**Machine Learning Approaches for Supernovae Classification**

Surbhi Agrawal Kakoli Bora Swati Routh

*PESIT-BSC, India PESIT-BSC, India Jain University, India*

**ABSTRACT**

*In this chapter, authors have discussed few machine learning techniques and their application to perform the supernovae classification. Supernovae has various types, mainly categorized into two important types. Here, focus is given on the classification of Type-Ia supernova. Astronomers use Type-Ia supernovae as “standard candles” to measure distances in the Universe. Classification of supernovae is mainly a matter of concern for the astronomers in the absence of spectra. Through the application of different machine learning techniques on the data set authors have tried to check how well classification of supernovae can be performed using these techniques. Data set used is available at Riess et al.2007 (astro-ph/0611572).*

Keywords: Cosmology, Core Collapse, Distance Modulus, Explosion, Redshift, Spectra, Thermonuclear, Type Ia, LDA, Naïve Bayes, K-Nearest Neighbor, SVM, Decision Tree, Random Forest

# INTRODUCTION

Cosmology is a data starved science. With the advancement of technology and new advanced technological telescopes and other such instruments, here we have a flood of data. Data which is not easy as well to be interpreted, very complex data. So, astronomical area requires various techniques which help in dealing with the problem of interpretation and analysis of such vast complex data. Out of several astronomical problems, here we have taken one such problem i.e. the problem of supernovae (SNe) classification using certain machine learning algorithms. But the question is why we need to classify supernovae or why is it important?

A supernova is a violent explosion of a star, whose brightness for an amazingly short period of time, matches that of the galaxy in which it occurs. This explosion can be due to the nuclear fusion in a degenerated star or by the collapse of the core of a massive star, both leads in the generation of massive amount of energy. The shock waves due to explosion can lead to the formation of new stars and also helps astronomers indicate the astronomical distances. Supernovae are classified according to the presence or absence of certain features in their orbital spectra. According to Rudolph Minkowski there are two main classes of supernova, the Type-I and the Type-II. Type-I is further subdivided into three classes i.e. the Type-Ia, the Type-Ib and the Type-Ic. Similarly, Type II supernova are further sub-classified as Type II-P, Type II-L and Type IIn. The detail classification of these two types of supernova is discussed in the following section. Astronomers face lot of problem in classifying them because a supernova changes itself over the time. At one instance a supernovae belonging to a particular type, may get transformed into the supernovae of other type. Hence, at different time of observation, it may belong to different type. Also, when this spectra is not available, it poses a great challenge to classify them. They have to rely only on photometric measurements for their classification. This poses a big challenge in front of astronomers to do their studies. Figure 1 shows the supernova classification from their light curves.

*Figure 1. Supernova light curves*

*courtesy: www.astro.princeton.edu/~burrows/classes/403/****supernovae****.pdf*

Machine learning methods help researchers to analyze the data in real time. Here, we build a model from the input data. A learning algorithm is used to discover and learn knowledge from the data. These methods can be supervised (that rely on training set of objects for which target property is known) or unsupervised (require some kind of initial input data but unknown class).

In this chapter, classification of Type Ia supernova are taking in considerations from a supernova dataset defined in (Davis et.al,2007), (Reiss et al.,2007) and (Wood Vessey et al.,2007) using several machine learning algorithms. To solve this problem, the dataset is classified in two classes which may aid astronomers in the classification of new supernovae with high accuracy. The chapter is further organized as - background, Machine learning techniques, results and conclusion.

## BACKGROUND

Current models of the universe posit the existence of a ubiquitous energy field of unknown composition that comprises about 73% of all mass-energy and yet that can only be detected through subtle effects. Cosmologists have dubbed this mysterious field dark energy, and over the past decade, it has become an accepted part of the standard cosmology and a focus of observational efforts. More than that: it is fair to say that understanding dark energy has become the central problem in modern cosmology. In (Christopher et.al.,2009) describe two classes of methods for making sharp statistical inferences about the equation of state from observations of Type-Ia Supernovae (SNe). The dark energy pressure and density are expressed in terms of co-moving distance, r, from which they calculated reconstruction equation, w, which is very important to express various cosmological models. First, they derive a technique for testing hypotheses about w, the equation of state that requires no assumptions about its form and can distinguish among competing theories. The technique is based on combining shape constraints on r, features of the functions in the null hypothesis, and any desired cosmological assumptions. Second, they develop a framework for nonparametric estimation of w with corresponding assessment of uncertainty. Given a sequence of parametric models for w of increasing dimension, we use the forward operator T(·) to convert it to a sequence of models for r and use the data to select among them.

Kernel principal component analysis (KPCA) with 1NN (K=1 for K nearest neighbor) was proposed in (Emile E.O.Ishida,2012), in order to perform the supernovae photometry classification. Dimensionality reduction was done using PCA and KNN algorithm was applied thereafter. The study concluded that for a dark energy survey sample, 15% of the original set will be classified with the purity of >=90%.

In (Djorgovski et.al.,2012), an automatic classification method was proposed for astronomical catalog with missing data. Bayesian networks (BNs), a probabilistic graphical model that is able to predict missing values in the observed data and dependency relationships between variables is used. To learn a Bayesian network from incomplete data, an iterative algorithm that utilizes sampling methods and expectation maximization algorithm to estimate the distributions and probabilistic dependencies of variables from data with missing values, was deployed. The goal was to extrapolate values of missing features from the observed ones. In this work, the authors used Gaussian node inference which is commonly used for continuous data. Each variable is modeled with a Gaussian distribution where its parameters are linear combinations of the parameters of the parent nodes in the Bayesian. In (Joseph W. R et. al.,2012), a semi-supervised method to classify photometric supernova typing was used. The nonlinear dimensionality reduction was performed on the supernova light curves using diffusion map followed by random forest classification on a spectroscopically confirmed trained set to learn a model that can predict types of each newly observed data. It was observed that despite collecting data on a smaller number of supernovae, deeper magnitude-limited spectroscopic surveys are better for producing training sets. For Type-Ia supernovae it was observed that there was 44% increase in purity and 30% increase in efficiency. When redshift is incorporated, it leads to a 5% improvement in Type Ia purity and 13% improvement in Type Ia efficiency. Next, they used K2 algorithm, a greedy search strategy to learn the structure of the BN. Finally, Random Forest (RF) classifier was implemented which produced reasonably accurate results. In (Karpenka N.V et.al.,,2013), authors proposed another classification method, which was mainly a two-step method. Initially, binary classification was performed to check, whether supernova is Type-Ia or not. They used neural network approach to identify whether the supernova is Type-Ia or not. In the first phase, SN light curve flux measurements were fitted individually using Gaussian likelihood function. Then from each fit the parameter vector having mean & standard deviations along with flux measurements. Maximum likelihood value and Bayesian evidence were used as input for neural network training.

**CATEGORIZATION OF SUPERNOVA**

The basic classification of supernova is done depending upon the shape of their light curves and the nature of their spectra. But there are different ways of classifying the supernovae-

1. **Based on presence of hydrogen in spectra.** If hydrogen is not present in the spectra then it belongs to the Type I supernova; otherwise, it is the Type II.
2. **Based on type of explosion.** There are two types of explosions that may takes place in the star- *thermonuclear* and *core-collapse* . Core collapse, happens at the final phase in the evolution of a massive star, whereas thermonuclear explosions are found in white dwarfs.

The detailed classification of supernova is given below where both types are discussed in correspondence to each other. The classification is the basic classification depending on Type I & Type II . According to (Phillip,P.,2013), it was considered that the two classes (Type I & Type II) are categorized based on the explosion method. According to that belief, Type I belongs to thermonuclear explosion & Type II belongs to core-collapse. But, with recent research studies, it has been observed that both types may belong to both categories.

**Type I supernova -**

Supernova are classified as Type I if their light curves exhibit sharp maxima and then die away smoothly and gradually. The spectra of Type I supernovae are hydrogen poor. As discussed earlier they have three more types- Type-Ia, Type-Ib and Type-Ic.

According to (Fraser,2016) and (TypeI and TypeII Supernovae,2016), Type Ia supernova are created when we have binary star where one star is a white dwarf and the companion can be any other type of star, like a red giant, main sequence star, or even another white dwarf. The white dwarf pulls off matter from its companion star and the process continues till the mass exceeds the **Chandrasekhar limit** of 1.4 solar masses.(Phillip P,2013),The Chandrashekahar limit/mass is the maximum mass at which a self-gravitating object with zero temperature can be supported by electron degeneracy method). This causes it to explode. Type-Ia is due to the thermonuclear explosion and has strong silicon absorption lines at 615 nm and this type is mainly used to measure the astronomical distances. This is the only supernova that appears in all type of galaxies. Type-Ib has strong helium absorption lines and no silicon lines; Type-Ic has no silicon and no helium absorption lines. Type-Ib and Type-Ic are core collapse supernova like Type II without hydrogen lines. The reason of Type-Ib and Type-Ic to fall in core collapse is that they produce little Ni (Phillips M.M.,1993) and are found within or near star formation regions. Core collapse explosion mechanism happens in massive stars for which hydrogen is exhausted and sometimes even He (as in case of Type-Ic). Both the mechanisms are shown in figure 2a & 2b.

*Figure 2a. Core Collapse Mechanism Figure 2b. Thermonuclear Mechanism*

***courtsey : http://www-astro.physics.ox.ac.uk/~podsi/sn\_podsi.pdf***

Some of the supernova are determined using their spectroscopic properties and some using light curves.

**Type II supernova**

Type-II is generally due to core collapse explosion mechanism. These supernovae are modeled as implosion-explosion events of a massive star. An evolved massive star is organized in the manner of an onion, with layers of different elements undergoing fusion. The outermost layer consists of hydrogen, followed by helium, carbon, oxygen, and so forth. According to (Fraser,2016), a massive star, with 8-25 times the mass of the Sun, can fuse heavier elements at its core. When it runs out of hydrogen, it switches to helium, and then carbon, oxygen, etc., all the way up the periodic table of elements. When it reaches iron, however, the fusion reaction takes more energy than it produces. The outer layer of the star collapses inward in a fraction of a second, and then detonates as a Type II supernova. Finally the process left with a dense neutron star as a remnant. This show a characteristic plateau in their light curves a few months after initiation. They have less sharp peaks at maxima and peak at about 1 billion solar luminosities. They die away more sharply than the Type I. It has visible strong hydrogen and helium absorption lines. If the massive star have more than 25 times mass of the Sun, the force of the material falling inward collapses the core into a black hole. The main characteristics of Type II supernova is the presence of hydrogen lines in its spectra. These lines have P Cygni profiles and are usually very broad, which indicates rapid expansion velocities for the material in the supernova.

Type II supernova are sub-divided based on the shape of their light curves. Type II-Linear (Type II-L) supernova has fairly rapid, linear decay after maximum light. Type II-plateau (Type II-P) remains bright for a certain period of time after maximum light i.e. they shows a long phase that lasts approximately 100d and here light curves are almost constant(plateau phase). Type II-L is rarely found and doesn’t show the plateau phase, but decreases logarithmically after their light curve is peaked. As they drop on logarithmic scale, more or less linearly, hence L stands for “Linear”. In Type II-narrow (Type IIn) supernova, hydrogen lines had a vague or no P Cygni profile, and instead displayed a narrow component superimposed on a much broader base. Some type Ib/Ic and IIn supernova with explosion energies E >1052 erg are often called ***hypernovae***.

The classification of supernova is shown in figure 3 with the following flowchart as-

## *Figure 3. Classification of Supernova*

## 4 Machine Learning Techniques

Machine learning is a discipline that constructs and study algorithms to build a model from input data. The type and the volume of the dataset will affect the learning and prediction performance. Machine learning algorithms are classified into ***supervised*** and ***unsupervised*** methods, also known as predictive and descriptive, respectively. Supervised methods are also known as classification methods. For them class labels or category is known. Through the data set for which labels are known, machine is made to learn using a learning strategy, which uses parametric or non-parametric approach to get the data. In *parametric* model, there are fixed number of parameters and the probability density function is specified as p(x|) which determines the probability of pattern x for the given parameter (generally a parameter vector). In *nonparametric* model, there are no fixed number of parameters, hence cannot be parameterized. Parametric models are basically probabilistic models like Bayesian model, Maximum Aposteriori Classifiers etc. and non- parametric where directly decision boundaries are determined like Decision Trees, KNN etc. These models ( parametric and nonparametric mainly talks about the distribution of data in the data set, which helps to take the decision upon the use of appropriate classifiers.

If class labels are not known (unsupervised case), and data is taken from different distributions it is hard to assess. In these cases, some distance measure, like Euclidian distance, is considered between two data points, and if this distance is 0 or nearly 0, the two points are considered as similar. All the similar points are kept in the same group, which is called as cluster. Likewise the clusters are devised. While clustering main aim is to keep high intracluster similarity and low intercluster similarity. There are several ways in which clustering can be done. It can be density based, distance based, grid based etc. Shapes of the cluster also can be spherical, ellipsoidal or any other based on the type of clustering being performed. Most basic type of clustering is distance based, on the basis of which K-means algorithm is devised which is most popular algorithm. Other clustering algorithms to name a few are K- medoids, DBScan, Denclue etc. Each has its own advantages and limitations. They have to be selected based on the dataset for which categorization has to be performed. Data analytics uses machine learning methods to make decision for a system.

According to (Nicholas,2010), supervised methods rely on a *training set* of objects for which the target property, for example a classification, is known with confidence. The method is trained on this set of objects, and the resulting mapping is applied to further objects for which the target property is not available. These additional objects constitute the *testing set*. Typically in astronomy, the target property is spectroscopic, and the input attributes are photometric, thus one can predict properties that would normally require a spectrum for the generally much larger sample of photometric objects.

On the other hand, unsupervised methods do not require a *training set.* These algorithms usually require some prior information of one or more of the adjustable parameters, and the solution obtained can depend on this input.

In between supervised & unsupervised algorithms there is one more type of model- ***semi-supervised*** method is there that aims to capture the best from both of the above methods by retaining the ability to discover new classes within the data, and also incorporating information from a training set when available.

Below are few supervised learning techniques which have been used for classification of supernovae.

## Decision Tree

A decision tree classifier is a machine learning approach which constructs a tree that can be used for classification or regression. Each of the nodes are based on a feature (attribute) of the data set, the first node is called as **root node**, which can be any important feature and hence considered as best predictor. Every other node of the tree is then split into child nodes based on certain splitting criteria or decision rule, which identify the allegiance of the particular object (data) to the feature class. The final result is a tree with **decision nodes** and **leaf nodes** where leaf nodes represent the classes of classification. Typically an impurity measure is defined for each node and the criterion for splitting a node is based on increase in purity of child nodes as compared to the parent node i.e. splits that produce child nodes which have significantly less impurity as compared to the parent node are favored. The Gini index (for CART) and entropy are two popular impurity measures. Entropy is used to interpret as a descriptor of information gain from that node. One significant advantage of decision tree is that both categorical and numerical data can be handled, a disadvantage is that decision trees tend to overfit the training data.

The core algorithm for building decision trees is known as **ID3** given by J. R. Quinlan. ID3 employs a top-down, greedy search through the space of possible branches with no backtracking and uses *Entropy* and *Information Gain* to construct a decision tree. Below is the ID3 algorithm(Decision Trees,2015)-

**ID3 Algorithm**

Characterization of the model:

* X is a set of feature vectors, also called feature space.
* C is a set of classes.
* c : X --> C is the ideal classifier for X.
*  is a set of examples.

Construction of a decision tree T to approximate c., based on D

**Algorithm ID3(D, Attributes, Target)**

* Create a node t for the tree.
* Label t with the most common value of Target in D.
* If all examples in D are positive, return the single-node tree t, with label say, “+”.
* If all examples in D are negative, return the single-node tree t, with label say, “–”.
* If Attributes is empty, return the single-node tree t.
* Otherwise:\*
  + Let A\* be the attribute from Attributes that best classifies examples in D.
  + Assign t the decision attribute A\*.
  + For each possible value “a” in A\* do:
    - Add a new tree branch below t, corresponding to the test A\* = “a”.
    - Let D\_a be the subset of D that has value “a” for A\*.
    - **If** D\_a is empty:

Then add a leaf node with label of the most common value of Target in D.

**Else** add the subtree ID3(D\_a, Attributes {A\*}, Target).

* Return t.

## K-Nearest Neighbor

K-NN classifies an unknown instance with the most common class among K closest instances. It is an instance-based classifier that compares new incoming instance with the data already stored in memory. K-Nearest Neighbors algorithm (or K-NN for short) is a non-parametric method used for classification and regression. Using a suitable distance or similarity function, K-NN relates new problem instances to the existing ones in the memory. K neighbors are located and majority vote outcome decides the classification. Occasionally, the high degree of local sensitivity makes the method susceptible to noise in the training data. If K = 1, then the object is assigned to the class of that single nearest neighbor. If K > 1, then the object is assigned to the class which has K number of similar nearest neighbor. According to (Veksler, 2013), in theory if infinite number of samples available, the larger the K values, the better is the classification, which gives smoother boundary. As per theoretical properties if K < sqrt(n), where n is number of objects to be classified, classification produces better results. One can choose K through cross validation also. The algorithm usually uses *Euclidian distance, Manhattan distance* or *Minkowski distance* to find the nearest neighbor for continuous variables. In the instance of categorical variables the Hamming distance must be used. This algorithm works well for larger dataset. It can be applied to the data from any distribution. A shortcoming of the K-NN algorithm is its sensitivity to the local structure of the data.

## Linear Discriminant Analysis

LDA is commonly used for both dimensionality reduction and data classification. The basic LDA classifier attempts to find a linear boundary that best separates the data. LDA maximizes the ratio of between-class variance to the within-class variance in any dataset guarantying maximal separability. This yields the optimal Bayes’ classification (i.e. under the rule of assigning the class having highest posterior probability) under the assumption that covariance is same for all classes. According to (Balakrishnama S. et. Al,1998), there are two different approaches used to classify test vectors. They are *Class dependent transformation* and *class independent transformation.*

**Class dependent transformation:** This approach involves maximizing the ratio of between-class variance to within-class variance so that adequate class separability is obtained. This approach uses two optimizing criterion to transform the data sets.

**Class independent transformation:** This approach involves maximizing the ratio of overall variance to within-class variance. This approach uses only one optimizing criterion to transform the data sets. In this type of LDA, each class is considered as a separate class against all other classes.

The choice of type of LDA depends on the dataset as well as the goal of classification. For good discrimination of the classes, we prefer class dependent transformation and for general classification problem, class independent transformation is preferred.

In present implementation, an enhanced version of LDA (often called Regularized Discriminant Analysis) is used. This involves Eigen-decomposition of the sample covariance matrices and transformation of the data and class centroid. Finally the classification is performed using the nearest centroid in the transformed space also taking into account prior probabilities.

## Naïve Bayes

Naïve Bayes classifier is based on Bayes theorem. It is a supervised learning method as well as a statistical method for classification. It allows us to capture uncertainty about a probabilistic model in a better way by determining probabilities of the outcome. It can solve both diagnostic and predictive problems. It can perform the classification of arbitrary number of independent variables and is generally used when data is high-dimensional. Data to be classified can be either categorical or numerical. A small amount of training data is sufficient to estimate necessary parameters. The method assumes independent distribution for attributes and thus estimates ; where  are ‘n’ input variables and *Yi* are ‘i’ different classes. Although this assumption is often violated in practice (hence the name Naïve), Naïve Bayes often performs well. This classifier provides practical learning algorithm and prior knowledge and observed data can be combined. It calculates explicit probabilities for hypothesis and it is robust to noise in input data. It is computationally fast and space efficient.

**Random Forest**

Random forest is an ensemble of various decision trees. Each tree enunciates a classification and decision is taken based upon mean prediction on them (regression) or majority voting (classification). When a new object from the data set needs to be classified, data is kept down at each of the trees. Classification implies a tree voting for that class. Random forest works efficiently with large datasets. It gives accurate results even in the cases of missing data. The training algorithm for random forests applies the general technique of bootstrap aggregating, or bagging, to tree learners. Grow each tree in the forest on an independent bootstrap sample from the training data. At each node select ‘m’ variables at random out of all ‘M’ possible variables; find the best possible split on the selected ‘m’ variables.

Given a training set with responses , bagging selects a random sample of the training set with replacement iteratively and fits trees to these samples. For b = 1, …, B: Sample, with replacement, n training examples from X, Y; call these . Next, we train a decision or regression tree on . Post-training, predictions for unseen samples x' can be made by averaging the predictions from all the individual regression trees on x': or by considering the majority votes in the case of decision trees.

* 1. **Support Vector Machine**

SVM classifiers are candidate classifiers for binary class discrimination. The basic formulation is designed for the linear classification problem; the algorithm yields an optimal hyperplane i.e. one that maintains largest minimum distance from the training data, defined as the margin. It can also perform non-linear classification via the use of kernels, which involves the computation of inner products of all pairs of data in the feature space; this implicitly transforms the data into a different space where a separating hyperplane can be found. One advantage of the SVM is that the optimization problem is convex. The result may not be transparent always which is a drawback of this method.

**SUPERNOVAE DATA SOURCE AND CLASSIFICATION**

The selection of classification algorithm not only depends on the dataset, but also the application for which it is employed. There is, therefore, no simple method to select the best optimal algorithm. Our problem is to identify Type Ia supernova from the given dataset in (Davis M.S. et al.,2007) which contains 292 different supernova information. Since the classification is binary classification, as one need to identify Type Ia supernova from the list of 292 supernovas, the best resulting algorithms are used for this purpose. The algorithms used for classification are Naïve Bayes,LDA, SVM, KNN, Random Forest and Decision Tree.

The dataset used is retrieved from (Davis M.S. et al.,2007). These data are a combination of the ESSENCE, SNLS and nearby supernova data reported in Wood-Vasey et al. (2007) and the new Gold dataset from Riess et al.(2007). The final dataset used is combination of ESSENCE / SNLS / nearby dataset from Table 4 of Wood-Vasey et al. (2007), using only the supernova that passed the light-curve-fit quality criteria. It has also considered the HST data from Table 6 of Riess et al. (2007), using only the supernovae classified as gold. These were combined for Davis et al. (2007) and the data are provided in 4 columns: redshift, distance modulus, uncertainty in the distance modulus and quality as “Gold” or “Silver”. The supernova with quality labeled as “Gold” are Type Ia with *high confidence* and those with label “Silver” are *Likely but uncertain SNe Ia.* In the dataset, all the supernova with redshift value less than **0.023** and quality value **Silver** are discarded.

**RESULTS AND ANALYSIS**

The experimental study was setup to evaluate performance of various machine learning algorithms to identify Type-Ia supernova from the above mentioned dataset. The data set mentioned above is tested on 6 major classification algorithms namely Naïve Bayes, Decision tree, LDA, KNN, Random Forest and SVM respectively. A ten-fold cross validation procedure was carried out to make the best use of data, that is, the entire data was divided into ten bins in which one of the bins was considered as test-bin while the remaining 9 bins were taken as training data. We observe the following results and conclude that the outcome of the experiment is encouraging, considering the complex nature of the data. Table 1 shows the result of classification.

*Table 1. Results of Type -Ia supernova classification*

|  |  |
| --- | --- |
| **Algorithm** | **Accuracy (%)** |
| Naïve Bayes | 98.86 |
| Decision Tree | 98.86 |
| LDA | 65.90 |
| KNN | 96.59 |
| Random Forest | 97.72 |
| SVM | 65.90 |

Performance analysis of the algorithms on the dataset is as follows.

1. Naïve Bayes’ and Decision Tree top the accuracy table with the accuracy of 98.86%.

2. Random Forest ranks 2 with accuracy of 97.72% and KNN occupies 3rd position with 96.59% accuracy.

3. The dramatic change was observed in the case of SVM, which occupied the last position with LDA with an accuracy of 65.9%. The geometric boundary constraints inhibit the performance of the two classifiers.

Overall, we can conclude Naïve Bayes’, Decision Tree and Random Forest perform exceptionally well with the datasets, while KNN acts as an average case.

**FUTURE RESEARCH DIRECTIONS**

Supernova classification is an emerging problem that scientists, astronomers and astrophysicists are working on to solve using various statistical techniques. In the absence of spectra, how this problem can be solved. In this chapter, Type-Ia supernova are classified using machine learning techniques based on redshift value and distance modulus. The same techniques can be applied to solve the overall supernova classification problem. It can help us to differentiate Type I supernova from Type II, Type Ib from Type Ic or so on. Machine learning techniques along with various statistical methods help us to solve such problems.

**CONCLUSION**

In this chapter, we have compared few classification techniques to identify Type Ia supernova. Here it is seen that Naive Bayes, Decision Tree and Random Forest algorithms gave best result among all. This work is relevant to astroinformatics, especially for classification of supernova, star-galaxy classification etc. The dataset used is a well-known which is the combination of ESSENCE, SNLS and nearby supernova data.

Supernovas are discovered in regular intervals. A supernova will occur about once every 50 years in a galaxy the size of the Milky Way. In other words, we can say a star explodes every second or so somewhere in the universe, and some of those aren’t too far from Earth. A huge task of categorizing those manually may be translated into a simple automated system using this work. The new data can be appended to the dataset with discovered but non-categorized supernova. The machine learning algorithms could then perform the task of classification, as demonstrated earlier with reasonably acceptable accuracy. A significant portion of time could thus be saved.

**REFERENCES**

[1] Nicholas M. Ball & Robert J. Brunner (2010),*Overview of Data Mining and Machine Learning methods.* Retrieved on 25-04-16, from http://ned.ipac.caltech.edu/level5/March11/Ball/Ball2.html

[2] Davis et al. (2007), *“Scrutinizing Exotic Cosmological Models Using ESSENCE Supernova Data Combined with Other Cosmological Probes”***,** Astrophys.J.666:716-725, DOI: [10.1086/519988](http://arxiv.org/ct?url=http%3A%2F%2Fdx.doi.org%2F10%252E1086%2F519988&v=6ba26310)

[3] Riess et al. (2007), “*New Hubble Space Telescope Discoveries of Type Ia Supernovae at z > 1: Narrowing Constraints on the Early Behavior of Dark Energy*” Astrophys.J.659:98-121, DOI: [10.1086/510378](http://arxiv.org/ct?url=http%3A%2F%2Fdx.doi.org%2F10%252E1086%2F510378&v=29a2c69b)

[4] Wood-Vassey et al. (2007), “*Observational Constraints on the Nature of the Dark Energy: First Cosmological Results from the ESSENCE Supernova Survey*”, Astrophys.J.666:694-715, DOI: [10.1086/518642](http://arxiv.org/ct?url=http%3A%2F%2Fdx.doi.org%2F10%252E1086%2F518642&v=2a2aff08)

[5] *Type Ia supernova data* used by Davis, Mörtsell, Sollerman, et al. (2007)**,** Retrieved from <http://dark.dark-cosmology.dk/~tamarad/SN/>

[6] Fraser Cain (2016)**,** *What are the Different Kinds of Supernovae?*, retrieved on 20-04-2016, from <http://www.universetoday.com/127865/what-are-the-different-kinds-of-supernovae/>

[7]*Type I and Type II Supernovae****,*** 2016, from [http://hyperphysics.phy-astr.gsu.edu/hbase/astro/snovcn.html#c3](http://hyperphysics.phy-astr.gsu.edu/hbase/astro/snovcn.html" \l "c3) , retrieved on 20-04-2016

[8] *Supernovae and Gamma-Ray Bursts,*  Philipp P,2013, Dept. of Astrophysics, University of Oxford, Oxford, OX1 3RH, UK <http://www-astro.physics.ox.ac.uk/~podsi/sn_podsi.pdf>

[9] Phillips (1993), M.M.: Astrophys. J. , 413 , L105

[10] *Decision Trees*,2015, retrieved on 03-05-16 from [www.uni-weimar.de/medien/.../unit-en-](http://www.uni-weimar.de/medien/.../unit-en-decision-trees-algorithms.pdf)**[decision](http://www.uni-weimar.de/medien/.../unit-en-decision-trees-algorithms.pdf)**[-](http://www.uni-weimar.de/medien/.../unit-en-decision-trees-algorithms.pdf)**[trees](http://www.uni-weimar.de/medien/.../unit-en-decision-trees-algorithms.pdf)**[-](http://www.uni-weimar.de/medien/.../unit-en-decision-trees-algorithms.pdf)**[algorithms](http://www.uni-weimar.de/medien/.../unit-en-decision-trees-algorithms.pdf)**[.pdf](http://www.uni-weimar.de/medien/.../unit-en-decision-trees-algorithms.pdf)

[11] Olga Veksler,2013, “*k Nearest Neigbors*”, retrieved on 04-05-16 from [www.csd.uwo.ca/courses/CS9840a/Lecture2\_](http://www.csd.uwo.ca/courses/CS9840a/Lecture2_knn.pdf)**[knn](http://www.csd.uwo.ca/courses/CS9840a/Lecture2_knn.pdf)**[.pdf](http://www.csd.uwo.ca/courses/CS9840a/Lecture2_knn.pdf)

[12] S. Balakrishnama & A. Ganapathiraju, “*Linear Discriminant Analysis- A brief tutorial*”, 1998, retrived on 04-05-16 from <https://www.isip.piconepress.com/publications/reports/1998/isip/lda/lda_theory.pdf>

[13] Christopher R. Genovese, Peter Freeman, Larry Wasserman, Robert C. Nichol and Christopher Miller, 2009, “*Inference for the Dark Energy equation of state using Type Ia Supernova data” – The Annals of applied statistics*, vol.3, No. 1, 144-178 .

[14] Emille E. O. Ishida ,2012., *Kernel PCA for Supernovae Photometric Classification. Proceedings of the International Astronomical Union*, 10, pp 683-684. doi:10.1017/S1743921314012897.

[15]S. G. Djorgovski, A. A. Mahabal, C. Donalek, M. J. Graham, A. J. Drak, B. Moghaddam, M. Turmon, 2012, *Flashes in a Star Stream: Automated Classification of Astronomical Transient Events*; <https://arxiv.org/ftp/arxiv/papers/1209/1209.1681.pdf>

[16] Joseph W. Richards, Darren Homrighausen, Peter E. Freeman, Chad M. Schafer, Dovi Poznanski, 2012, *Semi-supervised learning for Photometric Supernova Classification;* MNRAS, 419, 1121–1135 (2012) doi:10.1111/j.1365-2966.2011.19768.x

[17] N. V. Karpenka, F. Feroz and M. P. Hobson,2013, *A simple and robust method for automated photometric classification of supernovae using neural networks*, MNRAS429,1278–1285 (2013)