Research Article

Automatic Question Paper Generation using ML: A Review

Kanchan Babaji Dhomse[1], Ishan Ranjan[2]

[1]Research Scholars, SGVU, Rajsthan, India.
[2]Dean of SGVU, Rajsthan, India.
kdhomse@gmail.com[1], iranjan63@gmail.com[2]

Article History: Received: 10 November 2020; Revised: 12 January 2021; Accepted: 27 January 2021; Published online: 05 April 2021

Abstract: In any instructive field educational programs, the courses are characterized with learning goals. It is exceptionally challenging for the instructors to create Question Papers with changed questions. There's no standardizing strategy to guarantee quality of question paper. Consequently, there arises a have to be have a framework which can consequently produce it from we entered information inside few seconds. It's basic to specify tests and exams, it plans understudies in their journey for information, so having a legitimate examination paper & organize is essential. Now the traditional strategy of producing address paper has been manual and it is exceptionally time and cost expending to diminish these things we developed this framework which is able make the programmed address paper and this it can be valuable to numerous instructive establishing.

Keywords: Bloom's Taxonomy algorithm, Colony algorithm, Science, Map reduce techniques, Sensor etc

1. Introduction

A understudy appraisal could be a pivotal portion of guideline and is done through the process of examination and arrangement of address papers. Show day innovations help the instructor to stock the questions in a computer databases but the issue which happened is how the display day advances would moreover help the instructor to consequently make assortment sets of questions without stress almost replication and duplication from past exams. Instruction is the foremost imperative way of accomplishing victory. When we examine instruction having an appropriate examination paper and arrange is mandatory. Now the routine strategy of creating address paper has been manual. We have proposed a robotized handle of address paper generation which is quick, elegant, disarranged & secure. We proposed this framework to induce effectively working on making programmed address papers with less exertion and with less man control. We have proposed a taking after algorithm.

- Bloom's algorithm: - N. Omar, S. Haris, R. Hassan, H. Arshad, M. Rahmat, N. Zainal and R. Zulkifli. [20] Bloom's scientific categorization may be a set of three progressive models utilized to classify educator learning goals into levels of complexity and specificity. There are 3 records cover the learning targets in cognitive, overall feeling & tactile spaces.
- Colony algorithm: - In computer science and operations investigate; the insect colony optimization calculation could be a probabilistic strategy for understanding computational issues which can be diminished to 21 finding great ways through charts. Manufactured Ants stand for multi-agent strategies motivated by the conduct of genuine ants.

Any item that can successfully diminish time and control utilization is acknowledged and acknowledged. In this way, we are showing a computerized address creator of paper framework which is able to decrease time utilization to supplanting the ordinary strategy to their address paper era framework. Our framework points to supply quick operations, information capacity and tall security for all its tasks. The proposed framework can be accommodating to numerous instructive establishing and competitive exam related organizing. The Address Paper Generator framework consequently creates paper, plans doc records as per chosen paper arrange. Send to the all institutes by email. After that address paper is changed over into Pdf format organize record and messaged to institutes by pressing the button.

2. Literature Survey

a) Ontology: - Philosophy is an express determination of a conceptualization. Metaphysics could be a formal, unequivocal determination of a shared conceptualization. A 'conceptualization' alludes to an unique demonstrate of a few marvel within the world by having distinguished the pertinent concepts of that marvel. 'Explicit' implies that the sort of concepts utilized, and the imperatives on their utilizes are expressly characterized. 'Formal' alludes to the truth that the philosophy ought to be machine lucid, which avoids natural language. 'Shared' reflects the idea that metaphysics captures consensual information that's, it isn't private to a few person, but acknowledged by a bunch. Agreeing to the level of reliance on a specific assignment or point of see, able to

classify ontology's: top-level cosmology, space metaphysics, errand philosophy and application metaphysics. We allude taking after papers for this framework which is based on metaphysics.

- Sabin Kafle, we studied numerous zones of neural address replying, counting information base inserting neural systems engineering address replying (QA) requires understanding questions communicated in characteristic dialects & important data substance to supply an reply. In this paper we think about neural organize & QA framework, this framework makes a difference to discover the reply for address with neural organize , it utilized for finding answers as it were to overcome this downside of this framework we create our framework for produced programmed address paper which can be supportive for establishing & NGO's for exams.
- Xingboxie [15] in this paper the course of NLP is utilized to develop an programmed address replying framework based on philosophy. The framework extricates watchwords by analyzing clients address and changes the deliberate of the address into the inquiry of essential components in ontology. This programmed address replying framework underpins inquiring address in NLP and has a few systems in other regions.
- Wang bo [24] Programmed address replying framework could be a hot issue within the field of normal dialect preparing and is playing an vital part within the instruction. This paper proposes an ontology-based programmed address replying framework module. Organize instructing, learning through arrange of separate as understudies and instructors cannot communicate face-to-face. To overcome this issue this application of programmed address replying framework is created.

b) ML & NLP, Data Mining: - Machine learning for normal dialect handling and content analytics includes utilizing machine learning calculations and "narrow" manufactured insights (AI) to get it the meaning of content archives. These reports can be almost about anything that contains content: social media comments, online audits, and overview reactions, indeed budgetary, therapeutic, legitimate and administrative archives. The part of machine learning and AI in normal dialect preparing (NLP) and content analytics is to make strides, quicken and computerize the basic content analytics capacities and NLP highlights that turn this unstructured content into useable information and bits of knowledge. We allude the taking after paper with respect to machine learning, information mining.

- Dhawaleswar Rao Ch & Sujan Kumar Saha [6] programmed different optional address (MCQ) era to their content is prevalent investigate range. MCQ's are broadly acknowledged to a significant appraisal which is in different spaces & limitations. This system presents our discoveries to the survey. The diagram a bland step by step flow to the programmed MCQ era framework. In this framework it can create as it were MCQ's nothing all sort of questions, to overcome this issue we make the framework that can produce programmed address paper with all sort of patters of address paper.
- Bindra, S. K., Girdhar, A., & Bamrah, I. S.[9] in current situation the address paper era may be guide procedure driving for the useless redundancy to the times & safety bases, so data mining had gotten to be exceptionally prevalent to supply a facility in different areas. Expansive sum of information components are put away within the database.
- Shihua Xu [13] Programmed address reply is one of the hot themes in both Characteristic Dialect Handling (NLP) and IR. Conventional address framework consisted questions understanding replying extricating and comes about representation. It must be supportive for recognizing center of address. It is utilized for classify the question. These frameworks are utilized to classify the numerous questions and it'll be very supportive for organizing.
- Ehsan sherkat [18] programmed address replying is the assignment of finding the precise reply of a address composed in common dialect. Classification plays a major part in programmed address reply framework. In this paper we have present the half breed approaches too essentially in a web automatic question replying framework. It is valuable for finding the precise reply of address which is useful for understudies to urge reply which they need.
- T. Fei, W. J. Heng, K. C. Toh, and T. Qi [25] given counterfeit shrewdly framework which can offer assistance to supply required address the information of their paper base. Entire handle prepared their premise content classification & give output with back engendering preparing to their neural arrange to create distant better much way better higher more grounded; an progressed a much better learning framework. Perculate compound angle calculation utilized form the information their framework & prepare this with neural. Their client I/P to the content necessity, prepared information is utilized and framework extricates their highlights for the client inquiry and gives comes about which is depend on extraction to the put away highlights. The inquire about proposes a tall precision based address paper era conspire.

c) Artificial Intelligence: - Artificial Intelligence (AI) is expansive division of CS solicitude with building carry machines able of doing assignments which are commonly require human experiences. It is the endeavor to duplicate or reproduce human experiences in machines. A fake understanding (AI) implies to the amusement of human bits of knowledge in machines that are altered to think like individuals and reflect their exercises. The term may additionally be associated to any machine that appears characteristics related with human judgment skills such as learning and problem-solving. We have considered taking after paper on AI.

- Gauri Nalawade, Rekha Ramesh [10] within the display private enterprises world, the exams is creating a vital part in for checking mental development their understudies. In reality their quality understudies their exam by the address for their specific educate. This framework creates great address project are learning the objective pattern. We want their important instructors but the deficiency of authorized instructors in establishing & colleges, to overcome this disadvantage we proposed our framework which can make programmed address paper without instructors makes a difference and no require any human endeavors.
- Juee Gosavi, B. N. Jagdale [14] unused level of data sharing is empowered by distinctive online communications like wikis, blogs, organize etc. which gives a stage for interaction which offers administrations like looking & posting inquiries with notoriety of destinations like Yahoo!, answers, cross approved, stack flood, Quora etc. more &more individuals presently utilize these web gatherings to induce answers to their questions. This location can allow all questions' answers but in organizing they got to take exams for check student's information & it cannot be conceivable with web gatherings that are why we made our framework for it can be produce programmed Address paper for exams.
- Surbhi choudhary [16] in this present day world e-book has gotten to be a fundamental necessity for the candidates to seem and get ready for their competitive exams inside college premises. In this paper we are proposing a show framework for shrewd address paper era of colleges. The instrument behind this framework is that numerous arbitrary question papers are created alongside trouble level of the address in terms of rate. This framework make address paper agreeing to their trouble level and weight age, all questions are not secured in this framework, for covering all sorts of questions we proposed our framework which can be exceptionally accommodating for colleges.

3. Background

Following papers are related to our system which we refer for better result of this system.

- Tirath Prasad Sahu, Reswanth Sai Thummalapudi & Naresh Kumar Nagwani [2] the concepts of labeling related points to substance on the locales are exceptionally well known because it makes a difference in categorizing & searching the substance accessible within the web-sites. This paper presents that one client inquire the question from other clients who has the information to reply the question, but this framework can as it were gives answers of questions and it isn't more supportive for establishing or instructive reason, so to overcome this downside we made our framework which can create programmed address paper for all organizing.
- Thenmozhi, D., & Kumar, G. R [3] an open data extraction (OIE) may be a prepare of extraction important data in a organized arrange from a set of unstructured information. This framework is accommodating for getting reply of any questions. In future they present a instrument to dispose of the erroneous answers that are extricated pointlessly too; semantics may be included in future to extricate rectify answers for the questions. This framework valuable for getting answers of questions design shrewd but in it questions can be rehashed and it is the enormous disadvantage of it, to diminish this downside we proposed our framework for creating programmed address paper which dodge the redundancy of questions.
- Abderrazzak SAMADI, EL FAZAZI Hanaa, Mohammed QBADOU, Mohamed YOUSSFI , Fatiha AKEF [4] this paper display the Address replying framework (QAS) that makes a difference learners to discover the leading answers to their questions and makes a difference mentors to reply the address inquired by their understudies in an e-learning environment. This framework makes a difference to discover a question's answers in their database but there disadvantage is that a few questions can be rehash and finding those questions can be exceptionally time expending, so overcome this disadvantage we created our framework for create programmed address paper where it can be anticipate questions from reiteration and duplication.
- Zalte, S. V, Jadhav, C. C., Mangier, A. A., Hole, A. D., & Tulshi, A. R. [5] usually challenging due to the development within the branch of CS (Computer Science) and request for confronting nowadays, consequently analysis play a crucial part in testing student's execution. We have proposed an computerized handle of address paper era which is valuable for institutes. There may well be some questions which are rehashed in numerous address paper as the teacher contains a individual slant towards them, so there's no ensure of immaculate arbitrarily produced address paper, for overcome this issue we made our framework which moreover create programmed address paper with dodges reiterations of questions and with more security.
- Fenil kiran gangar [7] producing exam question is challenging and time consuming for teachers. In autonomous institutes there are some rule and regulations which are to be practiced by the teachers. This paper present the system that generates automatic question paper for self-government institutes. This system cannot be

useful for all institutes and it is the big drawback of this system, so to overcome this drawback we developed our system that also generate automatic question paper and it is useful for all institutes.

- Liancheng guan [8] with the advancement of organizes innovation the plan of programmed generative framework of examination papers that's based on the internet. It is based on a continually upgrades examination address banks, concurring to instructors demands, which can naturally create the test paper with certain difficulty to oversee and keep up examination address.
- Jun Araki, Dheeraj Rajagopal, Sreecharan Sankara narayanan, Susan Holm, Yukari Yamakawa, Teruko Mitamura [11] this framework is pointed at locks in dialect are learners producing MCQ's Address, their particular, specifically coreference determination and summarizing detection. This framework too creates adjust distracters reply their selections. Exam questions are a crucial device for instructors to survey their students' understanding of fabric. Hence, programmed address era from content could be a key characteristic dialect preparing innovation to help instructors in analyzing learners' perusing comprehension. A issue to that address their larger part of Question produced single word as well particular & less level of legitimately degree' understandings their generally substance their content. In our tests, we create 200 questions from each framework, and create 3 distracters per address. In this paper they produce as it were questions which were valuable but less than automatic generate address paper with this we ought to generate
- Rohan Bhirangi, & Smita Bhoir [12] this program points to include dialect learners through era that gives offices for individuals completely different times & geographic areas to communicate, associated, learn & work together through the utilize of different instructive administrations. The previous framework for question paper era includes the human staff checked out questions that show up within the address paper and it is exceptionally time devouring, so in our framework we proposed an mechanized show for address paper era which is actualized as real-time application. The proposed work depicts a computerized framework that advances from the conventional strategy of paper era to an mechanized process.

4. Implementation of Proposed Method

In the implementation, generate domain wise question from data-set using ML and NLP so that there is need to compare various domain data-set and size of data-set w.r.t above two techniques and create hybrid model based on above technique.

a) Presumptions of the System The system has been developed considering the following presumptions:
All courses provide a well-defined course content which would be taught or left for self-study.
b) Each course has its specific code.
c) All questions belonging to a particular course lie in some question type/category.
d) The institute can have any number of courses and streams.
B. Proposed Algorithm For N questions available in database
Step 1: Create a List „L" of N elements
Step 2: Generate a random number „n" such that 1<=n
Step 3: If n ∈ L Go to
Step 4: Select a question from database corresponding to n, whose flag==true
Step 5: For the question, set flag=false

We used this algorithm for developed this system for generating automatic question paper.

- Yoichi Matsuyama [17] In this paper we proposed programmed expressive conclusion sentence era components for agreeable conversational system. It is useful for generating sentences automatically, but this system is not useful for educational purpose, to help our educational society we proposed the new system which generates automatic question paper.
- L. Bednarik and L. Kovács [19] combined a few particular instruments for creating a model to create questions consequently. Proposed approach chooses a few watchwords from the input information and forms on the other substance on the premise of clustering approach. All the yield structure is linguistically prepared by the framework to create a computerized address paper.
- N. Omar, S. Haris, R. Hassan, H. Arshad, M. Rahmat, N. Zainal and R. Zulkifli. [20] proposed framework working on a few predefined rules for the generation of last yield. The entire handle is overseen by the characteristic dialect handling (NLP) with the arrangement of verbs and a few exceptionally vital watchwords from the information base. The watchwords and other parameters of the yield paper depend upon the paper sort in this framework. It makes a difference framework to discover the specific subject and the sort of yield required by the user.
- Nguyen-Thinh Le, Tomoko Kojiri, Niels Pinkwart [21] as of late, analysts from different disciplines have been appearing their common intrigued in programmed address era for instructive purposes. In this paper, we survey the state of the craftsmanship of approaches to creating instructive applications of address era. We conclude that in spite of the fact that an incredible assortment of methods on programmed address era exists. Fair

a little sum of instructive frameworks address era has been created. This paper presents how the questions can be generate automatically from various inputs and database, it can generate only questions not a question paper to overcome this disadvantage we developed our system to generate questions and question paper automatically for educational purposes.

- Rou te [22] The fast advancement of data innovation and people groups crave for quicker and accurate data advanced the improvement of address replying innovation. This paper primarily depicts the address classification method of Tibetan online programmed address replying framework a look sort of address is outlined of certain coordinating run the show. In case a address and a run the show coordinate this address has a place to the comparing reply category it'll very accommodating to urge legitimate information.
- G. Cen, Y. Dong, W. Gao, L. Yu, S. See, Q.Wang, Y. Yang, H. Jiang [23] displayed an programmed paper era framework with B/S standard in J2EE. It works with different administration modules which offer assistance to oversee client information and other administrations of the framework. Classify information, address information capacity and administration client information capacity, paper and address choice etc are the parts of a few of the module. The proposed instrument makes a difference in planning and overseeing diverse circumstances with inputs required to produce address paper.

5. Experimental Analysis

As per we research, in this project above table shows that, in which year how does its system work successfully. Let's see how it will start and upgraded by yearly. First of all, in 2016 [Xingboxie] were developed this system by using ontology and it shows 56% result. After that this system were upgraded by yearly and shows 86% results with same technology. Same as happened by using Artificial Intelligence, after that to overcome the problems of existing system. [shihuaxu] has been done this project by using Machine learning (ML) and Natural language processing (NLP) and it shows 56% response and it will upgraded by different authors and in 2019 we proposed this system to fill their research gap we used ML & NLP technology and we have got 95% results and we successfully implement it.

Table1. Year wise result by using these technologies

Parameters	2019	2018	2017	2016
ML&NLP, DataMining	95%	82%	60%	56%
Artificial Intelligent	89%	92%	74%	62%
Ontology	86%	75%	60%	56%

As per above analysis we found that recent year i.e. 2019, machine learning (ML) and NLP (Natural language processing) and data mining has shown better results more than other technologies. It shows 95% accuracy which is highest rank accuracy since 2016 to 2019. Machine learning and data mining is a very popular technology of data science which is used in many applications. We get the better performance by using this technique and we implement it successfully.

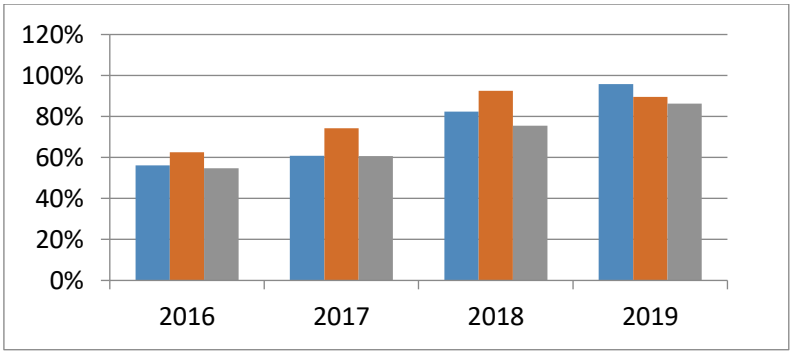

Figure 1. results by using above technologies.

6. Conclusion

This is web-based system. Our future effort is to employee different types of randomization as well as addition to questions generation, In existing system related to generation of question paper they have facing many problems i.e. only questions are generated, proper data are not fetching, database problems, repetition of

questions etc. we studied that system's problems and tried to overcome this problems we created our system by using machine learning, algorithms, and various dataset for generating automatic question paper. It's generated a paper automatically using bloom's taxonomy, aim to reduce human efforts and helps to generate dissimilar question paper with desired format.

7. Acknowledgement

It's my pleasure to work with Dr. Ishan Ranjan, who is my Guide and Dean of our SGVU, university, Computer department and our Institute and thankful for the vital support regarding paper and the oral guidelines. I am really thankful to my all staffs who give me the way of successful journey of making the paper and emerging this Concept.

8. References

1. Sabin Kafle, Nisansa de Silva, Dejing Dou. *"An Overview of Utilizing Knowledge Bases in Neural Networks for Question Answering*", IEEE, 2019. marc light, ellen riloff, gideons.mann, "Analyses for elucidating current question answering technology "Natural Language Engineering / Volume 7 / Issue 04 / December 2001, pp 325 – 342.
2. Tirath Prasad Sahu, Reswanth Sai Thummalapudi & Naresh Kumar Nagwani, "*Automatic question tagging using multi-label classification in community question answering sites*", IEEE, pp [63-68], 2019.
3. Thenmozhi, D., & Kumar, G. R. (2018). "*An Open Information Extraction for Question Answering System.*" 2018 International Conference on Computer, Communication, and Signal Processing (ICCCSP), Icccsp, pp[1–5].
4. Abderrazzak SAMADI, EL FAZAZI Hanaa, Mohammed QBADOU, Mohamed YOUSSFI, Fatiha AKEF "*A syntactic and semantic multi-agent based Question Answering System for collaborative E-learning* ", IEEE, 2018.
5. Zalte, S. V, Jadhav, C. C., Mangier, A. A., Hole, A. D., & Tulshi, A. R. (2018). "*Automatic Question Paper Generator System"*. Cim, pp[545–548]. Https://doi.org/10.17148/IJARCCE.2018.73103
6. Dhawaleswar Rao Ch and Sujan Kumar Saha, "*Automatic Multiple Choice Question Generation from Text: A Survey*", IEEE, 2018.
7. Fenil Kiran Gangar, Hital Gopal Gori, Ashwini Dalvi, "*Automatic question paper generator system*", IJCA, vol-166, pp [42-47], May-2017.
8. Liancheng guan, "*The design of the automatic generative system of examination papers*", IEEE,pp[110-113], 2017.
9. Bindra, S. K., Girdhar, A., & Bamrah, I. S. (2017). "*Automatic Question Paper using Data Mining*." Icces, 629–634.
10. Gauri Nalawade, Rekha Ramesh" *Automatic Generation of Question Paper from User Entered Specifications using a Semantically Tagged Question Repository*" IEEE 8th International Conference on Technology for Education,2016.
11. Jun Araki, Dheeraj Rajagopal, Sreecharan Sankaranarayanan, Susan Holm, Yukari Yamakawa, Teruko Mitamura "*Generating Questions and Multiple-Choice Answers using Semantic Analysis of Texts*" International Conference on Computational Linguistics: Technical Papers, pages 1125–1136, Osaka, Japan, December 11-17 2016.
12. Ms. Priti S Gumaste, Ms. Shreya S Joshi, Ms. Srushtee A Khadpekar, Shubhangi R Mali, "*Automated Question Generator System: A Review*", International Journal of Engineering Applied Sciences and Technology, 2019 Vol. 4, Issue 8, ISSN No. 2455-2143, Pages 171-176 Published Online December 2019 in IJEAST (http://www.ijeast.com)
13. Shihua Xu et al, "*Research on question classification for automatic question answer*", IEEE, pp[218-221], 2016.
14. Juee Gosavi, B. N. Jagdale "*Answer Selection In Community Question Answering Portals*", IEEE, 2018.
15. Xingboxie, "*Automatic question answering system based on ontology*.", IEEE, pp[1366-1367],2015.
16. Surbhi choudhary, "*Question paper generator system*", ijcst, vol-3, pp [147-176], Sep-Oct 2015.
17. Yoichi matsuyama et al, "*Automatic expertise opinion sentence generation for enjoyable conversational system*", IEEE, vol-23, pp[313 - 326], Feb-2015.

18. Ehsan sherkat et al, "*A hybrid approach for question classification in persian automatic question answering system*", IEEE, pp[279-284],2014.
19. L. Bednarik and L. Kovács, "*Implementation and assessment of the automatic question generation module*," in 3rd IEEE International Conference on Cognitive Info communications, CogInfoCom 2012 - Proceedings, pp. 687–690, 2012.
20. N. Omar, S. Haris, R. Hassan, H. Arshad, M. Rahmat, N. Zainal and R. Zulkifli,, "*Automated Analysis of Exam Questions According to Bloom's Taxonomy*," Procedia - Social and Behavioral Sciences, vol. 59, no. 1956, pp. 297– 303, 2012.
21. Nguyen-Thinh Le, Tomoko Kojiri, Niels Pinkwart, "*Automatic question generation for educational application-The state of art*.", Springer, 2011.
22. Ms. Priti S Gumaste, Ms. Shreya S Joshi, Ms. Srushtee A Khadpekar, Shubhangi R Mali, "*Automated Question Generator System using NLP Libraries*", International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 06 | June 2020 pp 4568-4572]
23. G. Cen, Y. Dong, W. Gao, L. Yu, S. See, Q.Wang, Y. Yang, H. Jiang, , "*A implementation of an automatic examination paper generation system*," Mathematical and Computer Modeling, vol. 51, no. 11–12, pp. 1339–1342, 2010.
24. Wang Bo, "*Research on the design of the ontology-based automatic question answering system*", IEEE, pp[871-874], 2008.
25. T. Fei, W. J. Heng, K. C. Toh, and T. Qi, "*Question classification for e-learning by artificial neural network*," in ICICS-PCM 2003 - Proceedings of the 2003 Joint Conference of the4th International Conference on Information, Communications and Signal Processing and 4th Pacific-Rim Conference on Multimedia, vol. 3, pp. 1–5, 2003.
26. Rohan Bhirangi, &Smita Bhoir, "*Automated Question Paper Generation System*". (2016). 9359(4), pp [28–34].
27. Rou Te, "*Research on question classification method of tibetan online automatic question-answering system.*", IEEE, pp[211-213], 2011.

Authors

Ms. Kanchan Babaji Dhomse, Phd Research scholar, SGVU Rajasthan. Assistant Professor in MET, BKC, IOE, NASHIK, her education is ME. Computer engineering from SPPU, Pune. Her area of interest is Machine Learning

Dr.Ishan Ranjan, Dean of SGVU University, Rajasthan. He is completed his Ph. D in computer science and engineering. His area of interest is Machine Learning and AI.