

**DISCOVER THE WORLD  
OF AUTOMATION**

**ACM TCET'S**

# **EZINE**

**2022 14<sup>th</sup> EDITION  
VOLUME 1**

**FIND INSIDE**

**DEEP DIVE IN  
MACHINE LEARNING**



**MEDDLE WITH APP  
DEVELOPMENT**



# GRADUATE ATTRIBUTES

1

**ENGINEERING KNOWLEDGE:** Apply Knowledge of Mathematics, Science, engineering fundamentals and an engineering specialization to the solution of complex engineering problems.

2

**PROBLEM ANALYSIS:** Identify, formulate, research literature and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences and engineering sciences.

3

**DESIGN / DEVELOPMENT OF SOLUTIONS:** Design solutions for complex engineering problems and design system components or processes that meet specified needs with appropriate consideration for public health and safety, cultural, societal and environmental considerations.

4

**CONDUCT INVESTIGATIONS OF COMPLEX PROBLEMS:** Using research based knowledge and research methods including design of experiments, analysis and interpretation of data and synthesis of information to provide valid conclusions.





# GRADUATE ATTRIBUTES

5

**MODERN TOOL USAGE:** Create, select and apply appropriate techniques, resources and modern engineering and IT tools including prediction and modelling to complex engineering activities with an understanding of the limitations.

6

**THE ENGINEER AND SOCIETY:** Apply reasoning informed by contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to professional engineering practice.

7

**ENVIRONMENT AND SUSTAINABILITY:** Understand the impact of professional engineering solutions in societal and environmental context and demonstrate knowledge of and need for sustainable development.

8

**ETHICS:** Apply ethical principles and commit to professional ethics and responsibilities and norms of engineering practice.



# GRADUATE ATTRIBUTES

9

**INDIVIDUAL AND TEAM WORK:** Function effectively as an individual, and as a member or leader in diverse teams and in multi-disciplinary settings.

10

**COMMUNICATION:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as being able to comprehend and write effective reports and design documentation, make effective presentations and give and receive clear instructions.

11

**LIFE-LONG LEARNING:** Recognize the need for and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

12

**PROJECT MANAGEMENT & FINANCE:** Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects in multidisciplinary environments.





# PROGRAM SPECIFIC OUTCOMES

---

**PS0-01**

To develop the culture of augmenting existing technologies to create scalable IT solutions.

**PS0-02**

To combine various technologies like IoT, Cloud and Analytics to provide integrated solutions to real time problems of government or industries.

**PS0-03**

To master in moulding any problem into a web or internet based solutions.



# PROGRAM EDUCATION OUTCOMES

## PEO-01

To prepare learners with a strong foundation in the area of information technology required to solve real life problems arising through software technology. [Knowledge]

## PEO-02

To prepare learners to be knowledgeable of the ethics, professionalism and cultural diversity in the work environment to meet applicable standards with continued motivation for research and development. [Skill & Professionalism]

## PEO-03

To prepare learners to understand the need for lifelong learning with effective written and oral communication skills and to be able to readily adapt to new software engineering environments. [Attitude, Presentation and Growth]





# INDEX

**1 ABOUT THE DEPARTMENT 10**

**2 FOREWORD 11**

## ARTICLES

**3 SENTIMENT ANALYSIS FOR CODE-  
MIXED LANGUAGES USING DEEP  
LEARNING 17**

**4 FERTILIZER PREDICTION AND COST  
ANALYSIS 20**

**5 SARCASM DETECTION FOR CODE-MIX  
LANGUAGE USING MACHINE LEARNING  
AND DEEP LEARNING 25**

**6 LACK OF STRAY SERVICES IN INDIA 31**

**7 ROUTE OPTIMIZATION FOR  
AMBULANCE 37**

# INDEX

8	AUTOMATED KYC	45
9	HEALTHTECH THE INNOVATORS IN HEALTHCARE	58
10	CUCUMBER (BDD FRAMEWORK)	63
11	DESIGN & DEVELOP APP FOR NUTRITIONAL SECURITY & ELIMINATING ANAEMIA	66
12	PREDICTION OF ANEMIA DISEASE USING MACHINE LEARNING ALGORITHMS	72
13	HIGH-RESOLUTION IMAGE SYNTHESIS WITH LATENT STABLE DIFFUSION MODELS	78
14	POSTERS	96



# INDEX

---

15

POEMS

106

16

EZINE COMMITTEE

109

18

CODE OF ETHICS

111

# DEPARTMENT OF INFORMATION TECHNOLOGY

## ABOUT THE DEPT:

The department of Information Technology, started its journey in the year 2002 and is committed to deliver the program with rigor and with active industry participation. Department has 120 seats intake at first year and 24 seats as lateral entry at 2nd year for engineering diploma students. The department believes in student centric approach. Its dedicated team of faculty members inculcate relevant knowledge, skills and attitude in students to become successful professionals. The U.G. programme is accredited by National Board of Accreditation (NBA), New Delhi for three years w.e.f. 16.09.2011. UG Programme has been re-accredited for 3 years b by NBA w.e.f 1st July 2016. Also t the programme is permanently affiliated with UOM since AY 2015-16 onwards.

## MISSION:

The IT department is committed to enrich students by rigorously implementing quality education with focus to make them industry ready, while imbibing in them professional ethics and social values to become responsible citizens.

## VISION

"The department of IT will strive to be at the top position among the renowned providers of IT education".



# FOREWORD



## **Dr. KAMAL SHAH**

**VICE-PRINCIPAL  
DEAN , R & D CELL**

In today's fast-paced world, rekindling the flame of invention and encouraging curiosity in young brains is critical. The Department of Information Technology publishes Ezine, which attempts to incorporate student ideas and encourage active engagement in the learning process. Ezine has established a significant benchmark in showcasing students' hidden inherent talent by providing them with an unrivalled opportunity and an excellent platform to not only express their ideas and creative potentials, but also to voice out their personal opinions on topics that are of utmost importance in the lives of students. A departmental magazine is designed to not only disseminate information, but also to introduce a whole new intriguing and thrilling arena of content in which students can explore their hobbies and feed their curiosities. Unlike previous technical journals, Ezine has expanded beyond science and technology to include other important fields, giving students the opportunity to investigate inter-disciplinary aspects of themes and to excite their natural curiosity. The Editorial Committee has made excellent use of the platform offered to them in harnessing the talents of all of the energetic students. I want to express my heartfelt gratitude to the entire Editorial Board for bringing us this much anticipated college magazine, which meets not only the stringent standards of punctuality but also curates information of the highest quality





## **Dr. Sangeeta Vhatkar**

**Associate Professor, I/C. HOD-IT  
ACM Branch Counsellor.**

**"All of us do not have equal talent. But, all of us have an equal opportunity to develop our talents" - APJ Abdul Kalam**

E-ZINE Magazine Published by Department of Information Technology. This is not a just Technical magazine but it speaks about overall growth of student personality. E-zine magazine published annually. This magazine has a great educative value. This help in encouraging the students to think and write and thus help them in developing their writing skills and talent. Magazine also help them in developing their power of thinking and strengthen their imagination as well.

This year E-zine magazine focus on Current Trends and technology like Cybersecurity, AI&ML. Apart from the department magazine E-zine we are also publishing Newsletter describing the events that occurred this semester and a Bulletin highlighting the achievements of students and faculty members. This E-Zine magazine also Motivate Non-Technical Articles, Sketches, Poems etc.

My Heartiest Congratulations to TCET-ACM Publication Head and Congratulations to entire editorial team for creative work. I hope that E-zine magazine will Provide platform for Overall development of Stakeholders.

I am thankful to Management of Thakur College of Engineering & Technology for providing State-of-Art Infrastructure and all Possible Support in caring out multidimensional activities and Event. I am also thankful to our Principal, Dr. B.K. Mishra, our Vice-Principal, Dr. Kamal Shah, for encouraging us and providing us with a fabulous platform like E-zine 2022 to express our ideas and thoughts.



## **Dr. Aaditya Desai**

**Faculty in-charge of E-zine  
Magazine 2022**

**“If you don’t like to read, you haven’t found the right book.”  
J.K. Rowling (Joanne Rowling) CH, OBE (British author and  
philanthropist – Harry Potter Fame)**

Dear readers,

Greetings of the season!

We welcome you to the 14th Edition of E-zine Magazine of Information Technology. This edition is special in the sense that it not only includes technical articles but also it includes poems written by the faculty.

The student community has showcased their creative side through technical posters. Most of these posters are not only informative but also are beautifully designed by the students.

There are contributions from our alumni and industry expert. All in all, this edition of E-zine is an interesting one, with lots of material to read and learn.

We thank the management of Thakur College of Engineering and Technology, the Principal of TCET, Dr. B.K. Mishra, Vice-Principal, Dr. Kamal Shah, I/C HOD-IT, Dr. Sangeeta Vhatkar and the entire publication committee of Information Technology Department.

Keep reading and all the best to you!





# **KRISHI CHAURASIYA**

## **ACM PUBLICATION HEAD**

Ezine's mission is to provide a platform for students, researchers, academicians, parents and other to share, showcase and exchange profound ideas and knowledge about technology, research, innovation and development. We are back with another new edition!

The 14th edition, Volume 1 has arrived!!

Ezine is published by the department of Information Technology, which incorporates student ideas and encourages active learning by participating in the process.

It not only provides information but also introduces a lot of new exciting content for students to satiate their curiosity.

Ezine is also an important platform for students to voice out their opinion on topics that are of utmost importance to the growth and learning of students. You'll be provided insights on various technical aspects along with a few additional non-technical works showcasing literary talents of our students.

I'd like to express my utmost gratitude to the Management of TCET, our HOD Dr. Sangeeta Vhatkar and faculty for providing us an transparent platform like Ezine to open up and express ourselves to the people.

I'd like to thank the authors for their time and articles we used to provide you all with plethora of bountiful knowledge.

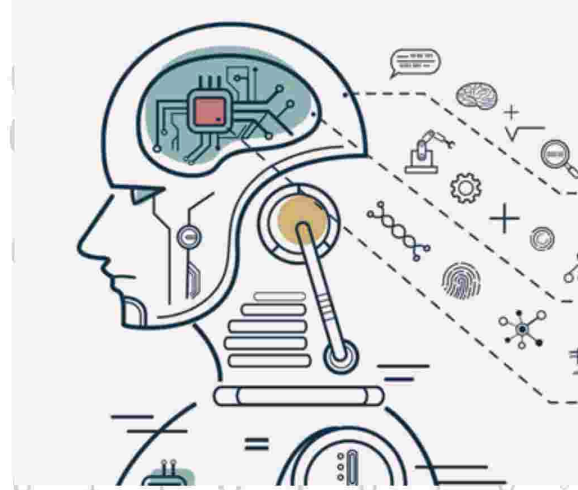
I'd also love to thank our Ezine Creative team who have worked hard for this magazine and completed the work with utmost precision and care.

Pleasant reading!

# ARTICLES



# SENTIMENT ANALYSIS FOR CODE-MIXED LANGUAGES USING DEEP LEARNING



**Mr. Preet Desai**  
Information Technology  
Thakur College of Engineering  
and  
Technology Mumbai, India  
desaipreetforever@gmail.com

**Mr. Ronak Dhakad**  
Information Technology  
Thakur College of Engineering  
and  
Technology  
Mumbai, India  
dhakadronak9967@gmail.com

**Mr. Somil Doshi**  
Information Technology  
Thakur College of Engineering  
and  
Technology  
Mumbai, India  
somil12doshi@gmail.com

## Abstract—

Due of how fundamentally different its properties are from the conventional structures of standard languages; code-mixed data poses a significant barrier to natural language processing. In this study, we present a unique contrastive learning-based method for categorizing sentences into those with positive, negative, or neutral connotations. This method is known as Sentiment Analysis of Code-Mixed Text (SACMT). To map the sentences of code-mixed and standard languages to a single sentiment space, we use the shared parameters of Siamese networks. We also offer a straightforward clustering-based preprocessing technique to identify word variants using code-mixed transliteration.

**Keywords—** Sentiment Analysis, Siamese Networks, Code- Mixed Text

By reducing the energy function connecting the twin networks, SACMT learns the shared model parameters and the similarity measure. If the sentiment of phrases on both individual Bi-LSTM networks is the same, parameter sharing and the Similarity Metric ensure that they are closer to each other in the sentiment space, otherwise, they are farther apart.

India match jit gayi (India won the match) and Diwali ke shubh kamnaye sabko (Happy Diwali to everyone) are thus seen as being closer to one another, while India match jit gayi (India won the match) and Bhai ki movie flop gayi (Bhai's movie was a fiasco) are depicted as being farther apart. The sentiment similarity of sentences is modeled into a shared sentiment space by the learned similarity metric. Numerous spellings of the same word are produced when phonetic languages like Hindi are transliterated into a roman script. As an illustration, "(more) can be written as *bahut,bohot, or bohot*. To overcome this difficulty, we implement a preprocessing phase that seeks to group various word variations using an empirical similarity measure. The remainder of the essay is structured as follows. The prior methods in the field are described in Section 2. The datasets are demonstrated in S section 3.

Identify the applicable funding agency here. If none, delete this text box.

XXX-X-XXXX-XXXX-X/XX/\$XX.00 ©20XX IEEE

Explaining SACMT's architecture is Section 4. The baselines are defined in Section 5. The experimental design and findings are presented in sections 6 and 7, respectively. Finally, Section 8 brings the essay to a close.

## I. INTRODUCTION

Social media platforms were widely accepted in multilingual societies with a respectable internet penetration rate. The use of code-mixed text increased as a result. Political campaigns, product reviews, marketing, and other social phenomena can all be closely examined using sentiment analysis of code-mixed data from social media platforms. Text that has been code-mixed frequently develops new structures dependent on its users and adopts the vocabulary and grammar of several different languages. Since standard approaches to semantic analysis do not capture the meaning of the sentences, this makes sentiment analysis difficult. The lack of annotated data accessible for sentiment analysis further hinders the field's development.



## II. RELATED WORK

A. The semantics of the words are captured by the distributional semantics technique, but the information about their order in the sentence is lost. The approach also believes a word to be unchangeable, which is one of its limitations. As a result, it cannot handle spelling mistakes or words that are not in the lexicon. polarity scores are given to specific words. The polarity of the sentence is determined by the sentiment score of all the individual words. As a result, the words' order and meaning relationship are lost, which causes inaccurate classification. N-grams restrict this issue but do not entirely cure it.

B. Another line of research uses character level LSTMs to extract information from social media text at the subword level. Using an annotated corpus and this information, the sentences are then classified. The model offers a useful strategy for embedding sentences. Abbreviations and Acronyms Siamese Networks .

Without being overly reliant on the input's attributes, siamese networks aid in the contrastive learning of a similarity metric. siamese networks were developed to address the issue of signature verification. Units

- Later, the discriminative loss function architecture was employed for face verification. Additionally, these networks successfully raise the standard of visual search. These networks have recently been used to address the issue of community question answering.

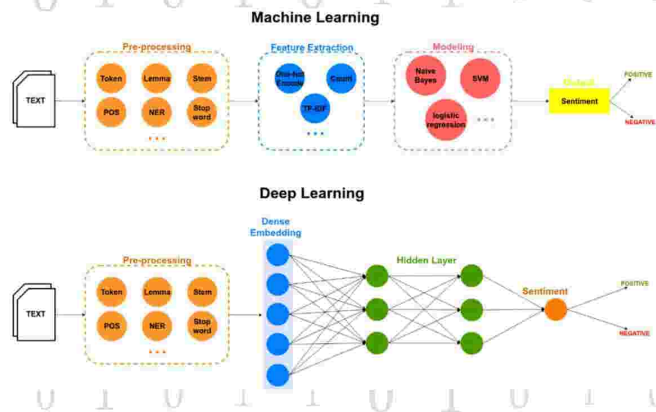
Let the family of functions with parameters,  $F(X)$ , be  $W$ . With regard to  $W$ ,  $F(X)$  is differentiable. The Siamese network seeks a value for the parameter such that, if  $X_1$  and  $X_2$  are members of the same category, the symmetric similarity metric is small, and if they are members of different categories, it is big. The scalar energy function  $S(C, R)$  can be written as  $\|F(C)F(R)\|$  to represent the sentiments relatedness between tweets of code-mixed (C) text and resource-rich (R) language.

### c. dataset

The equations are an exception to the prescribed specifications of this template. You will need to determine We use the datasets to evaluate the architecture on codemixed (Hindi-English) data as well as text from social media in a standard language (English). The datasets we used for our tests are listed below.

3879 annotated Hindi-English Code-Mixed utterances make up the dataset, which was proposed in.

103035 annotated English tweets make up the dataset, which was proposed in.



- SemEval 2013: The dataset for Task 2B1 of SemEval 2013 is made up of

### Architecture of SACMT

A completely connected layer is placed on top of a siamese network using twin character level Bi-LSTM networks. Sentences are projected on both ends of a common sentiment space by bi-LSTMs. We link the sentiment vectors that are produced to a layer that computes their similarity. The label and similarity measure are combined in the contrastive loss function. The gradient of the loss function is computed with respect to the weights and biases shared by the sub-networks via back-propagation across time.

### Handling Code-Mixed Word Variations

Transliteration from languages with phonetic script (like Hindi) leads to variation in word depending on the user. We solve this issue using clustering of skip-gram vectors [10]. Skip-gram vectors give the representation of a word in the semantic space based on their context. The variations belong to the same word with similar function implying a similar context. Also, the consonants of these variations in the cases are same (shown in table 2). Hence, we cluster the words based on a similarity metric that captures both these properties. The similarity metric is formally defined below:

$$f(v_1, v_2) = \{ \text{sim}(\text{vec}(v_1), \text{vec}(v_2)) \text{ if } v_1, v_2 \text{ have same consonants}$$

where  $f(v_1, v_2)$  denotes the overall similarity between  $v_1$  and  $v_2$ ,  $v_1$  and  $v_2$  are the two variations,  $\text{sim}$  is a similarity function (like cosine similarity),  $\text{vec}(v)$  yields the skip-gram vector of  $v$ , and  $v$  is the variation.

The closest word variations for the given term are provided by this metric. They collectively form a cluster, and the most frequent word in the cluster takes the place of all the others. Here, we make the supposition that the term that occurs the most frequently also has the greatest likelihood of being the right one.

### Primary Embeddings

Due to its informal nature, code-mixed text can encounter issues like spelling mistakes and vocabulary word shortages. These variances, which express passion, cannot be brushed aside as mistakes.



For instance, "Heey" expresses a good emotion whereas "Hey" is neutral. So, instead of treating words as immutable units, we treat character trigrams as such. The computational complexity is also reduced because there are more words than character trigrams. We provide the SACMT twin networks with two character-based term vectors representing the tweet and a label. The label specifies how close or how far apart the samples are in the sentiment space. Twin networks are fed with the tweets' vectors with the same information for positive samples (closer in the sentiment space).

The abbreviation "i.e." means "that is", and the abbreviation "e.g." means "for example".

An excellent style manual for science writers is [7].

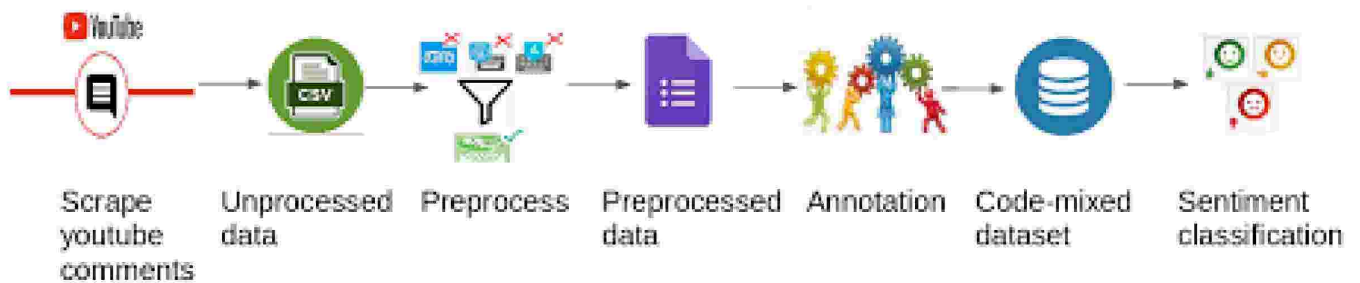
### III. USING THE TEMPLATE

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

### Conclusion -

In this study, we present SACMT, which solves the issue by projecting the sentences into a shared sentiment space, for sentiment analysis of code-mixed text. For the purpose of capturing a sentiment-based representation of the sentences, SACMT uses twin Bidirectional LSTM networks with shared parameters. Sentences with similar sentiment were grouped together using these sentimentbased representations and a similarity metric. Studies done on the datasets show that SACMT performs noticeably better than the cutting-edge techniques. SACMT makes use of resources from various languages to enhance sentiment analysis on material that has been code-mixed.

The management of word variants further improved the performance of all trained models, even baseline models. When used with SACMT, an emoji-based strategy improves the performance of overall sentiment.



A. An exercise in classification is being conducted. We take into account the Hindi-English CodeMixed (HECM) sentences and designate them 1 after aligning them with the English sentences from the Twitter datasets of the same mood (positive samples). Additionally, we randomly select an equal number of English lines with various moods (negative samples), assigning them the number -1. We examine the benefits of training HindiEnglish CodeMixed data alongside English phrases using this model (SACMT(HE-Eng)).

A. Additionally, we create the input data by labelling each HECM sentence 1 after aligning it with corresponding HECM sentences of the same emotion (positive samples). The same procedure is used to randomly select an equal number of HECM statements with various sentiments, labelling them as "-1."



# FERTILIZER PREDICTION AND COST ANALYSIS

## • INTRODUCTION:

**A**griculture is the backbone of Indian Economy. Plants require nutrients in order to grow and give more yields. Nutrients to the plants are provided through soil, manure and the fertilizers. There are 16 essential nutrients that are required for the plant growth viz., carbon, hydrogen, oxygen, nitrogen, phosphorous, potassium, Sulphur, calcium, magnesium, boron, iron, manganese, molybdenum, zinc, chlorine and copper.

These are either provided through soil or water in the form of micro-nutrients and macro-nutrients. The deficiency of these nutrients can be found out by using soil testing mechanism. There have been advancements in Science and Technology in India from past few decades. But the productivity of farming is declining and there are several reasons behind it: fragmented land holding illiteracy of farmer, lack of decision making in choosing good seed, manure and fertilizer. Several studies show that farmers have little knowledge over usage of optimum level of fertilizer for optimal crop yield. The plants require fertilizer when they are deficient of nutrients. Also, the usage of fertilizer more than the requisite level produces crops with increased fertilizer content which upon consumption leads to several harmful diseases. When fertilizers are added to the soil it decreases the microbial activity of soil. The factors affecting the fertilizer consumption are: extent of available land, cultivated land, rainfall, cropping pattern, soil characteristics. This proposed system takes soil data such as nitrogen level, phosphorous level, Potassium level, soil type, temperature, moisture, crop types as input and suggests you the best fertilizer which should use. It calculates price of fertilizer, water cost, amount of fertilizer required, and cost of other necessary instruments in the process and gives the total cost of fertilizing the land. If anyone was unaware about using new fertilizer it will show them ways and appropriate condition and time to use it.

## • ABSTRACT:

India is the second largest producer of wheat and rice, the world's major food staples. India is currently the world's second largest producer of several dry fruits, agriculture-based textile raw materials, roots and tuber crops, pulses, farmed fish, eggs, coconut, sugarcane and numerous vegetables. In order to produce more agricultural products good quality fertilizers should be used. But this fertilizer should be chosen by analyzing soil quality, amounts of minerals like nitrogen, phosphorous, potassium should take into consideration. Weather conditions, Geographical location also matters in such case. Farmer should know amount and cost of fertilizer need to be used.

## • RELATED WORK:

[1] Shivnath Ghosh, et al. (2014) In this paper machine learning system is divided into three steps, first sampling (Different soil with same number of properties with different parameters) second Back Propagation Algorithm and third Weight updating.

[2] P.Vinciya, et al. (2016) This paper mainly focused on analyzing the agriculture analysis of organic farming and inorganic farming, time cultivation of the plant, profit and loss of the data and analyzes the real estate business land in a specific area.

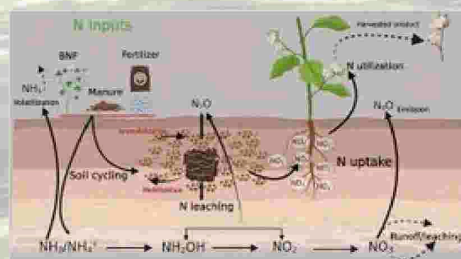




[3] Vaneesbeer Singh, et al. (2017) This work presents an approach which uses different Machine Learning techniques in order to predict the category of the yield based on macro-nutrients and micro-nutrients status in dataset. The dataset considered for the crop yield prediction was obtained from Krishi Bhawan (Talab-Tillo) Jammu. The parameters present in the data are Macro-Nutrients (ph, Oc, Ec, N, P, K, S) and Micro Nutrients (Zn, Fe, Mn, Cu) present in samples collected from different regions of Jammu District. After analysis Machine learning algorithms are applied to predict the category of yield. The category, thus predicted will specify the yield of crops. The problem of predicting the crop yield is formulated as Classification where different classifier algorithms are used.

[4] E. Manjula et al. (2017) This paper chooses Nitrogen, Phosphorus, Potassium, Calcium, Magnesium, Sulphur, Iron, Zinc, and so forth, nutrients for investigating the soil supplements utilizing Naïve Bayes, Decision Tree and hybrid approach of Naïve Bayes and Decision Tree. The performance of the classification algorithms are compared based on accuracy and execution time.

[5] Rohit Kumar Rajak et al. (2017) This method is characterized by a soil database collected from the farm, crop provided by agriculture experts, achievement of parameters such as soil through soil testing lab dataset. The data from soil testing lab dataset given to recommendation system it will use the collected data and do ensemble model with majority voting technique using support vector machine and ANN as learners to recommend a crop for a site-specific parameter with high accuracy and efficiency.



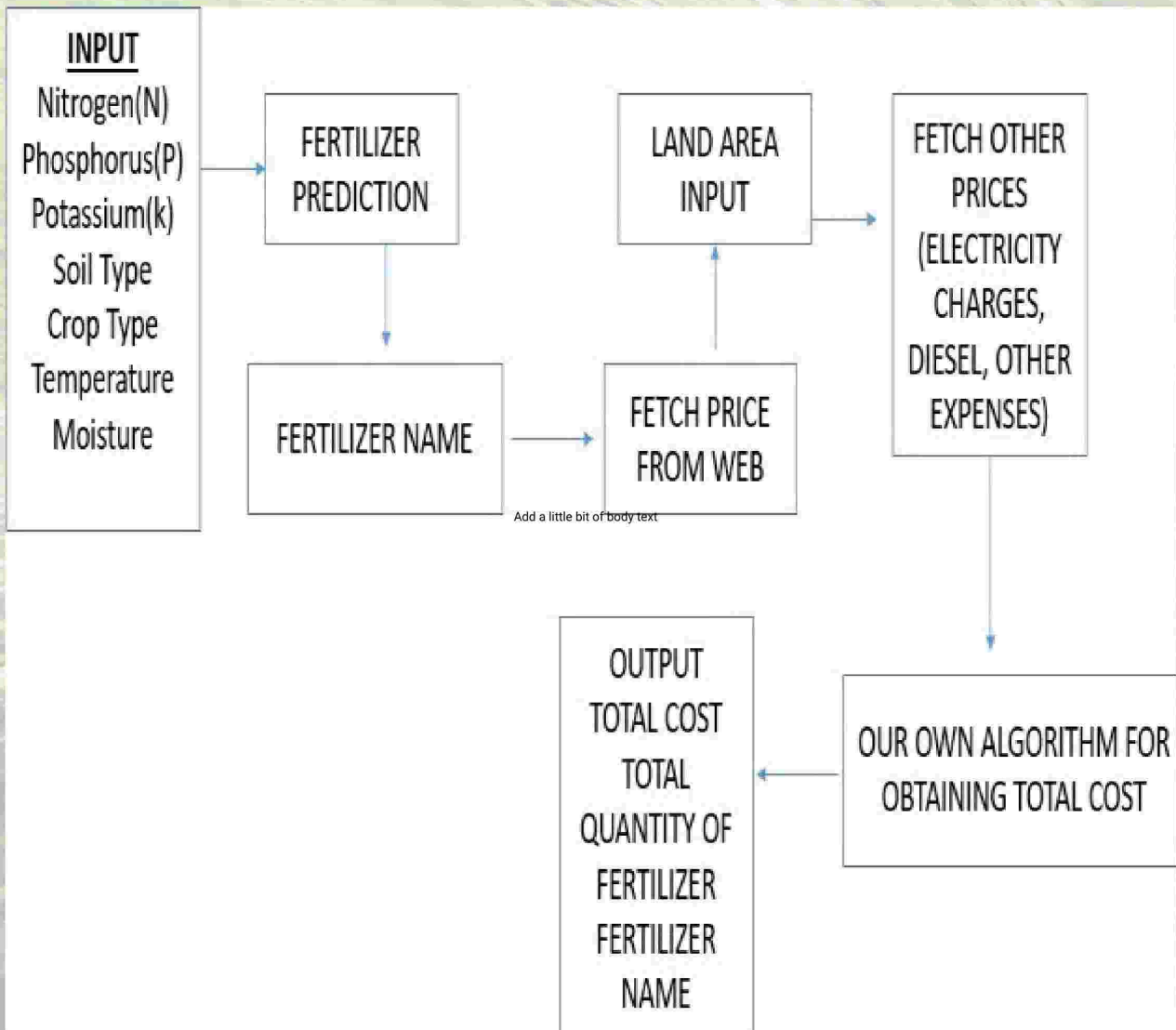
### Problem Definition:

The works done till now only predict fertilizer based on soil quality like nitrogen deficiency, potassium deficiency, phosphorus deficiency or the crop farmer is going to yield. It doesn't show how much different fertilizers need to use and weather conditions such as Temperature, humidity is not considered. So, it is necessary to provide more accurate result considering all parameters and suggest amount need to use and how much the total fertilization process cost.

• PROPOSED WORK:

Our Aim to suggest farmer the best choice on fertilizer, amount he required and predict the cost He will needed to complete the fertilization process. So, we planned things as follow to achieve optimum result.

A. *Flowchart of process:*



B. *Fertilizer Prediction:*

To achieve the goal of suggesting best fertilizer we take following parameters into consideration.

- Nitrogen
- Phosphorus
- Potassium
- Soil type
- Crop type
- Moisture
- Temperature
- Humidity



# We are using machine learning algorithm to predict the fertilizer.

## 1) Random Forest Algorithm:

- Random Forest is a powerful and versatile supervised machine learning algorithm that grows and combines multiple decision trees to create a "forest".
- Suppose there is a dataset that contains NPK, crop types, soil types. So, this dataset is given to the Random Forest classifier. The dataset is divided into subsets and given to each decision tree. During the training phase, each decision tree produces a prediction result, and when a new data point occurs, then based on the majority of results, the Random Forest classifier predicts the final decision.

## 2) SVM (Support Vector Machine):

- SVM develops a hyperplane or set of hyper planes in a high-or boundless dimensional space, which can be utilized for characterization, relapse, or different errands.
- Naturally, a great partition is accomplished by the hyperplane that has the biggest separation to the closest preparing information purpose of any class, since by and large the bigger the edge the lower the speculation blunder of the classifier.

- SVM calculation has a regularization parameter, which stays away from over-fitting.
- SVM calculation utilizes the portion trap, so you can construct master learning about the issue.

### A. Amount of fertilizer required:

By predicting How much amount of fertilizer a part of farm required we can predict total amount of fertilizer required. As if 1 acre of land require 13kg of fertilizer we can simply say how much fertilizer require by specific area of farm. This amount changes with respect to land properties.



### B. Cost analysis:

The algorithm is going to be precise and effective in calculating total amount, we also include the charge of fertilizer spraying machine, diesel for tractor, electricity charges and money to brought that fertilizer.

At last we will display the overall results to the farmer:

## CONCLUSION:

The prediction of fertilizer is based on parameters as mineral level, soil quality, crop type, moisture, temperature, rainfall which increases accuracy of prediction. By suggesting them approximate amount of fertilizer required prevents excess use and wastage of fertilizer. This system considers all expenses for fertilization process and gives total cost of fertilization process. It will save farmer's money, as farmers do not have to rely on fertilizer brokers, they will directly get the current real price of the fertilizer. It is cost effective and avoid wastage of fertilizers, as farmers sometimes brings more fertilizer than they needed, so we will tell them approx. quantity that is enough for their land, this will save their money as well avoids wastage of fertilizers. It saves time and energy of farmers as they are getting everything calculated, so they don't have to input their time and energy on such estimations

Harshit Narendra Bajaj  
Information Technology  
Thakur College Of Engineering and  
Technology  
Mumbai  
harshitbajaj48@gmail.com

Vedant Prabhakar Gosavi  
Information Technology  
Thakur College Of Engineering  
and Technology  
Mumbai vgosavi200@gmail.com

Sumit Vijaykumar Jain  
Information Technology  
Thakur College Of Engineering  
and Technology  
Mumbai Jain42507@gmail.com





# Sarcasm Detection for Code-Mix Language using Machine Learning and Deep Learning

## IT & TECH THINGS



### INTRODUCTION

A code-mix dataset consists of phrases and words from two or more languages in a single sentence. It is generally used by people who understand and use more than one language to communicate in their daily life. Code-Mix language is widely seen on social media and with so many social media applications available the amount of code-mix data available is tremendous. However, this data has various combinations of languages, especially in India where 22 different languages exist.

Sarcasm detection is an important processing problem in natural language processing (NLP), which is needed for better understanding to serve as an interface for mutual communication between machines and humans.

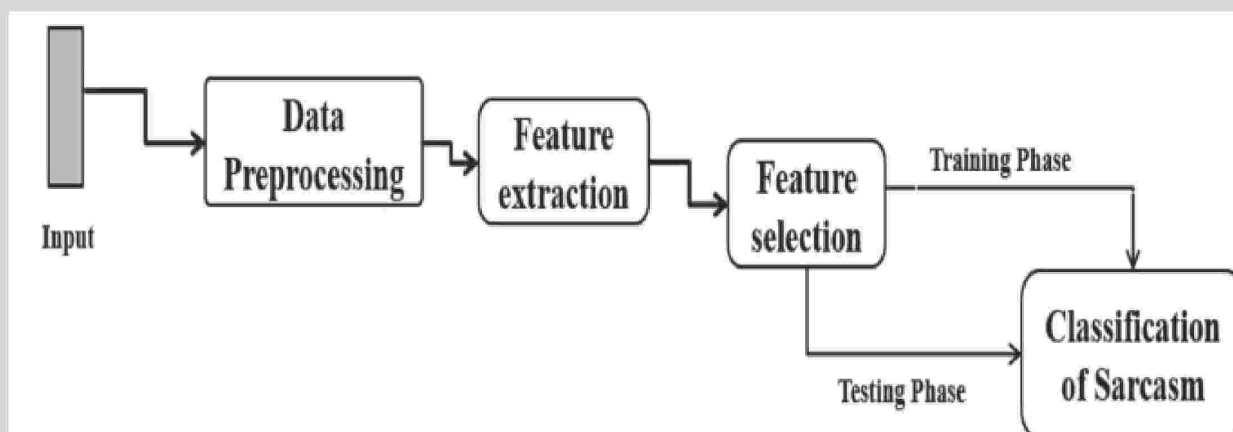


MACHINE LEARNING



To understand this is to underline the basic problem behind it - being able to detect the contradiction Time.

The sarcasm over social media can be of various types. Social media is an open platform to let people express their thoughts and feelings about something or someone. An individual may like-dislike, agree or disagree with another person's tweets, posts, or opinions.



It is however easier for humans to understand the gap between the underlying sarcastic nature of the comment, but the machine brain cannot easily differentiate. This is an addition to the various approaches in machine learning and natural language processing to resolve the issue. The paper focuses on code-mix data for the Hindi-English dataset abbreviated as the HI-EN dataset.

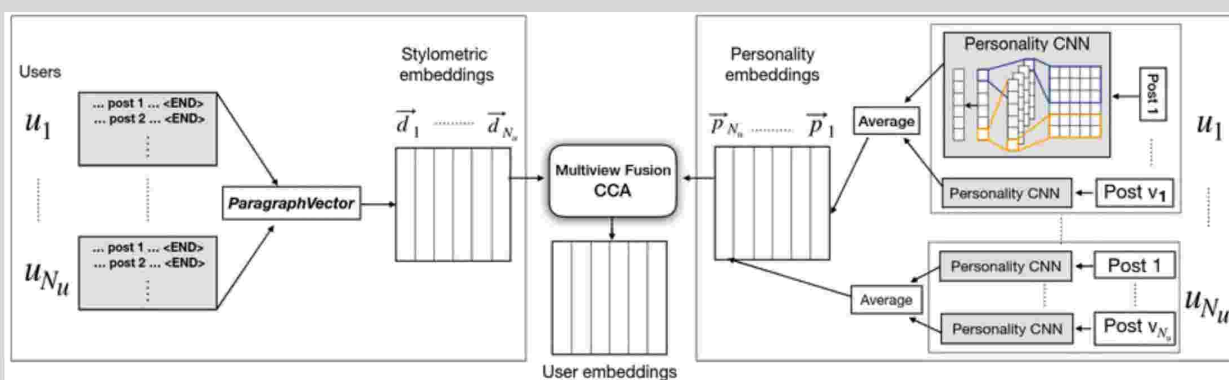
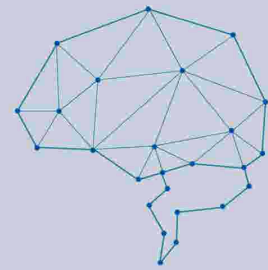


Figure 1: The figure describes the process of user profiling. Stylometric and personality embeddings are generated and then fused in a multi-view setting using CCA to get the user embeddings.

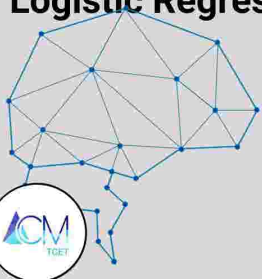




## LITERATURE SURVEY

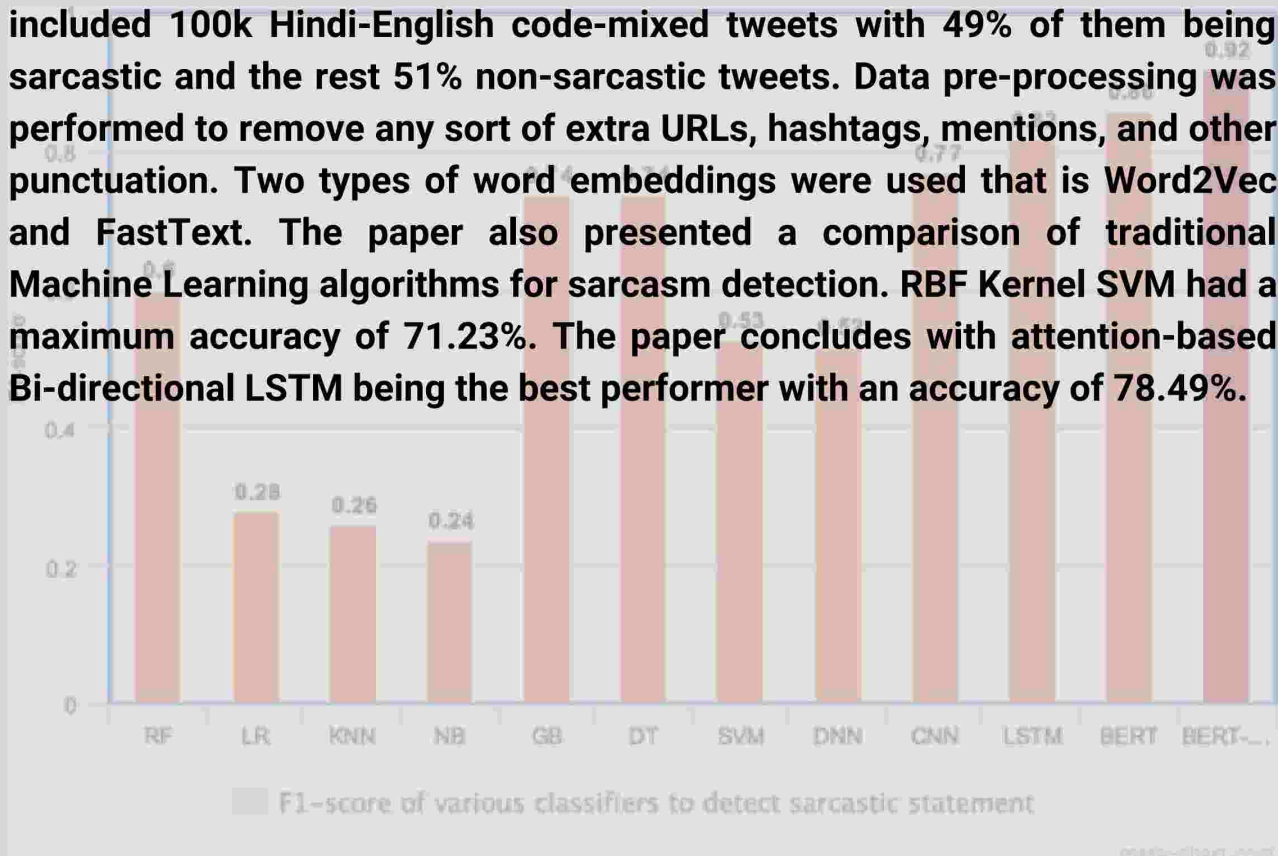
In the paper, **A Corpus of English-Hindi Code-Mixed Tweets for Sarcasm Detection** several experiments of sarcasm detection have been performed on tweets in English as well as in other languages such as Czech, Dutch and Italian but there have been no experiments on English-Hindi code-mixed texts mainly because of the lack of annotated resources. The main contribution of the paper is to provide a resource of English-Hindi code-mixed tweets which contain both sarcastic and non-sarcastic tweets. It provides tweet-level annotation for the presence of sarcasm and token-level language annotation. This corpus can be used to train, develop and also evaluate the performances of sarcasm detection and language identification techniques on a code-mixed corpus. In addition, the paper presents a baseline supervised classification system for sarcasm detection developed using the same corpus. Detection of sarcasm using machine learning classifiers and a rule-based approach shows a comparative study of various algorithms has been performed to identify which of these algorithms gives the best results. Machine Learning algorithms used to detect the sarcasm here are Support Vector Machine, Naïve Bayes, and Decision Tree for the SemEval 2018-T3-train-taskA.txt dataset and found the Support vector machine algorithm to be the best suited for the particular dataset. Algorithms like Random Forest and SVM are used for the dataset Sarcasm Detection.txt and found Random Forest algorithm gave the best results with an accuracy of 76%.

The paper "Sentimental analysis from imbalanced code-mixed data using machine learning approaches" throws light on class imbalance distribution in code-mixed Tamil-English data for sentiment analysis. Oversampling techniques namely Synthetic Minority Over-Sampling (SMOTE) and Adaptive Synthetic (ADASYN) are used to solve the class imbalance problem. An enhanced spell-checking algorithm is used for sentence classification. The Levenshtein distance metric is used to normalize the words with spelling variations. Term Frequency and Inverse Document Frequency (Tf-Idf) are used for feature extraction. Different machine learning algorithms that were experimented with include Random Forest Classifier, Logistic Regression, XGBoost classifier, SVM, and Naive Bayes. The macro average F1 score is used for evaluation. Logistic Regression performs better than other algorithms.



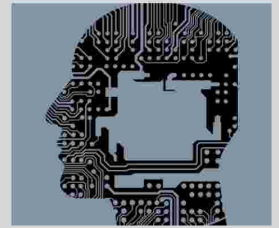
“A Dataset of Hindi-English Code-Mixed Social Media Text for Hate Speech Detection” paper addresses hate speech detection in code-mixed text. Using Twitter Python API tweets were mined by selecting certain hashtags and keywords from politics, public protests, riots, etc. A total of 1,12,718 tweets were retrieved in JSON format. After pre-processing, a dataset of 4574 code-mixed Hindi-English tweets was created. Different feature vectors were used to train the supervised machine-learning model. These vectors included Character N-Grams, Word N-Grams, Punctuations, Negation Words, and Lexicon. Support Vector Machine algorithms perform the best with an accuracy of 71.7%.

In the paper “Did you mean what you said?: Sarcasm Detection in Hindi-English Code-Mixed Data using Bilingual Word Embeddings” five different deep learning models were studied and compared for sarcasm detection in the code-mixed dataset. The corpus was scraped from Twitter. The dataset included 100k Hindi-English code-mixed tweets with 49% of them being sarcastic and the rest 51% non-sarcastic tweets. Data pre-processing was performed to remove any sort of extra URLs, hashtags, mentions, and other punctuation. Two types of word embeddings were used that is Word2Vec and FastText. The paper also presented a comparison of traditional Machine Learning algorithms for sarcasm detection. RBF Kernel SVM had a maximum accuracy of 71.23%. The paper concludes with attention-based Bi-directional LSTM being the best performer with an accuracy of 78.49%.





## OUR RESEARCH



Topics like sentiment analysis and sarcasm detection with machine learning and neural networks have gained attention in the past few years. Research has been done for monolingual datasets and certain code-mix datasets like Tamil-English, and Bengali-English. Research has been done on sarcasm detection in mainly the English dataset. However, very little research has been done on sarcasm detection for the HI-EN dataset.

Our approach is to research various algorithms, compare accuracy metrics and thus provide insight into the most suitable algorithm that can be implemented. We are aiming to do sarcasm detection on the code-mix dataset of Hindi-English by using machine learning algorithms as well as convolutional neural networks and we then compare the results.

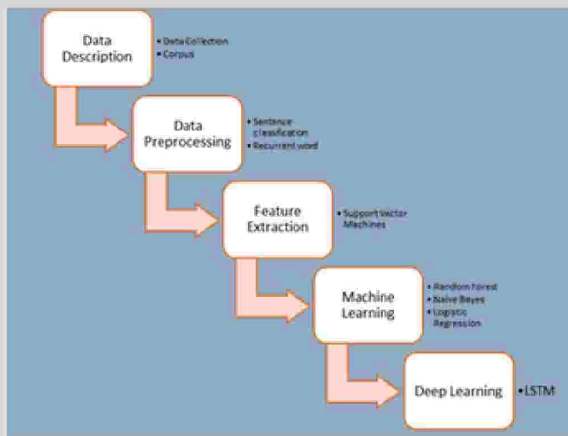


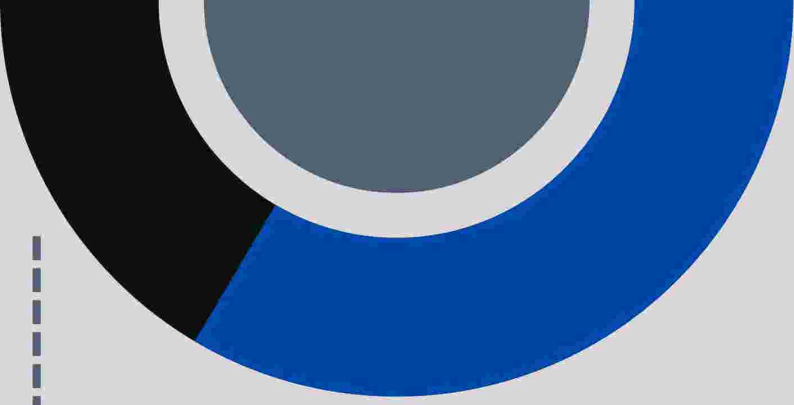
Fig 1: Project Pipeline

The machine learning algorithms that will be implemented are Bernoulli Naïve Bayes, Logistic Regression, and Support Vector Machines. The accuracy metrics that will be measured for analysis are F1 Score, Precision, and Recall. Deep learning techniques like Long-Short Term Memory and Convolutional Neural Networks will also be used.

## CONCLUSION

With the increase in the number of people using social media to express their views, tasks like opinion mining and sentiment analysis have gained a lot of importance. And using sarcasm in these social media texts make these tasks much more challenging.

In our project, we have presented an English-Hindi code-mixed dataset for sarcasm detection. We also presented a baseline supervised classification that is developed using the same dataset which uses three different machine learning techniques.

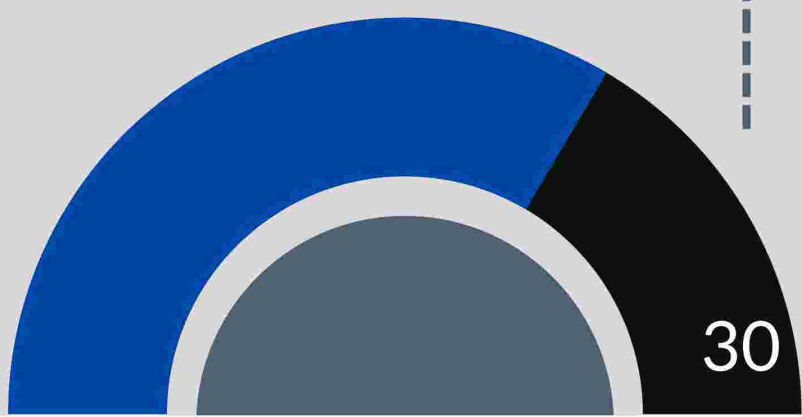


**Khushbu Ravaria**  
Department of Information  
Technology  
Thakur College of Engineering and  
Technology  
Kandivali (East), Mumbai-400101,  
India  
khushburavaria777@gmail.com

**Dhruti Seksaria**  
Department of Information  
Technology  
Thakur College of Engineering and  
Technology  
Kandivali (East), Mumbai-400101,  
India  
29dhruti@gmail.com

**Sanyuktha Udyavar**  
Department of Information  
Technology  
Thakur College of Engineering and  
Technology  
Kandivali (East), Mumbai-400101,  
India  
usanyuktha@gmail.com

**Shruti Mathur**  
Assistant Professor Department of  
Information Technology  
Thakur College of Engineering and  
Technology  
Kandivali (East), Mumbai-400101,  
India  
29dhruti@gmail.com







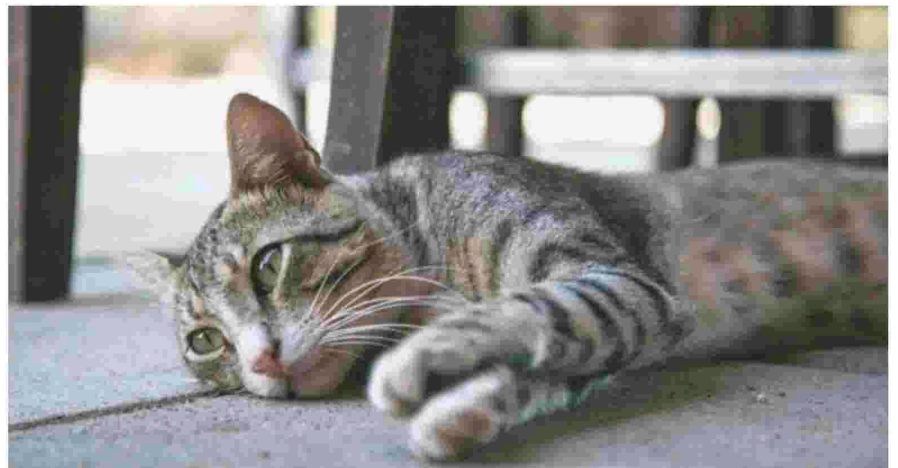
# LACK OF STRAY SERVICES IN INDIA

## INTRODUCTION

As the Indian population grows there is a lack of living spaces for humans. As humans struggle to find a roof above their heads there are millions of strays that have to bear the wrath of nature no matter what the season, condition or festival. Stray animals suffer in an unfriendly environment as they roam the streets for food and companionship. They are often victims of abuse, neglect, and fatal attacks from humans as well as diseases. Indiscriminate killing of stray animals often associated with high-speed road crashes, acts of terrorism and cruelty amongst humans is blamed as the main reason for the population explosion. Due to lack of shelters and treatment, many animals are forced to live in filthy conditions which often leads to chronic diseases and their death. People need to understand the importance of placing a shelter for these animals since they can not survive on the streets.

### ABSTRACT>

WITH EVER INCREASING NUMBERS OF STRAY ANIMALS IN INDIA THERE ARE COMPARATIVELY LESS SHELTERS TO SUPPORT THE INCREASING NUMBER OF STRAYS. THE WAYS OF HANDLING THIS SITUATION AND DISCUSSING THE OTHER SUSTAINABLE MEANS TO HELP BATTLE THIS PROBLEM.



## CURRENT SITUATION OF STRAYS IN INDIA

Today, strays are the most common animal in our neighborhood. Strays pose a serious threat to human health and safety as well as causing significant damage to community resources such as parks and public spaces. There are numerous factors that encourage stray dog populations including limited adoption programs, lack of maintenance on the community cat population and changes in other social factors such as increased poverty or unemployment.

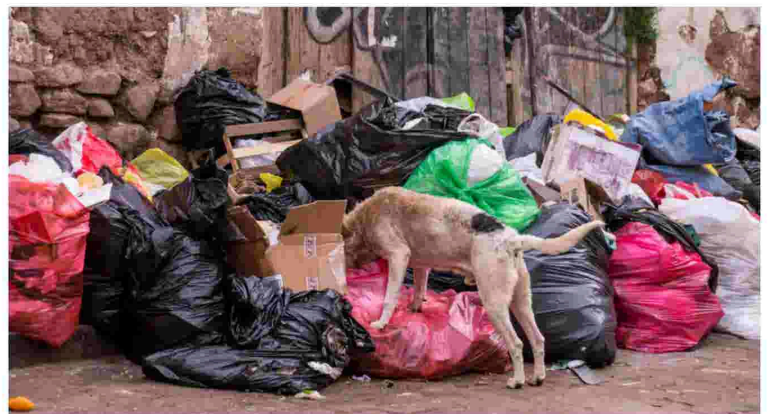
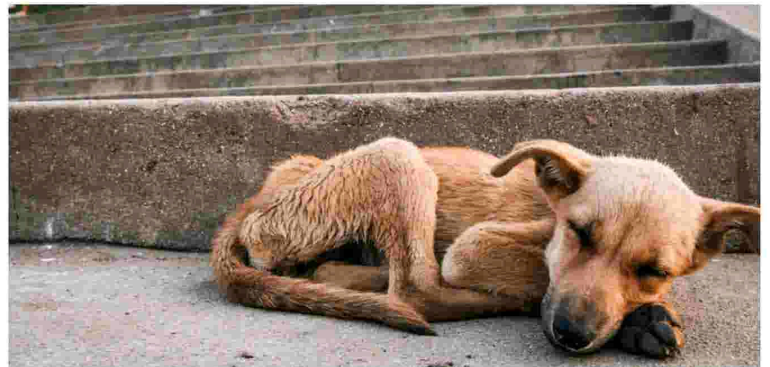
Many people think that strays come from dogs being abandoned by owners or not finding homes, but animals can also suffer cruelty at the hands of pet owners who do not want to take care of them anymore.

Today there are 6.2 crore stray dogs and 91 lakh street cats in India. They are the most vulnerable of all animals. These animals do not have a place to call their own and they have nowhere to go. The government is wasting millions of rands each year on trying to control these animals through round-ups, baiting and snaring.

As stray animals continue to proliferate across India, the plight of these animals has been neglected by the government and civil society. This problem has reached alarming proportions throughout the nation. The most common breed of stray animal is a Dog. Though stray dogs are commonly seen as cute creatures to be taken home and hugged, there is a need to come forward with proper support in order to save their lives

## PROBLEMS FACED BY STRAYS DUE TO HUMANS

Stray animals are facing problems as they have been deprived of the proper care and attention they need to grow up healthy. This can cause diseases and internal injuries, which ultimately lead to death. If there is no control on stray population, then cities will have a major problem. This may cause several issues like garbage disposal, landfill etc., which may damage or kill the ecosystem completely. Humans have encroached into their territory and caused irreparable harm to the environment. Coastal areas, wetlands and forests are being destroyed because of human activities which involve hunting and fishing.







They feed on fish and animals like rabbits, dogs and cats.

In some cases, these animals are poisoned or poisoned by garbage. In other, they lack space for living. Humans often ignore their responsibility to be responsible for stray dogs and cats by feeding them because they don't know where the best place to feed them is. Many people just want one thing in mind which is to get rid of those pesky creatures that take up space in their gardens or parks but not really considering the condition of those strays which can cause harm to human lives around them.

## PROBLEMS FACED BY HUMANS DUE TO STRAYS

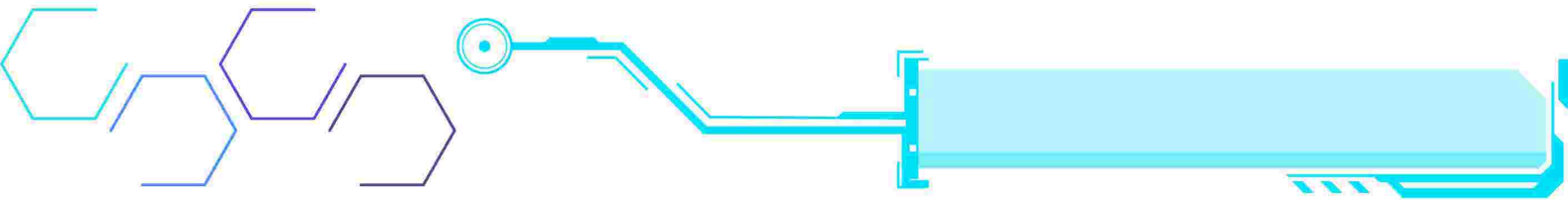
The problem of stray animals is very prominent in our society. People can spread deadly diseases to people and other animals when they are eaten. Also, there is a huge amount of food waste caused by strays who kill themselves and others out of hunger. Sadly, this leads to an increase in the amount of hunger around the world (Dr. Clifton, 2018)

Strays are a major social problem in many urban and rural areas around the world. They can be very dangerous for humans and animals alike. In addition, strays can create significant issues for individual homeowners, property owners, businesses and local governments. Stray dogs and cats also have an impact on wildlife populations, causing declines in native species of birds, mammals and insects.

Stray dogs cause more than 100,000 human deaths globally annually. There are an estimated 1 billion stray dogs worldwide. 99% of all dog bites involve a family pet or household pet. More than 300,000 people die annually from rabies. Rabid dogs have caused more than 40 million human deaths in the last 10 years. Human civilization has been unable to decrease this animal problem.

## LACK OF FACILITIES FOR THE STRAYS

In India, there are no safe places for dogs to be, and these dogs have no true place in our society. Although more people are adopting stray cats and dogs, there are still too many unadopted strays roaming around on the streets who have little chance of finding a loving home. Shelter is necessary for these homeless animals to give them a chance of having a life again. Many shelters in India don't take in the strays because they don't have enough resources or space to keep many dogs at the same time.

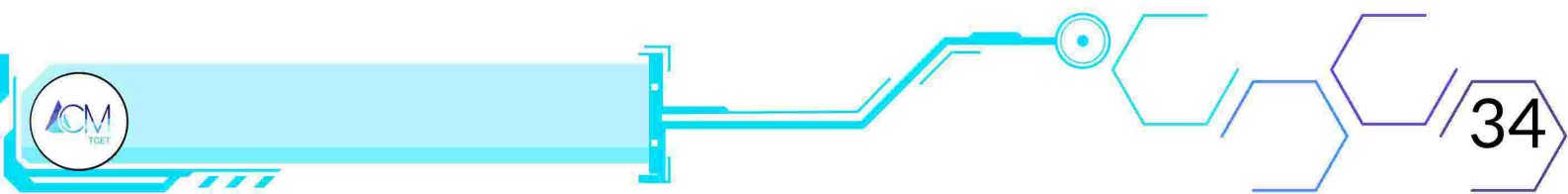


A few shelters do not even provide them with food or water, which is why most strays live on the street existing off scraps from dumpsters or rubbing shoulders with dangerous criminals who use them for target practice.

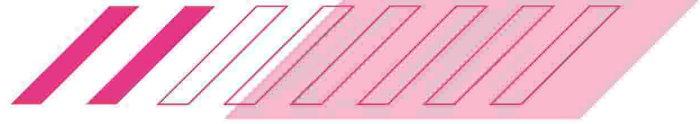
Stray animals are a major cause of animal cruelty and suffering with the right steps being taken, there is hope for strays in society. Stray animals lack shelter, food and water. To solve this issue we need to increase the number of shelters throughout the world by making each shelter take care of at least one stray animal every day. By doing this we can decrease animal cruelty, increase their chances of being adopted or returned home and hopefully rid our world of stray dogs and cats once and for all.



A lot of dogs are put down due to lack of shelter facilities. There are so many reasons as to why people adopt dogs and not strays. Firstly, most of them face a lot of problems such as negative behavior, they may suffer from skin diseases, or they may be left behind by the owners. Secondly, they are healthier than most free-roaming dogs because they have access to food and water supply regularly. Thirdly, they need vaccines and neutering/spaying appointments which are usually provided by shelters. Finally, it's easy for potential pet owners to choose a stray dog over a regular pet dog knowing that homeless dogs require little more than love and affection.

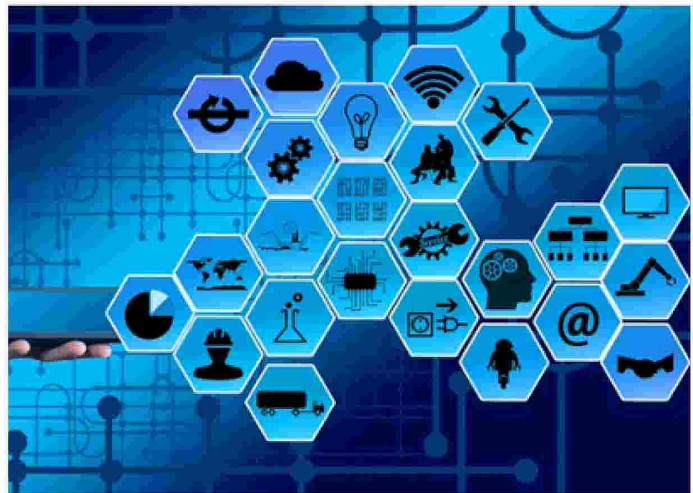




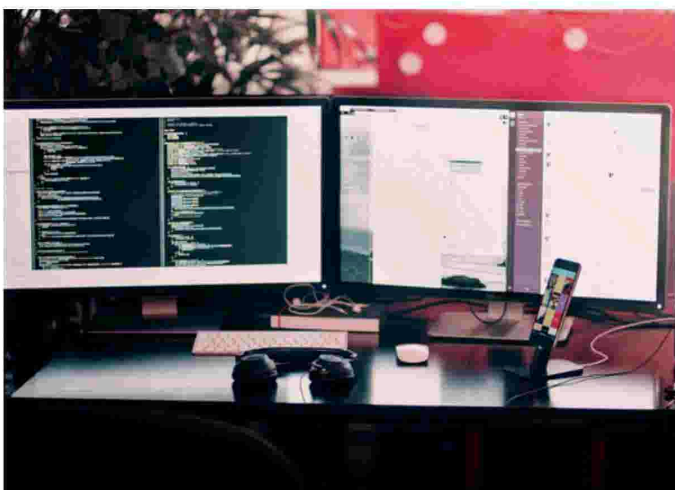


## PROPOSED WORK

The need of a web service for stray animals. Shelter management is hard work and every day, they deal with the issue of strays, which they cannot handle on their own. A simple yet powerful website that allows users to adopt stray animals is the solution to their problem. It will not only help them manage their shelter but also unite different communities in adopting animals together.



It's so simple for us to provide a social service through an easy-to-use web application. It can be in the form of a smartphone app, or a desktop app that can be used from any computer with internet. Another advantage is that it's free and the website has an option to give donations if someone likes it. We have several opportunities built into our service depending on what you need.



There are several reasons why you need to join the communities that go beyond just being a service. First, shelters can help you get involved with any animals at risk in your area. Secondly, many shelters provide a nearby adoption service, which will help you find an animal that matches your personality and price range. Also, some shelters participate in community events and fundraisers to raise awareness about the cause of pet overpopulation. Before joining any community based on stray animals, it is essential for you to know about their activities first so that you can make up your mind about whether you want to join them or not. Hence the website will constantly update on events and other services that we are looking to provide for the community.

## CONCLUSION

As we progress, we have started encroaching on the lands that do not belong to us. And in the process, we have somehow made it difficult for the strays to live out on the streets. The stray problem is something that is not well managed throughout the country and the shelter system is quite unreliable due to the lack of number of shelters. So, to promote the welfare of both humans and stray animals a website/community to help each and every one is needed. We also need to increase the number of shelters available so that they can sustain and help more animals through their service.

Kailash Ahir Information  
Technology (of Affiliation)  
Thakur College Of Engineering and  
Technology  
(of Affiliation)  
Mumbai, Maharashtra  
kailashahir2801@gmail.com

Hemin Dhamelia Information  
Technology (of Affiliation)  
Thakur College Of Engineering  
and Technology  
(of Affiliation)  
Mumbai, Maharashtra  
heminpatel2002@gmail.com

Dhruvi Jivani Information  
Technology (of Affiliation)  
Thakur College Of Engineering  
and Technology  
(of Affiliation)  
Mumbai, Maharashtra  
jivanidhruvi19@gmail.com



# ROUTE OPTIMIZATION FOR AMBULANCE

## **Abstract—**

The efficient management of ambulance routing for emergency requests is vital to save patients when a disaster response scenario occurs. In today's road traffic, the transportation of patients from emergency points becomes more difficult in such tragic situations. We consider an urgent situation where a lot of hurts/patients require simultaneously an urgent medical care. Patients are either (1) slightly injured which can be assisted on the spot or (2) seriously injured which should be transferred to hospitals. In this paper, we aim to enhance the response-time performance of emergency medical service providers by handling the ambulance routing problem (ARP). The problem can be modeled as either the Open Vehicle Routing Problem (OVRP)

or a Vehicle Routing Problem with Pickup and Delivery (VRPPD). We propose a cluster-first route-second algorithm based on the Petal algorithm and the particle swarm optimization (PSO) approach in order to efficiently handle the ARP.

**Keywords—** Emergency logistic Ambulance routing problem Particle swarm optimization Cluster-first Route-second



## Introduction

Ambulance routing problem(ARP) is one of the most important Emergency Medical Services (EMS) as it plays a vital role in saving injuries' lives and reducing the rate of mortality when a disaster occurs . The sensitivity of decision making in the EMS firstly attracted the attention of operations research experts who studied numerous classes of problems arising in the management of EMS systems. The ARP, called emergency logistic 6, is about managing and controlling the flow of ambulances to save people affected by disasters.

The response velocity of an EMS is a crucial feature since time is vital in emergency situations. An emergency is about either an ecological breakdown launched by the nature (earthquakes, floods,hurricanes) or a situation caused by man (accident, terrorist attacks). An emergency medical center (EMC) needs to decide how many ambulances should be sent to the call points. Thus, the emergency logistic network should be carefully modeled to provide reliable services to injuries needing help.

. In the recent literature, EMS is about establishing emergency facilities, rescuing survivors, providing medical assistance, transferring injuries to hospitals, and the coordination of these activities across organizations. 43 classified the EMS into five categories:



- **Transportation planning:** Programming the supplier's delivery without treating the vehicle itineraries .
- **Ambulance assignment:** Allocating the ambulances to the appropriate emergency points (40).
- **Routing:** Giving the ambulance routes to rescue the patients.
- **Roadway repair:** Repairing damaged roadways and rehabilitating the lifelines to demand areas.



- **Integrated problems:** Solving a set of above individual problems under one or more common objectives.

The remainder of this work is organized as follows. Section 2 is the literature review. The detailed description of the problem and its different modeling is given in Section 3. Section 4 explains the solution approach. The extensive experiments and the corresponding results based on the proposed methods are illustrated in Section 5. Finally, Section 6 concludes the results and discusses possible future works.



## **Cluster - First Route - Second approach**

Solving the ARP involves two kinds of decisions

(1) partitioning the injuries into clusters compatible with ambulance requirements and (2) sequencing the call points in each cluster to get an ambulance itinerary. A classical approach for the basic routing problems is based on the cluster-first route-second

principle, in which the partition is determined first. A traveling salesman problem(TSP) is then solved for each cluster.

The proposed algorithm, named PA-PSO, is based on the Petal Algorithm and the PSO approach.

## **Cluster - first method: Petal Algorithm**

The sweep algorithm is a method for clustering customers into groups so that customers in the same group are geographically close together and can be served by the same vehicle. Feasible clusters are initially formed by rotating a ray centered at

the depot. A vehicle route is then obtained for each cluster by solving a TSP.

The first mentions of this heuristic are found in 36 and 35, but the sweep algorithm is commonly attributed to 11, who popularized it. From each cluster a vehicle route is obtained by solving its corresponding TSP. Finally, exchanging vertices between adjacent routes can be performed if it turns out to be distance saving.

**Sweep Algorithm** is a two-phase heuristic that decomposes the VRP in two sub problems namely, clustering the vertices and constructing the routes with possible feedback information between the two subproblems.

## Algorithm 1 Sweep algorithm

**Required:** Distance matrix, polar coordinates, customer demands and vehicle capacity

**Ensure:** Best solution

- 1: Select an arbitrary vertex  $i^*$  to which assign a value  $\theta_{i^*} = 0$ .
- 2: Compute the remaining angles from  $(1, i^*)$  and rank vertices in increasing order of their  $\theta_i$ .
- 3: Select an unused vehicle  $k$ .
- 4: Start by the vertex having the smallest angle, then assign vertices to the vehicle as long as its capacity is not exceeded. If unrouted vertices remain, go back to 1.
- 5: Optimize each vehicle route by solving exactly or approximately the corresponding TSP. Re-optimize (using an improvement heuristic).





## **Extension of sweep algorithm:**

Petal Algorithm According to [16], the Petal Algorithm (PA) can be described as follows. The nodes are numbered in radial order about the depot. Each radially consecutive set of nodes forms a petal. A petal is feasible if the sum of customers' demands do not exceed the vehicle capacity. PA is a natural extension of the sweep algorithm to generate several routes, called petals. PA can be described as follows. Given  $S$  is the set of routes,  $x_k = 1$  if and only if route  $k$  belongs to the solution,  $a_{ik}$  is the binary parameter equal to 1 only if vertex  $i$  belongs to route  $k$ , and  $d_k$  is the cost of petal  $k$ . If routes correspond to contiguous sectors of

$$\text{Min } \sum_{k \in S} d_k x_k$$

$$\sum_{k \in S} a_{ik} x_k = 1; i \in N$$

$$x_k = 0 \text{ or } 1; k \in S$$

vertices, then this problem possesses the column circular property and can be solved in polynomial time. 352 Takwa Tlili et al. / Procedia Computer Science 112 (2017) 350–357 PA is about solving a set partitioning problem of the form.

## **Route-second method: Particle swarm optimization (PSO)**

Particle Swarm Optimization (PSO) is an evolutionary algorithm inspired by cooperation and social behavior of animals such as a flock of birds finding a food source or a school of fish protecting themselves from a predator. The good points of PSO are its high computational efficiency, rapid convergence and its ability to successfully avoid local minima [32]. This approach involves a population of particles, where each particle is a solution. Each particle has a related velocity, which is adjusted with an update equation that considers the history of individual and collective experiences. Mainly, it is about changing the position of each particle seeking for the best solution.

Given a d-dimensional search space and a swarm of S particles ( $p = 1, \dots, S$ ).

In PSO, to each particle p in generation t corresponds a position  $X_t p = (x_t p_1, x_t p_2, \dots, x_t p_d)$  and a velocity- vector

$V_t p = (v_t p_1, v_t p_2, \dots, v_t p_d)$ ,

where  $x_t p$  is the location and  $v_t p$  is the flying velocity of particle p in generation t in the nth dimension of the search space ( $n = 1, \dots, d$ ).

Then, particles memorize every reached position and save the one with the best fitness. This individual best position is denoted by  $B_t p = (b_t p_1, b_t p_2, \dots, b_t p_d)$ .

Particles record the whole best position's fitness until generation t. This global best position is referred to as  $G_t = (g_t 1, g_t 2, \dots, g_t d)$ .

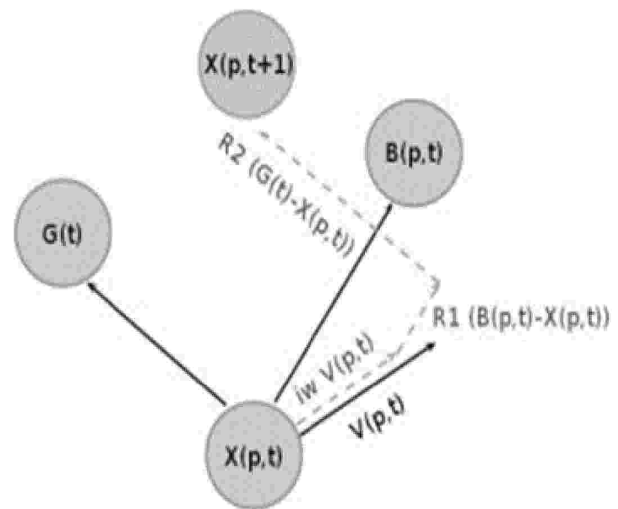
The position and velocity of particles in PSO can be updated with the following rules:

$$v_t p_n = iw v_{t-1} p_n + \lambda_1 (b_{t-1} p_n - x_{t-1} p_n) + \lambda_2 (g_{t-1} n - x_{t-1} p_n)$$

$$x_t p_n = x_{t-1} p_n + v_t p_n$$

$$p = 1, \dots, S; n = 1, \dots, d; t = 1, \dots, T$$

- **iw** The inertia weight that controls the influence of the precedent velocity on the current velocity
- **c1** The first acceleration coefficient given to the attraction to the previous best location of the current particle
- **c2** The second acceleration coefficient given to the attraction to the previous best location of the particle neighborhood
- **r1, r2** Two random variables uniformly distributed in [0,1], i.e.  $r_1, r_2 \sim U(0,1)$



- **Tmax** The maximum number of generations prefixed by the decision maker  $\lambda_1 = c_1 \times r_1$  and  $\lambda_2 = c_2 \times r_2$



## **Experimental results:**

To test our approach, we ran it on the 60 instance test bed composed of 27 problems from Augerat et al. (1995), 13 problems from Rochart and Taillard (1995) and 20 problems from Golden et al. (1998). As previously mentioned, the problem is modeled as OVRP and VRPPD. Thus, our aim is to test the performance of the proposed approach on both models. All computations were done on a PC with a 2.5 GHz Intel core i5 and 8 GB RAM running under Microsoft Windows 8. We programmed with Java language using NetBeans IDE 8.0.2.

## **Performance on open VRP instances :**

In order to evaluate the PA-PSO, its performance is compared to two stochastic-based approaches (Christiansen & Lysgaard, 2007 and Goodson et al., 2012) that were tested on the same set of instances. Table 6 shows that the PA-PSO could provide reasonably good solutions in testing the benchmark.

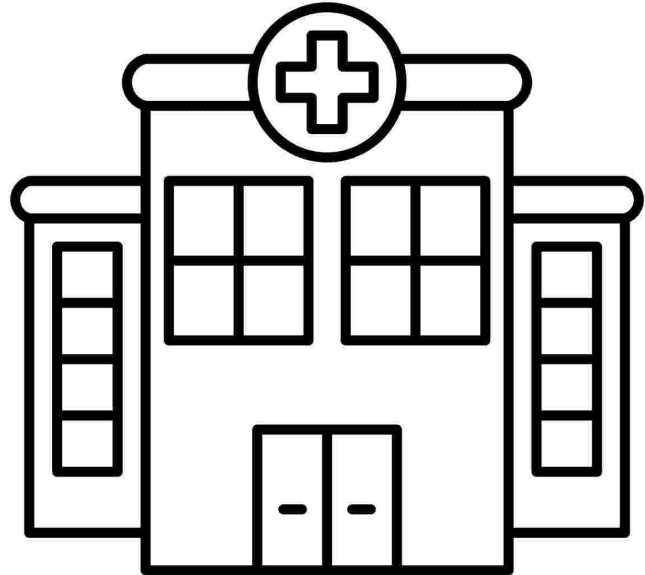


For the tested problems our approach outperforms in all cases the natural-based approach (Goodson et al., 2012) and the branch and price approach (Christiansen & Lysgaard, 2007). As it can be seen, the deviation between the PA-PSO solution and the best known solution is always under 5.52% while it is up to 10.37% and 12.45% for the other methods.

Tables 5 and 6 show that for the testbed instances, PA-PSO outperforms the SA approach

## **Conclusion:**

In this paper, we handled a patient transportation problem, named ambulance routing problem, derived from emergency medical services. The ambulance routing problem is modeled in this work as an open VRP and a VRP with pickup and delivery. The proposed approach is a cluster-first route-second method based on the petal algorithm and the particle swarm optimization. In order to test the performance of the algorithm, we ran it on test bed instances and compared the results with state-of-the-art methods.



*Name: Ekta Singh*

*Dept: TE-ITB-33*

*City: Mumbai, Maharashtra*

*Email: ektasingh16102002@gmail.com*

*Name: Dev Yadav*

*Dept: TE-ITB-58*

*City: Mumbai, Maharashtra*

*Email: devyadav270802@gmail.com*

*Name: Shaloo Singh*

*Dept: TE-ITB-66*

*City: Mumbai, Maharashtra*

*Email: shalusingh7170@gmail.com*

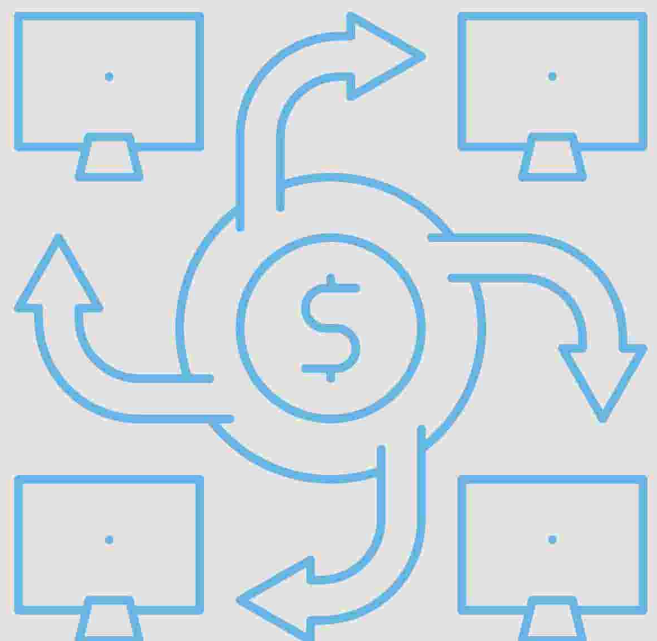


# AUTOMATED KYC

## **Abstract—**

Abstract—KYC (Know your Customer) entails several steps, including establishing customer identity, comprehending the nature of customers' activities, and ensuring the legitimacy of the source of funds. We have a portal in this project where clients can upload a video of themselves saying all of their information, and it then verifies the identity of clients from the database while also automating the KYC process. This is done to reduce the possibility of financial fraud and identity theft. It also ensures the veracity and authenticity of the information presented.

Keyword— Automated KYC, liveness, audio extraction, detection



## ***Introduction***

Financial organizations always struggle to identify their customers and implement policies to stop financial crime. A significant set of requirements known as KYC for customer identity verification must be followed by financial institutions (including banks, credit unions, and Fortune 500 financial businesses) to ensure security.

A series of procedures for confirming a customer's identification before or while conducting business with banks and other financial organizations is known as KYC, sometimes known as "Know Your Customer" or "Know Your Client." Following KYC regulations can assist prevent the funding of terrorism, money laundering, and other common fraud schemes. Financial institutions can more precisely identify suspicious actions by first confirming a customer's identity and intentions at the time of account opening, then understanding their transaction patterns.

When it comes to KYC laws, financial institutions must adhere to ever-higher criteria. To comply with KYC, they must spend more money, or face heavy penalties. Due to these restrictions, practically any company, platform, or group that interacts with a financial institution to open an account or conduct transactions is required to adhere to these requirements.

classified the EMS into five categories:

## ***Background***

Know Your Customer (KYC) is used for customers Administration and Identity Verification. This document was sent to your organization by a customer Build trust between the two. At first there was no way to verify the customer's identity. KYC was proposed in the US in 1990 and back then, the purpose of KYC was to stop terrorists Bank financing and money laundering. the most important KYC stakeholders are banks.



money laundering, terrorist financing, and financial fraud.

Therefore, banks currently do not allow account holders without KYC documentation. KYC documents include customer information, proof of identity, proof of address, and photo. Originally, a pen-and-paper approach was used to submit KYC documents, but problems with record keeping were evident. Banks were busy verifying identities each time through customer-filled forms. In such cases, documents are more likely to be misplaced.

Therefore, a digital KYC system called eKYC has been proposed. In this approach, the customer fills out the KYC document via her organization's web application. The transmitted data was stored in a central database.

Organizations can access customer information at any time through the customer ID. Overall costs were reduced because the system is paperless, but since data is stored in

a centralized database, there are no single points of failure, data redundancy, and third-party involvement in validation. There are still loopholes in centralized systems. Also, data stored in centralized servers can be compromised/attacked by hackers, so the existing centralized system architecture increases the chances of leaking sensitive customer data.

### ***Importance***

KYC stands for knowing your customer. The standards of KYC are designed to protect financial institutions against fraud, corruption, money laundering and terrorist financing.

KYC involves several steps.

Understand the nature of customer activity and ensure that funding sources are

legitimate. KYC policy was introduced by the Reserve

Bank of India. Financial Action Task Force (FATF)

Recommendations on Anti-Money Laundering

(AML) Standards and Counter-Terrorism Financing (CFT). anti-money laundering law Banks, financial institutions and intermediaries must ensure that they meet certain minimum requirements KYC and AML standards.

The KYC process has steps such as ID card verification, face verification, document verification such as utility bills as proof of address, and biometric verification. Banks must comply with KYC and anti-money laundering regulations to limit fraud. The process of KYC is the collection of information about the customer. They must complete an online KYC registration form on the portal they wish to conduct financial transactions with. Whether the information provided is accurate and up-to-date depends on the applicant's duty of care. Stage two is uploading of evidence which includes After the information is collected, the applicant must verify the information entered in the form with relevant documents.

These documents certify that the prior information entered by the user is genuine and not forged. The third step is verification in which After a document is uploaded, the document's template is identified and checked using various checks.

This ensures that the document has not been tampered with. Once the documents are verified data is extracted from the documents. This was the process of online KYC.

For offline the offline KYC process is like the online process. However, the key difference is that a physical copy of all documents and applications is required.





## **Objective**

The primary purpose of automated KYC is to reduce the use of physical documentation, storage space for such records, and to reduce costs and time required. E-KYC allows submission to take place in real-time. This makes the process streamlined and less prone to corruption in industries such as banking, finance, etc.

With online verification, customers can open new accounts, apply for loans, take out new life insurance policies, invest in new mutual fund companies, and trade with cryptocurrencies through electronic devices, especially smartphones, in an era when customers prefer to communicate through online channels.

Our system provides software that makes automation of the Know Your Customer processes easy. It reduces the time required for onboarding of a new customer and eliminates the possibility for human error.

Therefore, your company is on the safe side, when visited by the regulator. Moreover, the onboarding speed of the customer provides your company with business advantage among the competition. Faster onboarding means more customers.

## **Literature Survey**

### **1. Decentralize KYC system**

Independent third-party KYC systems can be used in several places to verify the identity of Individuals are the order of the day. Distributed immutability Ledger is a key concept to ensure that the data stored in it is stored. The system is tamper-proof. Data stored in distributed databases is encrypted Provides an extra layer of security. even if there is one Data stored in distributed databases is at risk. The data is encrypted so no harm is done.

## **2. Optimized and dynamic KYC system based on blockchain technology**

adoption via means of monetary establishments (FIs) is very difficult. device that reduces and stocks out several the monetary establishments that paintings with a patron the fees of the KYC method and additionally makes it feasible for FIs to dynamically replace statistics associated with clients and disseminates these statistics amongst taking part FIs. Additionally, our device addresses several the attributes that avoid the adoption of formerly proposed answers via way of means of FIs

## **3. Know Your Customer (KYC) Implementation with Smart Contracts on a Privacy-Oriented**

### **Decentralized Architecture**

As part of Know Your Customer (KYC) standardized, decentralized scheme to enable user privacy protection within the enterprise, Blockchain has been proposed by developing two types of smart contracts.

This system introduces simplicity of its scheme and efficiency and time efficiency of operation. Intelligent integration of various technology modules and components.

## **4. Hassle - Free and Secure e-KYC System Using**

### **Distributed Ledger Technology**

The blockchain technology is a prominent, reliable and secure technology which is getting into almost every The fundamental essence of blockchain technology offers features like transparency, decentralization, how the present banking industry, especially the KYC document verification process, can be impacted after using The current day banking KYC processes are highly reliable on paper a reliable and trustable technology like blockchain, that could withstand frauds, and resolve the scalability and In the proposed system, the use of blockchain in KYC process restricts the presence of middlemen. efficiency, reduction



in costs, enhanced customer rendezvous and end-to-end transparency during the process of integrating the customer documents into the bank database.

### **5. KYC Optimization using Blockchain Smart Contract Technology**

n organization for better understanding of customers and their financial operations. The current KYC mechanism is serious. Every institution has its own KYC process and sometimes involves third parties, which can cause problems. Increased maintenance costs, time, and redundancy. Starting at about \$27, including opportunity costs, maintenance costs, and know-your-customer costs

The current KYC process is very time consuming and

proposed an improved KYC system utilizing blockchain

Technology to improve existing KYC systems. Blockchain

mechanism, proposed model of KYC process, can be stored, updated,

Maintenance costs and many other costs that can affect your organization's performance.

### **6. KYC Optimization by Blockchain Based**

#### **Hyperledger Fabric Network**

financial institution with a long history and the future of financial institutions. Hyperledger Fabric network for KYC optimization. The proposed system can speed up KYC payment transfers. Challenge the inefficiencies that result from duplicative actions

### ***Problem Statement***

Know Your Customer is the process of cross-checking a customer's identity. KYC policies aim to prevent criminals from using banks for money laundering activities. It also helps banks better understand their customers and their financial operations, and manage risk prudently. The offline method requires a lot of documentation and on the other hand KYC online is not automated.

The details must be entered manually, which is time consuming as it involves many people and requires someone who can work and collect the information.

Online systems tend to consume both time and effort. The goal is to automate the system to reduce the time and effort involved. This can be done by recording a video that provides all the important details of a person which is required for KYC.

### **1. Contact with too many people within**

#### **Organization:**

The customer needs to contact corporate clients at least 4 times during the KYC compliance, while corporates on an average of eighth times.

### **2. Security concerns**

The documents required for KYC include the personal documents (such as passports) of company directors.

### **3. Compliance is taking longer**

The survey of more than 1,100 organization executives earlier this year highlighted several

problems with the KYC process globally. KYC regulations take 26 days .Corporate customers claim that they spend 32 days on KYC compliance.

### ***Solution***

KYC is a standard which is used globally by many sectors like banking, finance, business, Institutes, etc. KYC helps these organizations verify their client and protect the client's data from malicious attacks.

Basically, KYC

protects both the client and the organization and helps more clients to trust in them.

As the whole world is going Digital, so has these

organizations. Unlike before KYC verification is now

usually done online but still the process is very time

consuming as it involves long video call, long waiting

time, verification time, etc. The primary purpose of automated KYC

is to

reduce the use of physical documentation, storage space for

such records, and to reduce costs and time required. E-KYC



allows submission to take place in real-time. This makes the process streamlined and less prone to corruption in industries such as banking, finance, etc.

With online verification, customers can open new accounts, apply for loans, take out new life insurance policies, invest in new mutual fund companies, and trade with cryptocurrencies through electronic devices, especially smartphones, in an era when customers prefer to communicate through online channels.

Our system provides software that makes automation of the Know Your Customer processes easy. It reduces the time required for onboarding of a new customer and eliminates the possibility for human error. Therefore, your company is on the safe side, when visited by the regulator. Moreover, the onboarding speed of the customer provides your company with business advantage among the competition. Faster onboarding means more customers.

The solution we propose is removing the human part from this process. The client needs to upload a video then the verification happens within minutes through the portal. This not only solves the time-consuming issue but also does not need long waiting, or booking appointments. For a global Standard like KYC which is used almost in all sectors, the proposed solution is best for organizations.

## **Tech Stack**

Tech Stack:

1. Language:

Python  
SQL

2. Software:

Jupyter Notebook  
Phpmyadmin server  
Xampp

3. Libraries:

Face\_recognition  
Speech\_recognition  
Pyaudio  
Moviepy.editor

## Working

This project is a platform for implementing a digitalized KYC process that includes liveness and audio detection. To expedite the KYC process, users must upload a video containing identity proof such as a PAN card/ AADHAR. After you upload the video to the website, These organizations verify their client and protect the clients data from malicious attacks. Basically KYC protects both the client and the organization and helps more clients to trust in them.

As

### Modules

01

Face Recognition

02

Audio Extraction

03

Audio to text

04

Tally keywords with data

As the whole world is going Digital , So has these organizations. Unlike before KYC verification is now usually done online but still the process is very time consuming as it involves long video calls, long waiting time, verification time, etc.

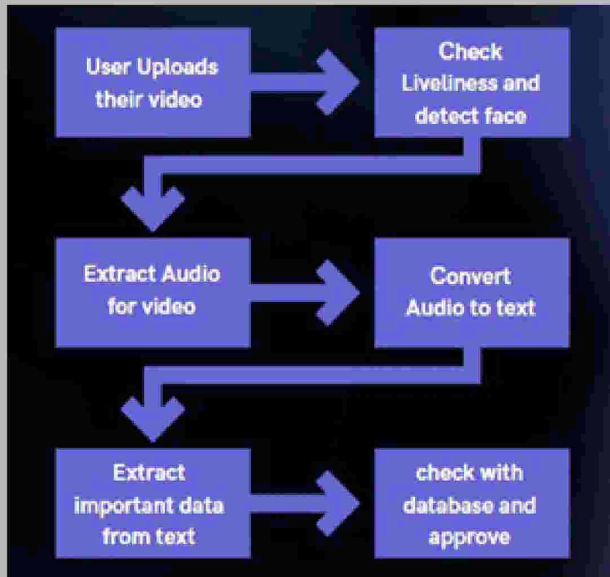
It does so by comparing the user's live video with the uploaded national identity document and making the KYC Procedure Complete. Our solutions use liveness detection technology to determine and validate customer's identity in real-time. In this way by uploading video, KYC helps in mitigating the risks of identity thefts and financial frauds

The solution we propose is removing the human part from this process. The client needs to upload a video then the verification happens within minutes through the portal.

This not only solves the time consuming issue but also does not need long waiting, or booking appointments.

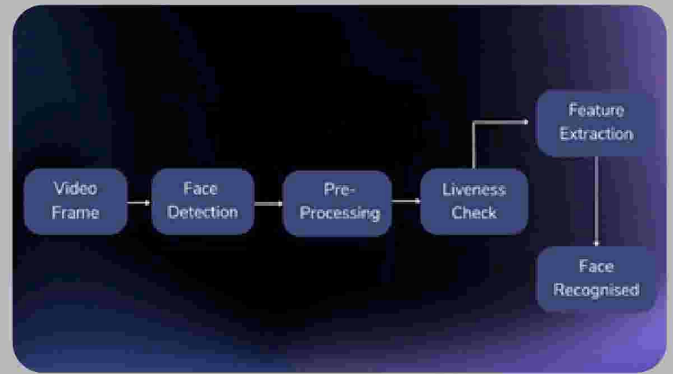


For a global Standard like KYC which is used almost in all sectors, the proposed solution is best for organizations. it will follow the following steps:



### Step 1: Recognising Face:

First step for KYC verification is ensuring the video belongs to the client. For this we use python library “face recognition” and match it to the database to ensure the face in the video is of the right person. This library follows a sequence of steps to achieve face recognition which is explained through the flowchart given below



As the whole world is going Digital , So has these organizations. Unlike before KYC verification is now usually done online but still the process is very time consuming as it involves long video calls, long waiting time, verification time, etc.

It does so by comparing the user's live video with the uploaded national identity document and making the KYC Procedure Complete. Our solutions use liveness detection technology to determine and validate customer's identity in real-time. In this way by uploading video, KYC helps in mitigating the risks of identity thefts and financial frauds. The solution we propose is removing the human part from this process.

For a global Standard like KYC which is used almost in all sectors, the proposed solution is best for orgonite client needs to upload a video then

the verification happens within minutes through the portal.

This not only solves the time consuming issue but also does not need long waiting, or booking appointments.

zations.

it will follow the following steps:

Step 2: Extract the audio from the video

For this we plan on using some inbuilt libraries of

Python which will help us extract the audio from video. The

Library we use will be "moviepy.editor" to achieve it.

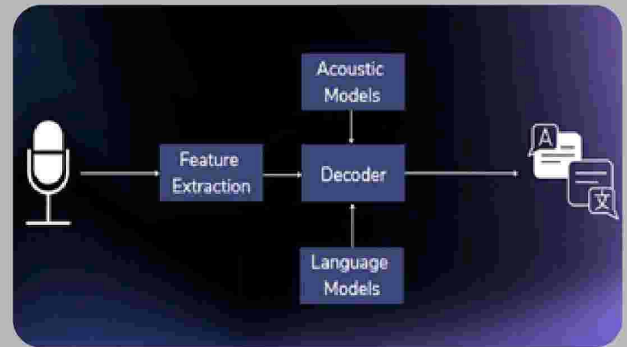
Step 3: Convert the audio to text

This Step involves using Python Library "speech

recognition" to convert speech to text. This python library

follows are sequence of steps to achieve speech recognition

which is explained through the flowchart given below



Step 4: Extract important data from the text

Once the Speech is converted to text, It is now time

to look for keywords like "PAN CARD" "AADHAR CARD" and note those details.

Step 5: Check the database

To ensure that all documents are valid. Once your

KYC verification is finished, the user will be alerted about it.



## **Conclusion**

The conventional way to complete the KYC process is to make an appointment or visit a nearby KYC outlet. This isn't always the wisest or most efficient option, though. The goal of this system is to provide a new KYC system that is simple to use and confirms identity using user data that is accurately mapped from a database. Although a lot of documents are still created on paper nowadays, it is obvious that automatic data recognition technologies are widely used.

The document has numerous copies since it is repeatedly copied and modified throughout future processing processes. They can successfully assist people in some situations, but they can also be ineffective in others. This system's goal is to successfully check data in text by precisely mapping each user's information to the form field from the database.

## **Result**

Our main aim of the project is to develop a system wherein information of people can be fetched by either a video or a photo and the admin can have easy access to this information. The system plays an important role during the pandemic period to mark the attendance for online classes so that a lot of time can be saved. It can also be used on airports, railways to make work easier, faster and safer. We have successfully built a prototype for this system which consists of six pages consisting of the home page, login/register page, Detection page and a profile page. We are trying to add more and more features to the project to make it user friendly and also give accurate results.

### **Credits**

1. *Aakanksha Maurya*  
*Information Technology*
2. *Janhavi Mishra*  
*Information Technology*
3. *Unnati Mistry*  
*Information Technology*

# HEALTHTECH

## THE INNOVATORS IN HEALTHCARE

The pandemic has given an opportunity to highlight the gaps in the Indian healthcare system and have asked for solution to look into gaps with innovation and given opportunity for everyone to rise up with innovations and out of the box thinking. This has forced the stakeholders in the industry to look at the technology be it cloud based Saas Application, virtual consultation, virtual monitoring of patients so that patient care is not affected which was need of the time.



Let me give a brief of healthcare application before the pandemic so that we understand how it has changed now in the post pandemic era. Before the pandemic the backbone of any hospital was the Hospital Information Management System (HIMS) which was used as an ERP in the hospital and hence was the major enabler of all the processes in the hospital and would basically cover these areas like

- **Patient registration:** The basic demographic of the patient by generation and unique hospital identity (UHID)
- **Consultation and Prescription :** The interaction of doctor and patients on the outpatient department (OPD)



- 
- **Pharmacy** : After the consultation the doctor would prescribe certain medication and that was brought from the hospital or local pharmacy.
  - **Diagnostic reports**: Diagnostic reports consists of the lab reports and the radiology reports like x ray , ct scan etc on OPD basis
  - **Admission and Discharge** : If the patient needs to be admitted then the patient will have to allocated a bed and other details need to be covered because in Inpatient department (IPD) the patient has a longer stay. Discharge is another important step when an IPD patient is due to leave the hospital after treatment.
  - **Billing** : Billing of patient for the services the patient has availed during his hospital visit be it OPD or IPD
  - **Feedback**: After the OPD or IPD visit the feedback of the experience was collected manually so as to analyse the

point which needs improvement and corrective actions.

These were major touch points which are covered either manually or some HIS which were used to digitise these and these were kept on servers in the local premise and normally were used before the pandemic but all these touch points had to undergo some or the other innovation so that patient care is not effected .The major challenge was to minimize the face to face interaction without compromising on patient care.



---

Enter the Health Tech companies which brought in the innovation on these touch points by introducing the latest tech and products which were being used by some other industry and helped in solving the issues to a major part. The major change that happened as part of the pandemic was the application which were kept on the premises moved to a cloud platform and thus enabling a cloud based HIS.

***Patient registration:***

Contactless or virtual registration using kiosk and web registration these methods made it easier for patient registration and helping in reducing the front desk crowd and minimizing the patient face to face interaction , Similarity these tools were used for booking appoint of doctors be it virtual or in person thereby helping in streamlining the crowds as well bring a process in the front desk of the hospital and these kiosk were adopted similar to kiosk in the airport.



***Consultation and prescription:***

In order to continue patient care in these crises and to sustain the hospital video consultation or virtual consultation became the need of the hour and forced the doctors and patient to meet virtually and the Healthtech companies were offering a platform to do this video consultation and also storing them in the application for future reference. Here the prescription also became online during the virtual consultation



---

## **Pharmacy:**

As all the processes were virtual so was the pharmacy . the hospital opens a virtual pharmacy with delivery of pharmacy items and online pharmacy also paid a bigger role during these and some health tech companies became the amazon of healthcare

## **Diagnostic reporting:**

Since consultation were online and so were diagnostic reporting in lab and radiology here the innovation was done by Health tech companies by introducing collection at home and then publishing the report virtually which could be uploaded on the application before the next follow up consultation.

## **Admission and Discharge :**

Here health tech companies introduced virtual hospital rooms in the comfort of the home using web camera for continuous monitoring and virtual consultation so that patient need not visit the hospital but



patient care was not effected all the necessary equipment were transferred to the patient home and set up before doing such an exercise

## **Billing and Payment:**

Instead of physical bills and payments, Virtual bills were submitted to patients through whatsapp and email and payment were also done virtually by using UPI or other online payment methods enabling the digitation of the entire process

## **Feedback:**

The physical feedback collection also became virtual by which the feedback was getting collected in a google form in the initial stage to Feedback application with analytics to improve upon the preventive and corrective actions

---

So as we can see the pandemic had exposed our healthcare vulnerable to various problems, but in that critical time as well our Healthtech companies innovated brought in solutions to tackle and made that gaps of necessity to opportunity to make the healthcare system relevant and tech enabled and helped in digitising the healthcare sector on a whole.

It made the patient, hospital and doctor to change their process and adapt to the new changes which they adopted even in this crisis period.

The pandemic made the government, the stakeholders and the patient adapt technology and new processes

and thereby embracing innovation by healthtech companies and making the Healthcare sector to move to enhanced version.

There are still many issues which needs innovation and out of the box thinking and more opportunities to build new healthtech giants in terms of business and future growth of this sector so whenever you next visit any hospital identify a group of issues which needs to be looked into with innovation and pandemic has shown us if the solution makes a huge impact the healthcare sector absorbs it you can become then Healthtech innovators as we always say that necessity is the mother of innovation



JOGESH NANDIKKARA



# CUCUMBER (BDD FRAMEWORK)

## Introduction:

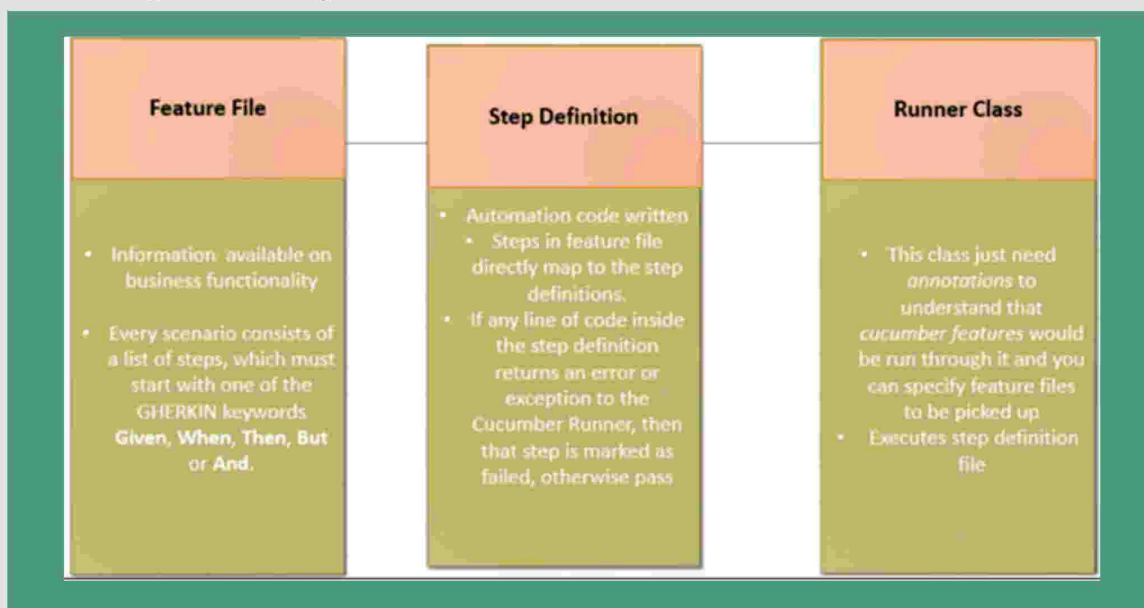
It's testing framework which supports Behavior Driven Development (BDD). It explains the behavior of the application in a simple English text using Gherkin language. Cucumber itself is written in Ruby. It supports various languages like ruby, java, .net etc.

## What is BDD?

BDD is abbreviated form of Behavior Driven development. In BDD we write tests first and the add application code. Tests are written in plain descriptive English. Tests are explained as behavior of application and are more user focused

## Description:

Cucumber executes feature file. Feature files are used to write test automation steps .The steps are application specific . Feature files contain executable specifications written in a language called Gherkin. In cucumber each line has to start with the gherkin keywords as described below.



## Feature:

This gives information about the high level business functionality and the purpose of Application under test. Provides high level description of software feature and groups related scenarios.

It has three basic elements –

- 1.Feature: keyword
- 2.Name
- 3.Description.

## Scenario:

Basically a scenario represents a particular functionality which is under test. It consists of a list of step . It can have as many steps, but its recommend you keep the number at 3-5 per scenario. Each scenario should follow given, when and then format.

## Steps:

**Given** - Given steps are used to describe the initial context of the system - specifies the pre-conditions.

**When** -When steps are used to describe an event, or an action.

**Then** -Then steps are used to describe an expected outcome, or result.

Example: Feature: Google Search

Scenario: User searches for "Accenture"

Given: user navigates to "http://www.google.co.uk"

And: user enters value "Accenture" in "Search Field"

When: user clicks on "Search Button"

## Step Definitions:

Step Definitions translates plain text Gherkin steps into actions that will interact with the system. When Cucumber executes a Step in a Scenario it will look for a matching Step Definition to execute. A Step Definition is a small piece of code with a pattern attached to it or in other words a step definition is a method in class with an annotation above it.



## Benefits:

- Involving stakeholders becomes easier regardless of their programming knowledge
- Testers can write Test scripts without having in-depth knowledge of programming
- Plugins are faster as compared to Selenium
- Supports various programming languages
- Code can be reused
- Simple and quick setup
- Flexible with different software platforms like selenium, Ruby on Rails, Watir, Spring framework, and so forth.

## Conclusion:

- Cucumber testing tool is a purely business-driven development tool written in Ruby
- The business-driven development approach is an advancement over test-driven development approach, which follows the
- 'Given-When-Then' steps for writing test cases
- Cucumber framework makes use of Gherkin ( A simple plain text language parser) to describe expected software behaviors in a logical language, which results in better communication and collaboration among technical and non-technical team members
- Cucumber provides compatibility with popular software platforms like Selenium, Watir Ruby, and other popular platforms.



# DESIGN & DEVELOP APP FOR NUTRITIONAL SECURITY & ELIMINATING ANAEMIA

## ABSTRACT:

**A**naemia is the most common Haematological disorder affecting humanity and is usually observed in chronic disease states such as non-specific anaemia, which may cause diagnostic difficulties. In chronically ill patients with anaemia, this has a negative impact on quality of life as well as survival. Mostly people don't know the nutritional food they are required what are the nutritional fruits and vegetables should have. The proposed work is an information portal which helps eradicate this problem. Our website is fully info based where we have given details of each food which are rich in protein, vitamins, carbohydrates etc. one can know the nutritional food they are required to consume every day. It serves as a gateway to reliable information on nutrition, healthy eating, physical activity, and food safety for consumers. This website also helps in building their diet and helping in maintaining their diet. The proposed work is being created using ReactJS, a JavaScript framework for frontend development which helps in creating a single page application. A single page application is comparatively faster and quicker to load as compared to traditional websites. For Backend, ExpressJS, a module for NodeJS is being used to handle the sever side scripting. MongoDB. A NoSQL database service handles the backend database in the proposed work.

## INTRODUCTION:

Anaemia is highly prevalent in India. The third National Family Health Study (NFHS-3) conducted during 2005–6 found that amongst children aged 6 to 59 months, the prevalence of anaemia is 69.5%; in rural India, the prevalence is 71.5%. The prevalence is maximal amongst younger children (12–17 months – 84.5%, 18– 23 months 81.6%). The prevalence of anaemia in rural areas appeared to have risen since the previous NFHS (in 1998–9) [1,2]. The proposed work serves as platform where a user can get info of food which are rich in nutrition components. When a user visits the portal, they can select a particular food item categorized in different sections like protein, vitamins, carbohydrates etc. It also shows dishes related to that food like paneer dish is palak paneer which is rich in iron and protein. Proposed Work Portal also shows information related to anaemia disease through which user can gain the information needed to diagnose himself/herself.

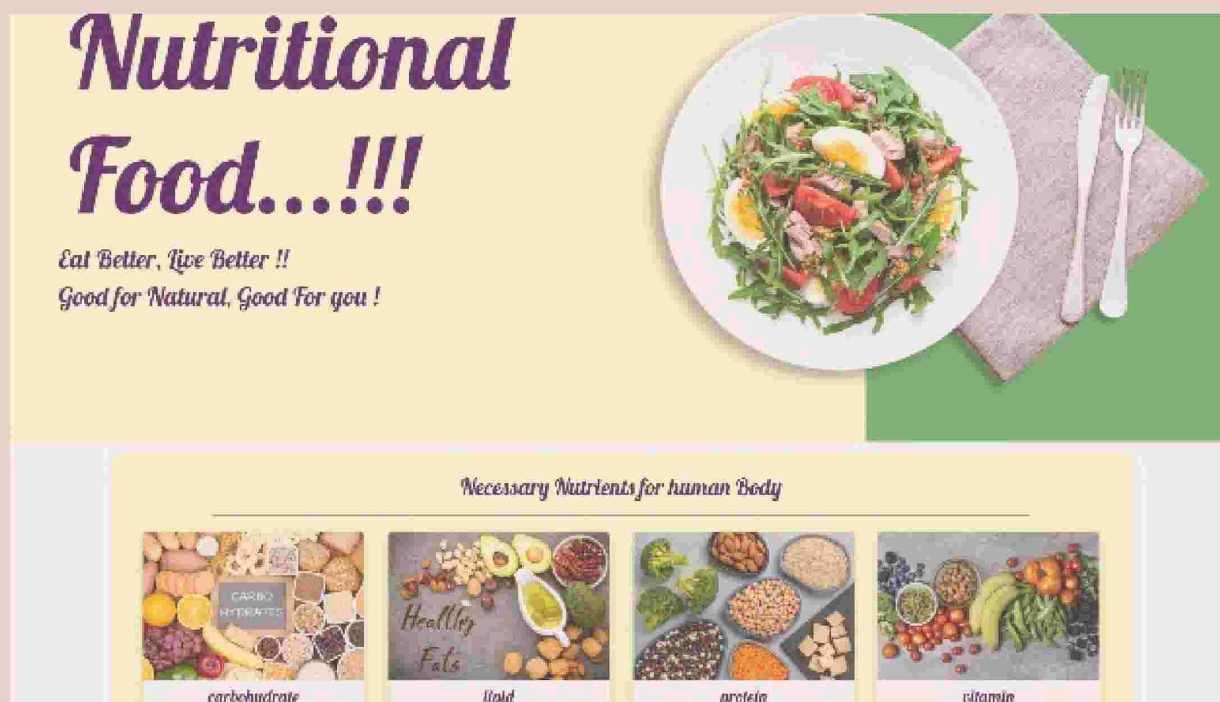
Suggestions and nutritional tips will also be provided for better self-treatment. For an enhanced treatment at home tutorial videos for consuming nutritional food and quality talks from doctors around the world in the field of anaemia is also a part of the proposed work. Facility to book an appointment with a doctor from top hospitals around the world is being included to help user in contacting doctors researching widely in the field of anaemia. As an addon a feature called “talk to us” is being provided so that a user can express his/her thoughts or can interact directly with experts in the field of anaemia is mostly caused due to lack of iron deficiency in children 11- 14 years.



## **LITERATURE SURVEY:**

Based on the survey [5,6] we found some research paper and found that there is no existing system for anaemia. There are some initiatives for the Anaemia projects, but the required system is not developed and there only research paper about the proposed work but no live system which helps user for guidance about anaemia and nutritional food.

Some Web apps are available for disease but the app is not specified to anaemia because it includes lots of disease. Similarly, some websites are there with old technologies which are not attractive and detailed about anaemia. So, after this survey we have taken initiative to develop an attractive and fully detailed and dynamic website for user regarding anaemia and nutritional food.



## **METHODOLOGY:**

### **Reason to created website: -**

Based on research what we found that Anaemia is most common disorder but there is no website in existing system to aware the people about anaemia. So, to make people aware about anaemia and advising them about symptoms, diagnosis, prevention about anaemia and guidance about the diet we initiated an approach to solve all issue regarding anomia using the website.



## How we built this project: -

At first, we created a design of all the component of websites using figma.

After completion of design, we started implementation of website using the MERN Stack. Here ReactJs is used for frontend implementation with some extra technologies like bootstrap, Sass, CSS, html. ExpressJs and Nodes are used for backend implementation and MongoDB used for Database Management .

All propped word is divided into 4 members where 2 members are involved in fronted development and other 2 members in backend development.

Side by Side members are involved in research and documentation for proposed work.

After implementation of each component the testing conducted each time and at last after final testing of whole project, the website is deployed on Heroku.

## Features of Website includes: -

- Anaemia Details: -

This component includes all types of Anaemia disease along with the particular anaemia details, which includes symptoms, diagnosis, prevention, treatment etc. Along with this there is recommendation for food diet, which is recommended based on anaemia type.

- Nutritional Guide:

This component includes all type of nutrients. Each nutrients contains many food items. Which contains all the details about the that food item. Which helps user to understand about nutrition diet.

- Hospital appointment: -

This component includes list of many hospitals where a user can take appointment in nearby hospital.

- Chatting System: -

When user is in doubt and want to ask any query to other user then they can use chatting system to take advice from other users.

- Services and guidance: -

Here we have shown services and also list of doctors which will be available for guidance if there is any query. A user can just call any doctor for is problems.

- Tutorials

We have added some tutorial about anaemia types with its all information and health tips.

- Parameters

This component includes all the parameter to check the Anaemia disease.



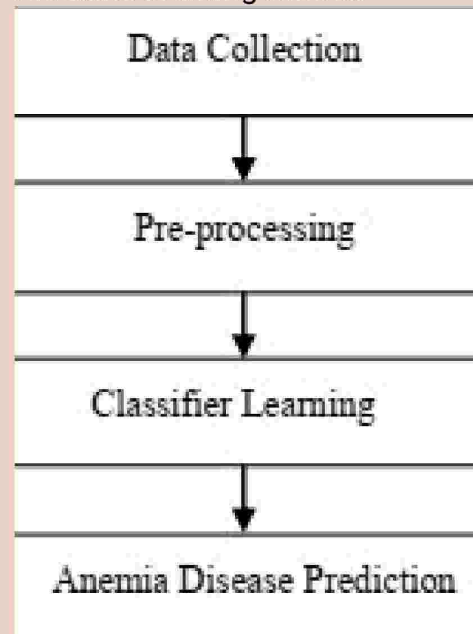
## PREDICTION:

As part of eliminating anaemia, a prediction model is necessary as it would help users to know if they are suffering from anaemia or not thereby also reducing the cost of visiting a pathogen lab for specialized test of anaemia which costs around 10 times than normal blood test. However, as anaemia is a disease which is caused due to decrease in the count of red blood cells (RBCs) or haemoglobin in the blood, it is required to have a blood report for the prediction model to work and predict the prevalence of this disease in a particular person.

When it comes to prediction model, there were three choices to choose from 1. Naïve Bayes Classifier 2. Random Forest 3. Decision Tree Classifier, out of which, Naïve Bayes Classifier yields the best result with the most accurate result among the three [7]. Therefore, Naïve Bayes classifier was chosen as the prediction algorithm to work with to create a machine learning model.

As far as data pre processing is concerned, training data is collected from various pathogen labs, as a result training data is new and fresh as opposed to getting old and inaccurate data from a data repository. The dataset collected from labs contains CBC test data which gives overall picture of the data that is required for the prediction algorithm i.e., Age, Gender, MCV, HCT, HGB, MCHC and RDW.

These are the steps taken for the successful creation of machine learning model:



Upon successful creation of the model, the model is then deployed for our API to hit and process the input data which would then predict the result and send back the response to the front end of the project which is then appropriately showed on the page.

Category Values	References
Men	> 13 g/dl
Women	> 12 g/dl
Pregnent Women	> 11 g/dl
Infants from 2-6 months	> 9.5 g/dl
Children from 6-24 months	> 10.5 g/dl
2 years to 11 years	> 11.5 g/dl



## RESULT:

The scope of the project is clear to give a simple and attractive application to simplify the work as well as to reduce the efforts while doing it offline or we can say by doing it with old methods. In this application we are able to save database of all patients present on the site. Login /registration Module: User can sign in to their account via login page or can create an account if they haven't registered before, by filling in the details asked on the registration page. Chat System Module: After login the use will be able to write a comment for asking any query and all user who are logged in can see and reply on the comment. Information Tree Module: a user can see the nutritional food card and after clicking on that card, he can see the all nutrients and after clicking on any nutrient user will be able to see all food rich in that nutrient then by clicking on any food user can see food details and the name of all recipe for that food. Similarly, user can follow same way to see anaemia types and their details. Appointment Module: A user can take appointment on any hospital after clicking on appointment button then user will be able to see official website for taking appointment. Hosting Server: A hosting server is generic term for a type of server that hosts or houses websites and/or related data, applications and services. It is a remotely accessible Internet server with complete Web server functionality and resources. Domain Name: A domain name is a string of text that maps to a numeric IP address, used to access a website from client software. In plain English, a domain name is the text that a user types into a browser window to reach a particular website. For instance, the domain name for Google is 'google.com'. The actual address of a website is a complex numerical IP address (e.g., 103.21.244.0), but thanks to DNS, users are able to enter human friendly domain names and be routed to the websites they are looking for. This process is known as a DNS lookup. Data Storage Facility: A database system provides an ultimate view of the stored data. However, data in the form of bits, bytes get stored in different storage devices. 4. Browser: A web browser is application software for accessing the World Wide Web. When a user follows the URL of a web page from a particular website, the web browser retrieves the necessary content from the website's web server and then displays the page on the user's device.

## FUTURE SCOPE:

The proposed work is being created using ReactJS, a JavaScript framework for frontend development which helps in creating a single page application. A single page application is comparatively faster and quicker to load as compared to traditional websites. For Backend, ExpressJS, a module for NodeJS is being used to handle the sever side scripting. MongoDB. A NoSQL database service handles the backend database in the proposed work. Eliminating anaemia is the main motto of our project which is also the major problem in this world deficiency of nutrients in red blood cell leads to anaemia. This is because lack of awareness people doesn't know what to eat and when to eat it's really important so that people have important and also defeat this anaemia. We have learnt lots of things about anaemia and how to make an effective website which help others to grab information.

In future, proposed work is planned to be shaped in an android app which will be made using Android Studio. A constructive design is already in creation which is planned to take on Material UI concept with dynamic theming support as present in android 13's Material You concept. Data will be fed in using the same API created for the website. Volley API will be used to fetch data from the API and will be populated accordingly. MVC architecture will be followed for the same. For now, Android Studio is in consideration but according to team's future decision Flutter might be used. If design of android app is not as we thought of then a switch to React Native will be required as designing is easier with React Native. An IOS app is also in consideration which will be made using X-Code. IOS app is still a decision which will require team's collective thinking and effort.



# CONCLUSION:

The proposed work is an information portal which helps eradicate this problem. Our website is fully info based where we have given details of each food which are rich in protein, vitamins, carbohydrates etc. one can know the nutritional food they are required to consume every day. It serves as a gateway to reliable information on nutrition, healthy eating, physical activity, and food safety for consumers. This website also helps in building their diet and helping in maintaining their diet for an enhanced treatment at home tutorial videos for consuming nutritional food and quality talks from doctors around the world in the field of anaemia is also a part of the proposed work. Facility to book an appointment with a doctor from top hospitals around the world is being included to help user in contacting doctors researching extensively in the field of anaemia. As an add-on a feature called "talk to us" is being provided so that a user can express his/her thoughts or can interact directly with experts in the field of anaemia. Also we are bringing the awareness among community to help understand the disease better and provide info regarding nutritional foods which plays a major role in applied field. Other additional features have been included in the proposed work to further improve the pace of eliminating this plaguing disease among the society.

**Farhan Patel**  
**Information Technology**  
**Thakur College of Engineering and**  
**Technology**  
**Mumbai, India**  
**farhanp9876@gmail.com**

---

**Jayraj Parki**  
**Information Technology**  
**Thakur College of Engineering and**  
**Technology**  
**Mumbai, India**  
**jayrajparkicet@gmail.com**

**Rohit Pandey**  
**Information Technology**  
**Thakur College of Engineering**  
**and Technology**  
**Mumbai, India**  
**rohitkp717@gmail.com**

---

**Eesha Pandit**  
**Information Technology**  
**Thakur College of Engineering and**  
**Technology**  
**Mumbai, India**  
**eeshapanditt@gmail.com**

# PREDICTION OF ANEMIA DISEASE USING MACHINE LEARNING ALGORITHMS

Aditya Dixit

IT Department Thakur College  
of Engineering and  
Technology Mumbai, India  
adityadixit01@gmail.com

Rahul Jha

IT Department Thakur  
College of Engineering and  
Technology Mumbai, India  
rjha27404@gmail.com

Raunak Mishra

IT Department Thakur College  
of Engineering and Technology  
Mumbai, India  
mishraraunak326@gmail.com

Dr. Sangeetha Vhatkar HOD - IT  
Department Thakur College of  
Engineering and Technology  
Mumbai, India  
sangeeta.vhatkar@thakureducat  
ion.org

## Abstract -

As we know, Red Blood Cells are the main part of blood that is responsible for the circulation of blood in the human body. Anemia is a well-known disease that is caused due to the deficiency of red blood cells. Due to anemia, red blood cells are unable to supply oxygen throughout the body. This sickness can be lethal to the human body if not treated promptly. We are creating machine learning techniques such as Random Forest, SVM, and others to detect anemia in a patient in this study. We can detect anemia in a patient using machine learning methods. As a result, we intend to create a classification-based ML model in which we provide the essential CBC test values for our model to predict whether a patient is anemic.

With the help of machine learning techniques, we are automating the process for detecting anemia in this study work. We compared the statistical analysis of all algorithms we've utilized to predict anemia in this paper. Keywords - Anemia, Machine Learning, Random Forest, SVM, Naive Bayes.



# I. INTRODUCTION

There has been an exponential increase in the data generated through the health industry because of the remarkable advances in Technology used. Using this data, We can extract all the useful information which can then be used for analysis, recommendation, prediction and decision making. In medical science, disease prediction at the right time is important for prevention and effective treatment plan. Anemia is a disease which is caused by the deficiency of red blood cells which are unable to deliver oxygen throughout the body.

Anemia is highly prevalent in India. The third National Family Health Study (NFHS-3) conducted during 2005–6 found that amongst children aged 6 to 59 months, the prevalence of anemia is 69.5%; in rural India, the prevalence is 71.5%. The prevalence of anemia is maximum among younger children between the age of 12-17 months and 18-23 months. The prevalence of anemia in rural areas appeared to have risen since the previous NFHS (in 1998–9).

Hence, it is important to take some measures to prevent the spread of anemia as much as possible using the latest advancements happening in the Tech Industry. In our study, we found out using various classifier algorithms like Random Forest, Decision Tree, Naïve Bayes etc. ,we can predict the early stage of anemia so that patients can take required medicine on time and prevent anemia. This project is important as, using the latest advancement in the field of machine learning we can

also make solutions in the field of medical science. This technology can be used in many areas like rural areas where health care systems are still not developed to the extent that of urban areas. Anemia is a disease, which needs timely treatment and early diagnosis, using machine learning we can achieve this. Machine Learning can help us overcome many different problems faced by our country in the field of medicine. Using this project, we will be able to detect whether a person or patient is suffering from anemia or not in a matter of seconds.

# II. PROBLEMS FACED

Anemia is a growing problem amongst young children living in rural India. In Rural areas, there is a lack of proper medical treatment and experienced doctors. This leads to patients traveling long distances to visit experienced doctors for treatment. This delay ultimately leads to the disease becoming more fatal. Also, many people avoid going to the doctor because they are scared or they can't afford it. Also, due to the lack of trained or experienced doctors in rural areas, they misdiagnose the symptoms resulting in Anemia becoming more fatal. Anemia, also goes quite unnoticed in many people especially children, which can go unnoticed at first but suddenly become fatal in nature. To identify this, a doctor needs to go through the CBC blood test report thoroughly to identify the early

stages of Anemia. Once identified, it is quite easy to cure the disease. To tackle all these problems, we are planning to create a Machine Learning Model, using which we would make use of multiple algorithms like Naive Bayes, Random Forest, SVM, etc. using which we will create a website, where the user can simply put in their blood test parameters in our machine learning model which would then predict whether the user is suffering from Anemia or not. Our machine learning model can predict and alert the user if the user is suffering from anemia and using which the user can be treated on time without the need of any experienced medical staff.

### III. METHODOLOGY

We followed the below methodologies to make our project: 1. Taking Input Data

- Firstly we collect the dataset
- Dataset should be in csv format.
- We import the dataset using various python libraries like Pandas.
- Above, in our dataset, we have considered five parameters - [3][5][6] 1. Gender - Gender is a very important parameter as the blood parameters and limits for both Male and Female are different and vary, so it is important to also consider this factor. 2. MCV - MCV stands for mean corpuscular volume. Basically this blood test measures the average size of the red blood cells. Using this test we can get to know whether our red blood cells are too small or too large which can depict any blood disorder such as anemia.

13. MCH - MCH is short for "mean corpuscular hemoglobin." It's the average amount in each of your red blood cells of a protein called hemoglobin, which carries oxygen around your body. [8] 4. MCHC - MCHC is a similar measure to MCH, MCHC stands for "mean corpuscular hemoglobin concentration". MCHC checks the average amount of hemoglobin in a group of red blood cells. A doctor might use both MCHC and MCH in order to diagnose Anemia. [9] 5. Hemoglobin - This parameter tells us about the amount of oxygen present in our blood. It is basically a protein which has the capacity to carry oxygen throughout the body from the lungs. It is also a very important parameter in prediction of anemia. For men, anemia is typically defined as a hemoglobin level of less than 13.5 g/dl and in women as hemoglobin of less than 12.0 g/dl.

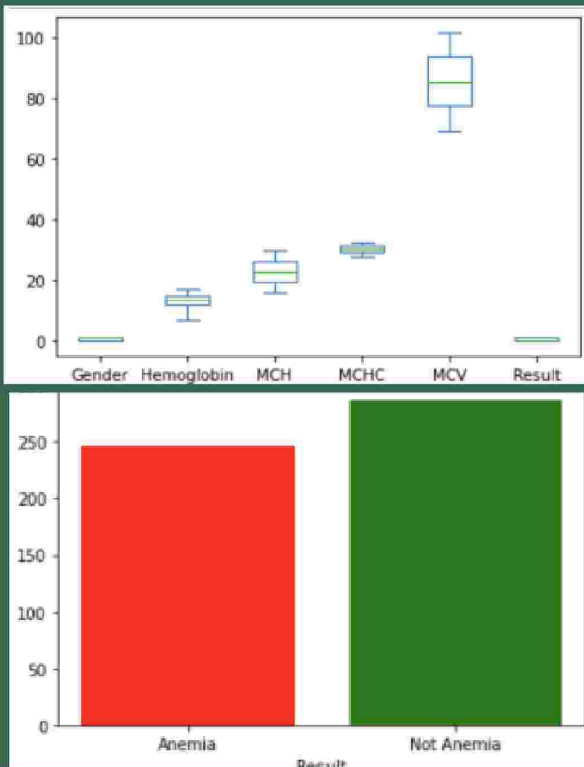
	Gender	Hemoglobin	MCH	MCHC	MCV	Result
0	Male	14.9	22.7	29.1	83.7	Not Anemia
1	Female	15.9	25.4	28.3	72.0	Not Anemia
2	Female	9.0	21.5	29.6	71.2	Anemia
3	Female	14.9	16.0	31.4	87.5	Not Anemia
4	Male	14.7	22.0	28.2	99.5	Not Anemia
5	Female	11.6	22.3	30.9	74.5	Anemia

2. Pre-Processing And Cleaning Dataset

- For data cleaning and preprocessing, we have imported the required dataset using the pandas dataset.
- After importing the dataset and making it a dataframe, we have first converted all values into integers. Checked for null values, we didn't find any null values in our dataset
- Next, we went ahead and checked all the number of entries and removed all duplicates.

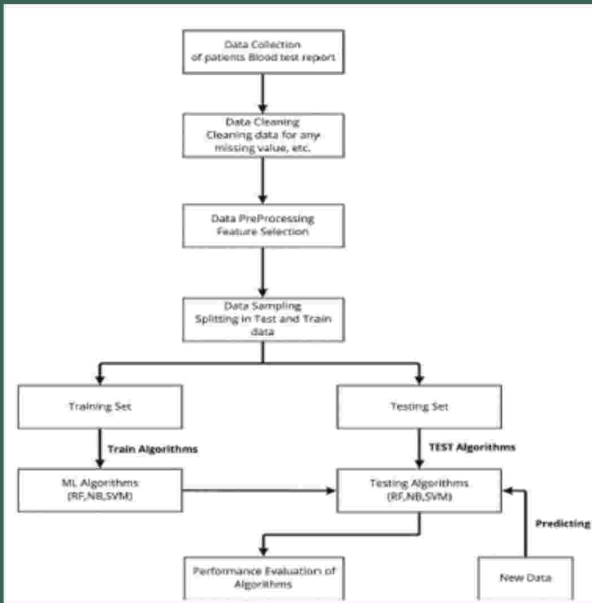


- Now, after cleaning the data, we went ahead for data visualization.



### 3. Feature Extraction/ Feature Selection

- As discussed above, we are using 5 features to predict whether a user/patient is suffering from anemia or not.
- We are using Gender, Hemoglobin, MCH, MCHC and MCV from the blood test reports to predict whether a user is suffering from anemia or not.



After cleaning all the data, we will then Normalize the data using MinMaxScaler. MinMaxScaler transforms all the features between 0 and 1. Here we extracted features that are required for model training.

### 4. Apply Classification Algorithms

- After feature extraction now comes to model training.
- First of all we have divided the dataset into training and testing using a method called `train_test_split()`. We have divided our dataset into a 75-25% train-test split.
- Now select the classification algorithm and import it from respective libraries.
- Algorithms that we are going to use are Random Forest, SVM, Naïve Bayes etc. [
- Below are the detailed study of our algorithms - 1. Naive Bayes Algorithm - Naive Bayes Algorithm is a supervised machine learning algorithm which is based on the famous bayes theorem and is used mostly to solve classification problems. It is one of the easiest and effective classification algorithms. It basically predicts the output based on the basis of the probability of the object. [12] Now, defining the formula as per our project - P(A|B) is Posterior probability: Probability of hypothesis A on the observed event B. P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true. P(A) is Prior Probability: Probability of hypothesis before observing the evidence.

P(B) is Marginal Probability: Probability of Evidence. [12]  $P(A|B) = P(B|A) P(A) / P(B)$  ... (1)

As per our problem, We define the formula: -  $P(\text{YES}|\text{Anemia})$  is Probability of having Anemia Disease in a person.  $P(\text{Anemia}|\text{YES})$  is the value of patients having parameters outside the normal range having anemia.  $P(\text{Anemia})$  is the value of people having Anemia.  $P(\text{YES})$  is the value of total people having blood parameters out of range. So, we can rewrite the Naive Bayes algorithm as -  $P(\text{YES}|\text{Anemia}) = P(\text{Anemia}|\text{YES}) * P(\text{Anemia}) / P(\text{YES})$  We will then compare this value with the normal or parameter of people not having Anemia -  $P(\text{NO}|\text{Anemia}) = P(\text{Anemia}|\text{NO}) * P(\text{Anemia}) / P(\text{NO})$  After this calculation, we will in the end compare both these, and the one greater will be the final answer - If  $P(\text{NO}|\text{Anemia})$  is greater than  $P(\text{YES}|\text{Anemia})$ , then the person is not suffering from Anemia, else vice-versa.

2. Random Forest - Random forest is a simple to use machine learning algorithm that delivers a good result much of the time, it also does not require us to use hyper parameter tuning. It is also one of the most commonly used algorithms due to its simplicity and versatility which can be used as both regression and classification algorithms. Why use Random Forest? Random Forest is one of the most popular machine learning algorithms used for both classification and Regression problems. It is used because of its speed, that is it works very fast even for very big datasets. It also provides a very high accuracy in comparison with the other machine learning algorithms. [11] How does the Random Forest algorithm work? Random Forest as the name suggests, is an algorithm created by the use of multiple decision trees.

In this, Random Forest algorithm creates multiple decision trees, and then as per the input, the decision tree shows the output. In random forest, the algorithm for a classification problem takes all the majority classes predicted by all decision trees and average of all predicted outputs for a Regression Problem.

Now, lets see the working of Random Forest Algorithm - Step-1: Firstly, we select random data points from the training data set. Step-2: Next, we build a decision tree for each of the respective data points. Step-3: Next, we decide the number of decision trees we want. Step-4: Repeat Step 1 & 2. Step-5: Now, for predicting, compile all the outputs of all decision trees and take the majority of all outputs for the final output.

3. SVM - Support Vector Machine is one of the best machine learning algorithms when it comes to classification problems. This is exactly what SVM does! It tries to find a line/hyperplane (in multidimensional space) that separates the two classes. It then classifies the new point as to whether it lies on the positive or negative side of the hyperplane, depending on the classes to be predicted. Steps to implement support vector regression in python :

- Import the library
  - Read the dataset
  - Feature Scaling
  - Fitting SVR to the dataset/
  - Predicting a new result
  - Visualizing the results of support vector regression
- Support vector regression is the counterpart of a support vector machine for regression problems. In our project we are using different attributes of the dataset and predicting the result using this support vector machine.



## 5. Real Time Implementation Of Project

- Here comes the main part where we have to map our project with the real world problems
- For this purpose we are trying to reach the various resource persons which are pathologists/doctors and provide them with solutions that our model is giving.
- We have decided to provide our service to NGOs or Social work bodies or organizations or medical bodies or rural clinics or hospitals where there is a lack of experienced medical staff.
- Patients can, on our website, just put in their blood test parameters and our machine learning model will predict whether the patient is suffering from anemia or not.

## IV. TECHNOLOGY USED

Technology and Tools that we are going to use in our project: We are using one of the most useful and powerful language ie. python. Python also has robust library support for Machine learning. 1. Google Collab – This is a jupyter notebook IDE where we can easily run and also see the output of each cell simultaneously. We will use Google Colab as it already has many of the required libraries installed. 2. Pandas - This is one of the most important libraries for data science applications. It is used for cleaning and perfecting our dataset before inserting it in the machine learning model. 3. Scikit – It is a machine learning library containing many models like classification, regression and clustering algorithms. It also has a metrics module which is used for checking the accuracy of the models.

4. Matplotlib – It is a library used for data analysis. It is a library used to create various types of graphs. 5. Seaborn – It is a library used for creating many types of graphs. 6. Flask - It is a library which we will use to create our website where the user enters their CBC parameters.

## V. RESULT AND DISCUSSION

- After implementation of all the above steps, we have come up with the accuracy we have achieved using the Random Forst, Naive Bayes and SVM algorithms -

Algorithm	Accuracy
Random Forest	99.38%
Naive Bayes	95.65%
SVM	97.52%

Table 1 - Algorithm Accuracy

- Above is the accuracy we have achieved from our algorithms after training them and then testing them with the test data.
- We have also below attached the True Positive, True Negative, False Positive and False Negative of each algorithm.

Algorithms	True Positive	True Negative	False Positive	False Negative
Random Forest	99	61	1	0
Naive Bayes	95	59	5	2
SVM	97	60	3	1

- Using TP, TN, FP, FN we have found the accuracy using the formula - Accuracy =  $(TP+TN)/(TP+TN+FP+FN) \dots(2)$  [14]
- As we see, our results are up to standards and the accuracy of each algorithm is very good, even exceeding our expectations.

# HIGH-RESOLUTION IMAGE SYNTHESIS WITH LATENT STABLE DIFFUSION MODELS

## *Introduction*

Image synthesis is one of the computer vision fields with the most spectacular recent development, but also among those with the greatest computational demands. Especially high-resolution synthesis of complex, natural scenes is presently dominated by scaling up likelihood-based models, potentially containing billions of parameters in autoregressive (AR) transformers. In contrast, the promising results of GANs have been revealed to be mostly confined to data with comparably limited variability as their adversarial learning procedure does not easily scale to modeling complex, multimodal distributions. Recently, diffusion models, which are built from a hierarchy of denoising autoencoders, have shown to achieve impressive

Figure 1. Boosting the upper bound on achievable quality with less aggressive downsampling.

Since diffusion models offer excellent inductive biases for spatial data, we do not need the heavy spatial downsampling of related generative models in latent space, but can still greatly reduce the dimensionality of the data via suitable autoencoding models, see Sec. 3. Images are from the DIV2K [1] validation set, evaluated at 5122 px. We denote the spatial downsampling factor by  $f$ . Reconstruction FIDs and PSNR are calculated on ImageNet-val.; see also Tab. 8. results in image synthesis and beyond, and define the state-of-the-art in class-conditional image synthesis and super-resolution. Moreover, even unconditional DMs can readily be applied to tasks such as inpainting and colorization or stroke-based synthesis, in contrast to other types of generative models. Being likelihood-based models, they do not exhibit mode-collapse and training instabilities as GANs and, by heavily



exploiting parameter sharing, they can model highly complex distributions of natural images without involving billions of parameters as in AR models. Democratizing High-Resolution Image Synthesis DMs belong to the class of likelihood-based models, whose mode-covering behavior makes them prone to spend excessive amounts of capacity (and thus compute resources) on modeling imperceptible details of the data. Although the reweighted variational objective aims to address this by under sampling the initial denoising steps, DMs are still computationally demanding, since training and evaluating such a model requires repeated function evaluations (and gradient computations) in the highdimensional space of RGB images. As an example, training the most powerful

from a hierarchy of denoising autoencoders, have shown to achieve impressive

Figure 1. Boosting the upper bound on achievable quality with less aggressive downsampling. Since diffusion models offer excellent inductive biases for spatial data, we do not need the heavy spatial downsampling of related generative models in latent space, but can still greatly reduce the dimensionality of the data via suitable autoencoding models, see Sec. 3. Images are from the DIV2K [1] validation set, evaluated at 5122 px. We denote the spatial downsampling factor by  $f$ . Reconstruction FIDs and PSNR are calculated on ImageNet-val.]; see also Tab. 8.

results in image synthesis and beyond, and define the state-of-the-art in class-conditional image synthesis and super-resolution. Moreover, even unconditional DMs can readily be applied to tasks such as inpainting and colorization

or stroke-based synthesis, in contrast to other types of generative models. Being likelihood-based models, they do not exhibit mode-collapse and training instabilities as GANs and, by heavily exploiting





sharing, they can model highly complex distributions of natural images without involving billions of parameters as in AR models. Democratizing High-Resolution Image Synthesis DMs belong to the class of likelihood-based models, whose mode-covering behavior makes them prone to spend excessive amounts of capacity (and thus compute resources) on modeling imperceptible details of the data. Although the reweighted variational objective aims to address this by under sampling the initial denoising steps, DMs are still computationally demanding, since training and evaluating such a model requires repeated function evaluations (and gradient computations) in the high dimensional space of RGB images. As an example, training the most powerful DMs often takes hundreds of GPU days (e.g. 150 - 1000 V100 days in) and repeated evaluations on a noisy version of the input space render also inference expensive, so that producing 50k samples takes approximately 5 days [15] on a single A100 GPU. This has two consequences for the research community and users in general:

Firstly, training such a model requires massive computational resources only available to a small fraction of the field, and leaves a huge carbon footprint. Secondly, evaluating an already trained model is also expensive in time and memory, since the same model architecture must run sequentially for a large number of steps

To increase the accessibility of this powerful model class and at the same time reduce its significant resource consumption, a method is needed that reduces the computational complexity for both training and sampling. Reducing the computational demands of DMs without impairing their performance is, therefore, key to enhance their accessibility.

**Departure to Latent Space** Our approach starts with the analysis of already trained diffusion models in pixel space: Fig. 2 shows the rate-distortion trade-off of a trained model. As with any likelihood-based model, learning can be roughly divided into two stages: First is a perceptual compression stage which removes high-frequency details but still learns little semantic variation. In the second stage, the



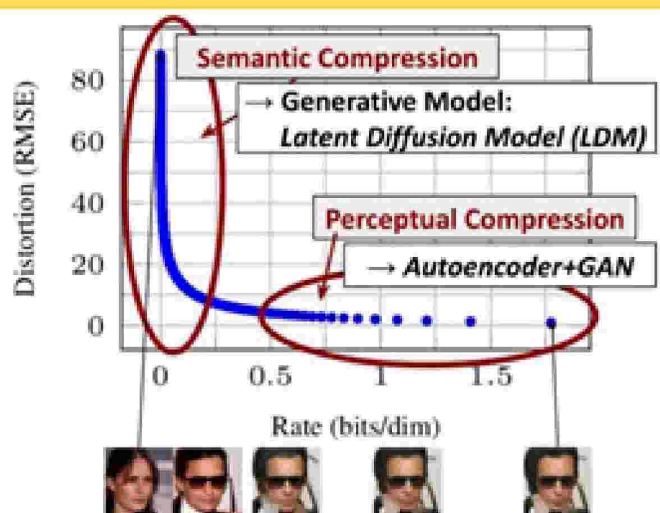
actual generative model learns the semantic and conceptual composition of the data (semantic compression). We thus aim to first find a perceptually equivalent, but computationally more suitable space, in which we will train diffusion models for high-resolution image synthesis.

Following common practice, we separate training into two distinct phases: First, we train an autoencoder which provides a lower-dimensional (and thereby efficient) representational space which is perceptually equivalent to the data space. Importantly, and in contrast to previous work, we do not need to rely on excessive spatial compression, as we train DMs in the learned latent space, which exhibits better scaling properties with respect to the spatial dimensionality. The reduced complexity also Following common practice, we separate training into two distinct phases: First, we train an autoencoder which provides a lower-dimensional (and thereby efficient) representational space which is perceptually equivalent to the data space. Importantly, and in contrast to previous work, we do not need to

rely on excessive spatial compression, as we train DMs in the learned latent space, which exhibits better scaling properties with respect to the spatial dimensionality. The reduced complexity also For the latter, we design an architecture that connects transformers to the DM's UNet backbone and enables arbitrary types of token-based conditioning mechanisms, see Sec. 3.3.

In sum, our work makes the following contributions:

- (i) In contrast to purely transformer-based approaches, our method scales more gracefully to higher dimensional data and can thus (a) work on a compression level which provides more faithful and detailed reconstructions than previous work (see Fig. 1) and (b) can be efficiently



**Figure 2. Illustrating perceptual and semantic compression: Most bits of a digital image correspond to imperceptible details. While DMs allow to suppress this semantically meaningless information by minimizing the responsible loss term, gradients (during training) and the neural network backbone (training and inference) still need to be evaluated on all pixels, leading to superfluous computations and unnecessarily expensive optimization and inference. We propose latent diffusion models (LDMs) as an effective generative model and a separate mild compression stage that only eliminates imperceptible details. Data and images from.**

**applied to high-resolution synthesis of megapixel images.**

**(i) We achieve competitive performance on multiple tasks (unconditional image synthesis, inpainting, stochastic super-resolution) and datasets while significantly lowering computational costs. Compared to pixel-based diffusion approaches, we also significantly decrease inference costs.**

**We show that, in contrast to previous work which learns both an encoder/decoder architecture and a score-based prior simultaneously, our approach does not require a delicate weighting of reconstruction and generative abilities. This ensures extremely faithful reconstructions and requires very little regularization of the latent space.**

**(i) We find that for densely conditioned tasks such as super-resolution, inpainting and semantic synthesis, our model can be applied in a convolutional fashion and render large, consistent images of 10242 px.**

**(ii) Moreover, we design a general-purpose conditioning mechanism based on cross-attention, enabling multi-modal training. We use it to train class-conditional, text-to-image and layout-to-image models. Finally, we release pretrained latent diffusion and autoencoding models at <https://github.com/CompVis/latent-diffusion> which might be reusable for a various tasks besides training of DMs**



## Method

To lower the computational demands of training diffusion models towards high-resolution image synthesis, we observe that although diffusion models allow to ignore perceptually irrelevant details by undersampling the corresponding loss terms, they still require costly function evaluations in pixel space, which causes huge demands in computation time and energy resources.

We propose to circumvent this drawback by introducing an explicit separation of the compressive from the generative learning phase (see Fig. 2). To achieve this, we utilize an autoencoding model which learns a space that is perceptually equivalent to the image space, but offers significantly reduced computational complexity. Such an approach offers several advantages: (i) By leaving the high-dimensional image space, we obtain DMs which are computationally much more efficient because sampling is performed on a low-dimensional space. (ii) We exploit the inductive bias of DMs inherited from their UNet architecture, which makes them particularly effective for data with spatial structure and therefore

alleviates the need for aggressive, quality-reducing compression levels as required by previous approaches. (iii) Finally, we obtain general-purpose compression models whose latent space can be used to train multiple generative models and which can also be utilized for other downstream applications such as single-image CLIP-guided synthesis.

## Perceptual Image Compression

Our perceptual compression model is based on previous work and consists of an autoencoder trained by combination of a perceptual loss and a patch-based adversarial objective. This ensures that the reconstructions are confined to the image manifold by enforcing local realism and avoids blurriness introduced by relying solely on pixel-space losses such as L2 or L1 objectives. More precisely, given an image  $x \in \mathbb{R}^{H \times W \times 3}$  in RGB space, the encoder  $E$  encodes  $x$  into a latent representation  $z = E(x)$ , and the decoder  $D$  reconstructs the image from the latent, giving  $\tilde{x} = D(z) = D(E(x))$ , where

different downsampling factors  $f = 2^m$ , with  $m \in \mathbb{N}$ .

In order to avoid arbitrarily high-variance latent spaces, we experiment with two different kinds of regularizations. The first variant, KL-reg., imposes a slight KL-penalty towards a standard normal on the learned latent, similar to a VAE, whereas VQ-reg. uses a vector quantization layer within the decoder. This model can be interpreted as a VQGAN but with the quantization layer absorbed by the decoder. Because our subsequent DM is designed to work with the two-dimensional structure of our learned latent space  $z = E(x)$ , we can use relatively mild compression rates and achieve very good reconstructions. This is in contrast to previous works, which relied on an arbitrary 1D ordering of the learned space  $z$  to model its distribution autoregressively and thereby ignored much of the inherent structure of  $z$ . Hence, our compression model preserves details of  $x$  better (see Tab. 8). The full objective and training details can be found in the supplement.

**3.2. Latent Diffusion Models**  
Diffusion Models are probabilistic

models designed to learn a data distribution  $p(x)$  by gradually denoising a normally distributed variable, which corresponds to learning the reverse process of a fixed Markov Chain of length  $T$ . For image synthesis, the most successful models rely on a reweighted variant of the variational lower bound on  $p(x)$ , which mirrors denoising score matching. These models can be interpreted as an equally weighted sequence of denoising autoencoders  $q_\theta(x_t, t)$ ;  $t = 1 \dots T$ , which are trained to predict a denoised variant of their input  $x_t$ , where  $x_t$  is a noisy version of the input  $x$ . The corresponding objective can be simplified to (Sec. A)

hi

$$\text{LDM} = \mathbb{E}_{x, q \sim N} \|q - q_\theta(x_t, t)\|_2^2, \quad (1)$$

(0,1),t  
with  $t$  uniformly sampled from  $\{1, \dots, T\}$ . Generative Modeling of Latent Representations With our trained perceptual compression models consisting of  $E$  and  $D$ , we now have access to an efficient, low-dimensional latent space in which high-frequency, imperceptible details are abstracted away. Compared to the high-dimensional pixel space,



this space is more suitable for likelihood-based generative models, as they can now (i) focus on the important, semantic bits of the data and (ii) train in a lower dimensional, computationally much more efficient space. Unlike previous work that relied on autoregressive, attention-based transformer models in a highly compressed, discrete latent space, we can take advantage of image-specific inductive biases that our model offers. This includes the ability to build the underlying UNet primarily from 2D convolutional layers, and further focusing the

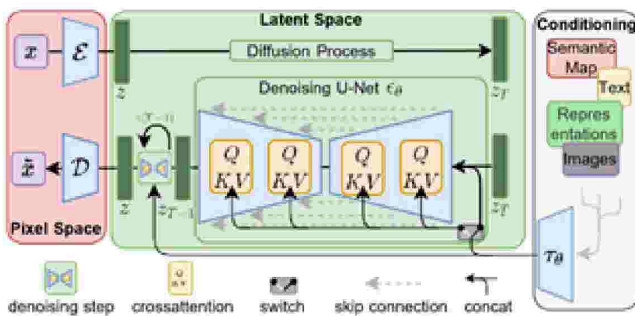


Figure 3. We condition LDMs either via concatenation or by a more general cross-attention mechanism. See Sec. 3.3 objective on the perceptually most relevant bits using the reweighted bound, which now reads

$$LLDM := \mathbb{E}_{\mathbb{E}(x), \varphi \sim \mathcal{N}(0,1), t} \left[ \| \varphi - \varphi_{\theta}(z_t, t) \|^2 \right] \quad (2)$$

The neural backbone  $\varphi_{\theta}(\cdot, t)$  of our model is realized as a time-conditional UNet. Since the forward process is fixed,  $z_t$  can be efficiently obtained from  $E$  during training, and samples from  $p(z)$  can be decoded to image space with a single pass through  $D$ .

### Conditioning Mechanism

Similar to other types of generative models, diffusion models are in principle capable of modeling conditional distributions of the form  $p(z|y)$ . This can be implemented with a conditional denoising autoencoder  $\varphi_{\theta}(z_t, t, y)$  and paves the way to controlling the synthesis process through inputs  $y$  such as text, semantic maps or other image-to-image translation tasks.

In the context of image synthesis, however, combining the generative power of DMs with other types of conditionings beyond class-labels or blurred variants of the input images is so far an under-explored area of research.

We turn DMs into more flexible conditional image generators by augmenting their underlying UNet backbone with the cross-attention mechanism, which is effective for learning attention-based models of various input modalities. To preprocess  $y$  from various modalities (such as language prompts) we introduce a domain specific encoder  $\tau_\theta$  that projects  $y$  to an intermediate representation  $\tau_\theta(y) \in \mathbb{R}^{M \times d_\tau}$ , which is then mapped to the intermediate layers of the UNet via a cross-attention layer implementing  $\text{Attention}(Q, K, V) = \text{softmax}(\sqrt{d} QK^T) \cdot V$ , with

$$Q = W_Q(i) \cdot \phi_i(z_t), K = W_K(i) \cdot \tau_\theta(y), V = W_V(i) \cdot \tau_\theta(y).$$

Here,  $\phi_i(z_t) \in \mathbb{R}^{N \times d_i}$  denotes a (flattened) intermediate representation of the UNet implementing  $\phi_\theta$  and  $W_Q(i) \in \mathbb{R}^{d \times d_i}$ ,  $W_K(i) \in \mathbb{R}^{d \times d_i}$  &  $W_V(i) \in \mathbb{R}^{d \times d_i}$  are learnable projection matrices. See Fig. 3 for a visual depiction

Figure 4. Samples from LDMs trained on CelebA HQ, FFHQ, LSUN-Churches, LSUN-Bedrooms and class-conditional ImageNet, each with a resolution of  $256 \times 256$ . Best viewed when zoomed in. For more

$$L_{LDM} := \mathbb{E}_{(x,y) \sim \mathcal{N}(0,1), t} \| \phi_\theta(z_t, t, \tau_\theta(y)) - x \|^2_2, \quad (3)$$

where both  $\tau_\theta$  and  $\phi_\theta$  are jointly optimized via Eq. 3. This conditioning mechanism is flexible as  $\tau_\theta$  can be parameterized with domain-specific experts, e.g. (unmasked) transformers when  $y$  are text prompts (see Sec. 4.3.1)

## Experiments

LDMs provide means to flexible and computationally tractable diffusion based image synthesis also including high-resolution generation of various image modalities, which we empirically show in the following. Firstly, however, we analyze the gains of our models compared to pixel-based diffusion models in both training and inference. Interestingly, we find that LDMs trained in VQ-regularized latent spaces achieve better sample quality, even though the reconstruction



capabilities of VQ-regularized first stage models slightly fall behind those of their continuous counterparts, cf. Tab. 8. Therefore, we evaluate VQ-regularized LDMs in the remainder of the paper, unless stated differently. A visual comparison between the effects of first stage regularization schemes on LDM training and their generalization abilities to resolutions higher than 2562 can be found in Appendix C.1. In D.2 we furthermore list details on architecture, implementation, training and evaluation for all results presented in this section.

#### 4.1. On Perceptual Compression Tradeoffs

This section analyzes the behavior of our LDMs with different downsampling factors  $f \in \{1, 2, 4, 8, 16, 32\}$  (abbreviated as LDM- $f$ , where LDM-1 corresponds to pixel-based DMs). To obtain a comparable test-field, we fix the computational resources to a single NVIDIA A100 for all experiments in this section and train all models for the same number of steps and with the same number of parameters.

Tab. 8 shows hyperparameters and reconstruction performance of the first stage models used for the

LDMs compared in this section. Fig. 5 shows sample quality as a function of training progress for 2M steps of class-Figure 5. Analyzing the training of class-conditional LDMs with different downsampling factors  $f$  over 2M train steps on the ImageNet dataset. Pixel-based LDM-1 requires substantially larger train times compared to models with larger downsampling factors (LDM- $\{4-16\}$ ). Too much perceptual compression as in LDM-32 limits the overall sample quality. All models are trained on a single NVIDIA A100 with the same computational budget. Results obtained with 100 DDIM steps and  $\kappa = 0$ .

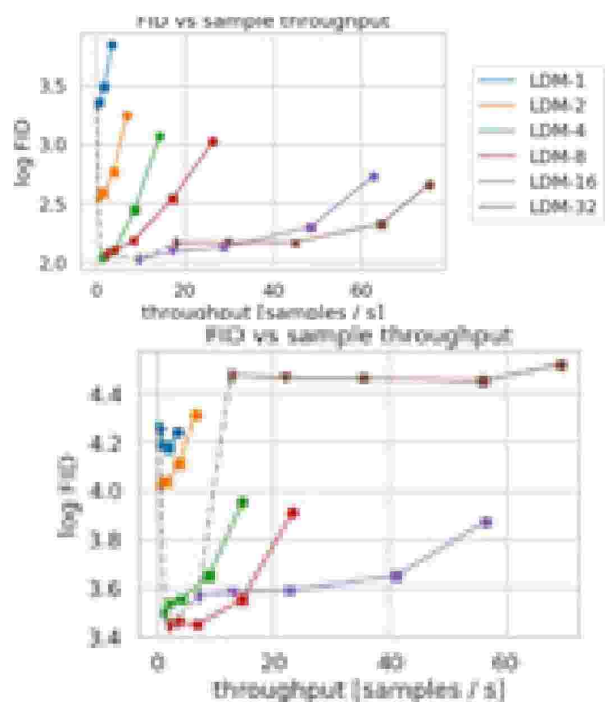


Figure 6. Inference speed vs sample quality: Comparing LDMs with different amounts of compression on the CelebA-HQ (left) and ImageNet (right) datasets. Different markers indicate {10, 20, 50, 100, 200} sampling steps with the DDIM sampler, counted from right to left along each line. The dashed line shows the FID scores for 200 steps, indicating the strong performance of LDM-{4-8} compared to models with different compression ratios. FID scores assessed on 5000 samples. All models were trained for 500k (CelebA) / 2M (ImageNet) steps on an A100.

models on the ImageNet [12] dataset. We see that, i) small downsampling factors for LDM-{1,2} result in slow training progress, whereas ii) overly large values of  $f$  cause stagnating fidelity after comparably few training steps. Revisiting the analysis above (Fig. 1 and 2) we attribute this to i) leaving most of perceptual compression to the diffusion model and ii) too strong first stage compression resulting in information loss and thus limiting the achievable quality. LDM-{4-16}

strike a good balance between efficiency and perceptually faithful results, which manifests in a sig-

CelebA-HQ 256 × 256				FFHQ 256 × 256			
Method	FID ↓	Prec. ↑	Recall ↑	Method	FID ↓	Prec. ↑	Recall ↑
DC-VAE [65]	15.8	-	-	StyleGAN2 [40]	5.57	-	-
VQGAN+T [23] (d=400)	10.2	-	-	StyleGAN2 [40]	10.9 (7.6)	-	-
PGGAN [18]	8.0	-	-	LDM [43]	5.54	-	-
LSGM [90]	7.22	-	-	StyleGAN2 [40]	4.16	-	-
LDM [43]	7.16	-	-	StyleGAN2 [40]	-	0.71	-
					3.08	0.65	-
LSUN-Churches (ours, 500k)	5.11	0.72	0.49	LSUN-Churches (ours, 200k)	4.98	0.73	0.50
LSUN-Churches 256 × 256				LSUN-Bedroom 256 × 256			
Method	FID ↓	Prec. ↑	Recall ↑	Method	FID ↓	Prec. ↑	Recall ↑
DDPM [39]	7.89	-	-	StyleGAN2 [40]	5.51	-	-
StyleGAN2 [40]	7.32	-	-	DDPM [39]	4.9	-	-
PGGAN [18]	6.42	-	-	LDM [43]	4.57	-	-
StyleGAN2 [40]	4.21	-	-	StyleGAN2 [40]	2.35	0.59	0.48
StyleGAN2 [40]	3.86	-	-	ADM [13]	1.90	0.68	0.52
StyleGAN2 [40]	1.59	0.61	0.44	StyleGAN2 [40]	1.52	0.61	0.34
LDM-F (ours, 200k)	4.02	0.64	0.52	LDM-F (ours, 200k)	2.95	0.66	0.48

Table 1. Evaluation metrics for unconditional image synthesis. CelebA-HQ results reproduced from [43], FFHQ from [39]. †: N -s refers to N sampling steps with the DDIM sampler. ‡: trained in KL-regularized latent space. Additional results can be found in the supplementary.

nificant FID gap of 38 between pixel-based diffusion (LDM-1) and LDM-8 after 2M training steps.

In Fig. 6, we compare models trained on CelebA-HQ and ImageNet in terms of sampling speed for different numbers of denoising steps with the DDIM sampler and plot it against FID-scores. LDM-{4-8} outperform models with unsuitable ratios of perceptual and conceptual compression. Especially compared



to pixel-based LDM-1, they achieve much lower FID scores while simultaneously significantly increasing sample throughput. Complex datasets such as ImageNet require reduced compression rates to avoid reducing quality. Summarized, we observe that LDM-4 and -8 lie in the best behaved regime for achieving high-quality synthesis results.

Text-Conditional Image Synthesis					
	DALL-E' [84]	CogVox' [17]	Latte' [105]	LDM-KL-8	LDM-KL-8-G
FID ↓	27.50	27.10	26.94	23.35	12.61
IS ↑	17.90	18.20	26.02	19.93 <sub>(0.35)</sub>	26.62 <sub>(0.38)</sub>

**Table 2. Evaluation of text-conditional image synthesis on the MS-COCO dataset:** Our model outperforms autoregressive and GAN-based methods by a significant margin when using 250 DDIM steps. †: Numbers taken from. : Classifier-free guidance, scale 1.5. **4.2. Image Generation with Latent Diffusion**

We train unconditional models of 2562 images on CelebA-HQ, FFHQ, LSUN-Churches and

-Bedroom and evaluate the i) sample quality and ii) their coverage of the data manifold using ii) FID and ii) Precision-and-Recall. Tab. 1 summarizes our results. On

CelebA-HQ, we report a new state-of-the-art FID of 5.11, outperforming previous likelihood-based models as well as GANs. We also outperform LSGM where a latent diffusion model is trained jointly together with the first stage. In contrast, we train diffusion models in a fixed space and avoid the difficulty of weighing reconstruction quality against learning the prior over the latent space, see Fig. 1-2. We outperform prior diffusion based approaches on all but the LSUN-Bedrooms dataset, where our score is close



**Figure 7. Top:** Samples from our LDM for layout-to-image synthesis on COCO]. Quantitative evaluation in the supplement. **Bottom:** Samples from our text-to-image LDM model for user-defined text prompts, which is trained on LAION-400M.

to ADM, despite utilizing half its parameters and requiring 4-times less train resources (see Appendix D.3.5). Moreover, LDMs consistently improve upon GAN-based methods in Precision and Recall, thus confirming the advantages of their mode-covering likelihood-based training objective over adversarial approaches. In Fig. 4 we also show qualitative results on each dataset.

### 4.3. Conditional Latent Diffusion

#### 4.3.1 Transformer Encoders for LDMs

By introducing cross-attention based conditioning into LDMs we open them up for various conditioning modalities previously unexplored for diffusion models. For text-to-image image modeling, we train a 1.45B parameter model conditioned on language prompts on LAION-400M. We employ the BERT-tokenizer and implement  $\tau\theta$  as a transformer to infer a latent code which is mapped into the UNet via cross-attention (Sec. 3.3). This combination of domain specific experts for learning a language representation and visual synthesis results in a powerful model, which generalizes well to complex, user-defined text prompts, cf. Fig. 7 and

. For quantitative analysis, we follow prior work and evaluate text-to-image generation on the MS-COCO validation set, where our model improves upon powerful ARand GAN-based methods, cf. Tab. 2. We note that applying classifier-free diffusion guidance greatly boosts sample quality. To further analyze the flexibility of the cross-attention based conditioning mechanism we also train models to synthesize images based on semantic layouts on OpenImages], and finetune on COCO, see Fig. 7. See Sec. C.4 for the quantitative evaluation and implementation details. Lastly, following prior work, we evaluate our best-performing class-conditional ImageNet models with  $f \in \{4, 8\}$  from Sec. 4.1 in Tab. 3, Fig. 4 and Sec. C.5. Here we outperform the state of the art diffusion model ADM while significantly reducing computational requirements and parameter count, cf. Tab 18.

Method	FID <sub>1</sub>	IS <sub>1</sub>	Precision	Recall	Params	
OpenAI-clip [1]	-	-	0.87	0.28	340M	-
ADM [10]	6.95	203.4 <sub>±2</sub>	0.69	0.83	554M	250 DDIM steps
ADM-G [10]	4.52	186.3	0.82	0.52	408M	250 DDIM steps
LDM-4 (ours)	10.56	103.9 <sub>±0.5</sub>	0.71	0.62	400M	250 DDIM steps
LDM-4-G (ours)	3.60	247.6 <sub>±0.5</sub>	0.87	0.48	400M	250 steps, classifier-free guidance [31], scale 1.5



Table 3. Comparison of a class-conditional ImageNet LDM with recent state-of-the-art methods for class-conditional image generation on the ImageNet dataset. A more detailed comparison with additional baselines can be found in C.5, Tab. 10 and E.

**4.3.2 Convolutional Sampling Beyond 2562** By concatenating spatially aligned conditioning information to the input of  $q\theta$ , LDMs can serve as efficient general-purpose image-to-image translation models. We use this to train models for semantic synthesis, super-resolution (Sec. 4.4) and inpainting (Sec. 4.5). For semantic synthesis, we use images of landscapes paired with semantic maps and concatenate downsampled versions of the semantic maps with the latent image representation of a  $f = 4$  model (VQ-reg., see Tab. 8). We train on an input resolution of 2562 (crops from 3842) but find that our model generalizes to larger resolutions and can generate images up to the megapixel regime when evaluated in a convolutional manner (see Fig. 8). We exploit this behavior to also apply the super-resolution models in Sec. 4.4 and

the inpainting models in Sec. 4.5 to generate large images between 5122 and 10242. For this application, the signal-to-noise ratio (induced by the scale of the latent space) significantly affects the results. In Sec. C.1 we illustrate this when learning an LDM on (i) the latent space as provided by a  $f = 4$  model (KL-reg., see Tab. 8), and (ii) a rescaled version, scaled by the component-wise standard deviation.

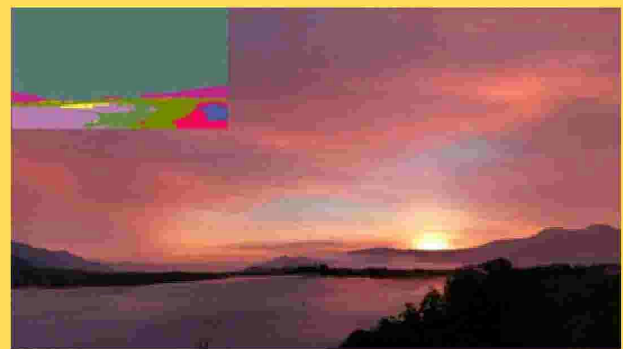


Figure 8. A LDM trained on 2562 resolution can generalize to larger resolution (here: 512×1024) for spatially conditioned tasks such as semantic synthesis of landscape images. See Sec. 4.3.2.

#### 4.4. Super Resolution with Latent Diffusion

LDMs can be efficiently trained for super-resolution by directly conditioning on low-resolution images via concatenated bicubic LDM-SR SR3



Figure 9. ImageNet 64→256 super-resolution on ImageNet-Val. LDM-SR has advantages at rendering realistic textures but SR3 can synthesize more coherent fine structures. See appendix for additional samples and cropouts. SR3 results from.

nation (cf . Sec. 3.3). In a first experiment, we follow SR3 and fix the image degradation to a bicubic interpolation with 4x-downsampling and train on ImageNet following SR3’s data processing pipeline. We use the  $f = 4$  autoencoding model pretrained on OpenImages (VQ-reg., cf . Tab. 8) and concatenate the low-resolution conditioning  $y$  and the inputs to the UNet, i.e.  $\tau\theta$  is the identity. Our qualitative and quantitative results (see Fig. 9 and Tab. 4) show competitive performance and LDM-SR outperforms SR3 in FID while SR3

has a better IS. A simple image regression model achieves the highest PSNR and SSIM scores; however these metrics do not align well with human perception and favor blurriness over imperfectly aligned high frequency details. Further, we conduct a user study comparing the pixel-baseline with LDM-SR. We follow SR3 where human subjects were shown a low-res image in between two high-res images and asked for preference. The results in Tab. 5 affirm the good performance of LDM-SR. PSNR and SSIM can be pushed by using a post-hoc guiding mechanism and we implement this image-based guider via a perceptual loss, see Sec. C.7. Since the bicubic degradation process does not generalize well to images which do not follow this pre-processing, we also train a generic model, LDM-BSR, by using more diverse degradation. The results are shown in Sec. C.7.1.

Method	FID ↓	IS ↑	PSNR ↑	SSIM ↑	$\Delta_{\text{perceptual}}$	Throughput (image/s)
Image Regression [10]	15.2	121.1	27.9	0.801	625M	N/A
SR3 [10]	5.3	180.1	26.4	0.762	625M	N/A
LDM-4 (ours, 100 steps)	2.8 $\pm$ 0.2	166.3	24.4 $\pm$ 1	0.69 $\pm$ 0.14	169M	4.62
LDM-4 (ours, big, 100 steps)	2.4 $\pm$ 0.3	174.9	24.7 $\pm$ 1	0.71 $\pm$ 0.15	552M	4.5
LDM-4 (ours, 50 steps, guiding)	4.4 $\pm$ 0.4	153.7	25.8 $\pm$ 1	0.74 $\pm$ 0.12	184M	0.38



Table 4.  $\times 4$  upscaling results on ImageNet-Val. (2562); † : FID features computed on validation split, ‡: FID features computed on train split; \*: Assessed on a NVIDIA A100

#### 4.5. Inpainting with Latent Diffusion

Inpainting is the task of filling masked regions of an image with new content either because parts of the image are

SR on ImageNet Inpainting on Places

User Study		LAMA [85]	LDM-4
Task 1: Preference vs GT †		13.6%	21.0%
	Pixel-DM (f1)		LDM-4

etween two generated images. More details in D.3.6

Model (reg.-type)	train throughput samples/sec.	sampling throughput @256	train val. hours/epoch	FID@2k epoch.6
LDM-1 (no first stage)	0.11	0.26	0.07	24.74
LDM-4 (KL, w/ attn)	0.32	0.97	0.34	7.66
LDM-4 (VQ, w/ attn)	0.33	0.97	0.34	7.04
LDM-4 (VQ, w/o attn)	0.35	0.99	0.36	6.66

Table 6. Assessing inpainting efficiency. †: Deviations from Fig. 6 due to varying GPU settings/batch sizes cf . the supplement.

are corrupted or to replace existing but undesired content within the image. We evaluate how our general approach for conditional image generation compares to more specialized, state-of-the-art approaches for this task. Our evaluation follows the protocol of LaMa, a recent inpainting model that introduces a specialized architecture

relying on Fast Fourier Convolutions. We describe the exact training & evaluation protocol on Places in Sec. D.2.2. We first analyze the effect of different design choices for the first stage.

We compare the inpainting efficiency of LDM-1 (i.e. a pixel-based conditional DM) with LDM-4, for both KL and VQ regularizations, as well as VQ-LDM-4 without any attention in the first stage (see Tab. 8), where the latter reduces GPU memory for decoding at high resolutions. For comparability, we fix the number of parameters for all models. Tab. 6 reports the training and sampling throughput at resolution 2562 and 5122, the total training time in hours per epoch and the FID score on the validation split after six epochs. Overall, we observe a

Task 2: Preference Score †	16.0%	30.4%	31.9%	68.1%
	29.4%	70.6%		

Table 5. Task 1: Subjects were shown ground truth and generated image and asked for preference. Task 2: Subjects had to decide

Method	40-50% masked		All samples	
	FID ↓	LPIPS ↓	FID ↓	LPIPS ↓
<i>LDM-4</i> (ours, big, w/ ft)	9.39	0.246± 0.042	1.50	0.137± 0.080
<i>LDM-4</i> (ours, big, w/o ft)	12.89	0.257± 0.047	2.40	0.142± 0.085
<i>LDM-4</i> (ours, w/ attn)	11.87	0.257± 0.042	2.15	0.144± 0.084
<i>LDM-4</i> (ours, w/o attn)	12.60	0.259± 0.041	2.37	0.145± 0.084

0speed-up of at least 2.7× between pixel- and latent-based diffusion models while improving FID scores by a factor of at least 1.6×.

The comparison with other inpainting approaches in Tab. 7 shows that our model with attention improves the overall image quality as measured by FID over that of. LPIPS between the unmasked images and our samples is slightly higher than that of. We attribute this to only producing a single result which tends to recover more of an average image compared to the diverse results produced by our LDM cf . Fig. 20. Additionally in a user study

(Tab. 5) human subjects favor our results over those of. Based on these initial results, we also trained a larger dif- fusion model (big in Tab. 7) in the latent space of the VQ- regularized first stage without attention. Following, the UNet of this diffusion model uses attention layers on three levels of its feature

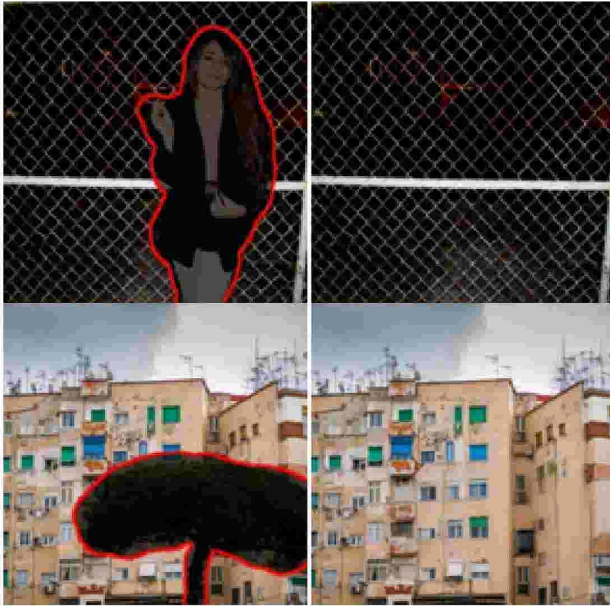
hierarchy, the BigGAN residual block for up- and downsampling and has 387M parameters instead of 215M. After training, we noticed a discrepancy in the quality of samples produced at resolutions 2562 and 5122, which we hypothesize to be caused by the additional attention modules. However, fine-tuning the model for half

LaMa +	12.31	0.243± 0.038	2.23	0.134± 0.080
LaMa	12.0	0.24	2.21	0.14
CoModGAN[	10.4	0.26	1.82	
0.15	21.3			
RegionWise	0.27	4.75	0.15	DeepFill v2 [22.1
EdgeConnect	30.5	0.28	8.37	0.16

Table 7. Comparison of inpainting performance on 30k crops of size 512 × 512 from test images of Places. The column 40- 50% reports metrics computed over hard examples where 40-50% of the image region have to be inpainted. Recomputed on our test set, since the original test set used in was not available.

an epoch at resolution 5122 allows the model to adjust to the new feature statistics and sets a new state of the art FID on image inpainting (big, w/o attn, w/ ft in Tab. 7, Fig. 10.).





**Figure 10.** Qualitative results on object removal with our big, w/ ft inpainting model. For more results, see Fig. 21.

## **Conclusion**

We have presented latent diffusion models, a simple and efficient way to significantly improve both the training and sampling efficiency of denoising diffusion models without degrading their quality. Based on this and our cross-attention conditioning mechanism, our experiments could demonstrate favorable results compared to state-of-the-art methods across a wide range of conditional image synthesis tasks without task-specific architectures.

## **CREDITS**

- 1. Saket Pradhan**
- 2. Raj Shah**
- 3. Ranveer Shah**

# POSTERS AND SKETCHES



# Route optimization for Ambulance!!

## Objective:-

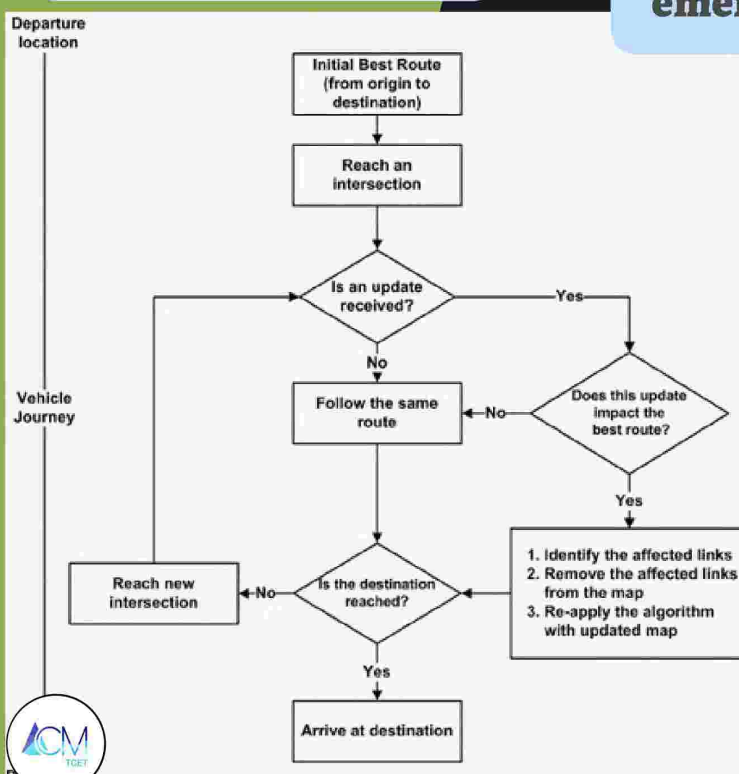
we aim to enhance the response-time performance of emergency medical service providers by handling the ambulance routing problem (ARP). The problem can be modelled as either the Open Vehicle Routing Problem (OVRP) or a Vehicle Routing Problem with Pickup and Delivery (VRPPD).

## Scope:-

The efficient management of ambulance routing for emergency requests is vital to save patients when a disaster response scenario occurs. In today's road traffic, the transportation of patients from emergency points becomes more difficult in such tragic situation.

## Algorithm:-

- Particle swarm
- Ant Colony
- OPTimization
- A\* bidirectional



BY:-ITB

33\_EKTA SINGH

58\_DEV YADAV

66\_SHALOO SINGH

## OBJECTIVE

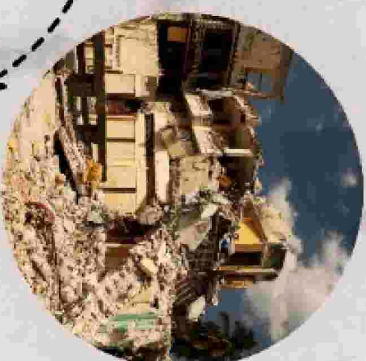
To minimize the event's impact, something that involves preparedness, response, recovery and mitigation.

## SCOPE

To build an automated system and to overcome First Come First Serve criteria. And to contact nearby NGOs as soon as possible.

## SOFTWARE USED

Frontend: Flutter/ React  
 Backend: Firebase  
 Database: MongoDB  
 Others: API



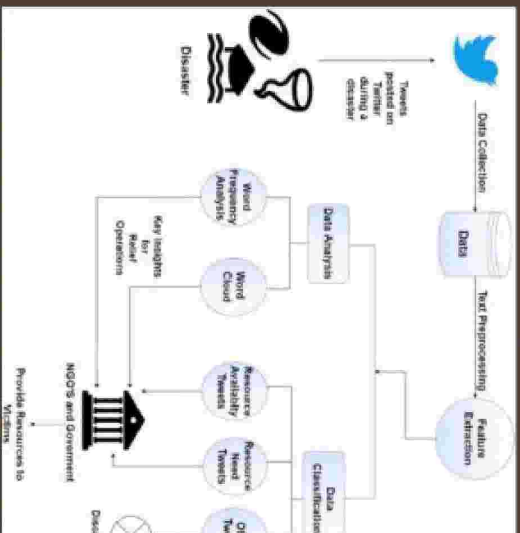
# DISASTER MANAGEMENT

## EARTHQUAKE

## FLOW



## METHODOLOGY



The task is to make an intelligence to attend calls and get details from call, analyse it and decide the action

## CONCLUSION

### GUIDE NAME

Anil Vasoya



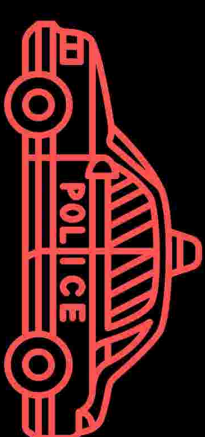
## NAME

42\_Mamishkaran Sundram  
 43\_Nikhil Suthar  
 44\_Chaitra Suvarna





# Crime Modelling to Interpret and Generate Behavioural Analysis of Crimes



## Objective

1. To build the predictive policing model/solution to help the police department to reduce the crime rates with the help of the FIR/Investigation Data.
2. To Show the potential threat/Crime areas in a different zones so that the precaution should be taken before hand.
3. To enhance the use of resources like patrolling and surveillance in the police department to fight against crime in a better way.

## Scope:

1. FIR Data/Investigation data on the full potential for the decision making process.
2. To promote Data Driven Decision making for patrolling which would be helpful to reduce crime rates

## Conclusion:

The main use case of the application is in the police department to make the better decision making and also to use the human resources in more efficient manner like for patrolling and surveillance purpose.

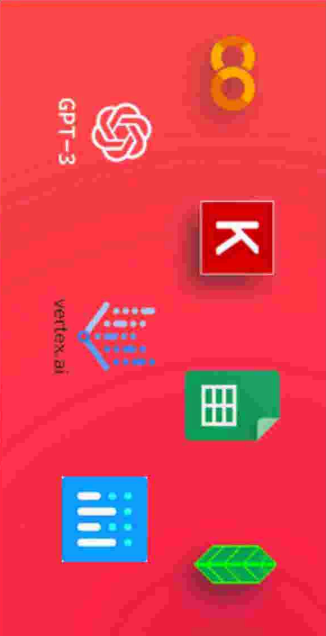
## Methodology

The main focus of methodology is using the previous crime records/FIR Data for predictive policing purpose which will ultimately help to reduce the crime rates.

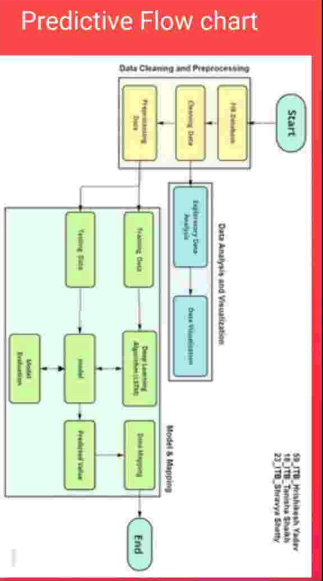
1. For the predictive policing, the model using is LSTM as the data of previous crime records have the timestamp which is helpful to instruct the future potential crime spots with different risk tonality.

2. For the better decision making according to the previous record, GPT 3 transformer is used to take out the intuition from the data which will guide police department for efficient decision making.

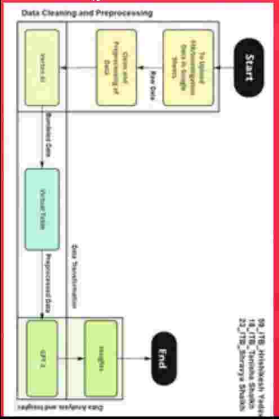
## Technology Used:



## Block Diagram



## Analysis Work Flow chart



## Results & Outcomes:

1. The predictive policing helps police to know which area are more prone to which crime also with the risk tonality so that the action can be taken according to that.
2. The analysis done on the previous crime records and converting it into the intuition would be helpful to understand the behaviour of the crime records and also for planning to reduce crime rates.

## Program Outcome Attained:

- PO 1 PO 2 PO3 PO4 PO5 PO6
- PO8 PO9 PO10 PO 11

## Guide Name:

Mr. Sudhir Dhekane



# DEPARTMENT OF INFORMATION TECHNOLOGY

## FERTILIZER PREDICTION & COST ANALYSIS

### INTRODUCTION

- Nutrients to the plants are provided through soil, manure and the fertilizers. There are 16 essential nutrients that are required for the plant growth viz., carbon, hydrogen, oxygen, nitrogen, phosphorous, potassium, Sulphur, calcium, magnesium, boron, iron, manganese, molybdenum, zinc, chlorine and copper.
- These are either provided through soil or water in the form of micronutrients and macronutrients. The deficiency of these nutrients can be found out by using soil testing mechanism

### OBJECTIVE

- The main aim of our project is to suggest farmers appropriate fertilizer considering soil conditions.
- It suggests amount of fertilizer required according to land.
- It predicts total cost require for fertilization.

### TECHNOLOGY USED

Machine Learning algorithm like Random forest, SVM and Webscraping.

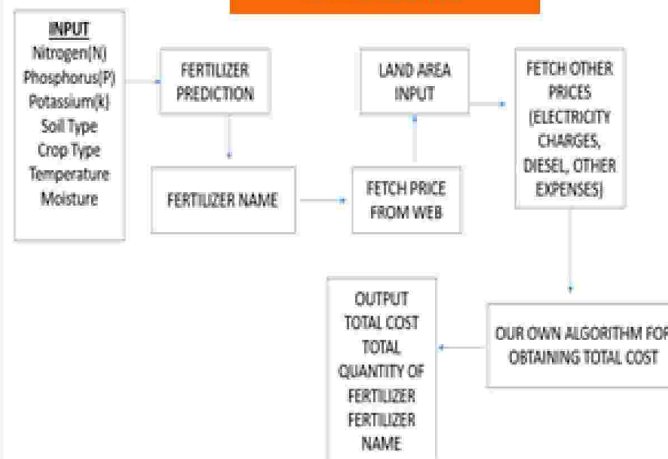
### METHODOLOGY

So, our project agendas are simple. Firstly, we will take seven input field and then we will predict best fertilizer and with the help of web scraping, we will determine the current price of that predicted fertilizer and we will also take input of farmer land so that the amount of fertilizer can be calculated. Also, we will fetch the current price of electricity, water and diesel and with the help of own algorithm that we are making will be able to calculate the total amount or price that farmer has to pay and we will display them.

### RESULT & OUTCOME

Based on the algorithm we will use, we are able to measure the best fertilizer and also its current price will be calculated using web scraping and also the total cost that farmer has to pay for specific amount of land and fertilizer.

### BLOCK DIAGRAM



### CONCLUSION

- The prediction of fertilizer is based on parameters as mineral level, soil quality, crop type, moisture, temperature, rainfall which increases accuracy of prediction.
- By suggesting them approximate amount of fertilizer required prevents excess use and wastage of fertilizer. This system considers all expenses for fertilization process and gives total cost of fertilization process. This helps farmers to manage their project.

### REFERENCES

[https://www.researchgate.net/publication/332675671\\_Prediction\\_of\\_Crop\\_Fertilizer\\_Consumption](https://www.researchgate.net/publication/332675671_Prediction_of_Crop_Fertilizer_Consumption)  
<https://www.ijeast.com/papers/371-376,Tesma405,IJEAST.pdf>



BY GROUP-8

03-IT-A-HARSHIT BAJAJ  
21-IT-A-VEDANT GOSAWI  
30-IT-A-SUMIT JAIN

GUIDE NAME: MR. SUDHIR DHEKANE

100





# DEPARTMENT OF INFORMATION TECHNOLOGY

## AI IN AGRICULTURE

### OBJECTIVES

The objective of this project is to implement the use of AI in agriculture so that we can use that in agriculture as it will help in predicting the yield of production, plant disease and even crop recommendation

### SCOPE

The scope of this project is vast as we can use this in India on a large scale as India's arable land area of 159.7 million hectares (39.4, 6 million acres) is the second largest in the world, after the United States.

### TECHNOLOGIES TO BE USED

1. Algorithm
2. Python
3. HTML/CSS
4. JAVASCRIPT
5. Backend

### APPLICATIONS

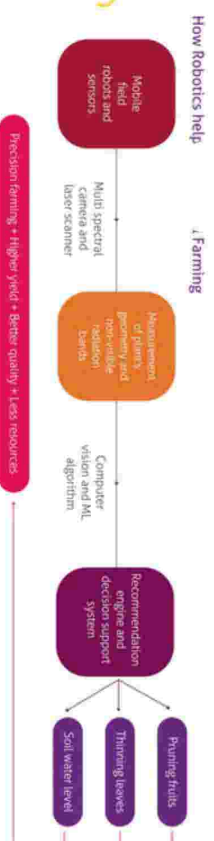
- Weather & price Forecasting
- Health Monitoring of Crops
- Agriculture Robotics
- Intelligent Spraying
- Disease Diagnosis
- Precision Farming

### STUDENT NAME

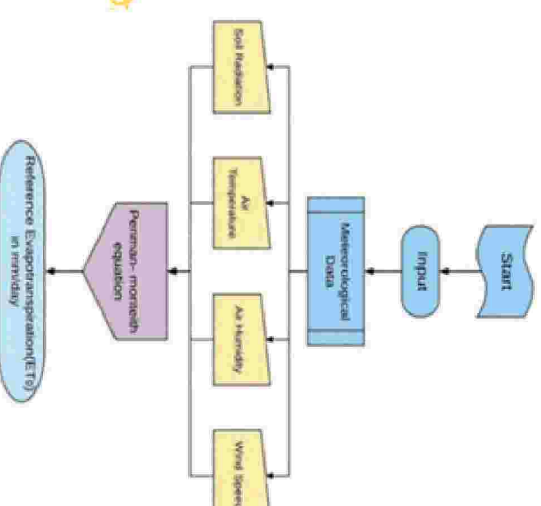
08\_JTB\_Brijraj Singh Rana  
21\_JTB\_Avaneesh Shetti  
35\_JTB\_Khushi Singh

**GUIDE NAME**  
KAMAL SHAH

### BLOCK DIAGRAM



### METHODOLOGY



### CONCLUSION

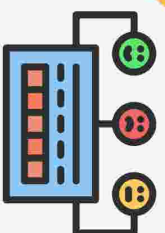
The agricultural industry faces various challenges such as lack of effective irrigation systems, weeds, issues with plant monitoring due to crop height and extreme weather conditions. But the performance can be increased with the aid of technology and thus these problems can be solved. It can be improved with different AI driven techniques like remote sensors for soil moisture content detection and a d d a sub heading to mated irrigation with the help of GPS.

### RESULT AND OUTCOMES

- Analyzing market demand
- Managing risk
- Breeding seeds
- Monitoring soil health
- Feeding crops
- Protecting crops
- Harvesting

# SENTIMENT ANALYSIS FOR CODE-MIXED LANGUAGES USING DEEP LEARNING

Keywords— Sentiment Analysis, Siamese Networks, Code- Mixed Text



**MEMBERS**  
11. TE, IFA, Preet Desai  
12. TE, IFA, Ronak Dhadak  
13. TE, IFA, Santosh Deshi

## 01. INTRODUCTION

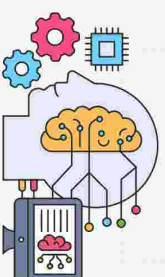
Social media platforms were widely accepted in multilingual societies with a respectable internet penetration rate. The use of code-mixed text increased as a result. Political campaigns, product reviews, marketing, and other social phenomena can all be closely examined using sentiment analysis of code-mixed data from social media platforms. Text that has been code-mixed frequently develops new structures dependent on its users and adopts the vocabulary and grammar of several different languages. Since standard approaches to semantic analysis do not capture the meaning of the sentences, this makes sentiment analysis difficult.

## 02. SIAMESE NETWORKS

Without being overly reliant on the inputs' attributes, siamese networks aid in the contrastive learning of a similarity metric. siamese networks were developed to address the issue of signature verification. Lints

## 05. ANALYSIS

- The semantics of the words are captured by the distributional semantics technique, but the information about their order in the sentence is lost. The approach also believes a word to be unchangeable, which is one of its limitations. As a result, it cannot handle spelling mistakes or words that are not in the lexicon. polarity scores are given to specific words. The polarity of the sentence is determined by the sentiment score of all the individual words. As a result, the words' order and meaning relationship are lost, which causes inaccurate classification. N-grams restrict this issue but do not entirely cure it.
- Another line of research uses character level LSTMs to extract information from social media text at the sub word level. Using an annotated corpus and this information, the sentences are then classified. The model offers a useful strategy for embedding sentences. Abbreviations and Acronyms



### Related Literature

We use the datasets to evaluate the architecture on code- mixed (Hindi-English) data as well as text from social media in a standard language (English). The datasets we used for our texts are listed below. 3879 annotated Hindi-English Code-Mixed utterances make up the dataset, which was proposed in [10303]. annotated English tweets make up the dataset, which was proposed in [10303]. Senteval 2013: The dataset for 1st, 2nd of SemEval 2013 is made up of Architecture of SAcMT

## 03. METHODOLOGY

- An exercise in classification is being conducted. We take into account the Hindi-English CodeMixed (HECM) sentences and designate them 1 after aligning them with the English sentences from the Twitter datasets of the same mood (positive samples). Additionally, we randomly select an equal number of English lines with various moods (negative samples), assigning them the number -1. We examine the benefits of training Hindi-English CodeMixed data alongside English phrases using this model (SACMT(Hi-Eng)).
- Additionally, we create the input data by labelling each HECM sentence 1 after aligning it with corresponding HECM sentences of the same emotion (positive samples). The same procedure is used to randomly select an equal number of HECM statements with various sentiments, labelling them as "-1."

## 04. RESULTS/FINDINGS

In this study, we present SAcMT, which solves the issue by projecting the sentences into a shared sentiment space. For sentiment analysis of code-mixed text. For the purpose of capturing a sentiment-based representation of the sentences, SAcMT uses twin Bidirectional LSTM networks with shared parameters. Sentences with similar sentiment were grouped together using these sentiment-based representations and a similarity metric. Studies done on the datasets show that SAcMT performs noticeably better than the cutting-edge techniques. SAcMT makes use of resources from various languages to enhance sentiment analysis on material that has been code-mixed.

## 06. CONCLUSION

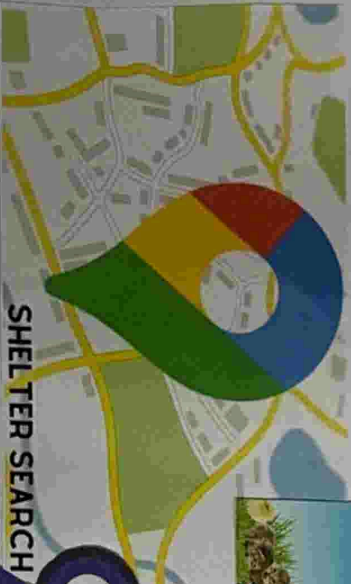
The management of word variants further improved the performance of all trained models, even baseline models. When used with SAcMT, an emoji-based strategy improves the performance of overall sentiment.





# OBJECTIVES

- Stray animals are facing problems as they have been deprived of the proper care and attention they need to grow up healthy.
- There is a need of a web service that is of the people, for the people, by the people that want to help the stray animals encountered in their daily lives.
- There is lack of a single platform to provide all services related to stray animals.
- To provide shelter, adoption, feeding and health care facilities for people that wish to help the strays.



SHELTER SEARCH

# SCOPE

- This web service can be used for all the stray animals in the future.
- It can be used to create a community that is dedicated to serve stray animals.
- Instant help can be provided by using this web service.



ADOPTION

# PAW PEACE

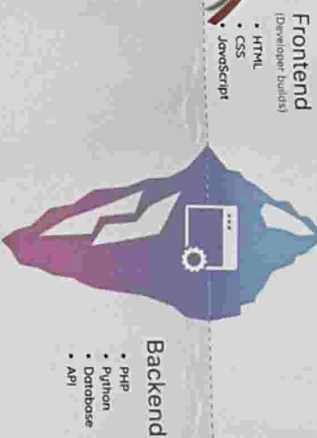
## RESEARCH METHODOLOGY

- The research methodology we used in this project is quantitative as well as qualitative methodology
- We studied about Factors affecting stray count, Are shelters our best option?
- We also analyzed the ratio of strays to shelters. Are the shelters outnumbered by the strays?

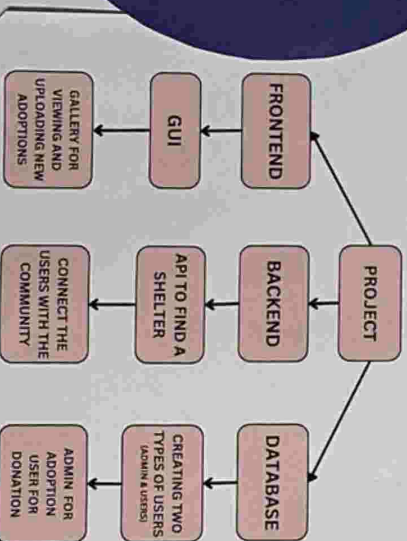


THE DOG COMMUNITY

## SOFTWARE USED



## BLOCK DIAGRAM



## RESULT

- There is a lack of shelters and service providers for stray animals
- Our project is a website for all the stray animals to receive food and shelter through the website.
- This project basically will bring the whole community of animal savors together.

## CONCLUSION

- This web service can be used for all the stray animals in the future.
- It can be used to create a community that is dedicated to serve stray animals.
- Instant help can be provided by using this web service.

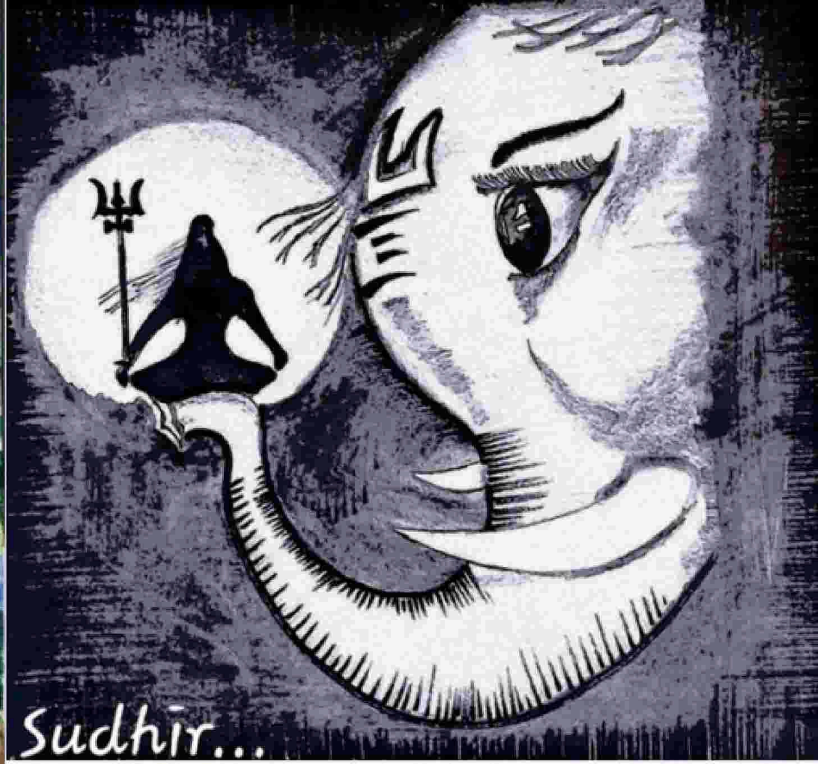
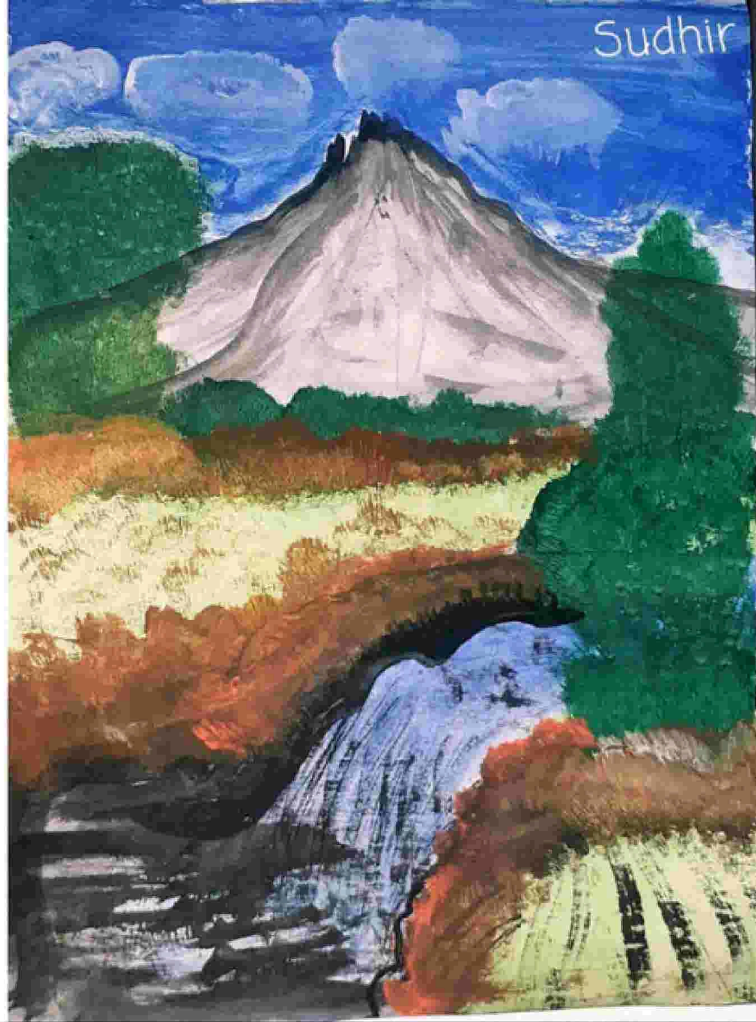
### TE-IT-A

- 02 - KHAIRAH AHIR
- 13 - HEMIN DHAMELJA
- 33 - DHRUVI TIVANI



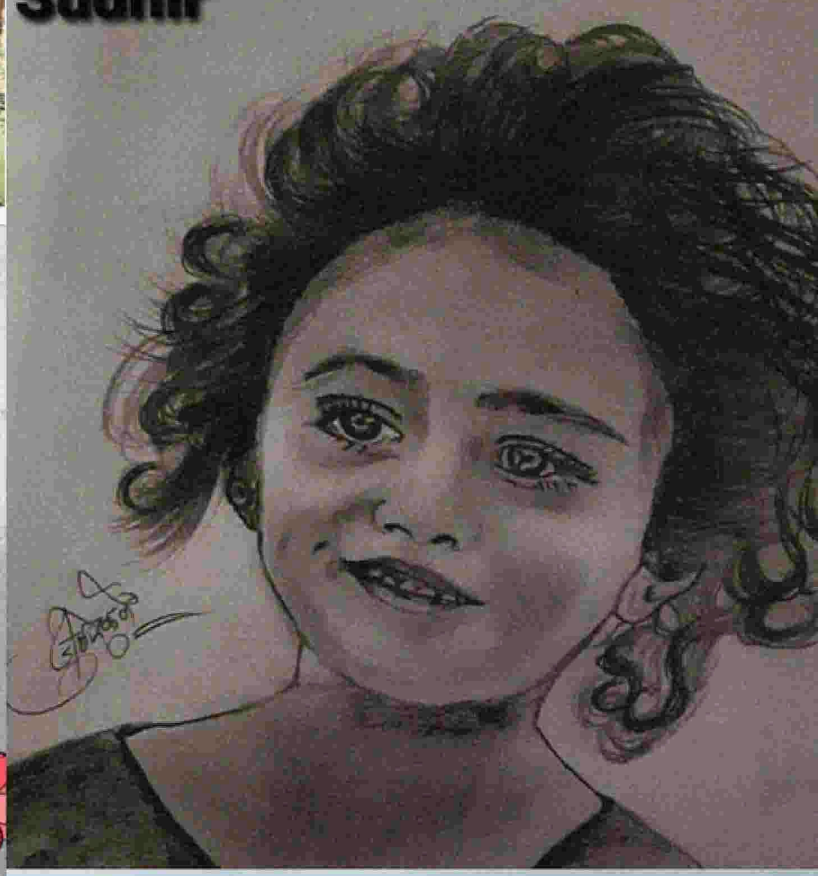


Sudhir

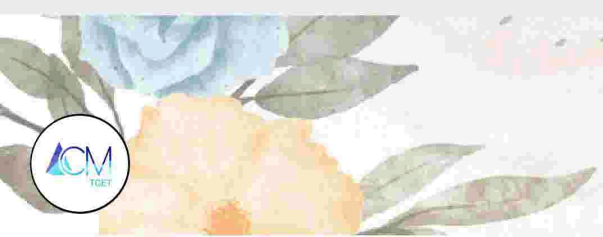
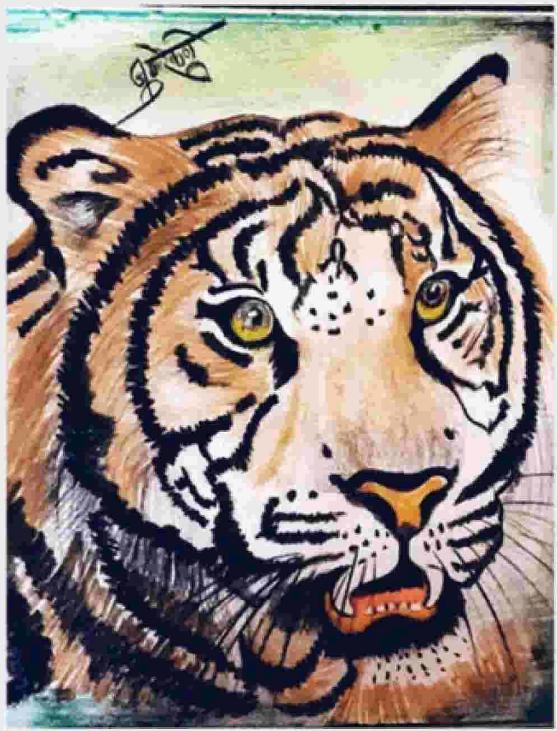
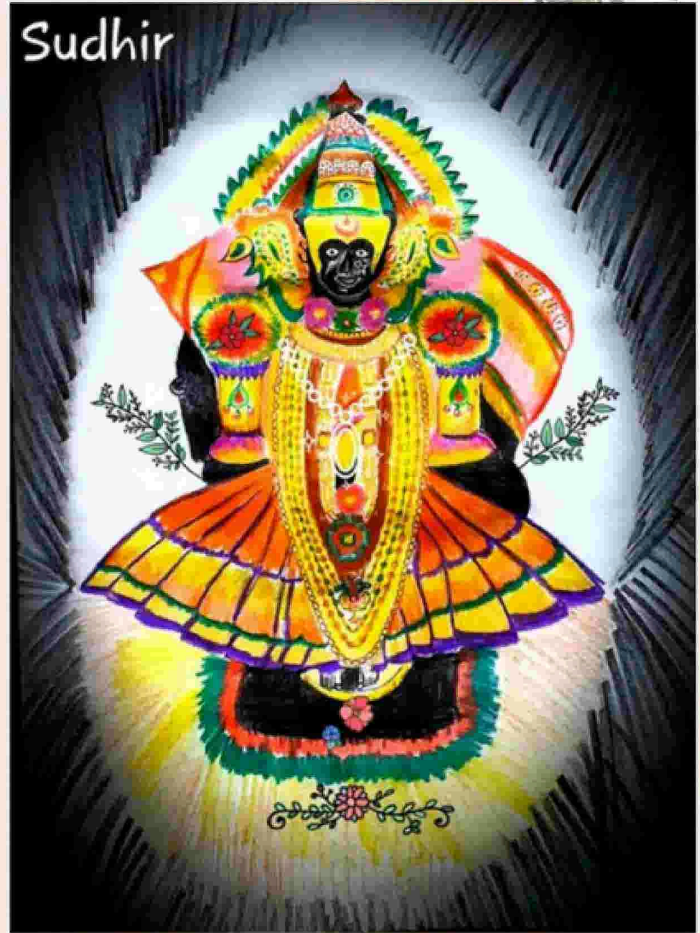


Sudhir...  
Sudhir

Sudhir







# POETRY



# TODAY'S NARI

Who am I?

Resilient, Independent, Calm or shy...

I wonder and think I am the sky...

I am beautiful and infinite.

Wiggling from caterpillar to a butterfly...

I choose my way and so alive...

Serene on the surface but turbulent inside.

With the sound of the raindrops, and sun setting by...

Reminding myself

Holding cup of Chai...

I am the Queen of my world,

Important, confident & bold...

Yes, I am today's independent Bhartiya Nari,

Sometimes Cultured, sometimes Not-So-Sanskari...

**BY AMATEUR**

**POETESS**

**MS.APEKSHA**

**WAGHMARE**

वेळ

जीवन जगाया लावशी तू कळ,  
कधी देशी, अन कधी काढून घेशी हे बळ  
विलक्षण वेळेचा कुणा न कळे हा खेळ .....

सुखाच्या क्षणी वेगात पळशी, दुखाःत मात्र मंद होशी  
म्हणून अवघड होते, घालणं तुझा मेळ  
विलक्षण वेळेचा कुणा न कळे हा खेळ .....

भूत, भविष्यात घालुनी पापपुण्यांचा मेळ  
वर्तमानात गाठायला लावशी तळ  
विलक्षण वेळेचा कुणा न कळे हा खेळ .....

सुधिर



# EZINE COMMITTEE

## PUBLICATIONS TEAM



**ATUL PAL**  
**VICE CHAIRPERSON**



**KRISHI CHAURASYA**  
**MAIN EDITOR**



**ADITYA DESAI**



**HARSH MISHRA**

# EZINE COMMITTEE

## CREATIVES TEAM



**DHYEY SWADIA**  
**DESIGN DIRECTOR**



**SEJAL NALAWADE**



**KEENJHAL GUPTA**



**SHAIVI PURANIK**



**AAYUSH TRIPATHI**



**VISHAL ANTIYA**



**RAMNARAYAN  
TIWARI**



# CODE OF ETHICS

The Department of Information and Technology of TCET believe's that Engineers make a direct impact on almost all aspects of Human Life for its betterment. IT Engineers should strictly adhere to the high principles of ethical conduct. In order to inculcate high standards in professional behaviour, the department advocates the following code of ethics for all the students, Faculty members & Staff of the department.

- Strive to be professionally competent to provide high quality product and services.
- To responsibly make decisions, minimising hazards to society and to disclose potential factors that maybe a threat to health and society.
- Be fair to all individuals and not discriminate between individual based on religion, race, sex, age, disability, national origin, etc.
- Give credits to contribution of other's viz. copyrights, patent, intellectual property, etc.
- Protect and respect privacy and ensure confidentiality of information whenever appropriate.
- The knowledge gained during the course of study will not be misused for carrying out any illegal activities, intruding and hacking of networks.