

SWELLING STUDIES OF PHYSICAL AND CHEMICAL CHITOSAN HYDROGELS

Gorriparthi Varalakshmi ^{1*}, and Anitha C. Kumar ²

¹ PhD Research Scholar, Department of Chemistry, Acharya Nagarjuna University, Nagarjuna Nagar, Guntur, AP, India.

¹ Assistant Professor, Department of Chemistry, Samskruti College of Engineering and Technology, Ghatkesar, Hyderabad, Telangana, India.

² Professor, School of Chemical Sciences, Mahatma Gandhi University, Kottayam, Kerala, India.

*Corresponding Author Email: vgoriparthi09@gmail.com

DOI: 10.5281/zenodo.11407444

Abstract

Hydrogels have been available for more than half a century and they have many uses in various practices ranging from industrial to biological. Hydrogels are a polymeric network that can absorb a large quantity of water and swells up to an equilibrium point. The water absorbing capacity of hydrogels depends on hydrophilic functional groups and space availability in the polymeric structure. Here in this work, we prepared hydrogels from an animal based natural polymer-Chitosan. Crosslinkers play an important role in the properties of hydrogels. Chitosan hydrogels were prepared without crosslinker and by using different concentrations of glutaraldehyde, which is the most commonly used crosslinker. Chitosan hydrogels can hold a large amount of water by swelling. Swelling nature of uncrosslinked and crosslinked hydrogels were studied in aqueous medium at various temperature and in different pH mediums at different intervals of time.

Keywords: Chitosan, Glutaraldehyde, Swelling properties, Hydrogels.

INTRODUCTION

Hydrogels are hydrophilic polymer materials that are cross-linked, three-dimensional and can retain, swell, and hold large volumes of water or aqueous fluids (1). The associations are comprised of homopolymers or co-polymers and are insoluble due to the presence of chemical cross-links (tie-points, junctions), or physical cross-links, such as embarrasments or crystallites (2). These gels are effective in day-to-day life and used extensively in food, husbandry manufacturing, cosmetic make up foundations, medicine, and medical discourse etc. Currently, Hydrogel get wide orbit of applications because of their lower berth price, nonpoisonous, and environment-sociable nature. Hence, these are used as biomaterials that reveal sensible purpose in several fields of biomedical or therapeutic industry (3).

Chitosan, is a linear polysaccharide collected of β -(1-4)-linked d-glucosamine and N-acetyl-d glucosamine. Chitosan is derived from Chitin, which is extracted from hard skeleton matter of shellfish, lobster, shrimp, and crabs. The structure of Chitosan is shown in figure 1. It has lots of applications in the medical, health, and pharmaceutical industry because of its biocompatible nature. Any form of chitosan (solution, dryflakes, and fine powder) can dissolve in organic acids such as acetic acid, formic acid, tartaric acid, and citric acid of $P^H < 6$. Crosslinked chitosan is frequently used to prepare microspheres of their form, they have been widely used for drug delivery such as antibiotics, anticancer agents, vaccines, etc. (4, 5). The positive amino group of chitosan is very active and readily binds with negatively charged surfaces such as mucosal membranes. The free amino group of chitosan can form crosslinked polymer networks with dicarboxylic acids to build up the material. It forms gels with several multivalent anions. The hydrogels of chitosan, like other hydrogels, contain much

water. Part of this water is tightly bound to the polymer and the rest is present as free water. Chitosan and its derivatives are one of the best-optimized material for medical and pharmaceutical industries, particularly in their swollen state but possess low mechanical strength, to overcome this disadvantage crosslinking is the best method most commonly used crosslinkers are aldehyde, epoxy, starch, guar-gum, epichlorohydrin, sodium tripolyphosphate etc.

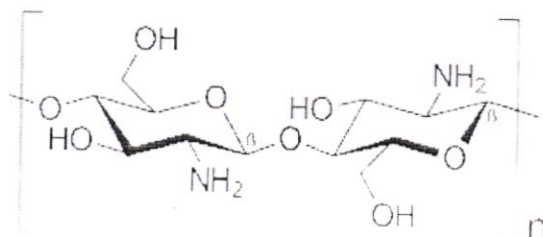


Figure 1: Structure of Chitosan

In the present study we have chosen glutaraldehyde as the crosslinker. We were prepared chemically crosslinked chitosan hydrogels with glutaraldehyde, which is an extreme crosslinker. The aldehyde group of glutaraldehyde interacts with the amino group of chitosan, thus forming Schiff's bases. The potentiality of hydrogel membrane depends on many factors. This paper also reports the variation of swelling ratio depends on temperature, P^H , and concentration of the crosslinking agent.

MATERIALS AND METHODS

Materials:

Chitosan (Ch), Glutaraldehyde (GA) & Acetic acid (Ac) are all chemicals purchased from Avra Synthesis Private Ltd, Hyderabad, Telangana, India. All obtained chemicals were used without any further purification. Buffer solutions of $P^H=4, 7$ & 9.2 purchased from Quality traders, Gorantla, Guntur.

Millipore water is used from the Merk Millipore system

Methods:

Hydrogel Preparation: Chitosan hydrogels were prepared with varying concentrations of glutaraldehyde from 0% to 0.125%. Initially chitosan allows for dissolution in a dilute acetic acid medium with stirring for 8-10 hours, after completion of dissolution chitosan is added with glutaraldehyde of various concentrations and continued stirring for 30-60 min.

Swelling experiments: For the swelling studies we prepared the dry films from the gels. The swelling property of dry film of hydrogels with and without cross-linker were examined by dropping 0.15gm of dry film in deionized water at room temperature at particular intervals of time until reached constant (stable weight) weight, simultaneously measuring the film weight in swollen state at each interval, In the same way, dry sample of known mass (wd) allow to float in a buffer solution of $P^H=4, 7$ & 9.2 . After a particular period (t) dry film remove from the buffer and their swollen sample weight (Wt) was noticed after whipping with filter paper to remove excess water, the water-retaining capacity of the film known Experimentally, the percentage swelling ratio of a polymeric hydrogel can be defined by mass difference process and is characterized by the following equation.

$$\text{Percent of Inflammation} = [(Wt-Wd)/Wd] \times 100$$

Where, Wt= mass of inflamed gel and
Wd = mass of dehydrated gel (dry gel).

FTIR Studies: Cary 630FTIR with diamond ATR, Agilent Technologies Ltd., was used to study the hydrogel samples within the wavelength range of 4000-600 cm^{-1} . Each spectrum was measured on an average of 64 scans at a spectral resolution of 4 cm^{-1} .

RESULTS AND DISCUSSION

Polymeric gels are inter-linked either physical and chemical interconnecting or thus the cross-linking implicit as a molecule at any rate of its size. The minor modifications are takes place in a polymeric gel by varying upon ecological factors such as temperature, pH, electric charge, and enzyme or ionic species, which lead to different physical texture of the gel. The inflammation property of hydrogels is due to degree of ionization of the cross-linking of monomers with the functional groups present in a polymer chain (6). Chitosan-based hydrogels are of two types (a) physically crosslinked gel and (b) chemically crosslinked gels. The first one is obtained by the weak wander Waals interaction and the second one is obtained via covalent crosslinking agents or ionic crosslinking agents. Figure 2 shows the schematic representation of the crosslinked hydrogel formation.

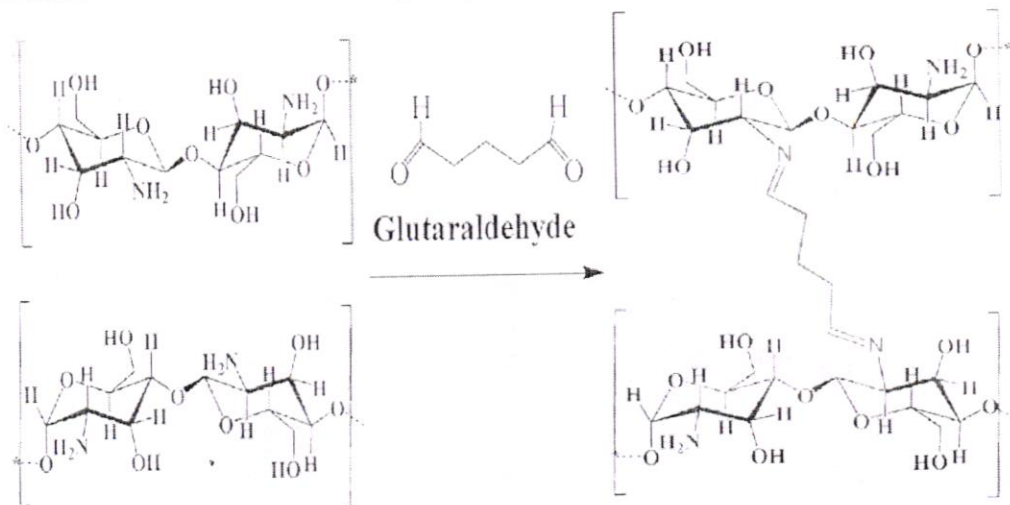


Figure 2: Schematic representation of the crosslinked hydrogel (cross linker – glutaraldehyde)

Preparation of chitosan hydrogel without crosslinker(Physical Hydrogel): Prepared 0.5, 1, 1.5, 2, 2.5, 3 and 3.5 percentages of chitosan in 10ml of 2% aqueous acetic acid at room temperature with continuous stirring for complete dissolution (8-10 hrs) and results a clear and viscous fluid. Transfer it to a sample bottle, gel nature was observed with concentrations (%) of 2.5, 3 & 3.5. These gels were a stable minimum of 2 days at room temperature, no gel was observed with remaining concentrations even after 7-10days. In this case, gels formed due to self-crosslinking of chitosan polymer, but at low concentrations, it fails. Decaying of gels nature was observed while keeping samples at room temperature with time and mentioned in table 1.

Table 1: The chitosan gel preparation at various concentrations of chitosan

s.no	%chitosan(ch)	volume of ch(ml)	Status of Gel formation	Gel formation time(hrs)	Gel decay time (days)
1	0.5	10	No gel	-----	-----
2	1	10	No gel	-----	-----
3	1.5	10	No gel	-----	-----
4	2	10	No gel	-----	-----
5	2.5	10	Gel	6-8	1
6	3	10	Gel	6-8	2
7	3.5	10	Gel	6-8	2

Preparation of chitosan hydrogel with varying concentrations of glutaraldehyde crosslinker: Since we got comparatively stable hydrogels with 3% of chitosan, it was selected for crosslinking. Solutions were prepared in the similar method and we prepared five different hydrogel samples by kept constant concentration of chitosan and varied the glutaraldehyde concentration from 0% to 0.125%. The details are given in the table 2.

Table 2: The chitosan gel preparation at various concentrations of glutaraldehyde crosslinker (Gha)

s.no	%ch	Vol. of ch (ml)	%Gha	Vol. of Gha(ml)	Status of Gel formation	Gel formation time(mins)	Gel(5ml) decay time (days)
1	3	10	0.025	0.33	No gel	30	-----
2	3	10	0.05	0.33	No gel	30	1
3	3	10	0.075	0.33	Gel	30	2
4	3	10	0.1	0.33	Gel	30	3
5	3	10	0.125	0.33	Gel	30	4

As have seen in table 2, glutaraldehyde 0.125 % gels were stable up to 4 days. The photographs of the same gel on day 3 and day 5 are shown in figure 3.

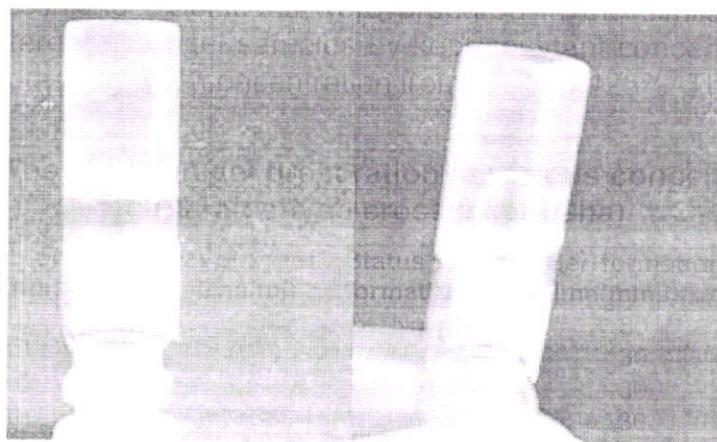


Figure 3: Hydrogel on day 3 followed by day 5(3% chitosan with 0.125%gluteraldehyde)

Swelling studies: These studies help to know the gel's stability in a water medium. Dry films were prepared from the same gel combination solutions. To study the swelling nature, a piece of film is weighed and dropped in deionized water (PH-7). The weight of the swollen film is measured in a regular interval. The swelling nature of chitosan physical gel at different temperature were studied and the results are shown in figure 4. The swelling ratio increases from with increase in temperature. It might be

due to the temperature help to break the internal bonds ie; hydrogen bonds in hydrogel and freely allow the water and attained an equilibrium point in approximately 100 mins.

The crosslinked gel with crosslinker 0.125% is also analyzed the same way at varying temperatures in deionized water (PH=7) for 100minutes in an interval of 10mins. The results are shown in the same figure4. In this case equilibrium was attained so early at 50 to70min and noted maximum swelling ratio (%S) is 250 at 46^oc which is low compared to physical gels. In the case of physical gel, the value is 965. This low value in crosslinked hydrogel might be due to strong covalent bonds in the internal structure, it does not allow the hydrogels to swell by restricting the free movement of the chain and water absorption capacity.

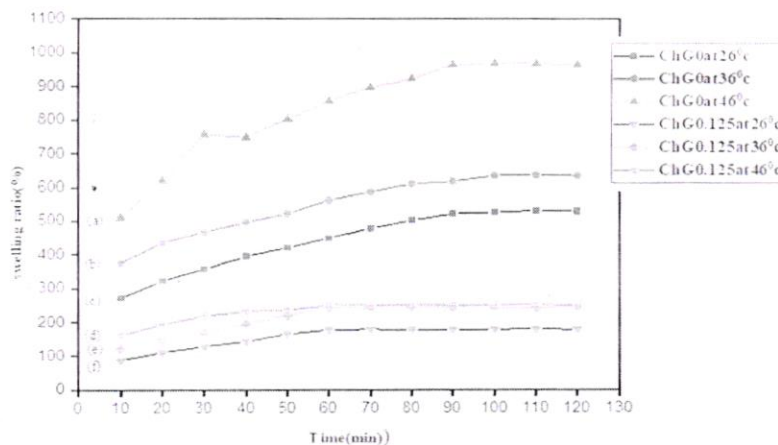


Figure 4: Swelling studies of noncrosslinked chitosan and crosslinked chitosan in water (PH=7) at various temperatures.

Hydrogels are normally formed by the crosslinking of linear polymers, depending on the type of crosslinker, hydrogel membranes are of two types Covalent and ionic, glutaraldehyde is one of the best covalent cross-linker which leads to the formation of covalent linkages between the polymeric chain. Crosslinked hydrogels vary in properties compared to noncrosslinked, particularly in their molecular structure, ionizing capacity, hydrophilicity, and degree of crosslinking. Owing to a change in crystallinity, swelling behavior differs from crosslinked to noncrosslinked.

The swelling behavior of covalent/ionic hydrogel membranes depends on external factors including the nature of the crosslinker, degree of crosslinking, the effect of time, temperature, the proportion of the crosslinker, the volume of the cross-linker, and the PH of medium. During the swelling amino group of chitosan polymer undergoes protonation, which describes the rate of swelling. Interactions of polymer to crosslinker and degree of swelling are determined by many factors, PH of the swelling medium is one among them.

The swelling behavior of the hydrogels in different PH were also studied. Figure 5 shows the swelling behavior of noncrosslinked hydrogel in buffer solutions of PH=4, 7 and 9.2 at room temperature. It was shown high values at PH=4 and low values at PH=7 and 9.2, which indicates the swelling properties of hydrogel films elevated at low values and not significant with increasing values of PH. Figure 6 shows the behavior of crosslinked hydrogels with crosslinker 0.125% at various PH values (4,7&9.2). The results shows that the swelling ratio is predominately high at lower PH=4 and gradually decreases with increasing PH (7&9.2). Noncrosslinked Gels swollen more in acidic

media ($P^H=4$) and shows a high swelling ratio than crosslinked gels. Usually, we are expecting a high swelling ratio with increasing densities of crosslinker $0.125 > 0.1 > 0$, but practically it has proven the reverse order. The swelling ratio %S is dependence on P^H in the case of crosslinked membranes and they are high at $P^H=4$, but low at P^H (7&9.2). The change in P^H causes for protonation of an amino group of chitosan and various concentrations of crosslinker were responsible for obtaining different values of the swelling ratio. The swelling behavior of crosslinked gels at various P^H are shown in figure 6.

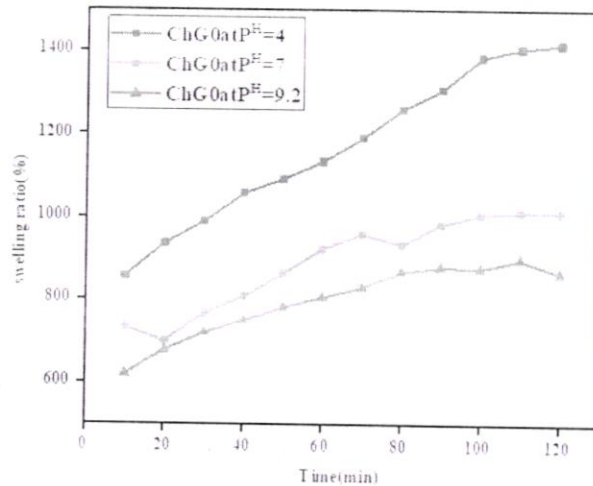


Figure 5: Swelling behavior of uncrosslinked hydrogel at various $P^H = 4, 7 \& 9.2$

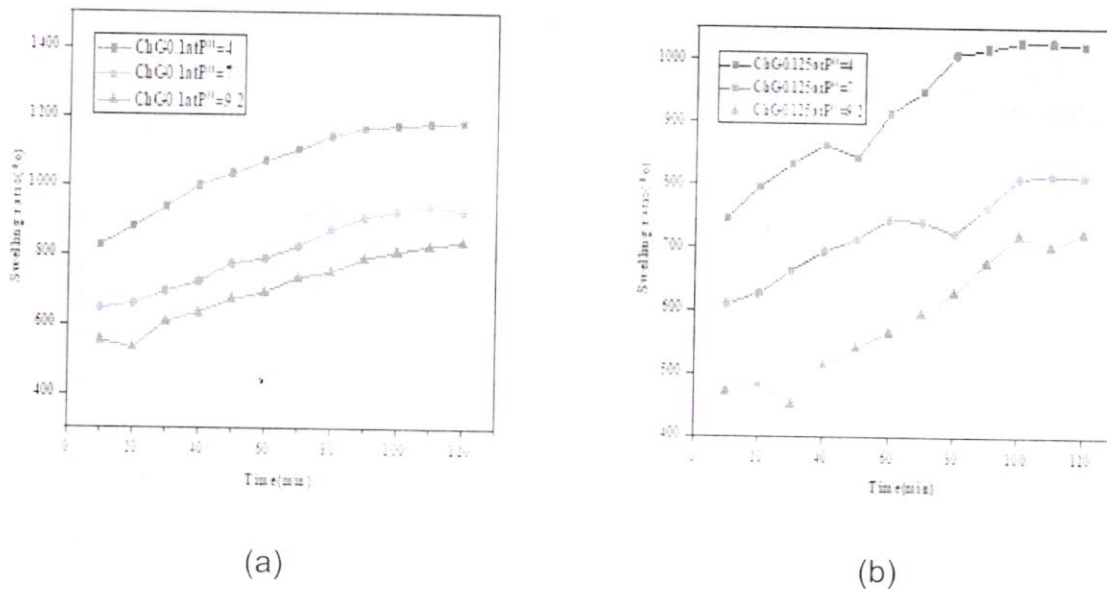


Figure 6: Swelling behavior of crosslinked hydrogels at various P^H (a) 0.1% and (b) 0.125%.

FTIR Studies: Commercial chitosan, 3%chitosan solution in acetic acid, which is noncrosslinked and crosslinked chitosan with glutaraldehyde (0.125%Gha) were analyzed and the results are shown in figure 7a,b and c.

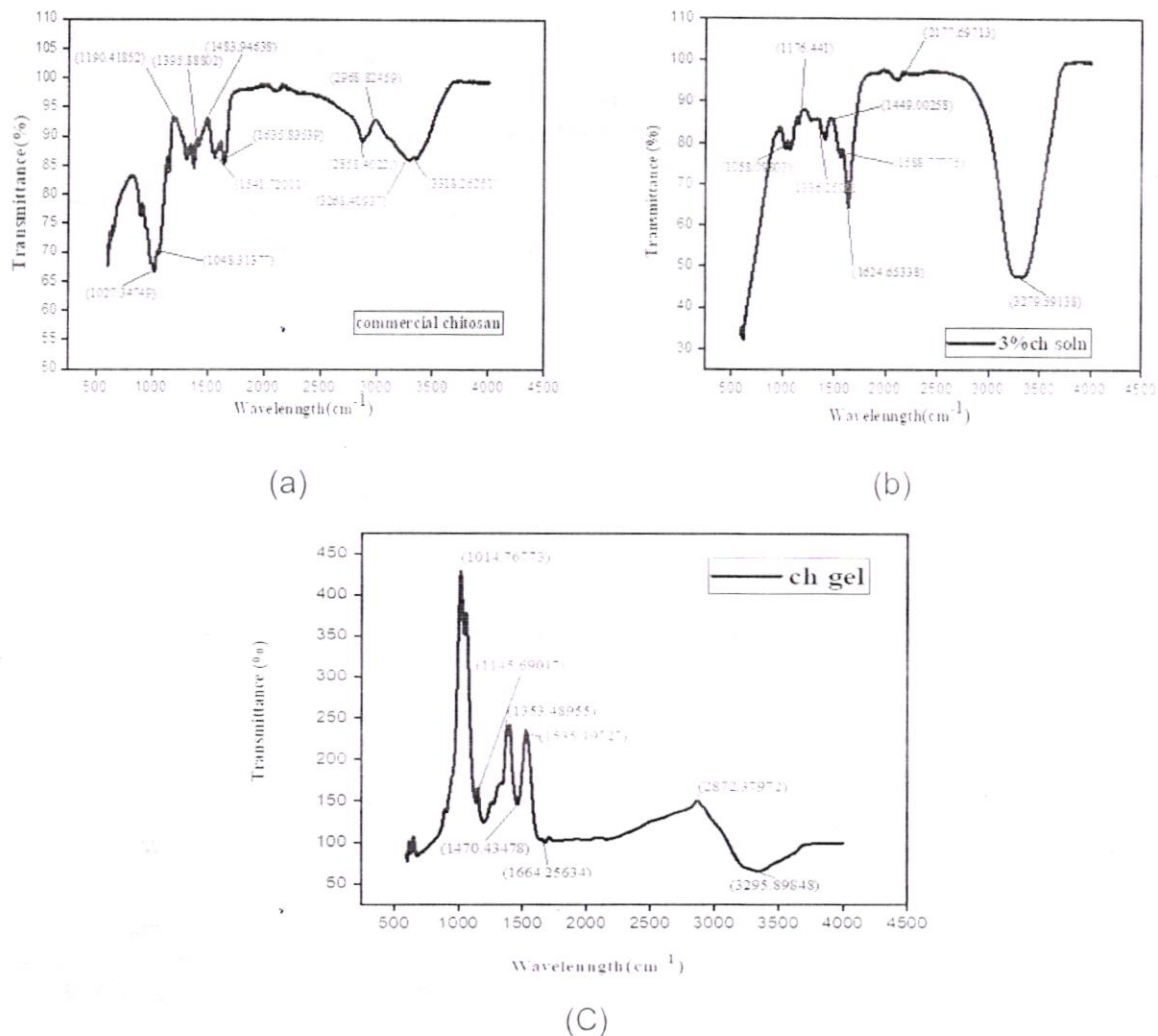


Figure 7: FTIR Spectrum of commercial chitosan (a), 3%chitosan solution in acetic acid(noncrosslinked) (b) and crosslinked chitosan with glutaraldehyde (c).

With commercial chitosan, N-H and O-H stretching vibration bands were observed at corresponding wavelengths were found at 3268 cm^{-1} , 3318 cm^{-1} respectively, same stretching vibrations were found at 3279 cm^{-1} , 3295 cm^{-1} for solution and gel. 2968 cm^{-1} , 2858 cm^{-1} can be attributed to C-H stretching vibrations of symmetric and asymmetric forms and the same (C-H) stretching vibration of the corresponding wavelength reported as 2177 cm^{-1} for solution and 2873 cm^{-1} for gels. 1635 cm^{-1} intensity band observed due to $\text{C}=\text{O}$ stretching of amide I which is slightly varied as 1624 cm^{-1} in solution and 1664 cm^{-1} in gel. The characteristics peak of N-H bending usually appears in the range of $1650\text{--}1550\text{ cm}^{-1}$ (7), which is recorded at 1540 cm^{-1} with commercial chitosan after protonation with acetic acid peak shifted to 1588 cm^{-1} & 1535 cm^{-1} . This can be attributed to the modification of chitosan structure with glutaraldehyde by crosslinking. 1483 cm^{-1} , 1449 cm^{-1} , 1470 cm^{-1} & 1395 cm^{-1} , 1336 cm^{-1} , 1353 cm^{-1} from the figures indicates C-H bending vibrations of alkyl and methyl groups consecutively. The peak range of $1300\text{--}1000\text{ cm}^{-1}$ characteristic of Ethers(C-O-) observed as 1190 cm^{-1} (Fig.7a), 1176 cm^{-1} (Fig.7b) & 1145 cm^{-1} (Fig.7c). Usually wavelength range of $1690\text{--}1640\text{ cm}^{-1}$ significance of imine bond ($\text{C}=\text{N}$), the sharp peak of 1664 cm^{-1} noticed in the case of gel which confirms Schiff's base structure (8).

CONCLUSION

Chitosan hydrogels were prepared without crosslinker and with varying concentrations of crosslinker. These gels were examined for swelling behavior at various temperatures, in water, and at different ranges of buffer solution. Crosslinked hydrogels reported a high swelling ratio concerning the above conditions because crosslinked hydrogels are internally linked with each other, acidic medium ($P^H=4$) more favor to swelling behavior which is comparatively less in the case of ($P^H=7$) & ($P^H=9.2$). Swelling behavior of noncrosslinked hydrogels varies considerably with varying temperatures, but it is negligible in the case of the crosslinked hydrogel.

FTIR analysis attributed to confirm the formation of the amide linkage, in the crosslinked hydrogels.

References

- 1) Klouda L, Mikos AG, Thermo responsive hydrogels in biomedical applications. *Eur J Pharm Biopharm* 68:34–45, 2008
- 2) P.K. Prashant, B.R. Vivek, N.D. Deepashree, P.P. Pranav. "Hydrogels as a drug delivery system and applications: a review", *International Journal of Pharmacy and Pharmaceutical Sciences*, vol. 4, pp. 1-7, 2012.
- 3) L. Takashi, T. Hatsumi, M. Makoto, I. Takashi, G. Takehiko, S. Shuji. "Synthesis of porous poly (N-isopropylacrylamide) gel beads by sedimentation polymerization and their morphology", *Journal of Applied Polymer Science*, vol. 104, pp. 842, 2007.
- 4) Sheng Lin-Gibson, Howard J. Walls, Scott B. Kennedy, and Eric R. Welsh. Chitosan Hydrogels: Crosslink Kinetics and Gel Properties. *Journal of Polymeric Materials: Science & Engineering*. 2003, Vol.88, PP 199-200.
- 5) Sushmit S. Rithe, Pravin G. Kadam, and Shashank T. Mhaske. preparation and analysis of novel hydrogels prepared from the blend of guar gum and chitosan: cross-linked with glutaraldehyde. *Journal of Advances in Materials Science and Engineering*. December 2014, Vol.1, No.2.
- 6) Zhang Y, Zhu W, Ding J. Preparation of thermosensitive microgels via suspension polymerization using different temperature protocols. *J. Biomed. Mater. Res. Part A*. 2005;75A:342–349
- 7) C.Yuen, S. Ku, P. Choi, C. Kan, S. Tsang. Determining Functional Groups of Commercially Available Ink-Jet Printing Reactive Dyes Using Infrared Spectroscopy. *Research Journal of Textile and Apparel*. 2005, Vol.9, No.2, PP 26–38. Doi: 10.1108/rjta-09-02-2005-b004.
- 8) Ragaa EL-sheikh, Moustafa M. S. Abo-ELfadi, Mohamed E. A. Ali, Khalid Khader, Impact of sludge produced from drinking water plants on groundwater and its treatment by a natural polymer. *Journal of American Science*. 2017, Vol.13, No.2.

Prediction Of Titanic Survival Using Machine Learning Techniques

¹Mr.K.Vamshee Krishna, ²Jampally Chandu Kumar Yadav, ³Boddu Dayakar, ⁴ Palagiri Anu Priyanka, ⁵paramshetty Satya Sai

¹ Assistant Professor, Dept. Of CSE, Samskruti College of Engineering & Technology, TS.

^{2,3,4,5}B. Tech Student, Dept. Of CSE, Samskruti College of Engineering & Technology, TS.

Abstract: *It can be very important to uncover the root causes of past human suffering so that future crises can be eliminated. The incident of April 15, 1912 is an example of a human tragedy in which approximately 1,500 passengers and crew members lost their lives. Today, continuing research shows that if the right steps are taken, it is possible to reduce human harm. Nowadays there are many new and effective technologies, with the help of data analysis the truth can be established. In these studies, reviewed, Titanic survivors were studied primarily as a tool to learn techniques. In monitoring, out of all the organizations, 891 organizations were used for learning and 418 organizations were used for testing and comparisons were examined between different learning systems, which gave importance to this research.*

Keywords—Machine Learning, Prediction, Pattern Recognition, Statistical Analysis.

I INTRODUCTION

Machine learning, a great sub field of artificial intelligence, evaluates important tasks that include prediction using the truth or, it can be said, specifically using mathematics to discover hidden patterns in data. It is important to note that where the traditional methods require appropriate methods, machine learning presents better decisions in the output.[1] Look at 1.

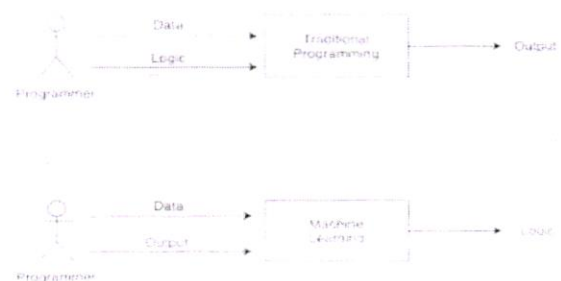


Figure 1: Difference between Traditional Programming and Machine Learning

Fig1: In the traditional system, the programmer feeds this system with the facts, logic and results for the program. But,

Samskruti College of Engg. & Technology
Kondapur (V), Ghatkesar (M), Medchal Dist.

with the knowledge acquired by the equipment, the programmer provides the version with data and output, and the model then generates logic or software. In this research work, a

The Titanic data set is used to evaluate the survival of the Titanic based on all observational analyzes of observational learning techniques such as logistic regression, random forest, decision making, K - nearest neighbor, etc.

According to the information, the aliens are the ship's gift, so all the predictions are passed on by the lovers so that it can maximize its survival. Before starting, to train the model, it is important to first record the data in all aspects such as missing values, similar formatting, outliers, etc. [2] [3]. For more information, a clear picture of how to make a cartoon is given in parent 2.

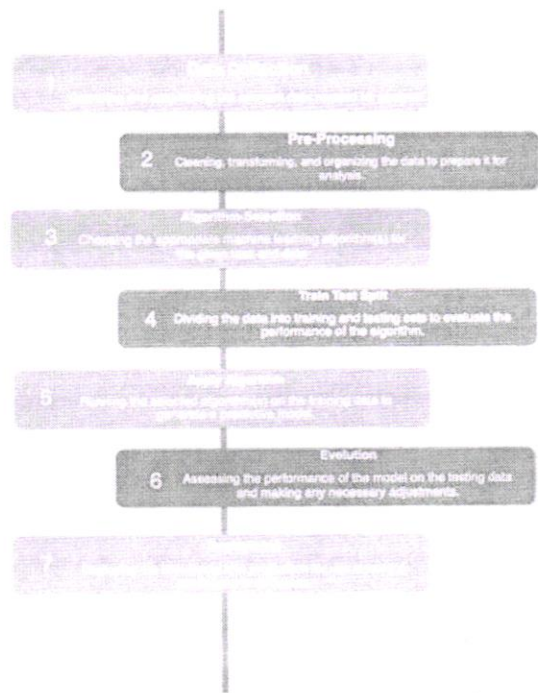


Figure 2: Workflow of Data analysis

Figure 2 describes the process of data analysis and how the author arrived at the belief. It is best to calculate all statistical methods properly

so it is easy to know the truth of the real way.

The purpose of this paper is to explore and analyze the Titanic data using learning tools to predict passenger survival. Special articles in the value of monitoring acquisition methods such as logistic regression, random forest, stochastic gradient descent, decision tree and nearest neighbor Divide the passengers into groups, survivors or not. The purpose of the evaluation is to evaluate the overall performance of the algorithm based on specific evaluation measures, including

Principal
 Samskruti College of Engg. & Technology
 Yashpur (V), Ghalesar (M), Medchal Dist.

accuracy, F1 score, recall, and accuracy. In addition, the article discusses data preparation techniques, feature engineering, and data visualization to better understand data. The results obtained from this study can be useful for the development of prevention methods and emergency strategies of maritime transport in the future.

II LITERATURE REVIEW

The titanic DATA-SET has been widely used in many studies to explore modeling strategies, task selection strategies, and machine learning algorithms. The DATA SET served as a benchmark for competition and training, and its evaluation confirmed the importance of predicting survival. The studies reviewed as part of this data assessment demonstrate the simplicity and importance of the titanic data set in advancing reconnaissance and predictive modeling. From the study of data mining techniques to the study of the sigma trait Parameters in synthetic neural networks

Table 1 Summarized view of Literature Review

Ref.	Year	Method Used	Assessment
1	2019	Data Set Provider	Kaggle.com provides the Titanic dataset and platform for the Machine Learning from Disaster competition, which serves as a popular benchmark for predictive modeling.
2	2013	Data Mining	This paper provides a comprehensive survey of data mining techniques, including supervised and unsupervised learning, and their applications in various fields.
3	2007	Feature Selection	The paper proposes a spectral feature selection method for both supervised and unsupervised learning tasks, which can improve the accuracy and efficiency of predictive modeling.
4	2018	Predictive Modeling	This study uses the Titanic dataset to predict the survivors of the disaster and compares the performance of various machine learning algorithms.
5	2012	Predictive Modeling	It proposed a predictive modelling approach using the Titanic data set and offers a comprehensive review of related concepts and methods.
6	2017	Predictive Modeling	This GitHub repository contains a predictive model for the Jack Dies competition, which is based on the Titanic data set and serves as a similar benchmark for machine learning.
7	2017	Predictive Modeling	This study uses various machine learning algorithms to analyze the Titanic disaster and identifies the most important factors for survival prediction.
8	1993	Neural Networks	The paper investigates the impact of sigma function parameters on the backpropagation learning algorithm in artificial neural networks.
9	2009	Decision Trees	This paper presents an implementation of the ID3 decision tree learning algorithm and provides a tutorial on how to apply it to predictive modeling tasks.
10	2018	Predictive Modeling	This study compares the performance of different machine learning techniques on the Titanic data set and identifies the most accurate

III DESCRIPTION OF DATA AND EXPERIMENTAL SET UP

The data series consists of two columns and eleven columns. Parch ticket, P-

Principal
Samskruti College of Engg. & Technology
Kondapur (V), Ghatkesar (M), Maharashtra 4254

elegance, call, passenger ID, life expectancy, gender, age, Sib Sp, price,

cabin, and onboard are lines. Besides ability, our goal is survival. The data thus categorizes certain family ties. Blood brothers, followers, spouses and children (wives and fiances are not considered spouses). This is how the data set defines family relationships. Mother and father are parents. A child refers to a son, daughter, granddaughter, or half-brother. Because some young people are better off traveling with a babysitter, the value of their parchment has gone down.

Before any form of analysis of the data is done, the author

should smooth the data set. Some important missing points are also found in the literature and should be addressed. Missing values in the variables such as Embarked, Cabin and Age are filled with the option selected by the main age. In this state, the Cabin column is deleted and updated with the value type of the Onboard Exclusion Value column. Along the middle line, fill in the missing value in the age column.

Research and evaluation of data

First, we can do an analysis of the information about our problem. The data were analyzed by clinical analysis of the data to identify the factors that affect the survival rate. By correlating each behavior with survival, the data is analyzed carefully. Figure 3 shows how sex affects survival rates.

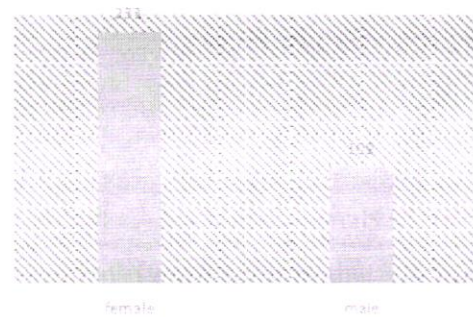


Figure 3: SUM of Survived by Sex

Figure 3 shows that women have a better survival rate than older men, as illustrated in Figure 6. The survival rates of women and men were calculated as 74.20382% and 18.89081%, respectively. Other attributes, including car price, cabin, title, family, P-perfect, auto, and survival, have a similar relationship. The name was created using the call property.' Sibbs and Parch were united. We are able to determine the importance of each character in the passenger's survival on this journey.

IV MACHINE LEARNING MODELS

Many learning algorithms are used to be accurate and help to survive.

Principal
Sanskriti College of Engg. & Technology
Kondapur (V), Ghatkesar (M), Medchal Dist.

a. logistic regression is a version of the distribution in place of a version of regression and is a simple distribution model that produces very good results with linearly separated classes [4] [5]. we use the logistic regression version here because our results are both survival and non-survival. logistic regression is a reliable and useful method for binary and linear class problems.

b.random forest is the owner manager. it can be used in learning gadgets to get answers in regression and distribution mode. "a random forest is a distributed system that has multiple tree selections of specific properties of the given data and uses the average to improve the expected results of that data" [6] according to as the name suggests. as part of the main data, instead of relying on a single selected tree, random forest collects the predictions from each tree and predicts the final result using that prediction. which received the most votes.

c. stochastic gradient descent (sgd), the author uses massive data and a method that optimizes the descent for the duration of each search as quickly as possible when selecting the weight vector. gradual descent is a method of searching in continuous or infinite space where 1) the assumptions are constant and 2) the error

varies depending on the parameters. the weights are initialized in sgd from the given data (titanic DATA-SET) and the code modifies the load vector with a statistic. when a calculation error is made, gradient descent gradually adjusts it to improve convergence.

d. decision tree is a method to obtain supervised knowledge that can be used to solve problems of type and regression such as titanic datasets, however, it is mainly used to solving classroom problems [7]. it has a tree structure, with internals for data set locations, branches for input options, and leaves for results. the order node and the leaf node are 2 nodes that form a selection tree. while leaf nodes are the result of such choices and do not have branches, decision nodes are used as choices and include more branches. the examination or assessment is usually based on the characteristics of the given data.

e. k-neighbors neighbors (knn) is a monitoring system that can be used for all types and returns. by calculating the distance between the test data and all the training points present in the big data, knn tries to predict the best classes for the test data. then determine where k is the maximum, just like the test [8] [9]. the knn algorithm determines which classes have the best results by counting the number of

times that control data is available for each "k" statistical class. the value in the regression condition is the average of the "k" decisions of the training program.

V RESULT

The first step in engaging in a survey is data collection. Exploratory statistics analysis makes understanding data and relationships between capabilities less problematic. Use various graphic techniques. A reference above uses histograms and ggplot. Some conclusions were drawn and data was discovered using a case study. Based on the research data analysis process, the need to construct schools and forecast versions is recognized in engineering works. The mastery of modes by machines presupposes the good quality of passengers who survive. To make predictions in class problems, the Random Forest method is used. With a precision of zero.827261504, a return of 0.813453456, an F1 score of zero.8237261504, and a precision of 0.827261504 in line with the confusion matrix, Random Forest appears to be the true version. This shows that Random Forest has an overall overestimation of the prediction skill in this data set using the selection function. For a detailed

picture of the statistical analysis, see Table 2.

Table 2. Performance Matrix Representation

Algorithm	Accuracy	F1 Score	Recall	Precision
Logistic Regression	78%	0.78	0.78	0.79
Random Forest	82%	0.82	0.81	0.82
Stochastic Gradient Descent	58%	0.45	0.58	0.63
Decision Tree	79%	0.79	0.79	0.79
K-nearest neighbor	66%	0.64	0.66	0.67

It is very obvious that when using a specific design process, the accuracy of the model can also be affected. The perfect models for this type of problem are Random Forest and Decision Tree because they provide a high level of accuracy. The results of our experiment, as evidenced in Figure 4, show the performance of many machine learning algorithms used to estimate the survival of the Titanic. We evaluate the performance of the algorithms using accuracy, F1 score, recall, and precision. The Random Forest algorithm came out on top with an accuracy of 82%, an F1 index of zero.82, a return of 0.8 one, and a precision of 0.82. The logistic regression and decision tree algorithms also performed well with 78% and 79% accuracy, respectively. However, the stochastic gradient descent algorithm guarantees a perfect result with an accuracy of only 58%. The K-nearest neighbor rule performed slightly better

Sanskriti College of Engg. & Technology, Kondapur (V), Ghatkesar (M), Medchal Dist.

with sixty-six percent accuracy. These results suggest that Random Forest rules are the most suitable for predicting the survival of Titanic passengers using learning techniques.

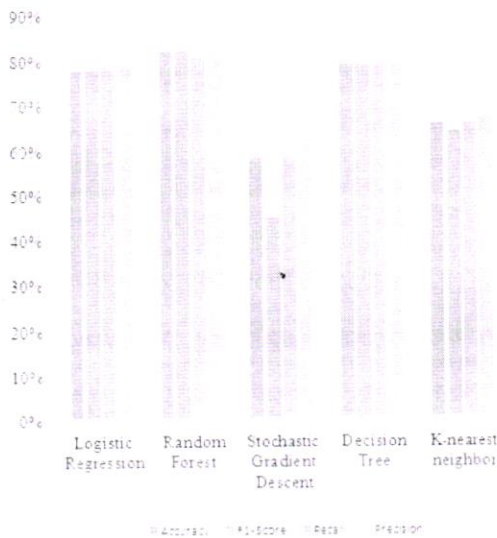


Figure 3: Performance Measures

Fig3: Display the results of the algorithms. This graph demonstrates the algorithm's performance in relation to accuracy and other factors.

VI CONCLUSION

The design using the system brought awareness to the perceived value of passengers who survived. The random forest method is used to make predictions about target species. The accuracy of each model is determined using the confusion matrix, and the Random Forest model comes out on top with an accuracy of zero.82. This shows that the Random

Woodland estimator performs in this data with very good selection ability. It is clear that when using the best modeling methods, the accuracy of the models can also change. The modes that provide the best level of accuracy for classification problems are random forest modes. Machine learning and statistical analysis were used in this work. This diagram can be used as a template for learning how to integrate EDA and core knowledge tools. With the use of more bookstore libraries, notably Vibrant in R, the concept can be improved in the future to create more graphical user interfaces. It is necessary to create interactive pages, where the same values as the chart attribute (like plot or histogram) will also change if the attribute value is changed at length. By combining our effects, we can obtain more comprehensive conclusions.

REFERENCES

1. Kaggle.com. (n.d.). Titanic: Machine Learning for Disaster. Retrieved October 29, 2019, from <http://www.kaggle.com/>
2. Jain, N., & Srivastava, V. (2013). Data mining techniques: A survey paper. IJRET: International Journal of Research in Engineering and Technology, 2(11), 2319-116

Principal
Samskruti College of Engg. & Technology
Kondapur (V), Ghatkesar (M), Medchal Dist.