6. Supernovae sampler

Along the ninety years, since their discovery, Type IA supernovae (SNIa) have been the proof of the current cosmic acceleration. The surveys have been changing given us a large population of observations, from Union 2.1[3] to the Joint LightCurve Analysis [21, 22], the data sets have been incrementing observations and also their redshift range. Currently, the Pantheon sampler, which consists of a total 1048 Type Ia supernovae (SNIa) in 40 bins [14] compressed, is the largest spectroscopically confirmed SNIa sample to date. This latter characteristic makes this sample attractive to constrain with considerably precision the free cosmological parameters of a specific model.

SNIa can give determinations of the distance modulus μ, whose theoretical prediction is related to the luminosity distance d_L according to:

$$\mu(z) = 5\log\left[\frac{d_L(z)}{1\text{Mpc}}\right] + 25, \tag{15}$$

where the luminosity distance is given in units of Mpc. In the standard statistical analysis, one adds to the distance modulus the nuisance parameter μ_0, an unknown offset sum of the supernovae absolute magnitude (and other possible systematics), which is degenerate with H_0.

Now, the statistical analysis of the this sample rests on the definition of the modulus distance as:

$$\mu(z_j, \mu_0) = 5\log_{10}\left[d_L(z_j, \Omega_m; \boldsymbol{\theta})\right] + \mu_0, \tag{16}$$

where $d_L(z_j, \Omega_m; \boldsymbol{\theta})$ is the Hubble-free luminosity distance:

$$d_L(z, \Omega_m; \boldsymbol{\theta}) = (1+z)\int_0^z \mathrm{d}z' \frac{1}{E(z', \Omega_m; \boldsymbol{\theta})} \quad . \tag{17}$$

With this notation, we expose the different roles of the several cosmological parameters appearing in the equations: the matter density parameter Ω_m appears separated as it is assumed to be fixed to a prior value, while $\boldsymbol{\theta}$ is the EoS parameters w_i. These later are the parameters that we will be constraining by the data. The best fits will be obtained by minimising the quantity [46–50]:

$$\chi^2_{\text{SN}}(\mu_0, \boldsymbol{\theta}) = \sum_{j=1}^{\mathcal{N}_{\text{SN}}} \frac{\left(\mu(z_j, \Omega_m; \mu_0, \boldsymbol{\theta})\right\} - \mu_{obs}(z_j))^2}{\sigma^2_{\mu,j}}, \tag{18}$$

where $\sigma^2_{\mu,j}$ are the measurement variances. And nuisance parameter μ_0 encodes the Hubble parameter and the absolute magnitude M and has to be marginalised over.

From now on, we will assume spatial flatness; therefore, the luminosity distance is related to the comoving distance D via the equation

$$d_L(z) = \frac{c}{H_0}(1+z)D(z), \tag{19}$$

where c is the speed of light, so that, using Eq. (15) we can obtain

[3] http://supernova.lbl.gov/Union/

$$D(z) = \frac{H_0}{c}(1+z)^{-1}10^{\frac{\mu(z)}{5}-5}. \tag{20}$$

The normalised Hubble function $E(z)$ can be obtained by taking the inverse of the derivative of $D(z)$ with respect to the redshift $D(z) = \int_0^z H_0 d\tilde{z}/H(\tilde{z})$. An usual alternative, instead of using the full set of parameters for this sampler, is to use the Pantheon plugin for CosmoMC to constrains cosmological models (something similar as in the case of Joint Light Curve Analysis sampler [22]).

Since we are taking nuisance parameter M in the sample, we choose the respective values of μ_0 from a statistical analysis of the Λ CDM model with Pantheon sample obtained by fixing H_0 to the Planck value given in [51]. It is common to perform this kind of fit using computational tools that can run a standard MCMC chains. In cosmology—at least at the moment this text is writing—several codes have been implemented in order to perform the statistical fit of this parameter. The lector can explore the tool called MontePython code[4] and run a standard MCMC for M using the model of their preference. As an example, if we run a Λ CDM model with this supernovae sample, the mean value obtained will be $\mu_0 = -19.63$.

7. Baryon acoustic oscillation sampler

As a standard ruler, these astrophysical observations can contribute important features by comparing the data of the sound horizon today to the sound horizon at the time of recombination (extracted from the CMB anisotropy data). Usually, the baryon acoustic distances are given as a combination of the angular scale and the redshift separation.

To define these quantities we require a relationship via the ratio:

$$d_z \equiv \frac{r_s(z_d)}{D_V(z)}, \quad \text{with} \quad r_s(z_d) = \frac{c}{H_0}\int_{z_d}^{\infty}\frac{c_s(z)}{E(z)}\mathrm{d}z \tag{21}$$

where $r_s(z_d)$ is the comoving sound horizon at the baryon dragging epoch,

$$r_s(z_d) = \frac{c}{H_0}\int_{z_d}^{\infty}\frac{c_s(z)}{E(z)}\mathrm{d}z\ , \tag{22}$$

and z_d is the drag epoch redshift with $c_s^2 = c^2/3\left[1+\left(3\Omega_{b0}/4\Omega_{\gamma 0}\right)(1+z)^{-1}\right]$ as the sound speed with Ω_{b0} and $\Omega_{\gamma 0}$, which are the present values of baryon and photon parameters, respectively.

We define the dilation scale as:

$$D_V(z,\Omega_m;w_0,w_1) = \left[(1+z)^2 D_A^2\frac{cz}{H(z,\Omega_m;w_0,w_1)}\right]^{1/3}, \tag{23}$$

where D_A is the angular diameter distance given by

$$D_A(z,\Omega_m;w_0,w_1) = \frac{1}{1+z}\int_0^z \frac{c\,\mathrm{d}\tilde{z}}{H(\tilde{z},\Omega_m;w_0,w_1)}. \tag{24}$$

[4] https://monte-python.readthedocs.io/en/latest/

Using the comoving sound horizon, we can relate the distance ratio d_z with the expansion parameter h (defined such that $H \doteq 100h$) and the physical densities Ω_m and Ω_b. Therefore, we have

$$r_s(z_d) = 153.5\left(\frac{\Omega_b h^2}{0.02273}\right)^{-0.134}\left(\frac{\Omega_m h^2}{0.1326}\right)^{-0.255} \text{ Mpc}, \tag{25}$$

with $\Omega_m = 0.295 \pm 0.304$ and $\Omega_b = 0.045 \pm 0.00054$ [22]. As we mentioned above, unfortunately, so far we have a very low data population of this sampler. Moreover, as an example for this text, we employed compilations of three current surveys: $d_z(z = 0.106) = 0.336 \pm 0.015$ from six-degree Field Galaxy Survey (6dFGS) [52], $d_z(z = 0.35) = 0.1126 \pm 0.0022$ from Sloan Digital Sky Survey (SDSS) [53] and $d_z(z = 0.57) = 0.0726 \pm 0.0007$ from Baryon Oscillation Spectroscopic Survey (BOSS) with high-redshift CMASS [54].

We can also, add to the full sample three correlated measurements of $d_z(z = 0.44) = 0.073$, $d_z(z = 0.6) = 0.0726$ and $d_z(z = 0.73) = 0.0592$ from the WiggleZ survey [55], which has the inverse covariance matrix:

$$\mathbf{C}^{-1}_{\mathbf{WiggleZ}} = \begin{pmatrix} 1040.3 & -807.5 & 336.8 \\ -807.5 & 3720.3 & -1551.9 \\ 336.8 & -1551.9 & 2914.9 \end{pmatrix} \tag{26}$$

In order to perform the χ^2-statistic, we define the proper χ^2 function for the BAO data as

$$\chi^2_{\text{BAO}}(\boldsymbol{\theta}) = \mathbf{X}^T_{\mathbf{BAO}}\mathbf{C}^{-1}_{\mathbf{BAO}}\mathbf{X}_{\mathbf{BAO}} \tag{27}$$

where $\mathbf{X}_{\mathbf{BAO}}$ is given as

$$\mathbf{X}_{\mathbf{BAO}} = \left(\frac{r_s(z_d)}{D_V(z, \Omega_m; w_0, w_1)}) - d_z(z)\right) \tag{28}$$

Then, the total χ^2_{BAO} is directly obtained by the sum of the individual quantity by using Eq. (27) in

$$\chi^2_{\text{BAO}-\text{total}} = \chi^2_{\text{6dFGS}} + \chi^2_{\text{SDSS}} + \chi^2_{\text{BOSSCMASS}} + \chi^2_{\text{WiggleZ.}} \tag{29}$$

8. How to deal with Bayesian statistics

Now, we are ready to introduce how to extrapolate the above frequentist analyses to the Bayesian field [56]. The important difference between both statistics is that in the first one we are dedicated in work with a standard χ^2 fit, while in the second one, we are taking into account the following idea: given a specific set of cosmological values (the priors), which are the probability of a second set of values to fit the hypothesis [57–60].

The above idea is what we call a Bayesian model selection, which methodology consist in describe the relationship between the cosmological model, the astrophysical data and the prior information about the free parameters. Using Bayes theorem [61], we can update the prior model probability to the posterior model probability. However, when we compare models, the evidence function is used to evaluate the model's evolution using the data at hand.

We define the evidence function as:

$$\mathscr{E} = \int \mathcal{L}(\theta) P(\theta) d\theta, \tag{30}$$

where θ is the vector of free parameters (which for the dark energy models presented in the above sections, will be given by the w_i free parameters). $P(\theta)$ is the prior distribution of these parameters.

From a computational point of view, and due to the large population of data and the model used, Eq. (30) can be difficult to calculate due that the integrations can consume to much computational time when the parametric phase space is large. Nevertheless, even when several methods exist [62, 63], in this text, we present test with a nested sampling algorithm [64] which has proven practicable in cosmology applications [65].

Once we obtain the evidence, we can therefore calculated the logarithm of the Bayes factor between two models $\mathcal{B}_{ij} = \mathscr{E}_i / \mathscr{E}_j$, where the reference model ($\mathscr{E}_i$) with highest evidence can be the Λ CDM model and impose a flat prior on H_0, i.e., we can use an exactly value of this parameter.

The interpretation of the results of this ratio can be described by a scale known as Jeffreys's scale [66], which easily can be explained as follows:

- if $\ln B_{ij} < 1$, there is no significant preference for the model with the highest evidence;
- if $1 < \ln B_{ij} < 2.5$, the preference is substantial;
- and, if $2.5 < \ln B_{ij} < 5$, it is strong; if $\ln B_{ij} > 5$, it is decisive.

9. About deep learning in cosmology

Although Bayesian evidence remains the preferred method compared with information criterions and Gaussian processes on the literature, a complete Bayesian inference for model selection—this to have a scenario where we can discriminate a pivot model from a hypothesis—is very computationally expensive and often suffers from multi-modal posteriors and parameter degeneracies. As we pointed out in the later section, the calculation of the evidence leads to large time consumption to obtain the final result.

As the study of the Large Scale Structure (LSS) of the universe indicates, all our knowledge relies on state-of-the art cosmological simulations to address a number of questions by constraining the cosmological parameters at hand using Bayesian techniques. Moreover, due to the computational complexity of these simulations, some studies look remains computationally infeasible for the foreseeable future. It is at this point where computational techniques as machine learning can have a number of important uses, even for trying to understand our universe.

The idea behind the machine learning is based on considering a neural network with a complex combination of neurons organised in nested layers. Each of these neuron implements a function that is parameterised by a set of weights W. And every layer of a neural network thus transforms one input vector—or tensor depending the dimension—to another through a differentiable function. Theoretically, given a neuron n, it will receive an input vector and the choice of an activation function A_n, the output of the neuron can be computed as:

$$h^{<t>} = A_n\left(h^{<t-1>} \cdot W_h + x^{<t>} \cdot W_x + b_a\right), \tag{31}$$

$$y^t = A_n\left(h^t \cdot W_y + b_y\right), \tag{32}$$

where $h^{<t>}$ is called the hidden state, A_n is the activation function, and y^t is the output.

The goal is to introduce a set of data in order to train this array, and therefore, the architecture can learn to finally give an output set of data. For example, the network can learn the distribution of the distance moduli in the dark energy models, then feed the astrophysical samplers (surveys) to the network to reconstruct the dark energy model and then discriminate the most probable model. [5]

Moreover, while neural networks can learn complex nested representations of the data, allowing them to achieve impressive performance results, it also limits our understanding of the model learned by the network itself. The choice of an architecture [67] can have an important influence on the performance of the neural network. Some designs have to made concerning the number and the type of layers, as well as the number and the size of the filters used in each layer. A convenient way to select these choices is typically through experimentation—which for our universe, we will need these to happen first—as it is, we can select the size of the network, which depends on the number of training test as networks with a large number of cosmological parameters likely to overfit if not enough training tests are available.

At the moment these lines are writing, a strong interest over this kind of algorithm is not only bringing new opportunities for data-driven cosmological discovery but will also present new challenges for adopting machine learning—or, in our case, a subset of this field, deep learning—methodologies and understanding the results when the data are too complex for traditional model development and fitting with statistics. A few proposals in this area have been done to explode the deep learning methods for measurements of cosmological parameters from density fields [68] and for future large-scale photometric surveys [69].

10. Deep learning for dark energy

The first target in order to start training an astrophysical survey is to design an architecture with an objective function of neural networks that can have many unstable points and local minima. This architecture makes the optimisation process very difficult, but in real scenarios [70, 71], high levels of noise degrade the training data and typically result in optimisation scenarios with more local minima and therefore increase the difficulty in training the neural network. It can thus be desirable to start optimising the neural network using noise-free data which typically yield smoother scenarios. As an example, in **Figure 1**, we present a standard network using an image of a cosmological simulation (the data) and then divided an array of several layers to finally extract the output cosmological parameters value [72, 73]. Each neuron use a Bayesian process to compute the error propagation as it is done in the standard inference analyses.

We can describe a quickly, but effective, recipe to develop a Recurrent Neural Network with a Bayesian computation training [29, 74–78] in the following steps:

[5] In this text we are employing a Recurrent Neural Network. There are several in this machine learning field e.g. in [57] and references therein.

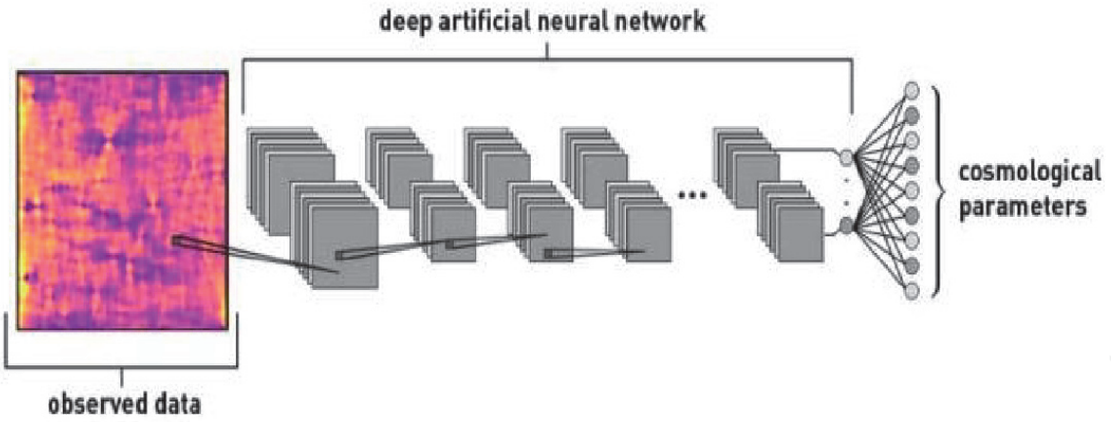


Figure 1.
A deep learning architecture for dark energy.

- Step 1. Construction of the neural network. For a Recurrent Neural Network method, we can choose values that have one layer and a certain number of neurons (e.g., you can start with 100 for a supernovae sampler).

- Step 2. Organising the data. We need to sort the sampler from lower to higher redshift in the observations. Afterwards, we re-arrange our data using the number of steps (e.g., try with four steps numbered as x_i for a supernovae sampler).

- Step 3. Computing the Bayesian training. Due to the easiness of neural networks to overfit, it is important to choose a mode of regularisation. With a Bayesian standard method to compute the evidence, the algorithm can calculate errors via regularisation methods [74]. Finally, over the cost function we can use Adam optimiser.

- Step 4. Training the entire architecture. It is suitable to consider a high number of epochs (e.g., for a sampler as Pantheon, you can try with 1000 epoch per layer). After the training, it is necessarily to read the model and apply more times the same dropout to the initial model. The result of this step is the construction of the confidence regions.

- Step 5. Computing modulus distance $\mu(z)$ for each cosmological model. Using the definitions of $E(z)$, we can compute $\mu(z)$ by using a specific dark energy equation of state in terms of z and then integrating them.

- Step 6. Computing the best fits. Finally, the output values can be obtained by using the training data as a simulated sample. We use the publicly codes CLASS[6] and Monte Python[7] to constrain the models as it is standard for usual Bayesian cosmology.

- The results of this recipe can be seeing in **Figure 2**.

[6] https://github.com/lesgourg/class_public

[7] https://github.com/baudren/montepython_public

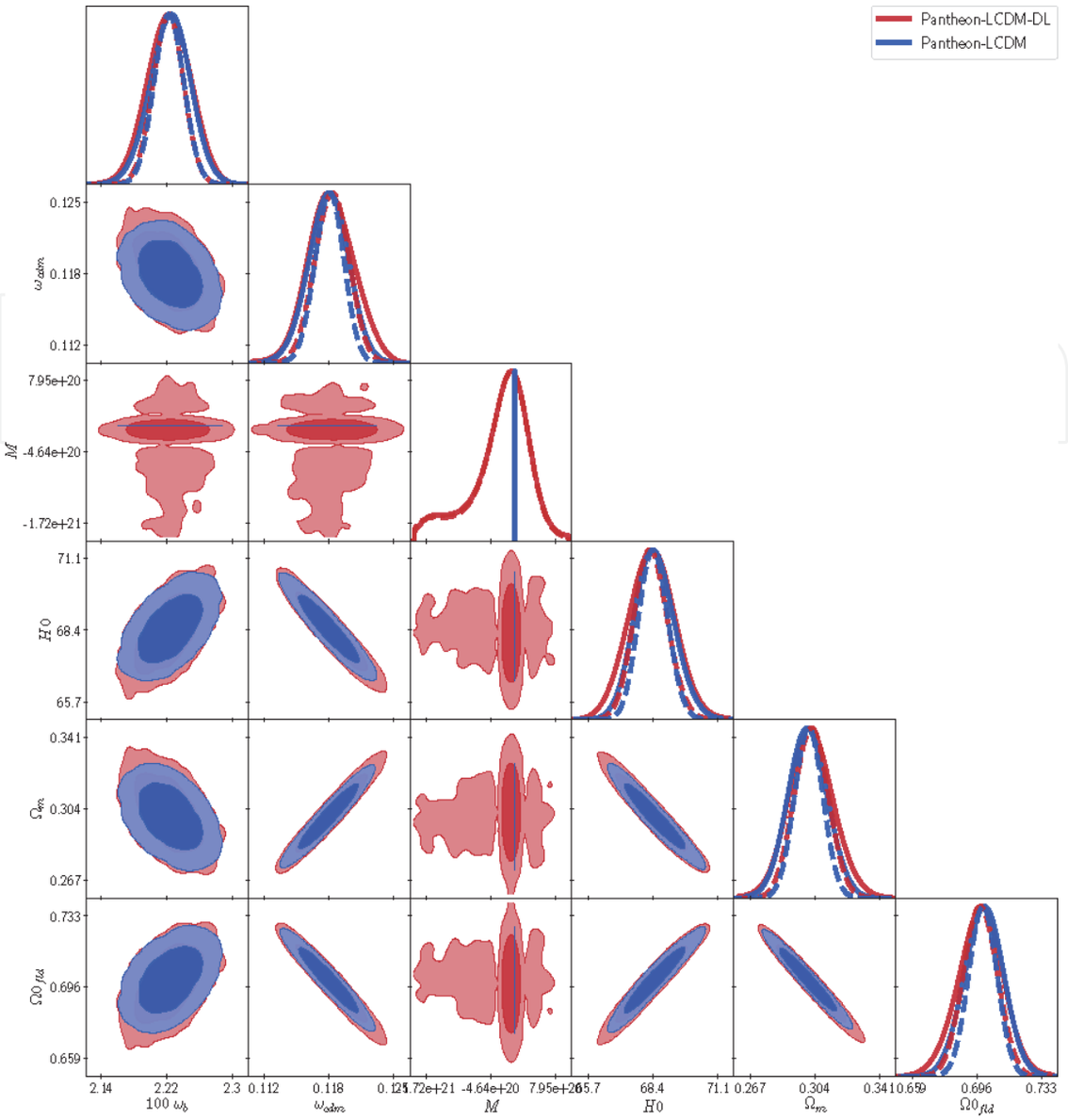


Figure 2.
Statistical contours levels for Λ CDM using observational data (red colour) and training deep learning data (blue colour).

11. Conclusions

In this chapter, we discuss how to derive the equations of state for a specific dark energy model. Also, we studied the standard models of dark energy in order to project the cosmic acceleration according to the current data available in the literature. It is important to remark that each Bayesian statistics performed will depend solely on the data used to develop them. More the data, better the statistics. So we expect that future surveys will improve the constrains over the cosmological parameters, not only at background level, but also at perturbative level.

The exploration of these astrophysical surveys has reached a new scenario in regards to the machine learning techniques. These kind of techniques allow to explore—without technical problems in the astrophysical devices—scenarios where the pivot model of cosmology, ΛCDM, a theoretical framework that accurately describes a large variety of cosmological observables, from the temperature anisotropies of the cosmic microwave background to the spatial distribution of galaxies. This model has a few free parameters representing fundament quantities, like the

geometry and expansion rate of the Universe, the amount and nature of dark energy, and the sum of neutrino masses. Knowing the value of these parameters will improve our knowledge on the fundamental constituents and laws governing our universe. Thus, one of most important goals of modern cosmology is to constrain the value of these parameters with the highest accuracy. Therefore, as an extrapolation between the ideas of the standard cosmostatistics and the use of machine learning techniques will improve even better the constrain of the cosmological parameters without to be worried about the intrinsic uncertainties of the data [79].

Acknowledgements

CE-R is supported by the *Royal Astronomical Society* as FRAS 10147, *PAPIIT* Project IA100220 and ICN-UNAM projects.

Author details

Celia Escamilla-Rivera
Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, México D.F., México

*Address all correspondence to: celia.escamilla@nucleares.unam.mx

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Weinberg S. The Cosmological Constant Problems. 2000, arXiv:astro-ph/0005265

[2] Sahni V, Starobinsky AA. The case for a positive cosmological lambda term. International Journal of Modern Physics D: Gravitation; Astrophysics and Cosmology. 2000;**9**:373-443

[3] Feng L, Lu T. A new equation of state for dark energy model. Journal of Cosmology and Astroparticle Physics. 2011;**2011**:034

[4] Stefancic H. Equation of state description of the dark energy transition between quintessence and phantom regimes. Journal of Physics Conference Series. 2006;**39**:182

[5] Wang Y, Tegmark M. Uncorrelated measurements of the cosmic expansion history and dark energy from supernovae. Physical Review D. 2005; **71**:103513

[6] Barboza EM, Alcaniz JS, Zhu Z-H, Silva R. A generalized equation of state for dark energy. Physical Review D. 2009;**80**:043521

[7] Pantazis G, Nesseris S, Perivolaropoulos L. Comparison of thawing and freezing dark energy parametrizations. Physical Review D. 2016;**93**:103503

[8] Jassal HK, Bagla JS, Padmanabhan T. WMAP constraints on low redshift evolution of dark energy. Monthly Notices of the Royal Astronomical Society. 2005;**356**:L11

[9] Wang Y. Physical Review D. 2008;**77**: 123525. DOI: 10.1103/PhysRevD. 77.123525 [arXiv:0803.4295 [astro-ph]]

[10] Escamilla-Rivera C, Capozziello S. Unveiling cosmography from the dark energy equation of state. International Journal of Modern Physics D. 2019. DOI: 10.1142/S021827181950 1542 [arXiv: 1905.04602 [gr-qc]]

[11] Jaime LG, Patiño L, Salgado M. Note on the equation of state of geometric dark-energy in f(R) gravity. Physical Review D. 2014;**89**(8):084010. DOI: 10.1103/PhysRevD.89.084010 [arXiv: 1312.5428 [gr-qc]]

[12] Lazkoz R, Ortiz-Baños M, Salzano V. f(R) gravity modifications: From the action to the data. European Physical Journal C. 2018;**78**(3):213. DOI: 10.1140/ epjc/s10052-018-5711-6 [arXiv: 1803.05638 [astro-ph.CO]]

[13] Capozziello S, D'Agostino R, Luongo O. International Journal of Modern Physics D: Gravitation; Astrophysics and Cosmology. Extended Gravity Cosmography. 2019;**28**(10): 1930016. DOI: 10.1142/ S0218271819300167 [arXiv:1904.01427 [gr-qc]]

[14] Scolnic DM et al. The Astrophysical Journal. 2018;**859**:101

[15] Busca NG, Delubac T, Rich J, Bailey S, Font-Ribera A, Kirkby D, et al. Baryon acoustic oscillations in the Ly- α forest of BOSS quasars. Astronomy and Astrophysics. 2013;**552**:A96

[16] Alberto Vazquez J, Bridges M, Hobson MP, Lasenby AN. Reconstruction of the dark energy equation of state. JCAP. 2012;**1209**:020. DOI: 10.1088/1475-7516/2012/09/020 [arXiv:1205.0847 [astro-ph.CO]]

[17] Seikel M, Clarkson C, Smith M. Reconstruction of dark energy and expansion dynamics using Gaussian processes. JCAP. 2012;**06**:036. DOI: 10.1088/1475-7516/2012/06/036 [arXiv: 1204.2832]

[18] Montiel A, Lazkoz R, Sendra I, Escamilla-Rivera C, Salzano V.

Nonparametric reconstruction of the cosmic expansion with local regression smoothing and simulation extrapolation. Physical Review D. 2014;**89**(4):043007. DOI: 10.1103/PhysRevD.89.043007 [arXiv:1401.4188 [astro-ph.CO]]

[19] Zhao GB et al. Dynamical dark energy in light of the latest observations. Nature Astronomy. 2017;**1**(9):627. DOI: 10.1038/s41550-017-0216-z [arXiv: 1701.08165 [astro-ph.CO]]

[20] Jaime LG, Jaber M, Escamilla-Rivera C. Physical Review D. 2018;**98**(8): 083530. DOI: 10.1103/ PhysRevD.98.083530 [arXiv: 1804.04284 [astro-ph.CO]]

[21] Ade PAR, Aghanim N, Arnaud M, Ashdown M, Aumont J, Baccigalupi C, et al. Planck 2015 Results. XIII. Cosmological Parameters. 2015, arXiv: astro-ph.CO/1502.01589

[22] Betoule M et al. [SDSS Collaboration]Astronomy and Astrophysics. 2014;**568**:A22. DOI: 10.1051/0004-6361/201423413 [arXiv: 1401.4064 [astro-ph.CO]]

[23] Huterer D, Turner MS. Probing the dark energy: Methods and strategies. Physical Review D. 2001;**64**:123527

[24] Lazkoz R, Nesseris S, Perivolaropoulos L. Exploring cosmological expansion Parametrizations with the gold SnIa dataset. Journal of Cosmology and Astroparticle Physics. 2005;**2005**:010

[25] Barboza EM Jr, Alcaniz JS. A parametric model for dark energy. Physics Letters B. 2008;**666**:415-419

[26] Wetterich C. Phenomenological parameterization of quintessence. Physics Letters B. 2004;**594**:17-22

[27] Wetterich C. Cosmology with Varying Scales and Couplings. 2003, arXiv:hep-ph/0302116

[28] Escamilla-Rivera C, Casarini L, Fabris JC, Alcaniz JS. Linear and non-linear perturbations in dark energy models. 2016, arXiv:1605.01475

[29] Escamilla-Rivera C, Fabris JC. Galaxies MPDI. Galaxies. 2016;**4**(4):76. DOI: 10.3390/galaxies4040076 [arXiv: 1511.07066 [astro-ph.CO]]

[30] Weller J, Albrecht A. Future supernovae observations as a probe of dark energy. Physical Review D. 2002; **65**:103512

[31] Huterer D, Turner MS. Physical Review D. 2001;**64**:123527. DOI: 10.1103/PhysRevD.64.123527 [astro-ph/0012510]

[32] Wang FY, Dai ZG. Constraining dark energy and cosmological transition redshift with type Ia supernovae. Chinese Journal of Astronomy and Astrophysics. 2006;**6**:561

[33] Linder EV. The dynamics of quintessence, the quintessence of dynamics. General Relativity and Gravitation. 2008;**40**:329-356

[34] Chevallier M, Polarski D. Accelerating universes with scaling dark matter. International Journal of Modern Physics D: Gravitation; Astrophysics and Cosmology. 2001;**10**:213-223

[35] Albrecht A, Amendola L, Bernstein G, Clowe D, Eisenstein D, Guzzo L, et al. Findings of the Joint Dark Energy Mission Figure of Merit Science Working Group. 2009, arXiv:0901.0721

[36] Liddle AR. How many cosmological parameters? Monthly Notices of the Royal Astronomical Society. 2004;**351**: L49-L53

[37] Riess AG et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. [Supernova Search Team] Astronomy Journal. 1998;**116**:1009.

DOI: 10.1086/300499 [astro-ph/9805201]

[38] Perlmutter S et al. Measurements of Omega and Lambda from 42 high-redshift supernovae. [Supernova Cosmology Project Collaboration]The Astrophysical Journal. 1999;**517**:565. DOI: 10.1086/307221 [astro-ph/9812133]

[39] Available from: http://desi.lbl.gov/

[40] Available from: https://www.darkenergysurvey.org/

[41] Available from: https://www.lsst.org/

[42] Available from: https://wfirst.gsfc.nasa.gov/

[43] Takada M, Jain B. The three-point correlation function in cosmology. Monthly Notices of the Royal Astronomical Society. 2003;**340**:580. DOI: 10.1046/j.1365-8711.2003.06321.x [astro-ph/0209167]

[44] Marin FA et al. The WiggleZ Dark energy survey: Constraining galaxy bias and cosmic growth with 3-point correlation functions. [WiggleZ Collaboration] Monthly Notices of the Royal Astronomical Society. 2013;**432**:2654. DOI: 10.1093/mnras/stt520 [arXiv:1303.6644 [astro-ph.CO]]

[45] Tsujikawa S. Dark energy: Investigation and modeling. 2010. DOI: 10.1007/978-90-481-8685-3_8, arXiv:1004.1493 [astro-ph.CO]

[46] Press WH, Teukolsky A, Vetterling W, Flannery B. Numerical Recipes. 3rd ed. New York, USA: Cambridge Press; 1994

[47] Escamilla-Rivera C, Lazkoz R, Salzano V, Sendra I. Tension between SN and BAO: Current status and future forecasts. Journal of Cosmology and Astroparticle Physics. 2011. DOI: 10.1088/1475-7516/2011/09/003

[48] Burigana C, Destri C, de Vega HJ, Gruppuso A, Mandolesi N, Natoli P, et al. Forecast for the Planck precision on the tensor to scalar ratio and other cosmological parameters. The Astrophysical Journal. 2010;**724**:588

[49] Bull P et al. Physics in the Dark Universe. 2016;**12, 56**. DOI: 10.1016/j.dark.2016.02.001 [arXiv:1512.05356 [astro-ph.CO]]

[50] Sendra I, Lazkoz R. SN and BAO constraints on (new) polynomial dark energy parametrizations: Current results and forecasts. Monthly Notices of the Royal Astronomical Society. 2012;**422**:776. DOI: 10.1111/j.1365-2966.2012.20661.x [arXiv:1105.4943 [astro-ph.CO]]

[51] Aghanim N, et al. [Planck Collaboration], arXiv:1807.06209 [astro-ph.CO].

[52] Beutler F, Blake C, Colless M, Jones DH, Staveley-Smith L, Campbell L, et al. The 6dF galaxy survey: Baryon acoustic oscillations and the local Hubble constant. Monthly Notices of the Royal Astronomical Society. 2011;**416**:3017-3032

[53] Anderson L, Aubourg É, Bailey S, Beutler F, Bhardwaj V, Blanton M, et al. The clustering of galaxies in the SDSS-III baryon oscillation spectroscopic survey: Baryon acoustic oscillations in the data releases 10 and 11 galaxy samples. Monthly Notices of the Royal Astronomical Society. 2014;**441**:24-62

[54] Xu X, Padmanabhan N, Eisenstein DJ, Mehta KT, Cuesta AJ. A 2% distance to $z = 0.35$ by reconstructing baryon acoustic oscillations–II: Fitting techniques. Monthly Notices of the Royal Astronomical Society. 2012;**427**:2146-2167

[55] Blake C, Brough S, Colless M, Contreras C, Couch W, Croom S, et al.

The WiggleZ dark energy survey: Joint measurements of the expansion and growth history at z ¡ 1. Monthly Notices of the Royal Astronomical Society. 2012; **425**:405-414

[56] Escamilla-Rivera C, Lazkoz R, Salzano V, Sendra I. JCAP. 2011;**1109**: 003. DOI: 10.1088/1475-7516/2011/09/003 [arXiv:1103.2386 [astro-ph.CO]]

[57] Verde L,Treu T, Riess AG, arXiv: 1907.10625 [astro-ph.CO]

[58] Ratra B, Peebles PJE. Cosmological consequences of a rolling homogeneous scalar field. Physical Review D. 1988;**37**: 3406. DOI: 10.1103/PhysRevD.37.3406

[59] Armendariz-Picon C, Mukhanov VF, Steinhardt PJ. Physical Review Letters. 2000;**85**:4438. DOI: 10.1103/PhysRevLett.85.4438 [astro-ph/0004134]

[60] Escamilla-Rivera C. Status on bidimensional dark energy parameterizations using SNe Ia JLA and BAO datasets. Galaxies. 2016;**4**(3):8. DOI: 10.3390/galaxies4030008 [arXiv: 1605.02702 [astro-ph.CO]]

[61] Bayes RT. An essay toward solving a problem in the doctrine of chances. Philosophical Transactions. Royal Society of London. 1764;**53**:370-418

[62] Gregory P. Bayesian Logical Data Analysis for the Physical Sciences. New York, USA: Cambridge University Press; 2005

[63] Trotta R. Applications of Bayesian model selection to cosmological parameters. Monthly Notices of the Royal Astronomical Society. 2007;**378**: 72-82

[64] Skilling J. Bayesian Annal. 2006:833. Available from: http://www.mrao.cam.ac.uk/steve/maxent2009/images/skilling.pdf

[65] Liddle AR, Mukherjee P, Parkinson D, Wang Y. Present and future evidence for evolving dark energy. Physical Review D. 2006;**74**:123506

[66] Jeffreys H. Theory of Probability. 3rd ed. Oxford, United Kingdom: Oxford University Press; 1998

[67] Ntampaka M, et al. arXiv: 1902.10159 [astro-ph.IM]

[68] Schmelzle J, Lucchi A, Kacprzak T, Amara A, Sgier R, Réfrégier A, et al. arXiv:1707.05167 [astro-ph.CO]

[69] Charnock T, Moss A. The Astrophysical Journal. 2017;**837**(2):L28. DOI: 10.3847/2041-8213/aa603d [arXiv: 1606.07442 [astro-ph.IM]]

[70] Moss A, arXiv:1810.06441 [astro-ph.IM]

[71] Moss A, arXiv:1903.10860 [astro-ph.CO]

[72] Aurelien G. Hands-On Machine Learning with Scikit-Learn and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media; 2017. https://www.oreilly.com/conferences/

[73] Kessler R, Conley A, Jha S, Kuhlmann S, arXiv:1001.5210 [astro-ph.IM]

[74] Géron A. Hands-On Machine Learning with Scikit-Learn & TensorFlow. O'REILLY; 2017. https://www.oreilly.com/conferences/

[75] Goodfellow I, Bengio Y, Courville A. Deep Learning. USA: MIT Press; 2016. Available from: http://www.deeplearningbook.org

[76] Zaremba W, Sutskever I. arXiv: 1505.00521 [cs.LG]

[77] Pedamonti D. arXiv:1804.02763 [cs.LG]

[78] Gal Y, Ghahramani Z. A Theoretically Grounded Application of Dropout in Recurrent Neural Networks. NIPS. 2016 [arXiv:1512.05287v5]

[79] Escamilla-Rivera C, Quintero MAC, Capozziello S. A deep learning approach to cosmological dark energy models. JCAP. 2019;(3). DOI: 10.1088/1475-7516/2020/03/008. arXiv:1910.02788 [astro-ph.CO]

IntechOpen

IntechOpen