

BPIT

Bhagwan Parshuram Institute of Technology

A Unit of Bhartiya Brahmin Charitable Trust (Regd.)

(Approved by Ministry of HRD Govt. of India, AICTE & DTTE, Govt. of Delhi)

Affiliated to Guru Gobind Singh Indraprastha University, Delhi

PSP-4, Sec.-17, (Opp. Sec.-11), Rohini, Delhi-89 Tel.: 27571080, 27572900

Fax : 27574642 E-mail : bpitindia@yahoo.com Website : www.bpitindia.com



BBIJTM

Chief Patron

Sh.Vinod Vats Chairman, BPIT

Patrons

Sh. Surender Sharma (Vice President)

Sh. Ram Babu Sharma (General Secretary)

Sh. Shambhu Sharma (Secretary)

Prof. A.K Tandon (Founder Director)

Prof. Y.D Gaur (Executive Director)

Editor-in-Chief

Prof. Payal Pahwa

Principal, BPIT

Associate Editor

Prof. Achal Kaushik

Assistant Editors

Dr. Shweta Taneja

Editorial Board Members

1. Professor Chin-Chen Chang
Chair Professor, Department of information Engineering and Computer Science, Feng Chia University, Taiwan
2. Dr. Rana Singh
Assistant Director, Institutional Research and Planning Gulf Medical University, Ajman, UAE
3. Prof. Arif Suhail
Professor, Department of Mechanical Engineering, Aligarh Muslim University
4. Prof. Sanjeev Mittal
Professor,
University School of Management Studies, Guru Gobind Singh Indraprastha University, Delhi
5. Dr. Rajeev Jain
Professor & Director of Chemistry, Institute of Engineering, Gwalior
6. Dr. Neeraj Kaushik
Associate Professor,
NIT, Kurukshetra
7. Dr. Chitra Sharma
(Formerly at IGTUW, Delhi)
Philadelphia (U S A)
8. Prof. Rajiv Sharma
Professor, Deputy Director & HOD (ECE),
BPIT, Delhi.
9. Prof. Abhishek Swaroop
Professor & HOD (IT),
BPIT , Delhi.

Advisory Body

1. Prof. Oscar Castillo Tijuana Institute of Technology, Mexico
2. Prof. Frank Wang University of Kent, England
3. Dr. Rajkumar Buyya The University of Melbourne, Australia
4. Prof. Anil Kumar Tripathi
Banaras Hindu University, India
5. Prof. R K Agarwal
Jawaharlal Nehru University, India
6. Prof. D P Vidyarthi
Jawaharlal Nehru University,
New Delhi, India
7. Prof. Jemal H. Abawajy Deakin University, Australia
8. Dr. João Manuel R. S. Tavares University of Porto, Portugal
9. Dr. M^a Cristina Rodríguez Sánchez King Juan Carlos University, Spain
10. Dr. Alex Khang University of Texas at Austin
11. Dr. Amit Prakash Singh Guru Gobind Singh Indraprastha University, Delhi, India
12. Dr. Lalit Singh NPCIL-BARC , Dept. of Atomic
13. Prof. M. N. Hoda BVICAM, Delhi, India
14. Dr. Neeraj Kumar National Institute of Technology, Jalandhar, India
15. Prof. Sushma Nagpal Netaji Subhas University of Technology, Delhi, India
16. Prof. Ela Kumar Indira Gandhi Delhi Technical University for Women, Delhi, India
17. Mr. Udayan Khurana Research Staff Member, IBM Research
18. Dr. Pinaki Chakraborty Netaji Subhas University of Technology, Delhi, India
19. Dr. Rahul Johri Guru Gobind Singh Indraprastha University, Delhi, India
20. Prof. Shailesh Tiwari Krishna Engineering College, Uttar Pradesh, India
21. Dr. M P Mishra Indira Gandhi National Open University (IGNOU), India

Vol 8 Issue2 December 2022
Copyright © Principal, BPIT, Delhi
PSP-4, Dr KN Katju Marg, Sector 17
Rohini, New Delhi 110089.

Disclaimer: The publisher and Editor assume no responsibility for any statement of the facts or options expressed in the published papers.

From the Chief Patron Desk



It is rightly said:

“Investment in knowledge pays the best interest.”

At BPIT, we believe that an investment in education is an investment in people and their promising future. In 2005, we set ourselves an ambitious goal of establishing a quality education institution in India with a mission to offer outstanding learning and educational services to the community. We are deeply grateful to those who have assisted us and participated in the accomplishment of this goal, i.e. the establishment of BPIT.

BPIT is committed to develop and enhance the offered accredited academic programs, services, and facilities to enable BPIT stakeholders to achieve their goals. During the previous years, though we faced many challenges, we were determined to continue the fulfillment of BPIT mission of providing an accessible and valued education in India.

BPIT has demonstrated quality teaching. Its faculty members have proved their commitments to share their knowledge, talent, and experience into learning and teaching. At the BPIT, we have proved that the institute is capable of pursuing the never-ending quest for quality technical education and research. As a means of striving towards this goal of excellence, I feel proud to present the issue (Volume 8, Issue 02) of BODH: BPIT International Journal of Technology and Management.

This journal is a sincere and dedicated effort of the learned faculty and management to disseminate the wealth of knowledge in the field of technology and management. This journal would provide a great opportunity to make them more influential, dynamic and innovative. All this could be possible with the hard work done by college staff under the guidance of visionary management. I feel proud to acknowledge the contribution of highly qualified, dynamic and multi-talented faculty and the college management for working tirelessly to enrich the content of this journal.

Sh. Vinod Vats
Chairman, BPIT



From the Editor-in-Chief Desk

BPIT has always looked forward to upgrade itself to fulfil its mission and vision to provide quality education. The function of education is to teach one to think intensively and to think critically.

It gives me immense joy to present the issue (Volume 8, Issue02) of BODH: BPIT International Journal of Technology and Management, which includes research papers from Technology and Management. This journal is a tool for building knowledge and for facilitating learning. The journal is a good means of extending knowledge to the research scholars, professors, scientists and students regarding the current issues related to Engineering and management. The journal is an amalgam of the efforts of the highly dedicated faculty who have worked day in and day out for the successful launch of this issue of BODH. I congratulate all for this sustained effort that has produced such outstanding results.

On behalf of our editorial team, I would like to offer a word of thanks to our readers, contributors, authors, editors and anonymous reviewers, all of whom have volunteered to contribute to the success of the journal and also for its mission to improve the quality of research in the form of publication. We are publishing our journal twice a year with a particular emphasis on quality, novelty and better outcomes of research.

I extend my whole-hearted thanks to Hon'ble General Secretary Sh. Vinod Vats ji and the entire management of BPIT for their untiring support and worthy guidance.

Thank you. We hope you will find BBIJTM more informative in the future endeavor.

Prof. Payal Pahwa
Principal BPIT

From the Associate Editor Desk



It gives immense pleasure to present the issue (Volume 8, Issue 02) of “Bodh”, BPIT’s International Journal of Technology & Management of Bhagwan Parshuram Institute of Technology (BPIT) Delhi. I am equally elated to inform you all that BBIJTM has been contributing tremendously to improve the quality of research by publishing its issues regularly since 2015. An enormous amount of work has been done towards the development of this journal in the past days.

The current issue presents a compilation of eight research papers in the emerging fields of engineering, science, management and technology. The first paper proposes a deep learning-based model to classify and detect pneumonia and tuberculosis using Convolutional Neural Networks. The second paper proposed Fuzzy C-Mean Clustering algorithm to improve the web page recommendations. The next paper presents understanding the importance and impact of digital financial services as a route to better financial inclusion and economic development in India, especially pre and post COVID periods. The fourth paper aimed at the evolution and the benefits of FACTS devices with the purpose of improving the operation of an electrical power system. A Wavelet transform based image informatics is presented in Fifth paper. The next paper, presents Accident-Avoidance system, aimed to reduce the risks of accidents at the hairpin bend on a Hilly track, Ghats, or other Zero visibility turns. The seventh paper proposed a smart garbage monitoring system which gets a signal using GSM whenever garbage gets filled and results will be monitored on the Think Speak website using IOT. The last paper presents PV applications in battery charging.

I wish to express my sincere gratitude to the entire editorial advisory board, editorial board and the panel of reviewers to bring out the best of knowledge from the authors. Also, I thank the authors for showing their interest in the Journal. I would also like to thank our Patron, Hon’ble Chairman Sh. Vinod Vats ji, the complete advisory board and editorial team for their motivation and support in bringing out yet another edition of the journal.

Prof. Achal Kaushik
Associate Editor

S. No.	Title of the Article	Authors	Page
1.	Modeling and Control of Photo Voltaic Emulator by Robust Controller	Simmi Sharma and Dheeraj Joshi	1
2.	XPEN – A Voice Powered Expense Tracker Full Stack Web Application	By Kritika Balihar, Himani Sharma and Shikha Bhalla	9
3.	Optimization: Background and Existing Challenges	Neha Khanduja	18
4.	DRD-CNN: Diabetic Retinopathy detection using CNN	Aman Singh, Raman Chola, Ashish Negi, Keshav Agarwal and Neha Sharma	28
5.	FRecS - FINANCIAL RECOMMENDER SYSTEM	Charu Gupta and Palak Girdhar	38
6.	Development of Student Interaction Platform	Suman Arora, Siddhant Chadha and Tushar Jain	77
7.	E- voting system using Blockchain	Ms. Akanksha Dhamija, Mugdha Sharma and Richa Sharma	45
8.	DICGM: A Deep Learning based Novel Image Caption Generator Model	Bhawna Suri, Shweta Taneja and Savneet Kaur	53
9.	Bulletin Cognizance Scorn Analysis using NLP	Preeti Arora, Tanisha Madan and Nikita Nijhawan	65
10	A Study on Suicidal Rate Detection with social media using Deep Learning	Abhishek Singh, Akshay Khare, Anand Kumar Jha, Shikhar Kannoja and Shimpy Goyal	70



Modeling and Control of Photo Voltaic Emulator by Robust Controller

Simmi Sharma, Dheeraj Joshi

simmisharma@bpitindia.com, joshidheeraj@dce.ac.in

Abstract-This research presents signal flow graph method for transfer function calculation and emulating V-I and P-V characteristic of PV panels. The usage of robust sliding mode controller, a diode model and a lookup table add to this. If several panels are to be replicated on the same model, the system does not require complicated calculations. Circuits with non-ideal components will also benefit from the proposed method, which will make mathematical modeling easier. The testing and installation process for solar electricity in rural and distant areas will thus be accelerated. To achieve a viable solution for this Nonlinear model, a sliding mode controller is implemented with stability constraint. The use of a diode model will incorporate variables such as temperature change and partial shading. The suggested system provides a simple and stable testing environment 24 hours a day, seven days a week for Non-linear PV panels before installation to save time and money. The MATLAB results match the actual shell 115 PV panel properties, demonstrating model and control correctness.

Keywords: Diode model, lookup table, mathematical modeling, PV Emulator, graph theory, robust sliding mode controller, MATLAB Simulation.

I. INTRODUCTION

The world's energy consumption is rapidly rising and research reveals that electricity generation solely from non-renewable energy resources will be unable to meet these rising demands. This ushered in a revolution in the energy sector and established the renewable energy sector on a global scale. Traditional and hybrid types of solar, wind and hydro energy have reframed research topics due to technological improvements and advantages linked with renewable energy. Solar and wind power account for the majority of renewable energy capacity in India. Wind energy's reliance on geographical locations and meteorological conditions limits its applications in countries like India, where harsh weather occurs throughout the year. India is positioned on the equator, which receives the most sunlight, making it the best location for solar energy development. To improve the usage of solar energy, factors such as setting up 24hour labs for testing, space for panel installation, changing environmental conditions and panel shading require substantial research. As India is partly tropical and subtropical in climate, it receives sunlight all year. Every year, close to 5000 trillion kW-hr of solar energy is received. With the growing population, the potential for trappable solar power has increased to 750 GW. The goal of the National Solar Mission, which began in 2010, was to reach 20 GW of solar installed capacity by 2022, but that goal was then raised to 100 GW by 2022. This aim of 20 GW installed capacity was

met in 2018, resulting in a 370 percent [1] increase in India's solar capacity. Energy demand has risen dramatically in recent decades and is likely to continue to rise in the coming years as the world's population and industrialization grow. However, because of policy incentives, simplicity in structure and cost effectiveness, the use of clean technologies in the power sector is a top priority for global consumers. Solar PV and wind energy are the most cost-effective electricity generation options in many far-off areas [2]. The introduction of PV Emulators can improve laboratory-based testing of solar panels of various ratings [3]. A wide range of solar panels are available for purchase in the market. Purchasing each type of panel for testing separately is prohibitively expensive and not feasible. Under variable temperatures and shadowing situations, a PVE will be able to simulate these many types of solar panels. It protects itself from overloading and short circuit events that occur frequently during panel testing because it is a programmable device. PVE is used in both commercial and academic research. Since PV systems are connected to dynamic loads via inverters and maximum power point (MPPT) trackers, a lot of work can be carried out to optimize the quality of the emulated output. The most common power electronic topologies for emulation circuits include buck, boost, and buck-boost converters that operate as DC sources. According to research, the cost of production and installation, the area covered, the panel's shading condition and the maximum power tracked all play a role in solar panel selection, and a PVE will be a faster and more efficient technique to determine the best alternative. When planning a solar plant, an emulator can be used to estimate how a panel or module will perform. It should be able to change its operating point in response to changing loads. A controller, a reference model, and a PWM circuit are the other crucial components. Diode model [4] and lookup table [5] are employed as reference models. To finalize the single diode model, five parameters are required: load current, panel current, shunt resistance, series resistance, and ideality factor. The model is also confirmed by replacing the look-up table for one of the diode models. Because of its great efficiency, switched-mode power supplies are widely applied in PVEs. Signal Flow Graph is used to estimate the converter's transfer function in transient states (SFG). Mason's gain formula replaces the complex equations in this method. The details will be discussed in the following section. The robustness and stability of sliding mode controllers (SMC) are well established. P, PI, and PID controllers used for power converter control do not provide the best results in dynamic situations [6] of non-linear systems. SMC was conceived for variable structure systems, and it is well known that dc-dc converters are inherently variable structured, enabling SMC implementation easier [7- 9]. It aids in the design of a PV system with dynamic loading that is both resilient and accurate. The paper is divided into three sections, the first describing the mathematical derivation of a converter using SFG at switching states, the second explaining the construction of a PVE model including a buck converter in a closed loop configuration and the third explaining SMC implementation in a feedback loop using one diode and LUT as reference models. The characteristic plots obtained from designed models are detailed in the following part. Finally, the conclusion and scope of future research are discussed.

II. SIGNAL FLOW GRAPH (SFG)

A signal flow graph, also known as a switching flow graph, is a graphical approach for calculating a system's transfer function. The benefits include a simplified approach for determining the transfer function without having to draw the system repeatedly and a reduction in the number of components by replacing them with nodes and loops. The whole system's transfer function is determined using Mason's Gain formula. Due to its simple structure, a DC-DC buck converter was used to imitate the electrical behavior of a PV panel under various situations. The internal resistance of the components is either zero or ignored in an ideal buck converter model. To obtain the transient equations characterizing the switch's behavior. the ON-OFF states are studied. The

equations for transient circuits are as follows:

$$L \frac{di_L}{dt} = -v_o + V_{in}$$

$$C \frac{dv_o}{dt} = i_L - \frac{v_o}{R} \tag{1}$$

$$L \frac{di_L}{dt} = -v_o$$

$$C \frac{dv_o}{dt} = i_L - \frac{v_o}{R} \tag{2}$$

On averaging equations, (3) and (4) becomes:
and it is well known that dc-dc converters are inherently variable structured, enabling SMC implementation easier [7-

$$L \frac{d\bar{i}_L}{dt} = -\bar{v}_o + \bar{d}_1 \bar{V}_{in} \tag{3}$$

$$C \frac{d\bar{v}_o}{dt} = \bar{i}_L - \frac{\bar{v}_o}{R} \tag{4}$$

Small Signal Analysis of above two average equations (3) and (4) gives:

$$sL\bar{i}_L = -\bar{v}_o + \bar{D}_1 \bar{v}_{in} + \bar{d}_1 \bar{v}_{in} + \bar{d}_1 \bar{V}_{in} \tag{5}$$

$$sC(\bar{v}_o) = (\bar{i}_L) - \left(\frac{\bar{v}_o}{R}\right) \tag{6}$$

$$s\bar{i}_L = \left[-\frac{\bar{v}_o}{L} + \frac{\bar{D}_1}{L} \bar{v}_{in} + \bar{d} \frac{\bar{V}_{in}}{L}\right] \tag{7}$$

$$s\bar{v}_o = \frac{\bar{i}_L}{C} - \frac{\bar{v}_o}{RC} \tag{8}$$

Equations (5), (6), (7) and (8) represent AC component expressions. From DC component expression;

The duty cycle equation is derived as $\bar{V}_o / \bar{V}_{in} = \bar{D}_{in}$ (9)

$$\bar{i}_L = \frac{\bar{D}_1 \bar{V}_{in}}{R} \tag{10}$$

Using equations (5) and (6), following signal flow graph also termed as switching flow graph in this case is obtained [10-11] as shown in Fig.1.

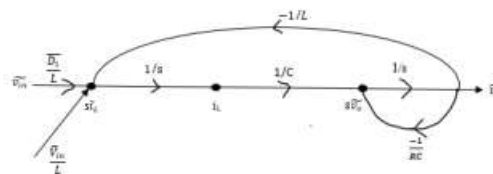


Fig. 1 SFG of ideal Buck converter

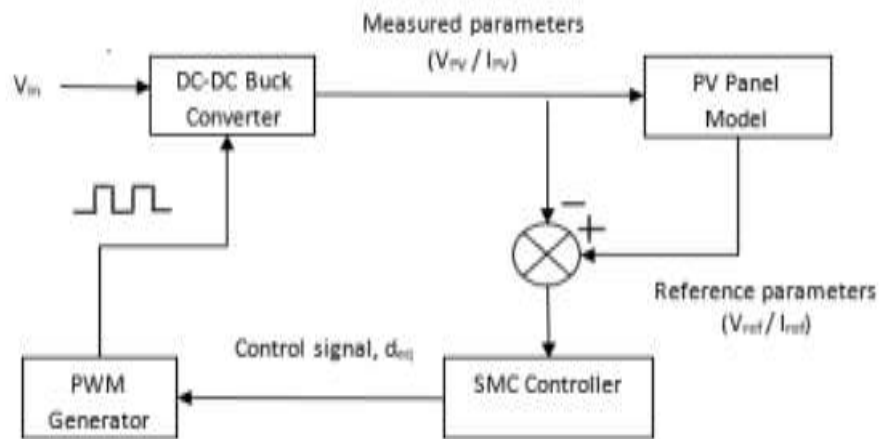


Fig.2 Schematic Diagram of PV Emulator

Fig. 2 shows schematic diagram of PVE showcasing the role of reference model and controller in PVE model. Further, equations (5) to (6) are used to calculate the transfer function of the buck converter using Mason’s Gain formula and transformed to a signal flow graph as shown in figure 1. Small signal transfer functions, as well as big signal and steady-state models, can be calculated using the SFG approach [12]. The transfer function is produced by substituting the values of forward path gain and loop gains for individual and nontouching loops in Mason's gain calculation, according to the signal flow graph. This strategy is simple to devise and put into practice. It reduces calculation time, reducing the overall duration of the emulator design process. PVE considers converters in a closed loop circuitry. Mason's Gain Formula gives:

$$\text{Forward Path Gain} = \frac{\bar{D}_1}{L} * \frac{1}{s} * \frac{1}{C} * \frac{1}{s}$$

$$\text{Loop Gain, } L_1 = \frac{-1}{sRC}$$

$$\text{Loop Gain, } L_2 = \frac{-1}{s^2RC}$$

$$\Delta = 1 + \frac{1}{sRC} + \frac{1}{s^2RC}$$

$$\text{Transfer function; } \frac{\bar{V}_o}{\bar{V}_{in}} = \frac{\bar{D}_1 * R}{RLCs^2 + Ls + R} \quad (11)$$

$$\text{Transfer function; } \frac{\bar{V}_o}{\bar{d}} = \frac{V_{in}}{LC(s^2 + \frac{1}{RC}s + \frac{1}{LC})} \quad (12)$$

III. SLIDING MODE CONTROLLER (SMC)

The goal of PVE's control strategy is to properly follow PV model signals, decrease processing load, generate reliable emulator output and simulate a range of PV modules without having to rethink the entire control strategy and influencing the power converter system and load. Nonlinear control approaches are investigated in order to maintain stability and effective control in massive signal conditioning, as well as to increase the model's dynamic response or robustness. SMC provides robustness and quick reaction to fluctuations in supply, load, and circuit parameters [13][14]. PV emulator configuration uses a buck converter in closed loop with SMC and diode model. T and d are the switching interval and duty cycle, respectively.

From above equations (1) and (2)

$$X = f(X, t, d) \quad (13)$$

When the system approaches the sliding surface, the controller aims for the output to reach the desired value (given by either a diode model or a LUT). Partial or complete state variables can reach this surface. A new portion akin to an integral with a constant is added to the surface equation to minimize the error between the actual and intended value. As a result, the new sliding surface looks like this:

$$S = a_{in} + b_{vo} + m \int e dt \quad t=0 = 0 \quad (14)$$

$S > 0$ in open state and $S < 0$ in closed state. In steady state, when the system is kept on sliding surface, the derivative of S becomes zero and the stable point is when $e=0$, $v_r = v_o$. So, sliding surface equation becomes; $a_{in} = - b_{vo}$. For buck converter;

$$\dot{S} = [a \quad b] * \dot{X} + m \int_0^t e dt = 0 \quad (15)$$

$$\dot{X} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix}$$

$$\dot{x}_1 = \frac{v_i}{L} d_{eq} - \frac{x_2}{L}$$

$$\dot{x}_2 = \frac{x_1}{C} - \frac{x_2}{RC}$$

(16)

At steady state, placing $\dot{S} = 0$ gives;

$$d_{eq} = \frac{aRCx_2 - Lb(x_2 - Rx_1) + RLCm(V_r - x_2)}{aRCV_i} \quad (17)$$

For above defined system to be stable; $v_0 = v_r$, the error must be zero. And $v_0 = x_2$ therefore $v_0 = v_r = x_2$

$$\text{Also, } i_c = x_1 - \frac{x_2}{R} = i_L - \frac{v_o}{R} \text{ and } d_{eq} = \frac{v_r}{V_i}$$

Thus, a new control input independent of load, R . The new control equation is represented below:

$$d_{eq}(t) = \frac{x_2(aC - LCm) + Lbi_c + LCm(V_r)}{aCV_i} \quad (18)$$

With d_{eq} , the system's closed loop dynamics converge towards a sliding surface. The system will reach the sliding surface in finite time if $s = 0$. In the presence of parametric uncertainties, a d_{eq} estimate is available. The solution of the system simulated and checked by routh Hurwitz criteria indicates that the system is stable. As a result, the inferred control law can be used to achieve desired results in ambiguous environments. The dynamic behaviour of the system is due to changing load conditions and irradiance due to partial shading of panel. Thus, it is proved that the main advantage of choosing SMC controller over the traditional controllers like P, PI and PID [15-16] is existence of stability and robustness performance in uncertain system, where these controllers fail specially in the case of uncertain environment.

IV. RESULTS AND DISCUSSION

Simulation is used to evaluate the system performance under dynamic situations. The converter's SFG methodology resulted in detailed small signal, large signal and steady state analysis with the minimum possible changes. As previously stated in section 1, SFG-based methodology assisted in obtaining transfer functions for various circuit variables in a more transparent and concise way. In the following section, the corresponding sliding coefficients are calculated. Figure 3 shows the V and I plots Vs Load resistance using a LUT-PID controller, which depicts the PVE performance of the PV model. Figure 4 depicts the viability of the PV model in conjunction with the diode model. Figures 3 and 4 demonstrate that the characteristics of the diode-SMC model also yield continuous values near MPP. LUT-PID and SMC models, on the other hand, show discontinuity at specific load points. As a corollary, the diode-SMC model is preferred for PVE implementation. For dynamic loading, the system VI and PV plots (shown in figures 5(a)(b) and 6(a)(b)) are obtained. The diode-SMC model's characteristics show a high level of accuracy for the planned solar panel. They also produce good results in conditions such as open circuit, maximum power point and short circuit. Figure 7 illustrates a comparison of the finalized model under standard test conditions with 60% irradiation. Implementation of robust diode-SFG-SMC achieves the goal of obtaining an emulator that follows the actual panel behaviour, and the results are made more efficient by computing transfer function at switching states. The suggested controller's performance is evaluated using derivations derived from mathematical modelling with SFG and simulations performed in MATLAB. The analysis in this work is based on ideal buck converter components and the results are very close to the actual panel results

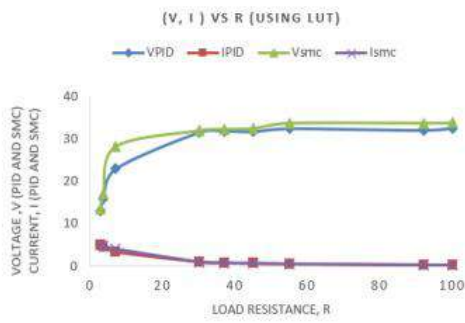


Fig 3. Voltage and Current Vs load Resistance using LUT

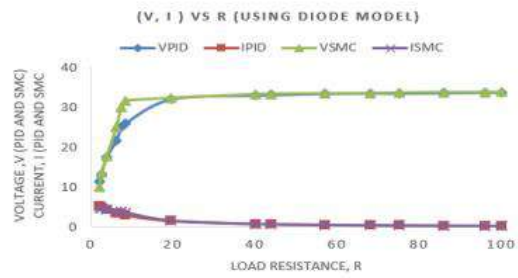


Fig 4. Voltage and Current Vs load Resistance using Diode

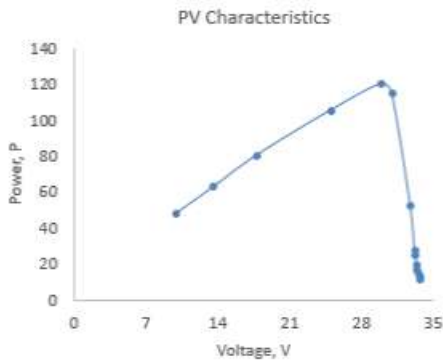


Fig. 5(a). V-I Characteristics of PVE using LUT

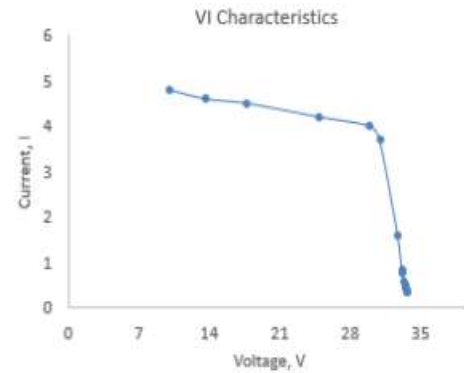


Fig. 5(b) P-V Characteristics of PVE using LUT

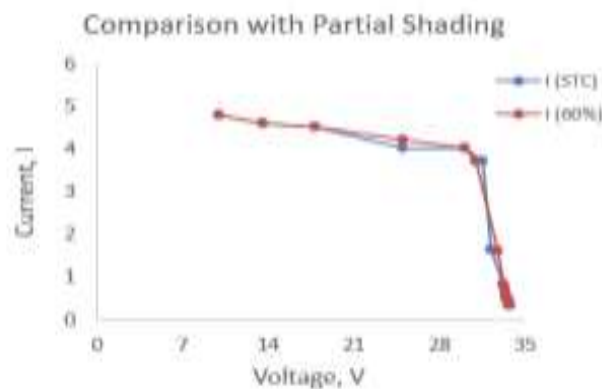
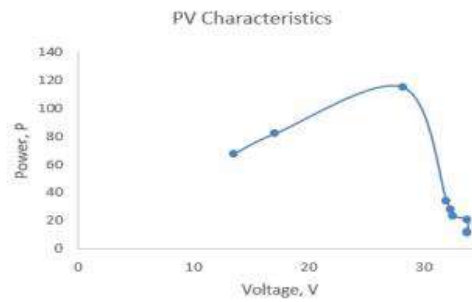
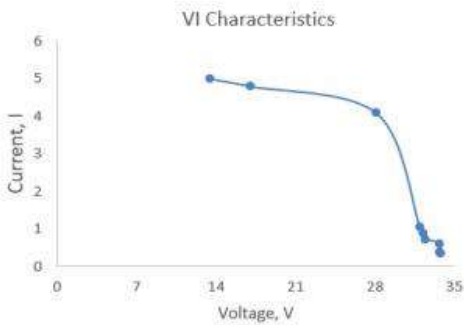


Fig. 7 V-I Characteristics of PV Emulator using diode-SMC model at STC and partial shading. For stability study, the model's time domain performance for PID and SMC controllers is studied. Because the model is based on actual PV panel parameters, both stability and accuracy are important. Figure 6 shows the time domain analysis of the PID controller under various load conditions. Figures 8(a) and 8(b) depict PID and SMC controllers under dynamic loading (b). PID controller overshoots in various places, whereas SMC offers overshoot values that are acceptable

for stability.

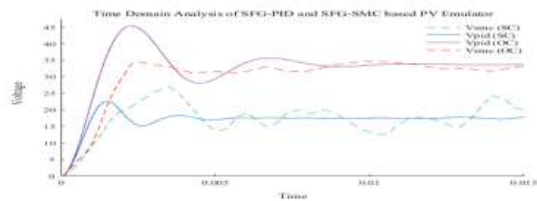


Fig 8(a) Time Domain comparison of SFG-PID and SFG-SMC at open circuit and short circuit

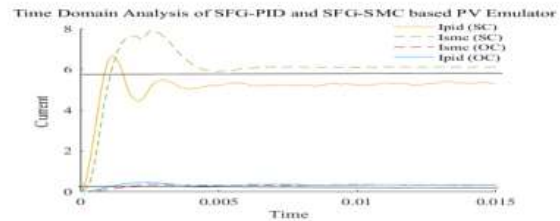


Fig 8(b) Time Domain comparison of SFG-PID and SMC at open circuit and short circuit

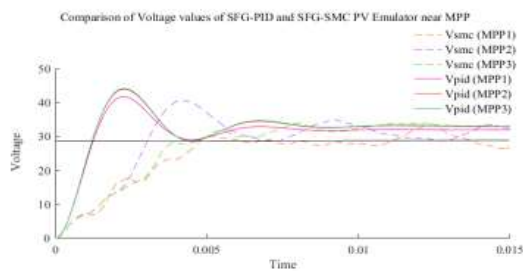


Fig 9 (a) Comparison of voltage values of SFG-PID and SFG-SMC based PVE near MPP

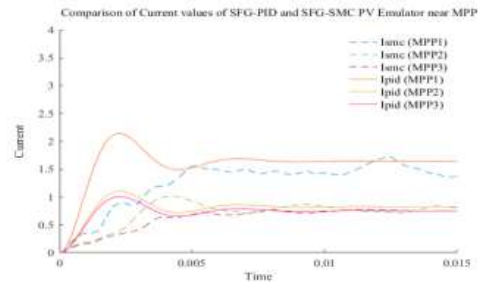


Fig 9(b) Comparison of current values of SFG-PID and SFG-SMC based PVE near MPP

The modeling and design process is described in detail, taking into account the various features of PVE operating conditions. Modeling of reference values is one of the most important factors determining computation efficiency. The main task was to estimate a non-linear characteristics curve. As a result, they are produced using a LUT and a single diode model. The SFGPID and SFG-SMC approaches use equations developed from an analysis of the converter's dynamic behaviour during sliding mode operation, as well as the system's stability criteria. SMC can be used in PVE design, as shown in Figures 6 (a) and (b), especially where the value of short circuit parameters is perhaps the most essential factor. Figure 7 depicts a comparison of the Emulator model under standard test conditions and partial shading (with only one colour) (60 percent of irradiance). The stability and accuracy of voltage and current measurements are considered in Figures 8(a) and (b). Overshoots observed at various operating ranges are used to determine system stability. The open circuit, short circuit and maximum power points are specified on the data sheet of the PV panel, hence the final comparison in figures 8 and 9 is based on these points. Figure 9 reveals that in the case of PID, there are more overshoots (voltage and current waveforms). The results depicted in the graphs above are indeed very close to the actual panel used in the model. The goal of conducting a comprehensive analysis of a PV panel is completed and the simulation results are reviewed to illustrate a comparative comparison of non-linear and dynamic PV panel behaviour.

V. CONCLUSION

The proposed methodology has developed a method for solar panel emulation considering its Non-linear and dynamic constraints. The method would speed up the installation of solar panels in rural and remote areas. The proposed approach will aid in the continual testing of Non-Linear panels with minor modifications. Regardless of the irradiance conditions, the findings obtained by SFG mathematical modeling and simulation of the PVE model in MATLAB show close agreement with the actual PV panel. The graphs illustrate that the SFG-based model is a viable alternative for representing switching converters with complex structures. The findings of the reference diode model are satisfactory at short circuit values, but LUT results lag accuracy at specific loading points, according to the research. The SFG-SMC model can solve the accuracy, robustness and stability of the system in emulator design. In the future, there is a lot of scope for research in the current system. Advanced non-linear controller designs can be used to accomplish potential research. One of the most key aspects of this system is that the methodology established here may be adapted to different converter types for use in power electronics and control systems.

REFERENCES

- [1] Economic times Jan ,2018/industry/energy/power/
- [2] <https://www.iea.org/reports/world-energy-outlook2021/a-new-energy-economy-is-emerging/>.
- [3] M.C. Di Piazza, G. Vitale, “Photovoltaic Sources: Modeling and Emulation”, Springer Verlag London, 2013.
- [4] Kashif Ishaque, Zainal Salam, Hamid Taheri, Syafaruddin, “Modeling and simulation of photovoltaic (PV) system during partial shading based on a two-diode model”, J. Simulation Modelling Practice and Theory., Vol.19(7), pp. 1613-1626, Aug. 2011.
- [5] Iqbal MT, Tariq M, Ahmad MK, Arif MSB, “Modeling, analysis and control of buck converter and Z-source converter for photo voltaic emulator”, IEEE conf. proc. Power Electro. Intelligent Control and Energy Sys (ICPEICES), pp. 1–6, 2016.
- [6] A.J. Forsyth and S.V. Mollow, “Modelling and control of DC-DC converters,” IEEE Power Engg. J., Vol. 12(5), pp. 229–236, Oct 1998.
- [7] M. Ahmed, M. Kuisma, K. Tolsa, P. Silventoinen, “Implementing Sliding Mode Control for Buck Converter”, IEEE Annual conf. Power Electro. Specialist, PESC, pp. 634-637, June 2003.
- [8] J. Mahdavi A. Emadi, H.A. Toliyat, “Application of State Space Averaging Method to Sliding Mode Control of PWM DCDC Converters”, IEEE Industry Applications Society Annual Meeting, pp. 820-827, Oct.1997.
- [9] Hanifi Guldemir, “Study of Sliding Mode Control of Dc-Dc Buck Converter”, Energy and Power Eng., pp. 410-406, Sept. 2011.
- [10] K. Smedley, S. Cuk, “Switching Flow Graph nonlinear modeling technique”, IEEE Trans. Power Electro., Vol.9, no.4, pp 405-413, July 1994.
- [11] Jia Yao, Alexander Abramovitz, “Fast SFG Modeling of Integrated Converters”, IEEE J. Emerging and selected topics in Power Electronics, Vol.5(3), pp. 1008-1019, Sept 2017.
- [12] Majid Abbasi, Ahmad Afifi, Mohammad Reza Alizadeh Pahlavani, “Signal flow graph modelling of a switching converter with single inductor triple output DC–DC structure”, IET Power Electri, Vol. 11(7), pp. 1195-1204, April 2018
- [13] A. Gupta, D. Joshi, “Comparative Analysis of NonLinear SMC controller with linear PID controller for Flyback Converter”, Springer ESIEE book series, pp. 71-87, Sept.2021.
- [14] Y. He and F. L. Luo, “Study of Sliding-Mode Control for DC-DC Converters”, Int. conf. Power System Technology – POWERCON, pp.1969-1974, Nov. 2004.
- [15] S. J. Gambhire, D. Ravi Kishore, P. S. Londhe, S. N. Pawar, “Review of sliding mode-based control techniques for control system applications”, Int. J. Dynamics and Control, Springer-Verlag Germany, pp. 1-16, May 2020.
- [16] Simmi Sharma, Dheeraj Joshi, “PV Emulator Model Design using AI based controllers”, Springer Singapore, Proc. of 3 rd Int. conf. on Computing Informatics & Networks: ICCIN, pp. 439-448, July 2021.

XPEN – A Voice Powered Expense Tracker Full Stack Web Application

Ms. Kritika¹, Dr. Himani², Ms. Shikha³

^{1,2,3}Assistant Professor

Department of Computer Science and Engineering (CSE)

Bhagwan Parshuram Institute of Technology

Delhi, India

Abstract— Budgeting is an integral part of society. From the beginning of civilization humans have been selling and purchasing goods. Since then, it has become an essential and irreplaceable part of our lives. Many of us people live in middle class families. We have to track our daily budget to avoid the debt. Tracking our expenses daily can save our amount, but it can also help us set financial goals for the future. If we know exactly where our amount is going every month, we can easily see where some cutbacks and compromises can be made. Most of us have a fixed income and we get it on time (i.e., daily, monthly, annual, etc.). In addition, everyone follows a strict budget of spending. Generally, the budget is assembled according to category. Categories vary, for example, food, entertainment, transportation, education, health, clothing, and so on. However, spending is limited to budget revenue. For this reason, we need to keep track of our expenses so that they do not exceed our budget. People often track expenses using pen and paper methods or take notes on a mobile phone or a computer. These processes of storing expense require further computations and processing for these data to be used as a trackable record. In this work, we are proposing an automated system named XPEN to store and calculate these data. XPEN is a web application which aims to assist a user in managing personal or family finances by offering not only a basic expense check but also a brief analysis of incomes and expenses. All operations are performed in private through a personal account. Voice capabilities are also added to this application with the help of which users can save their expense by just speaking. The application solves all of the problems that occur with old ways of entering data. It uses the voice to traverse the app. This

application filters the keywords from the user’s voice and saves the amount and description for further processing. All transactions of the user are accumulated to the daily, weekly and monthly total sum and visualized as a histogram. The user then gets a better understanding of the weekly, monthly and yearly fluctuation of incomes and expenses. XPEN can manage daily expenses much faster than any other traditional app in the market which takes manual input. Overall, this is a smart automated solution for tracking expenses.

Keywords—Expenditure, Track Expense, VUI, Voice Recognition, Voice Interface, User Interface, Automatic Speech Recognition

I. INTRODUCTION

From the beginning of human civilization, people have exchanged their fortune with each other for buying or selling goods. It has become a crucial and unchangeable part of our daily life since then. Most of us have a fixed income and we get it in a timely basis (i.e. daily, monthly, yearly etc.). Moreover, everyone follows a strict budget of expense. Generally, the budget is assembled as per category. The categories are distinct, for example, food, entertainment, transportation, education, healthcare, clothing etc. However, the budget of expense is restricted to the income. For that reason, we need to track our expense so that it doesn’t exceed our budget. In old days, people used to track their expense manually i.e., using pen and paper system which takes a lot of effort and time.

Nowadays, the availability of electronic devices like smartphones, computers have made our life a lot easier and faster. We can use computers to track our daily expense by using the online and offline software available. But the computer is not accessible all the time. The smart solution to the problem is to use smartphones. Nearly 44% of the world population use smartphones [1]. Smartphones have become an irreplaceable part of our daily lives as they are always accessible on the go. There are some existing applications that can track daily expense [2]–[5]. These applications use a manual input system from the keyboard which is tiresome and time-consuming. To meet the challenge of avoiding manual input, we are proposing a smart method of doing the same work but in a more automated and efficient way which takes less time.

In the proposed approach, users can save their expense simply by just speaking. To capture the user’s voice, the ‘sounddevice’ library for python has been used. The ‘wavio’ package in python has been used for file I/O to save the audio in wav format. The ‘wav’ format provides lossless

compression for the audio which means that no detail from the input voice is left out. The python’s ‘Speech Recognizer’ library is used for the conversion of received audio file into text. This library takes in a ‘wav’ file as an input and predicts the spoken text/ transcript as an output. The Processed voice command after conversion to text needs to be interpreted as an executable command/sequence. This mapping will be done by wit.ai. The ‘wit.ai’ model is based on a probabilistic language model developed on artificial neural network.

The sections are structured as follows: in Section 2, the background analysis is covered. It includes some of the popular applications which are available. Section 3 defines our approach, Section 4 is the implementation and Section 5 includes a conclusion.

II. RELATED WORKS

Expense management apps are very common in the application market. Many of them offer exciting features. Different apps have taken different approaches to manage the daily expense.

Daily Expense 3 [2] is a system that can track income and expense and classify them into categories. The application shows reports grouped by periods. Users can also schedule their recurring records. The application also creates a backup of their records to restore information if necessary.

AndroMoney [5] supports multiple accounts to manage expense and income. It uses cloud storage so that the data is safe. Users can set a budget for the expense and the app will notify if they exceed the budget. It provides a number pad to calculate any record. It generates trend, pie and bar charts for cash flow.

Monefy [3] - money manager is an expense management application that has an intuitive interface. It can store a record faster than the others mentioned above. It provides widgets to enhance ease of access. It provides default categories and option to add customized categories. It also provides an onscreen calculator to calculate the expense.

Expense Manager [12] is a feature-rich application, which is interactive and well balanced. Besides tracking user’s income and expense it also saves the picture of a receipt. It also tracks tax, mileage, and debts. It provides some convenient tools such as currency converter, regular calculator, tip calculator, loan calculator, credit card payoff calculator etc.

The next application which is being used to track expense is Money Lover [4]. This application manages expense category wise. It also manages income, debts, and loans. Users can set events and save plans on this app. It saves the receipt. It also provides a currency converter and a calculator.

In our work, we have developed a system which has the functionalities of a traditional expense management web applications along with the voice capability to take automatic input. The system is more interactive and operational than all the other solutions mentioned above.

III. PROPOSED SYSTEM

The project is a MERN stack application, The front-end will be developed in HTML5, CSS3, and React.js while the back- end will be created in Express and Node.js. The data will be stored in MongoDB Atlas and users can only access their own data not others.

A. User Story

After login, the user can access the following key features:

- (i) *Add/ Edit/ Delete a transaction:* The user can create a new income or expense transaction which can be deleted or modified after the creation.
- (ii) *Check recent transactions:* The user will see all transactions took place in the last three days and the sum of income, expense and balance in the personal dashboard.
- (iii) *Check history transactions:* The user can also check the list of previous transactions along with the sum and ratio of income and expense in a day, a week, a month and a year.

(iv) *Get an overview of incomes and expenses:* All transactions of the user are accumulated to the daily, weekly and monthly total sum and visualized as a histogram. The user then gets a better understanding of the weekly, monthly and yearly fluctuation of incomes and expenses.

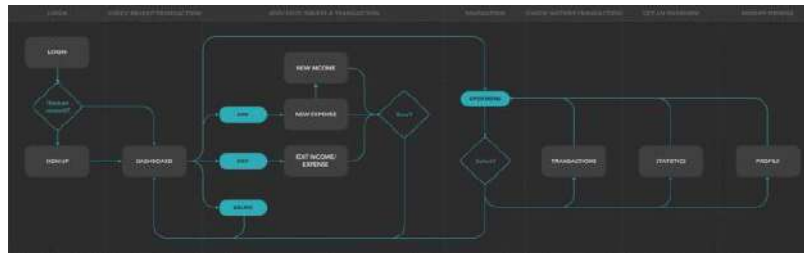


Fig. 1. User Stories

B. Web Development

The project is a MERN stack application, deployed on Heroku. Since personal financial records are sensitive information, it is necessary to keep personal data in private. Therefore, user authentication was required and verified with JSON Web Token (JWT). By doing so, a user can login and log out with a personal account.

(i) Front-end

The front-end was built with responsive modern technologies (HTML5, CSS3, React.js). The data analysis was achieved by visualizing transaction data with D3.js.

(ii) Back-end

A REST API for the project was created with Express, Node.js while the data was stored in MongoDB Atlas, a NoSQL and document-oriented database in the cloud. Then, a user can perform CRUD operations through this API such as create, read, update, and delete data.

C. VUI System

The architecture of the complete VUI system consists of modules: The details of various modules are described below:

A. Speech Recording

To capture the user's voice, the 'sounddevice' library for python has been used. The library provides recording of audio signals via single channel and dual channel mode. It also allows the users to select the primary recording device. This library allows the storage of sound in numpy arrays as digital sampled signal at the given frequency. The Numpy arrays are thus processed to get the text from the audio signal.

B. Audio File Handling

The 'wavio' package in python has been used for file I/O to save the audio in wav format. The 'wav' format provides lossless compression for the audio which means that no detail from the input voice is left out. The wavio package stores the output sound signal as an uncompressed 8 bit pcm format. The number of bits in the PCM format denote the depth at which the sound has been captured. Higher bits implies higher detail. However the 8 bit format is the best one to choose due to its relatively less space complexity and good enough quality for speech recognition.

C. Speech To Text API

The python's 'Speech Recognizer' library is used for the conversion of received audio file into text. This library takes in a 'wav' file as an input and predicts the spoken text/ transcript as an output. The SpeechRecognizer library can be used in 2 variants. online mode and offline mode. As ours is a web application, we are using online mode. This mode uses various online speech to text tools to convert speech to text. One API that is used is the Google's web search speech api. The advantage of this mode is that the major players in the speech to text grounds can be utilized and the speech recognition is accurate.

D. Natural Language Processing API

The Processed voice command after conversion to text needs to be interpreted as an executable command/sequence. This mapping will be done by wit.ai. Every voice command that is given to the computer has 2 parts:

- 1) Intent: This is the objective of the user to complete a specific work/task that is given to the computer.
- 2) Parameter: The parameter is the object on which the intent has to be performed.

The two part bifurcation of the NLP model make its highly easy to understand and interpretable. The 'wit.ai' model is based on a probabilistic language model developed on artificial neural network. This model is based on the theory of n-grams for probability prediction i.e. it can predict the probability of a given word in a sentence or a statement by usage of conditional probability on the statement from the statistics of the text corpus. The Model can predict the probabilities of intents and parameters from sentences similar to the trained sentence. Along with the use of probability, the model also makes use of text labelling also known as Named Entity Recognition (hereafter referred to as NER). The NER labels the parts of the text to the pre-defined labels. It is achieved by finding patterns in the text that has been used for the training of the model. For example, consider the sentence Add Expense, the intent identified is Add and the argument identified is Expense.

E. System Programming

The system programming module does the role of executing the actual commands on the host machine. For implementation, following system API operations were added: Add Expense, Add Income, Add Rs. <Amount>, Add in Category Travel, Add in Category Entertainment, Add in Category Food, Add in Category Clothing, Add in Category Health, Add in Category Education.

IV. IMPLEMENTATION

XPEN is divided into seven major sections. Those are Login, Check Recent Transaction, Add/Edit/ Delete a Transaction, Navigation, Check History Transaction, Get an Overview and Modify Profile. This system works as a one tap solution for tracking everyday expense. It also preserves yearly and monthly records. For the availability of the records, users can check their histories to keep track of their expense so that they do not exceed their pre-allocated budget. First of all we broke up the UI in components. We used context api. So, we have global state, global context api, we passed everything to the components. Basically we put all the transactions in the global state, and passed it to the components, then we did all the calculations stuff inside that. The transactions are stored in MongoDB database.

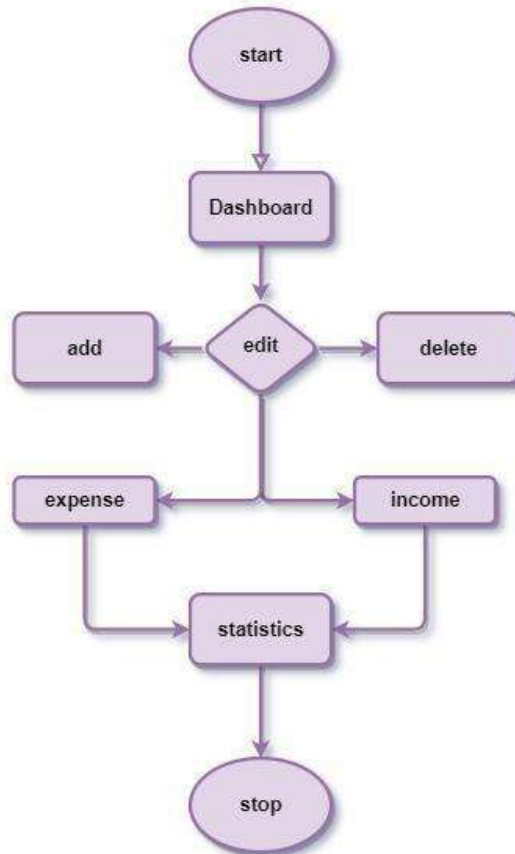


Fig. 2. Flow of UI Components of the App

In this project there is only one user. The user start the program on the user dashboard it show three option add,edit and delete after its calculate and add to database and show final result in statistics and user stop the program.

The screenshot shows a web form titled "Add new transaction". It contains two input fields: "Text" with a placeholder "Enter text..." and "Amount" with a placeholder "Enter amount...". Below the "Amount" field, there is a note: "(negative - expense, positive - income)". At the bottom of the form is a blue button labeled "Add transaction".

Fig. 3. Add Transaction

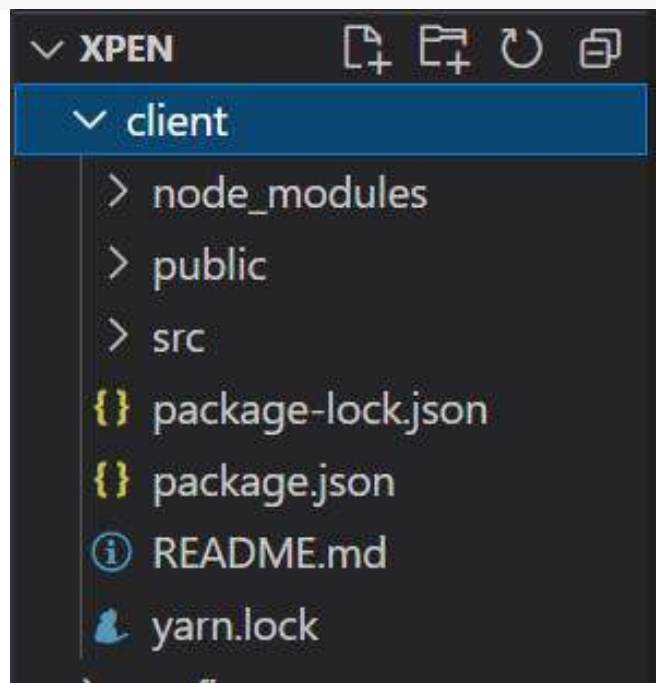


Fig. 4. Balance Section of the App



Fig. 5. History

Section of the App



Folder

Structure

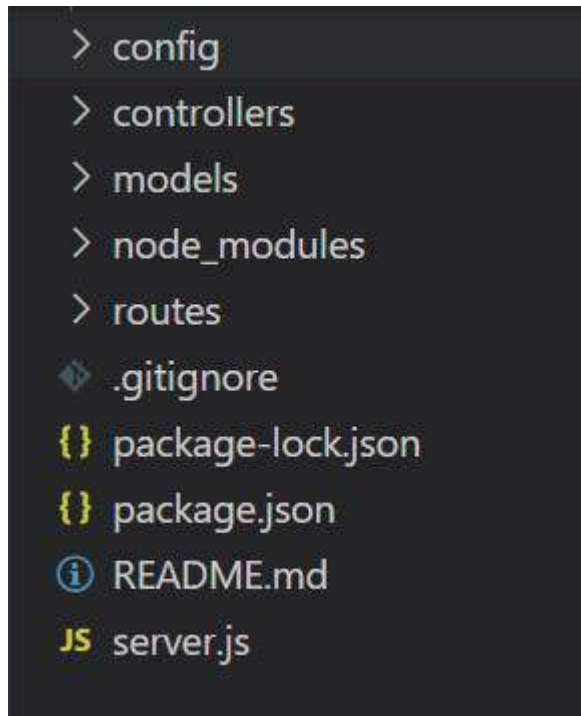


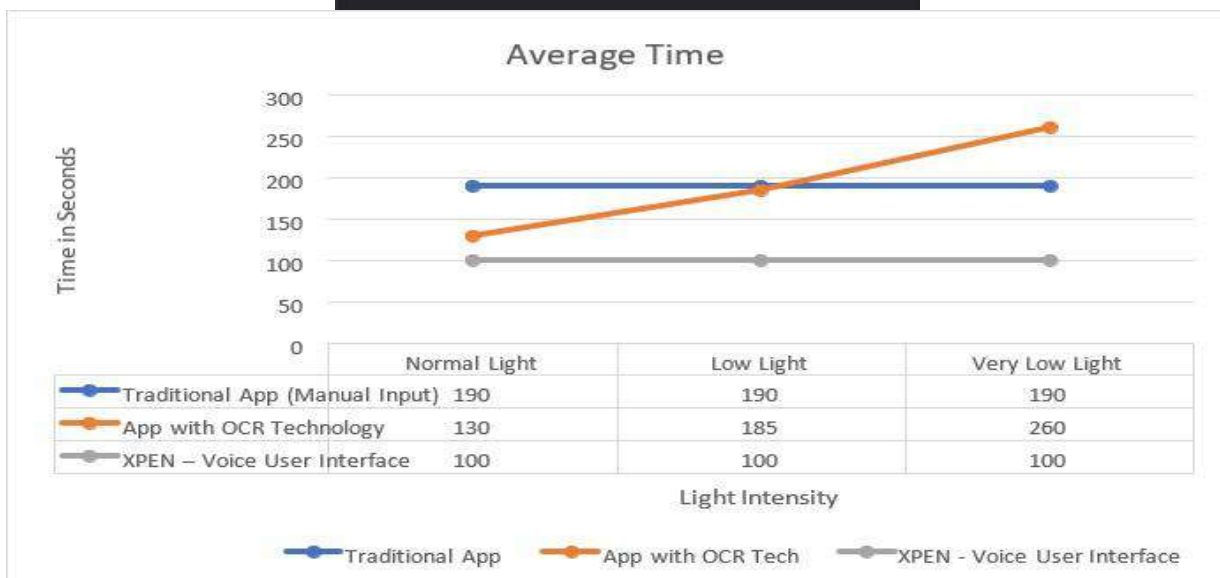
Fig. 7. Server side of the App

Side of

V. ANALYSIS

Fig. 8 Average time taken to add or delete a transaction

time taken to add or



There are mainly four tasks that can be performed by the user in the XPEN app. There are “add income”, “add expense”,

“delete income” and “delete expense”. The user can perform any one of these tasks. So, let’s say the user performs any one of these tasks repeatedly for several number of times. To find the average time taken to complete a task, divide the total time

Fig. 6. Client Side of the App taken to performed the tasks by the total number of tasks performed by the user. There are three series in the graph namely traditional app, app with OCR tech and XPEN. There are three categories namely normal light, low light and very low light.

There is no effect of light intensity on the traditional app and the XPEN [VUI] app. But the performance of app using OCR technology decreases when the intensity of light decreases. The character recognition from the receipt declines in low lights. XPEN performs better than both the traditional app and the app with OCR technology. It takes lesser time to complete all the tasks.

VI. CONCLUSION

In today’s world, time is the most valuable asset because people lack ample of it. People are

obsessed with completing tasks in lesser time and our system is an approach serving this purpose. XPEN can manage daily expense much faster than any other traditional app in the market which takes manual input. Nowadays, the world is leaning towards the one tap solution and our system is one of a kind. After all, automation is the way of future and XPEN can be a step towards it. In the future, the AI technologies of speech recognition and generation is of the most difficulties to break through. With developed AI technologies, VUI can make machines hear and talk humanly, so the convenience and accessibility of UI will come to another leap after GUI. In addition, from the aspect of ethic, new policies ensuring the fairness of access to VUI should be put forward so that people in the society could take advantage of VUI equally.

There are mainly four tasks that can be performed by the user in the XPEN app. There are “add income”, “add expense”, “delete income” and “delete expense”. The user can perform anyone of these tasks. So, let’s say the user performs anyone of these tasks repeatedly for several number of times. To find the average time taken to complete a task, divide the total time taken to performed the tasks by the total number of tasks performed by the user.

Therefore: There is no effect of light intensity on the traditional app and the XPEN [VUI] app.

VII. REFERENCES

- [1] “44% of World Population will Own Smartphones in 2017.” [Online]. Available: <https://www.strategyanalytics.com/strategyanalytics/bl ogs/smart-phones/2016/12/21/44-of-world- populationwill-own-smartphones-in- 2017#.Wq1BcehuY2y>. [Accessed: 17Mar-2018].
- [2] “Daily Expense 3 - Apps on Google Play.” [Online]. Available: <https://play.google.com/store/apps/details?id=mic.app. gastosdiarios>. [Accessed: 17-Mar-2018].
- [3] “Monefy - Money Manager - Apps on Google Play.” [Online]. Available: <https://play.google.com/store/apps/details?id=com.mo nefy.app.lite>. app performs poorly if the input image from the receipt includes a lot of noises. The app is unable to detect the region of interest automatically. On that note, the user has to set the region of the receipt to have a better result. Again, the performance of the [Accessed: 17-Mar-2018].
- [4] “Money Lover: Budget Planner, Expense Tracker - Apps on Google Play.” [Online]. Available: <https://play.google.com/store/apps/details?id=com.bookmark.money> . [Accessed: 17-Mar-2018].
- [5] “AndroMoney (Expense Track) - Apps on Google Play.” [Online]. Available: <https://play.google.com/store/apps/details?id=com.kpmoney.android> . [Accessed: 17- Mar-2018].
- [6] “Expense Manager - Apps on Google Play. [Online]. Available: <https://play.google.com/store/apps/details?id=com.expensemanager>. [Accessed: 17-Mar-2018].
- [7] Sabab, S. A., Islam, S. S., Rana, M. J., & Hossain, M. (2018, September). eExpense: A Smart Approach to Track Everyday Expense. 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (ICEEiCT).<https://doi.org/10.1109/ceeict.2018.8628070>
- [8] Yurochkin, D. E., Horoshiy, A. A., & Karpukhin, S.A. (2021, January). Development of an Application for Expense Accounting. 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). <https://doi.org/10.1109/elconrus51938.2021.9396201>
- [9] Dabre, P., Gonsalves, R., Chandvaniya, R., & Nimkar, A. V. (2020, April). A Framework for System Interfacing of Voice User Interface for Personal

Computers. 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA).

<https://doi.org/10.1109/cscita47329.2020.9137809>

[10] Begany, G. M., Sa, N., & Yuan, X. (2015). Factors Affecting User Perception of a Spoken Language vs. Textual Search Interface: A Content Analysis. *Interacting with Computers*, iww029.

<https://doi.org/10.1093/iwc/iww029>

[11] Zhang, X., Lu, C., Yin, J., Xie, H., & Liu, T. (2020).

The Study of Two Novel Speech-Based Selection Techniques in Voice-User Interfaces. *IEEE Access*, 8, 217024–217032.

<https://doi.org/10.1109/access.2020.3041649> [12] 2016 International Conference on Computational

Techniques in Information and Communication Technologies (ICCTICT)

[13] *International Journal of Commerce & Business Studies* Volume-4, Issue-1, January-March, 2016, www.iaster.com ISSN (O) 2347-2847 (P) 2347-8276

[14] 2020 International Conference on Computing and Data

Science (CDS) 978-1-7281-7106-7/20/\$31.00 ©2020 IEEE DOI 10.1109/CDS49703.2020.00045

[15] 2020 International Conference for Emerging

[16] Technology (INCET) Belgaum, India. Jun 5-7, 2020 978-1-7281-6221-8/20/ ©2020 IEEE 1
Authorized licensed use limited to: Cornell University Library. IEEE Xplore. Restrictions apply

T. Kita, C. Nagaoka, N. Hiraoka and M. Dougiamas, "Implementation of Voice User Interfaces to Enhance Users' Activities on Moodle," 2019 4th International Conference on Information Technology (InCIT), 2019, pp. 104-107, doi:10.1109/INCIT.2019.8912086.

Optimization: Background and Existing Challenges

Neha Khanduja

Dept. of Electrical & Electronics Engineering
Bhagwan Parshuram Institute of Technology

Delhi, India

orcid.org/0000-0001-9386-1173

Abstract—Researchers working on problems in engineering, computer science, biology, and the physical sciences are developing advanced mathematical methods for control. Technological advances have had a major impact on the use of new analytical methods for dealing with nonlinear problems. One of the most challenging parts of control theory is tuning the parameters of nonlinear systems for an optimum solution. In the past, metaheuristic methods were tried to address this problem. They have proved to be useful when dealing with complex systems. Metaheuristic optimization techniques, unlike deterministic algorithms, excel at addressing problems with uncertain search spaces. Optimization-based control is now favored over conventional or intelligent control.

Keywords—optimization, metaheuristic optimization, local search, global search

I. INTRODUCTION

Advanced mathematical techniques for control are being developed by researchers working on issues in engineering, computer science, biology, and the physical sciences. The application of novel analytical approaches for tackling nonlinear issues has been significantly influenced by technological advancements [1]. State may not be entirely quantifiable in most situations involving nonlinear control systems, making complicated control engineering problems difficult to address. The employment of a variety of distinct models and ideas, a lack of parameter standardization, a lack of suitable control approaches, external disruptions, and the greater level of nonlinearity of the equations that drive processes are all important challenges in the field of control technology. Another difficulty is a lack of understanding of critical variables, since the system's states might significantly affect the nature of the control design stage, allowing for excellent performance. As a result, it is clear that enhanced forecasting, control, and optimization approaches are required to ensure optimal nonlinear system performance. Understanding the system's control needs necessitates knowledge of the system; nevertheless, nonlinearities are frequently so complicated that control design for acceptable system performance is challenging [2]. New control techniques have developed over time to maintain optimal system performance that prevent interruptions, pauses, and design flaws.

Tuning the parameters of nonlinear systems for an optimal solution is one of the most difficult aspects of control theory. Metaheuristic strategies have been used to solve this challenge in the past. When dealing with complicated systems, they have proven to be beneficial. Unlike deterministic algorithms, metaheuristic optimization methods excel at solving problems with uncertain search spaces. These optimization approaches have been utilized in practically every sector of research, technology, and engineering to discover the best answer from a number of feasible solutions [3].

II. BACKGROUND AND EXISTING CHALLENGES

A. Optimization

An important paradigm which is everywhere along with wide range of utilizations is

optimization. In practically all application areas such as mathematics, computer science, operation research, industrial and engineering designs, we are continually attempting to upgrade something - regardless of whether to limit the expense and vitality utilization, or to expand. The benefit, yield, execution and effectiveness. In all actuality, assets, time and money are consistently restricted; thus, optimization is unmistakably progressively significant [4].

One of the most key standards in our reality is the quest for an ideal state. It starts in the microcosm where molecules in material science attempt to frame bonds so as to limit the vitality of their electrons. At the point when particles structures strong bodies during the freezing process, they attempt to accept vitality optimal crystal structures. These procedures, of course, are not driven by any higher aim yet simply result from the laws of material science. The equivalent goes for the natural guideline of natural selection which, along with the organic development, prompts better adjustment of the species to their environment. Here, a nearby (local) optimal is a very much adjusted animal groups that rules every single other creature in its environmental factors. Homo sapiens have arrived at this level, imparting it to ants, microorganisms, flies, cockroaches, and a wide range of other creatures. For whatever length of time that mankind exists, we take a stab at flawlessness in numerous territories. We need to reach a most extreme level of joy with minimal measure of exertion. In our economy, benefit and deals must be expanded and expenses ought to be as low as could be expected under the circumstances. In this way, optimization is one of the most established sciences which even stretches out into everyday life [5].

Optimization is the study of choosing the best choice among a debilitated hover of choices [1] or it tends to be seen as unitary of the major quantifiable mechanism in system of dynamic in which judgments must be employed to enhance single or more evaluations in some affirmed set of conditions [6].

Most of the engineering and industrial design problems are based on computer simulations, which results in added complications like non-linear constraints, interdependencies amongst variables and a large solution space to optimization. Any approach which can accelerate the time of simulation and optimization process results in saving of time and money. Thus, methods of optimization can be defined as mechanism specifically designed to attain the objective of minimizing or maximizing a fitness function (or objective function) subject to given set of constraints. It must give enough good solution in enough time frame [7][8].

Each problem of optimization accompanies some decision variables, certain objective (fitness) function and few constraints. Need of employing optimization techniques is to acquire the estimations of decision variables that optimize a fitness function subject to specific constraints [4]. Decision variables are inputting which can be controlled and thumb rule of any optimization problem is to choose minimum number of design variable. The following undertaking in the optimization is to locate the fitness or objective function in terms of the design variables and other problem parameters the constraints show functional relationships among the design variables and other design parameters satisfying certain physical phenomenon and certain resource limitations. The nature and number of constraints to be included in the formulation depend on the user. Constraints may have exact mathematical expressions or not [5].

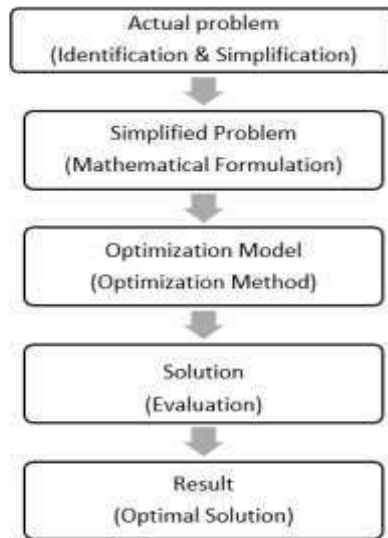


Figure 1.1 Flow chart of Optimization

This optimization vocabulary can be understood by this example:

A football coach is putting up a practice schedule for his defensive players.

- The main optimization will be to achieve maximum running yards, which will be his goal function.
- He may have his players spend time in the weight room, sprinting, or practicing ball protection during practice. A decision variable is the amount of time spent on each. However, the overall amount of time he has is limited. Also, if he totally foregoes ball protection, he may notice an increase in both rushing yards and fumbles, therefore he may set a restriction on how many fumbles he finds acceptable. These are boundaries.
- The objective function is influenced by the decision (or planned) variables, and the constraints restrict the scope of the variables[6]

An optimization algorithm is characterized by

1. the methodology through which it assigns fitness to individual
2. way of selecting an algorithm for future analysis
3. approach of applying search operations
4. the way it builds and treats its state information Literature review of optimization algorithm reveal that there is no systematic classification is available. But in a broader way optimization algorithm can be classified as

- Some problems have constraints but some problems do not have constraints.
- Variable can be one or more than one
- Variable can be continuous or discrete.
- Some problems are static while some are dynamic.
- System can be of either deterministic or stochastic.
- Mathematical equation of optimization problem can be linear or nonlinear.
- Design variable can be of different types[7].

5. Deterministic and stochastic optimization methods are two types of optimization algorithms. A deterministic algorithm is one that works in a physically definite way without any randomness. If we start with the same starting point, such an algorithm will arrive at the same ultimate answer. Deterministic algorithms include hill-climbing and downhill simplex. On the other hand, if the method has any

randomness, the algorithm will usually arrive at a different location each time it is run, even if the same starting point is utilized, Such Stochastic algorithms include things like genetic algorithms and PSO.

If a function's gradient is the emphasis, optimization techniques can be divided into derivativebased and gradient-free algorithms. Hill-climbing algorithms, for example, utilize derivative information and are frequently quite efficient. Derivative-free algorithms rely solely on the function's values rather than derivative information. Because some functions have discontinuities or it is costly to calculate derivatives precisely, derivative-free techniques like Nelder-Mead downhill simplex come in handy.

Optimization algorithms can be classed as trajectorybased or population-based from a separate standpoint. A trajectory-based algorithm usually works with a single agent or solution at a time, tracing out a path as the iterations progress. Hill climbing uses a piecewise zigzag pattern to connect the starting and ending points. Simulated annealing, a common metaheuristic algorithm, is another good example. Particle swarm optimization (PSO), for example, is a population-based technique that uses several agents to interact and trace multiple pathways (Kennedy and Eberhardt, 1995).

Here are two types of search algorithms: local and global search algorithms. Local search algorithms usually converge to a local optimum, not necessarily (and frequently not) the global optimum, and they are often deterministic and have no way of escaping local optima. Simple hill climbing is an example of this. Local search algorithms, on the design vector. the other hand, are ineffective for global optimization, and global search algorithms should be utilized instead. In most Here the segments of x are x_i termed as decision circumstances, modern metaheuristic algorithms are suitable for global optimization, however they are not always successful or efficient.

In general form an optimization problem is defined as :

methods to the stated problem are the three key challenges in simulation-driven optimization and modelling.

variables or design variables. These design variables can be real continuous, discrete or combination of two. The function $f_i(x)$ is termed as objective (or fitness, or cost, or energy) function. In equation 1.1 if $M=1$ then it is termed as single objective function, and if M is greater than 1 then it is termed as multi objective function. In equation 1.2 $\phi_j(x)$ is termed as equality constraint whereas in equation 1.3 $\varphi_k(x)$ is termed as inequality constraint. The space spanned by the values of objective function is termed as solution (or response) space and the space taken by the design variables is termed as search (or design) space [7].

$$\text{Minimize } f_i(x) \quad (i = 1, 2, \dots, M), \quad (1.1)$$

$$\text{subject to } \phi_j(x) = 0, \quad (j = 1, 2, \dots, J), \quad (1.2)$$

$$\varphi_k(x) \leq 0, \quad (k = 1, 2, \dots, K), \quad (1.3)$$

Where $f_i(x)$, $\phi_j(x)$ and $\varphi_k(x)$ are termed as function of the design vector.

$$x = (x_1, x_2, \dots, x_n)^T \quad (1.4)$$

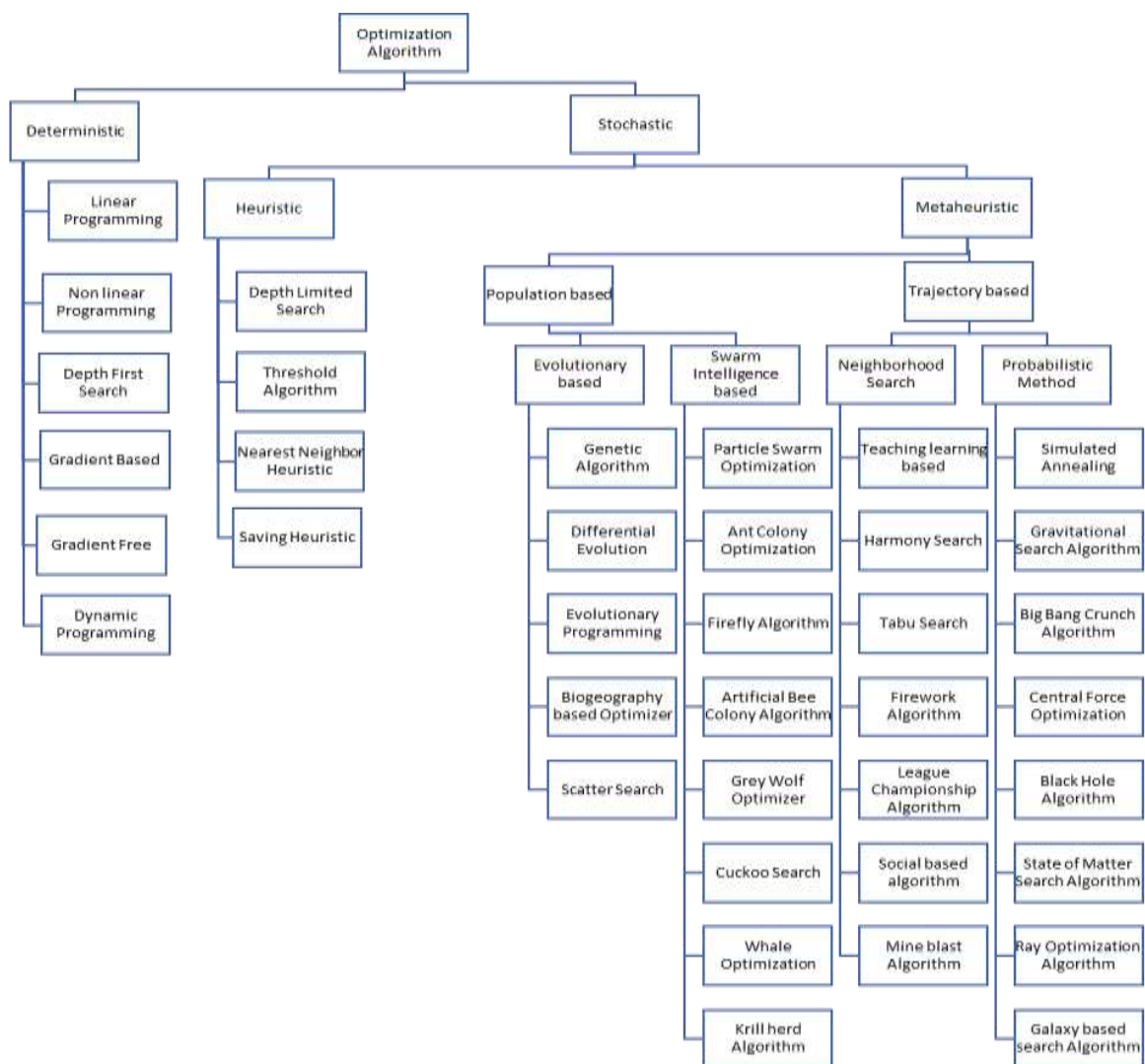


Figure 1.2 Taxonomy of optimization algorithm [8].

B. Existing Challenges in Optimization

The effectiveness of an algorithm, the effectiveness and precision of a statistical simulator, and assigning the precision of a statistical simulator, and assigning the correct methods to the stated problem are the three key challenges in simulation-driven optimization and modelling. Despite their importance, there are no adequate rules or norms in place. We certainly strive to employ the most appropriate methods feasible, but the actual efficiency of an approach depends on a variety of factors, including the method's internal workings, the information required (such as fitness functions and derivatives), and implementation concerns

1) Algorithm's Effectiveness

It's critical to have a good optimizer in order to get the best results. An optimizer is essentially an optimization technique that has been appropriately built to perform the required search. It may be connected and merged with other modelling elements. According to the No Free Lunch Theorem[9], there are several optimization methods in the literature, and no one solution is suited for all issues.

1. Algorithm's Correctness

The selection of the appropriate optimizer or method for a particular issue is critical from an optimization standpoint. The kind of issue, the structure of the methodology, the required quality of outcomes, the modern computing resource, timeframe, availability of the method implementation, and the selection' experience will all influence the algorithm selected for an optimization job [10][11].

2) Effectiveness of statistical Solver

The most computationally intensive element of solving an optimal control problem is usually evaluating the design objective to see if a preferred approach is viable and/or optimum. Typically, we must perform these evaluations hundreds, thousands, or even millions of times. As a result, any method for decreasing computing time, whether by limiting the number of assessments or enhancing the simulator's effectiveness, saves both time and money. The major approach to decrease the amount of objective assessments is to utilise an effective algorithm, so that only a minimal number of such evaluations are required [12].

C. Metaheuristic optimization

Meta- stands for "beyond" or "higher level" in metaheuristic algorithms. They outperform simple heuristics in most cases. Local search and global exploration are used in all metaheuristic algorithms in some way. Randomization is frequently used to achieve variety of solutions. Despite the prevalence of metaheuristics, the literature lacks an agreed-upon definition of heuristics and metaheuristics. The terms 'heuristics' and 'metaheuristics' are sometimes used interchangeably by scholars. However, a recent trend has been to label any stochastic algorithms that include randomization and global exploration as metaheuristics. Randomization is an effective strategy to move away from local search and toward global search. As a result, nearly all metaheuristic methods may be used for nonlinear modelling and global optimization. Metaheuristics can be an effective technique to provide acceptable solutions to a complicated problem through trial and error in a reasonable amount of time. Because of the complexity of the problem of interest, it is difficult to search for every possible alternative or combination; instead, the goal is to identify a good practicable solution in a reasonable amount of time. There's no guarantee that the best solutions will be found, and we don't even know if an algorithm will work or why it will work if it does[13]. The goal is to create an efficient and practical algorithm that works the majority of the time and produces high-quality results. It is reasonable to predict that some of the found quality solutions will be approximately ideal, however this cannot be guaranteed [14].

1) Characteristics of Metaheuristic Algorithms

The fundamental technique to problem-solving has always been heuristic or metaheuristic – via iterations – throughout history, especially at the beginnings of human history. Heuristics were used to make many scientific breakthroughs by thinking outside the box, and frequently by coincidence. The Eureka moment of Archimedes was a heuristic victory. In reality, our daily learning experiences (at least as children) are

mostly heuristic [15]. According to [16] “*Metaheuristic computing is an adaptive and/or autonomous methodology for computing that applies general heuristic rules, algorithms, and processes in solving a category of computational problems.*”

Metaheuristic algorithms' high performance is often due to their ability to mimic nature's best qualities. Metaheuristic algorithms have two main characteristics: intensification and diversification. The intensification phase, also known as exploitation, searches for and identifies the best candidates or solutions based on the present best approaches. The diversification phase, also known as exploration, guarantees that the algorithm efficiently traverses the search space. A tight balance between these two components has a significant impact on an algorithm's overall efficiency. If the exploration is insufficient and the exploitation is excessive, the system may become stuck in a local optimum. Finding the global optimum would be extremely difficult, if not impossible, in this instance. On the other hand, if there is too much exploration but not enough exploitation, the system may fail to converge. The overall search performance slows down in this instance. Balancing these two components is a huge optimization challenge in and of itself [17][15].

A good technique or criterion for selecting the best solutions should be explored during the search. A typical measure is "survival of the fittest." It is predicated on continuing to update the current best solution discovered so far though. Furthermore, a certain amount of elitism should be applied. This is to ensure that the finest or fittest solutions do not become extinct and are passed down to future generations.

Each algorithm and its variants employ various methods to achieve a balance of exploration and exploitation. Certain randomization in combination with a deterministic technique might be viewed as a cost-effective means of achieving exploration or diversification. This ensures that the freshly generated solutions are distributed as widely as possible within the search space available. From the standpoint of implementation, the method used to implement the algorithm has an impact on performance. As a result, any algorithm's implementation must be validated and tested [18].

2) No free lunch theorem

There are the so-called "No free lunch theorems," which can have considerable impacts in the optimization field (Wolpert and Macready 1997). According to this, “If algorithm A outperforms algorithm B for particular optimization functions, then B will outperform A for all other functions”. In other words, if both algorithms A and B are averaged over all potential function space, they will perform equally well. That is to say, there are no algorithms that are uniformly superior. Another point of view is that for a particular optimization issue, there is no need to average over all feasible functions. The most important goal in this situation is to locate the optimal solutions, which has nothing to do with the average over all potential function space. Other researchers argue that there is no universal tool and that some algorithms outperform others for specific sorts of optimization problems based on their expertise. As a result, the main goal would be to either choose the best algorithm for a given problem or to develop bigger algorithms for the majority of problems, not necessarily all of them [19].

According to [21], Metaheuristic algorithms share the following traits:

- The algorithms are based on natural events or behaviors, and they follow specific rules (e.g., biological evolution, physics, social behavior).
- Probability distributions and random processes are used in the selection phase, which contains random elements.
- They provide a number of control parameters to modify the search method, since they are intended to be general-purpose solvers they don't depend on a priori knowledge, which is information about the process that is accessible before the optimization run begins. Nonetheless, such knowledge may be beneficial to them (e.g., to set up control parameters).

3) Existing issues with Metaheuristic Optimization

Finding the optimal answer to a problem is the optimization process. As a result, the primary challenge for metaheuristics is figuring out how to cope with this issue. Despite the fact that many metaheuristics have been suggested, only a handful metaheuristics have consistently attained the required success rate. Population-based metaheuristics, in particular, are frequently employed because they can adapt to large-scale optimization issues. Metaheuristics, as previously stated, are problem-specific algorithms. As a result, the issue is, "What is the optimal algorithm parameter specification based on the kind and size of the problem search space?" Furthermore, selecting the proper metaheuristic algorithm is a complex thing. Recent developments seek to liberalize metaheuristic methods in order to overcome these problems.

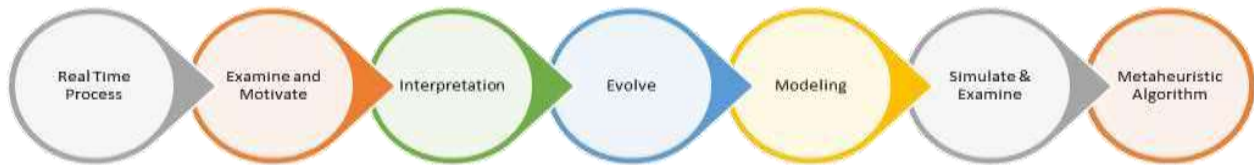


Figure 1.3 Development Procedure of Metaheuristic Algorithms [20]

REFERENCES

- [1] T. Kumar, A.; Jain, “Sub-optimal Control Design for Second Order Non-linear Systems using Krotov Su_ cient Conditions.,” *IFAC Pap.*, vol. 53, pp. 272–276, 2020.
- [2] P. benefits of hybrid control design for linear and Nonlinear and Systems, “Mcclamroch, N.H.; Kolmanovsky, I.,” *Proc. IEEE*, vol. 88, pp. 1083–1096, 2000.
- [3] R. Recio-Colmenares, K. J. Gurubel-Tun, and V. Zúñiga-Grajeda, “Optimal neural tracking control with metaheuristic parameter identification for uncertain nonlinear systems with disturbances,” *Appl. Sci.*, vol. 10, no. 20, pp. 1–18, 2020, doi: 10.3390/app10207073.
- [4] K. Hussain, M. N. Mohd Salleh, S. Cheng, and Y. Shi, “Metaheuristic research: a comprehensive survey,” *Artif. Intell. Rev.*, vol. 52, no. 4, pp. 2191–2233, 2019, doi: 10.1007/s10462017-9605-z.
- [5] P. G. Lelièvre, R. Bijani, and C. G. Farquharson, “Optimization methods,” *78th EAGE Conf. Exhib. 2016 - Work. Program.*, pp. 1– 62, 2016, doi: 10.7551/mitpress/2687.003.0008.
- [6] C. Outline, “Introduction to Mathematical Optimization.” X. S. Yang, *Engineering Optimization: An Introduction with Metaheuristic Applications*. 2010.
- [7] S. Desale, A. Rasool, S. Andhale, and P. Rane, “Heuristic and Meta-Heuristic Algorithms and Their Relevance to the Real World: A Survey,” *Int. J. Comput. Eng. Res. Trends*, vol. 351, no. 5, pp. 2349–7084, 2015, [Online]. Available: <http://www.ijcert.org>.
- [8] W. G. Wolpert, D. H. and Macready, “No free lunch theorems for optimization,” *IEEE Trans. Evol. Comput.*, vol. 1, pp. 67–82, 1997.

- [9] X. S. Koziel, S. and Yang, *Computational Optimization, Methods and Algorithms*, Springer, Germany. 2011.
- [10] X. S. Yang, *Engineering Optimization: An Introduction with Metaheuristic Applications*, John Wiley & Sons. .
- [11] X. S. Yang, *Nature-Inspired Metaheuristic Algorithms, First Edition*, Luniver Press, UK. 2008.
- [12] X. S. Yang, *Engineering Optimization: An Introduction with Metaheuristic Applications*, John Wiley and Sons, Hoboken, NJ, USA. 2010.
- [13] A. H. Gandomi, X. S. Yang, S. Talatahari, and A. H. Alavi, *Metaheuristic Algorithms in Modeling and Optimization*, no. December. 2013.
- [14] A. H. Gandomi, X. S. Yang, S. Talatahari, and A. H. Alavi, *Metaheuristic Algorithms in Modeling and Optimization*, no. May. 2013.
- [15] Y. Wang, “A Sociopsychological Perspective on Collective Intelligence in Metaheuristic Computing,” *Int. J. Appl. Metaheuristic Comput.*, vol. 1, no. 1, pp. 110–128, 2010, doi: 10.4018/jamc.2010102606.
- [16] X. S. Yang, “Metaheuristic optimization: algorithm analysis and open problems,” 2011.
- [17] E. G. Talbi, *Metaheuristics: From Design to Implementation*, Wiley, Hoboken, NJ, USA. 2009.
- [18] W. G. Wolpert, D.H., Macready, “No Free Lunch Theorems for Optimization,” *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67– 82, 1997.
- [19] J. Rajpurohit, T. K. Sharma, A. Abraham, and Vaishali, “Glossary of metaheuristic algorithms,” *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.*, vol. 9, pp. 181–205, 2017.
- [20] P. Boussaïd, I., Lepagnot, J., & Siarry, “A survey on optimization metaheuristics,” *Inf. Sci. (Ny)*, vol. 237, no. 82–117, 2013.
- [21] M. A. Ajwad SA, Iqbal J, Ullah MI, “A systematic review of current and emergent manipulator control approaches,” *Front Mech Eng China*, vol. 10, pp. 198–210, 2015.
- [22] L. W. Slotine J-JE, *Applied nonlinear control: Prentice-Hall Englewood Cliffs, NJ*. 1991.
- [23] M. F. Baños A, Lamnabhi-Lagarigue F, *Advances in the control of nonlinear systems: Springer Science & Business Media*. 2001.
- [24] A. C. Menini L, Zaccarian L, *Current trends in nonlinear systems and control - In Honor of Petar Kokotovic and Turi Nicosia: Springer*. 2006.

- [25] T. O. Ting, X. S. Yang, S. Cheng, and K. Huang, “Hybrid metaheuristic algorithms: Past, present, and future,” *Stud. Comput. Intell.*, vol. 585, no. December, pp. 71–83, 2015, doi: 10.1007/9783-319-13826-8_4.
- [26] R. Burke, E.K., Gendreau, M., Hyde, M., Kendall, G., Ochoa, G., Özcan, E., Qu, “Hyperheuristics: a survey of the state of the art.,” *J. Oper. Res. Soc.*, vol. 64, pp. 1695–1724, 2013.

DRD-CNN: Diabetic Retinopathy detection using CNN

Aman Singh
Department of Computer
Science & Engineering
Engineering
Bhagwan Parshuram Institute
of Technology
Delhi, India
amansingh.95409540@gmail.com

Ashish Negi
Department of Computer-
Science & Engineering
Bhagwan Parshuram Institute
of Technology
Delhi, India
aashishnegi0021@gmail.com

Raman Chola
Department of Computer
Science &
Bhagwan Parshuram
Institute of Technology
Delhi, India
ramanchola3@gmail.com

Keshav Agarwal
Department of Computer Science & Engineering
Bhagwan Parshuram Institute of Technology
Delhi, India
akeshav251@gmail.com

Neha Sharma
Department of Computer Science & Engineering
Bhagwan Parshuram Institute of Technology
Delhi, India
neha.sh.2689@gmail.com

Abstract- Diabetes, often known as diabetes mellitus, is a metabolic disorder in which the body generates insufficient amounts of insulin, resulting in elevated blood sugar levels. Up to 80% of persons with diabetes who have had it for 10 years or more will develop diabetic retinopathy (DR), an eye disease brought on by the disease. In this study, we use U-Net segmentation with region merging and Convolutional Neural Network (CNN) to automatically diagnose various stages of diabetic retinopathy. then group high-resolution retinal images according to the severity of the disease into 5 phases. The CNN model is trained using training datasets, and CNN will provide the likelihood that a diabetic has infected the eye. In order to effectively determine the severity of diabetic retinopathy disease, the initial goal of the model is to train it by providing the training datasets.

The EyePacs Dataset provided the testing dataset, which includes over 35,000 photos with an average of 6 million pixels per image and retinopathy scales. Images from patients representing a wide range of ages, ethnicities, and lighting conditions were included in this dataset. The Proposed technique is efficient than the existing techniques.

Keywords:Diabetes, retinopathy, Convolutional Neural Network (CNN), Image Classification

1. Introduction

Diabetic Retinopathy is an eye disease which can affect the retina and can further cause permanent vision loss. Detection of diabetic retinopathy in early stage is very important to prevent blindness. Many physical tests like visual sharpness/excellent ability test, pupil (expanding/enlarging), optical clearness tomography can be used to detect diabetic retinopathy but are time consuming and affects patients also. In India itself, more than 62 million people are suffering from diabetes.

[1] According to the International Diabetes Federation, the number of adults who are suffering from diabetes in the whole world is estimated to be 366 million in 2011 and by 2030 this would have risen to 552 million. The number of people with type 2 diabetes is increasing in every country. Most of people with diabetes live in low-and middle-income countries and don't treat diabetes seriously. India stands first with 195% (18 million in 1995 to 54 million in 2025). Previously, diabetes mellitus was considered to be

present, largely, among the urban population in india. Recent studies clearly show an increasing pervasiveness in rural areas as well. Indian studies shows that a 3-fold increase in the presence of diabetes among the rural population over the last decade or so (2.2% in 1989 to 6.3% in 2003) [2] In India, Study shows the estimated prevalence of type 2 diabetes mellitus and diabetic retinopathy in a rural population of south india are nearly 1 of 10 individuals in rural south india, above the age of 40 years. At least 90% of new cases could be reduced with proper medication as well as frequent monitoring of the eyes. It primarily affects the retinas of both the eyes, which can lead to vision loss if it is not treated. Poorly controlled blood sugars, high blood pressure, and high cholesterol increase the risk of developing Diabetic retinopathy. Due to development of Computer vision in recent times & availability of large dataset, it is now possible to use a deep Neural network for the detection & classification of Diabetic retinopathy. This paper is structured as follows. The Literature review is presented in Sec.2. Section 3 consists of Problem Formulated and Approached Followed in Sec.4. Section 5 presents the implementation. Section 6 shows the results & analysis. Section 7

contains conclusion and the references are in the Section 8.

II- Literature Review

In "Automated detection of diabetic retinopathy using SVM" research paper, Enrique V. Carrera, Andres Gonzalez, Ricardo Carrera automatically classified the non-proliferative Diabetic retinopathy grade of any fundus image by extracting blood vessels, micro-aneurysms and hard exudates. They obtained a maximum sensitivity of 94.6%, a predictive capacity value of 93.8%. Shailesh Kumar, Basant Kumar proposed a diabetic retinopathy detection scheme by extracting accurate area and ate number of microaneurysm from color fundus images. For detection of microaneurysms, principal component analysis (PCA), contrast limited adaptive histogram equalization (CLAHE), morphological process, averaging filtering have been used. Classification of DR has been done by linear Support vector machine (SVM). The sensitivity and speci_city of the Diabetic retinopathy detection system are observed as 96% and 92% respectively. In another research paper, hard exudates in retinal fundus images are employed to classify the moderate and severe non-proliferative diabetic retinopathy. The hard exudates are segmented using mathematical morphology and the extracted features are classified by using soft margin SVM. The classification model achieves accuracy of 90.54% for 75 training data and 74 testing data of retinal images. Mahendran Gandhi and R. Dhanasekaran's paper focuses on automatic detection of diabetic retinopathy through detecting exudates in color fundus retinal images and also classifies the rigorousness of the lesions. Decision making of the severity level of the disease was performed by SVM classifier.

Authors Name	Publication year	Solution Provided
Shailesh Kumar, Basant Kumar	2018	Extracting Area and Number of Microaneurysm
Enrique V. Carrera [3]	2017	SVM(Support Vector Machine)
Karan Bhatia [2]	2016	ensemble machine learning algorithms
Valliappan Raman [4]	2016	CAD (Computer Aided Detection)
Ömer Deperlioğlu	2018	image processing and deep learning algorithms
Toan Bui	2017	automated segmentation algorithm
Yuchen Wu	2019	transfer learning based approach

III-Problem Formulated and Approached Followed

When I (Aman Singh) was playing cricket with my buddies and an unexpected ball struck my right eye, I immediately went home and began warming my eyes and using natural remedies to reduce the redness. The red eye went away, but after 8–9 months, I suddenly realized that I

could not see clearly with either of my eyes. I initially believed that the sole change might have been an increase in the strength of my spectacles, but this was not the case. When we visited the doctor, they suggested that I get checked out at the AIIMS hospital since I may have diabetic retinopathy or cataracts in the early stages.

We learned at AIIMS that the ball that impacted my eye didn't actually touch my retina, detaching it instead. My retina was operated, but I still can't see clearly out of one eye. This gave me the insight that while we cannot identify someone's cataract or detached retina, we can, however, detect diabetic retinopathy. As a result, my team and I conducted some study and created a straightforward online platform that allows us to accomplish that using convolutional neural network techniques and deep learning methods in ML

IV- Proposed Approach

The detection of diabetic retinopathy through color fundus images requires experienced healthcare professional to identify the presence and significance of many small features which along with a complex grading system, makes this a difficult and time consuming task. Therefore in this paper we propose a CNN approach to diagnosing DR from digital fundus images. The main aim of this project is to detection of diabetic retinopathy disease. Convolutional Neural Network model can be trained by using training datasets and this Convolutional Neural Network will give the probability of the eye infected with diabetics. In order to effectively determine the severity of diabetic retinopathy disease, the initial goal of the model is to train it by providing the training datasets.

Following steps are needed to achieve maximum accuracy from the images dataset are:

- Initialization of Networks
- Data Augmentation
- Pre-processing
- Training
- Activation function selections
- Regularizations

Data Augmentation Data augmentation is the addition of new data artificially derived from existing training data. Data augmentation can be used for random rotation 0-90 degrees, random yes or no

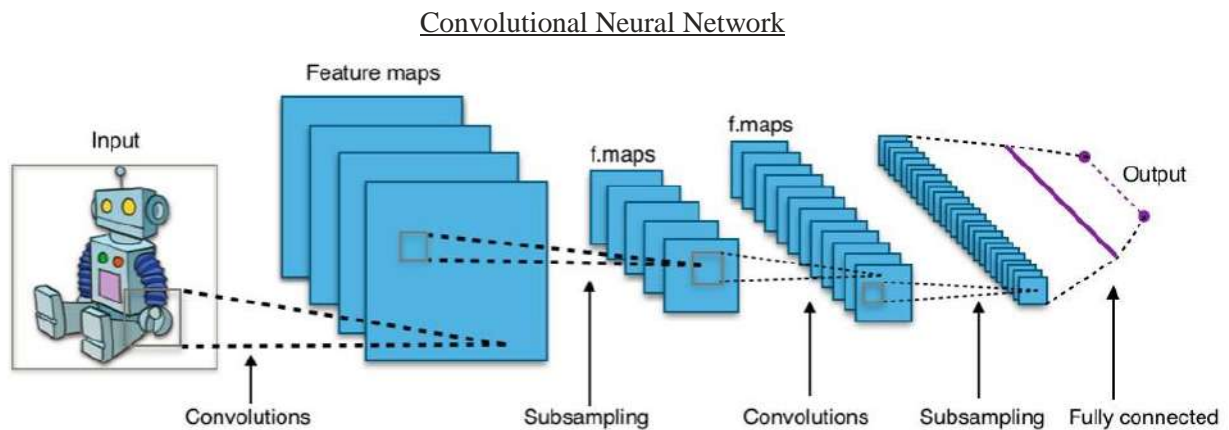


Fig. 0: Convolutional Neural Networks

horizontal and vertical flips and random horizontal and vertical shifts.

Pre-processing

Since the Dataset Contains fundus images of different ages, ethnicities, and various levels of lighting conditions were included in this dataset which can affect the pixel intensity of Fundus image. To restrain the color normalization was implemented on the images using the OpenCV package.

Training

The Convolutional Neural Network has reached a significant level with the help of more than 10,000 images. Several stages of the network. For every batch loaded for back-propagation in the class-weights were updated with a ratio respective to how many images in the training batch were distinguished as having no signs of Diabetic retinopathy.

The network was trained using stochastic gradient descent with Nesterov momentum. A low learning rate of 0.0001 was used for 5 times to stabilize the weights. This was then dataset the accuracy of the network had increased to over 70%. The learning rate was then lowered by a factor of 10 every time training loss and accuracy saturated.

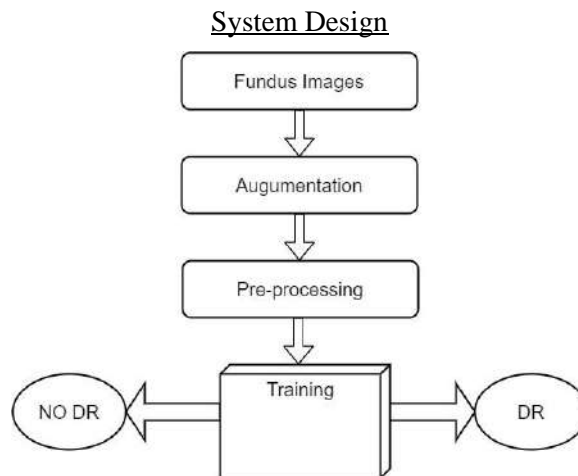


Fig.1: Detection method

Convolutional neural networks (CNNs) are a type of artificial neural network that is often used in computer vision tasks, such as image and video recognition. CNNs are well-suited to tasks where the input data is highly variable, such as recognizing objects in images.

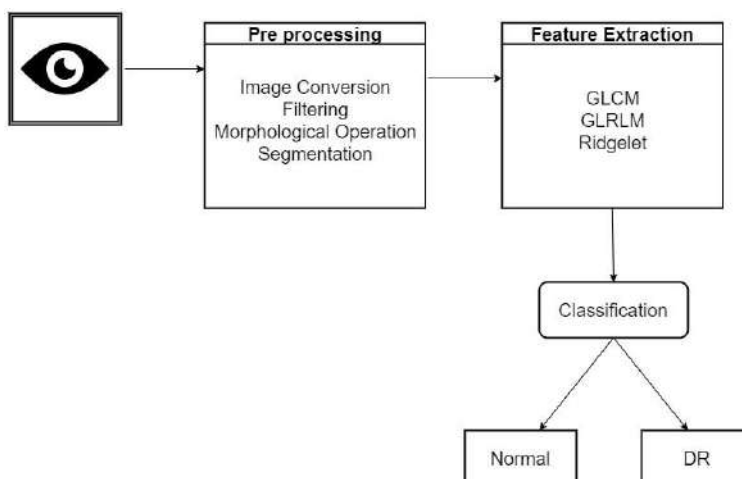
Some applications for CNNs include: Image and video recognition: CNNs can be used to classify objects that appear in the input data. For example, a CNN might be trained to recognize different types of animals in an image, such as cats, dogs, and birds.

Object detection: CNNs can be used to locate and classify objects in an image or video. This can be used, for example, to detect pedestrians in a video feed from a self-driving car.

Natural language processing: CNNs can be used to process text data and extract features that are relevant to the task at hand, such as sentiment analysis or topic modeling. There are many other potential applications for CNNs, such as medical image analysis, facial recognition, and speech recognition. The flexibility and power of CNNs make them a valuable tool for many different tasks.

The potential benefit of using our trained CNN is that it can classify thousands of images every minute allowing it to be used in real-time whenever a new image is acquired. In practice images are sent to clinicians for grading and not accurately graded when the patient is in for screening. The trained CNN makes a quick diagnosis and instant response to a patient possible. The network also achieved these results with only one image per eye. The network has no issue learning to detect an image of a healthy eye. This is likely due to the large number of healthy eyes within the dataset.

Convolutional Neural Networks steps:



Convolution Operation: This is the main part of the Convolution Neural Network. In this step we

basically act as the neural network's filters. we also decide how the patterns are detected and how the findings are mapped out.

Pooling: In this part we are using special type of pooling called max pooling and which were help in cover various approaches, though, including mean (or sum) pooling. This part will end with a demonstration made using a visual interactive tool that will definitely sort the whole concept out for you.

Flattening: This will be a brief breakdown of the flattening process and how we move from pooled to flattened layers when working with Convolutional Neural Networks.

Full connection: All the three things that we covered above will be merged in this part. you'll get to visualize a full picture of how Convolutional Neural Networks operate and how the "neurons" that are finally produced learn the classification of images. GLCM is a **second-order statistical texture analysis method**. It examines the spatial relationship among pixels and defines how frequently a combination of pixels are present in an image in a given direction and distance.

[13] The **gray level run length matrix (GLRLM)** whose entries are statistics recording distribution and relationship of images pixels is a widely used method for extracting statistical features for medical images.

[14] Ridgelet transform **can be used where images contain edges and straight lines**. Curvelet transform has been introduced to solve this problem

V- IMPLEMENTATION

1. INPUT IMAGES AND TRAINING

DATASET: First we train the machine using a training dataset then we give any retinal image as input. Fig 3 shows the sample images for testing.

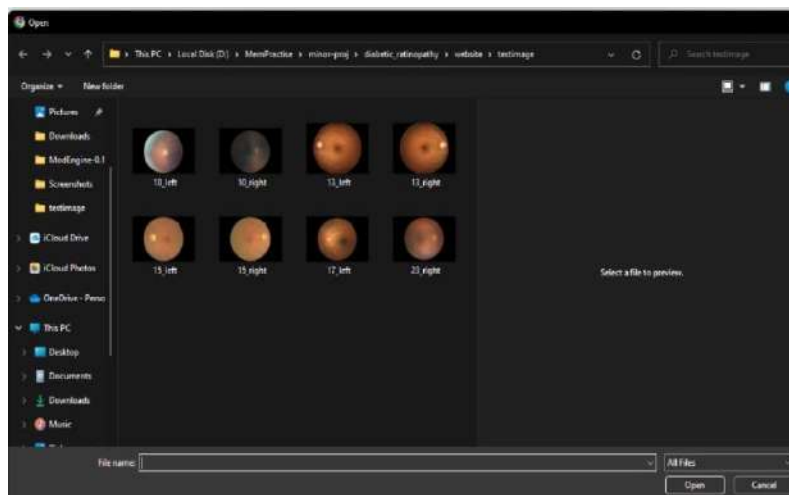


Fig 3: Sample Fundus images



Fig 4: Uploading Fundus image

2. PRE-PROCESSING

Filtering the input retinal image to find whether the given input retina image has any defects based on

features like colors and also noise reduction process(unwanted things) as shown in Fig 5.

3. SEGMENTATION

Segmentation is done based on Binarization, Median filtering, Morphological hole filling.

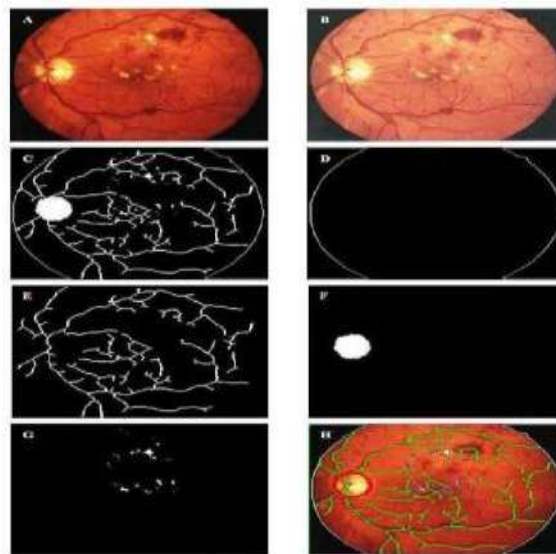


Fig 5: Image Preprocessing

VI- RESULT AND ANALYSIS

Image 1 Result	
Categoires Wise Prediciton	
Category	Chances(%)
NO_DR	98.91000390052795
Mild	0.20156078971922398
Moderate	0.8722278289496899
Severe	0.002444559140712954
Proliferate_DR	0.013760635920334607

Fig 6: Left eye result

Image 2 Result	
Categoires Wise Prediciton	
Category	Chances(%)
NO_DR	92.71862506866455
Mild	0.9499670937657356
Moderate	6.188832595944405
Severe	0.041095344931818545
Proliferate_DR	0.10148793226107955

Fig 7: Right eye result

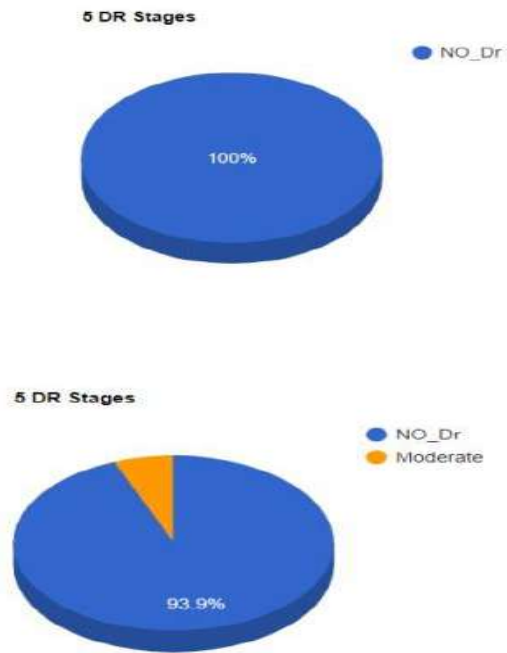


Fig 8: Pie chart for Left and right eye respectively Bar Chat

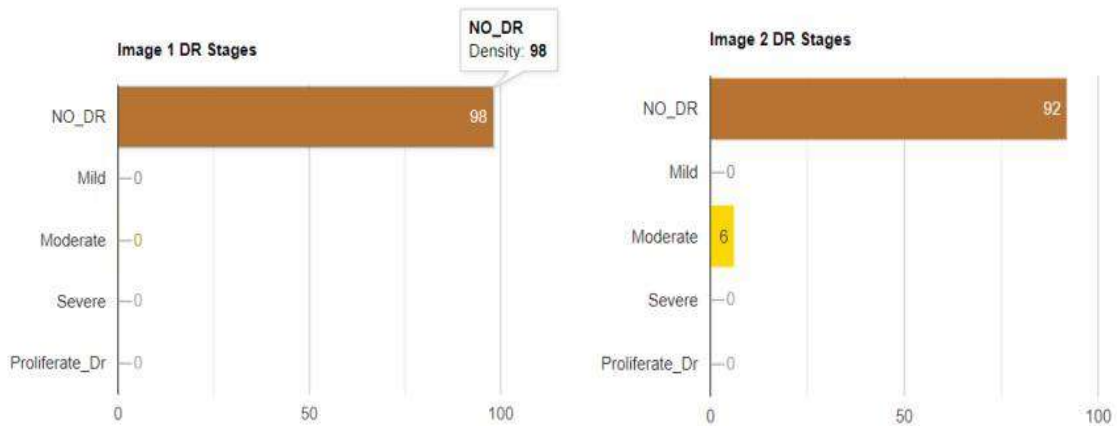


Fig 9: Left and right eye respectively

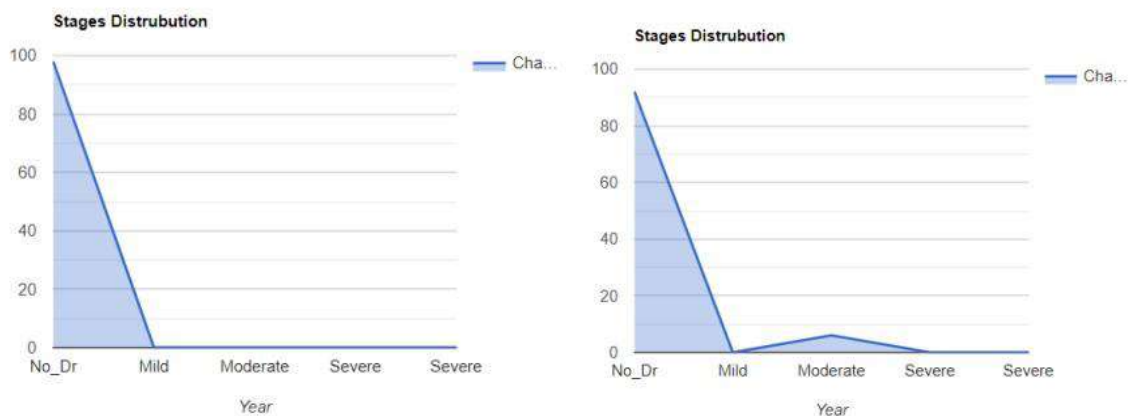


Fig 10 Left and right eye respectively

Fig.6 and Fig.7 show the result of the left and right right respectively. All 5 stages of Diabetic retinopathy are shown in percentage.

Fig.8 shows the pie chart for both eyes. Fig.9 and Fig.10 show the same result but in graphical representation for better understanding and clear view.

VII- Conclusion

This paper proposes that the five-class problem for screening of DR can be approached using a CNN method. Our network has shown promising signs of being able to learn the features required to classify the fundus images, accurately classifying the majority of proliferative cases and cases with no DR. The issues came in making the network to distinguish between the mild, moderate and severe cases of DR. The low sensitivity, mainly from the mild and moderate classes suggests the network struggled to learn deep enough features to

8.REFERENCES

- [1] Rishab Gargeya ,Theodore Leng MD,Automated Identification of Diabetic Retinopathy Using Deep Learning
- [2] Karan Bhatia, Shikhar Arora, Ravi Tomar, Diagnosis of diabetic retinopathy using machine learning classification algorithm.
- [3] Enrique V. Carrera; Andrés González; Ricardo Carrera, Automated detection of diabetic retinopathy using SVM.
- [4] Valliappan Raman; Patrick Then; Putra Sumari, Proposed retinal abnormality detection and classification approach: Computer aided detection for diabetic retinopathy by machine learning approaches.
- [5] R.Priya, P.Aruna, SVM and Neural Network based Diagnosis of Diabetic Retinopathy, International Journal of computer Applications(00975-8887), volume 41-No.1,(March 2012).
- [6] S.Giraddi, J Pujari, S.Seeri, Identifying Abnormalities in the Retinal Images using SVM Classifiers, International Journal of Computer Applications(0975-8887), Volume 111 – No.6,(2015).
- [7] G.Lim, M.L.Lee, W.hsu, Transformed Representations for Convolutional Neural Networks in DiabeticRetinopathy Screening, Modern Artificial Intelligence for Health Analytic Papers from the AAI(2014).
- [8] P Kulkarni, J Zepeda, F Jurie, P Perez, L Chevallier, Hybrid Multi-layer Deep CNN/Aggregator feature for Image Classification, Computer Vision and pattern recognition, ICASSP conference, (2015).
- [9] E M Shahin, T E Taha, W Al-Nuaimy, S.El Raaie, O F Zahran, F E Abd El-Samie, Automated Detection of Diabetic Retinopathy in Blurred Digital Fundus Images, IEEE International Computer Engineering Conference , pages-20-25,(2012).
- [10] Xiang chen et al, A novel method for automatic hard exudates detection in color retinal images,

Proceedings of the 2012 International Conference on Machine Learning and Cybernetics, Xian (2012).

[11] Vesna Zeljkovi et al, Classification Algorithm of Retina Images of Diabetic patients Based on Exudates Detection, 978-1-4673-2362-8/12, IEEE(2012).

[12] kaggle EyePacs Dataset is used for training last updated Dec 2021 [13] Dhanshree Gadkari, Image Quality Analysis Using GLCM

[13] Hari Wibawanto, Discriminating Cystic and Non Cystic Mass Using GLCM and GLRLM-based Texture Features

[14] Hari Wibawanto, Discriminating Cystic and Non Cystic Mass Using GLCM and GLRLM-based Texture Features.

FINANCIAL RECOMMENDER SYSTEM

Charu Gupta, Palak Girdhar

Department of Computer Science and Engineering, Bhagwan Parshuram Institute of Technology, Delhi, India

Abstract-To meet their daily needs and for future use, people invest their money in variety of plans. In today's era, recommender systems are getting lot of attention because it assist people in finding out the product and knowledge about the things they desire for. Recommender system provides ample number of similar options with various features. But no recommender system is available in the literature which recommends the product with the aim to save money. In this paper, we have designed an algorithm which will assist the people in smarter way. The algorithm is based on Collaborative Filtering (CF), which generates high quality and accurate recommendations to the users. In order to generate filtered suggestions for the current user, CF employs a subset of users it refers to as neighborhood users. Additionally, this system generates the outcome list using straightforward heuristics. This enables the user to receive better suggestions without disclosing a lot of personal information. The system's empirical evaluation is based on both the results of the online evaluation and the influence of recommender methods.

INTRODUCTION

The development of recommender systems technology has been fueled by the Web's growing significance as a platform for electronic and commercial transactions. The simplicity with which the Web enables users to provide information is a significant catalyst in this regard. People can provide the feedback about their likes and dislikes. For Example, considering a scenario of a content provider such as Hotstar. People can easily provide the feedback in terms of numerical values, showing the liking and disliking of the content, with just a single click of mouse. And this feedback can be advantageous to someone else.

Recommender system learns from the traces of user interactions and offers or recommends the personalized information to the user [2]. It aims to recommends the particulars items which are likely to be of interest of the user [7]. Amazon, Netflix, YouTube [3] are few examples of popular recommender systems. These platforms collect the data from the users and utilize the collected data for the recommendation. On the basis of gathered data, movies, playlists and videos are recommended to the user having same interest.

Recommender system suggest the items to user on the basis of their preferences. For this specific purpose, it uses various filtering techniques.

The filtering techniques of recommender system are categorized [15] into Collaborative, [12] Content-based [11] and hybrid [13].

Content based filtering suggests items of interest based on their distinguished features. For example, a person who is interested in reading news will be recommended next time with the similar words that are read before.

Collaborative filtering offers or recommends those particulars which are fetched on the basis of previous ratings of the similar interest's users.

Hybrid techniques combines the features of both mentioned techniques to cope up the specific limitation of a single solution.

In the proposed approach, Collaborative Filtering is used. Recommendations for users are made on how similar other users liked the item. Collaborative filtering-based recommender systems uses rating for items provided by various users. It recommends the items to the target user who have not yet considered the suggested item but likely to have interest in the recommendation. The ratings are stored in a form of matrix, where rows specify the user's ratings and columns denotes the item's rating.

This filtering technique is sometimes also referred to as social filtering. Recommendation can be made to the other person having same interest or genre.

This method predicts new interactions on the basis of historical data.

There are two types of approaches to obtain information in collaborative filtering: Item based filtering [9,10] and user-based filtering [14]

Item based filtering technique was developed by Amazon originally. It computes the relationship between the two items that are bought together. For example, if bread and butter appear more often in the shopping bag or user history, it can be used for generating recommendations. Next time, when someone adds bread to his cart, system automatically recommend butter by learning the previous history.

The user-based technique differs from item-based filtering in that it calculates the distance between users based on their ratings or likes rather than on the goods themselves. Facebook's algorithms, for instance, can suggest videos that Angela liked to Sam even though Sam hasn't watched them

before if Sam liked certain videos on Facebook and Angela liked the related videos.

In both the techniques, system has no prior information about the relationship between the items and the users.

The system's sole concern is that either these products appeared in the same cart together or that people who have similar interests like it.

It also sometimes referred as cognitive filtering, content-based filtering is all about assigning tags and attributes to the items, so the algorithm of the system knows something about the content of each item in the database pool. Netflix is the most appropriate demonstration of this approach. Based on its genre, each Netflix movie is assigned a few tags. Brian, for instance, recently watched the movie "Bojack Horseman." It might be identified with the tags "animated," "adult," and "comedy." The next time, Brian will be given recommendations for programmes with similar tags to "Family Guy."

Both of the collaborative and content-based filtering have their shortcomings. Therefore, the hybrid approach combines multiple filtering techniques like collaborative and content-based to get the desirable and much more accurate results. One can achieve the best of both the worlds and get a much more accurate and precise data for the recommendations.

Further, the organization of the paper is follows: In section II, the architecture of the financial recommender system is introduced profiling various recommendation techniques that help in getting the result list of recommendations.

Section III discusses about the geometry of the dataset used and the pre-processing done to refine the data so that it would work best with the algorithm. Section IV analyzes the performance and the quality of the recommendations made by the recommender system (FRecS). Finally, section V presents the conclusion and Section VI elaborates the future aspects.

I. FRecS Architecture and design

The FRecS presents a collaborative filtering recommender system that recommends shops and websites to the user. The overall architecture of the system is shown in the following figure 2. The data set that has been used (Expenditure.csv) is first pre-processed (e.g. disambiguation and data cleaning). At runtime, whenever a new recommendation is requested by a user, a similarity array is calculated between this user and other users in the system using some similarity measures. This similarity array is then sorted in descending order. Afterward, top N users that are closest to the user who has requested the recommendations are marked in a `sim_user_set`.

Every user in the `sim_user_set` are classified with the help of the ratings, all the users who gave

positive feedback (i.e. rating \geq 3) are then marked in the final_rec_list. Now with the available list, the top N similar users to the user that requested the recommendations are obtained. At last, the targeted user is shown all the shops and websites rated higher by the user in final_rec_list as the result.

For computing similarities, four attributes from the dataset were used which are Age, Occupation, Gender, and Location. These attributes are represented as vectors of features. The similarity between users can be calculated by using some distance metric between these vectors. Some commonly used similarity measures are cosine, Pearson, Euclidean [4] etc. The similarity measure used here is cosine.

Cosine similarity is a measure of calculating the similarity between two vectors of an inner product space that are non-zero that measures the cosine of the angle between them [6, 8]. The measure of the angle between two users, where users are represented as vectors gives out their cosine value [5] as shown in Equation 1. Vectors of two items with attributes are compared in cosine similarity function is shown in figure 1.

$$u(c, s) = \cos \cos (\vec{w}_c, \vec{w}_s) = \frac{\vec{w}_c \cdot \vec{w}_s}{\|\vec{w}_c\| \times \|\vec{w}_s\|} \quad [1]$$

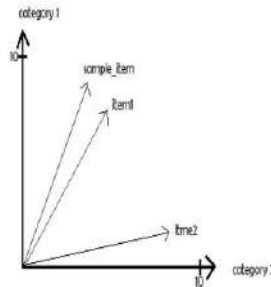


Figure 1: Cosine Similarity

1. Locality Similarity:

The user’s location (u_l) is given in the data provided by the user on the website. For finding the similarity between the current user and the other users in the system ($v_{l_i} \dots \dots N$) We use cosine similarity function given in equation 2.

$$l(u_l, v_{l_i}) = \cos \cos (\vec{u}_l, \vec{v}_{l_i}) = \frac{\vec{u}_l \cdot \vec{v}_{l_i}}{\|\vec{u}_l\| \times \|\vec{v}_{l_i}\|} \quad [2]$$

2. Age Similarity:

The age of the users is divided into groups with a gap of 4 years. The age of the current user (u_a) is similar to the age of other users ($v_{a_i} \dots \dots N$) In the system if they belong to the same age group ($A_i \dots \dots N$).

$$A(u_a, A_i) = \{if\ u_a\ A_i,\ then\ 1\ else\ 0\}$$

1. Sex Similarity:

If the current user’s sex (u_s) is same as the other user’s sex ($v_{s_i} \dots \dots N$) In the system then those users are considered to be similar.

$$S(u_s, v_{s_i}) = \{if\ u_s = v_{s_i},\ then\ 1\ else\ 0\}$$

2. Occupation Similarity:

If the current user’s occupation (u_o) is same as the other user’s occupation ($v_{o_i} \dots \dots N$) Present in the system then those users are considered to be similar.

$$O(u_o, v_{o_i}) = \{if\ u_o = v_{o_i},\ then\ 1\ else\ 0\}$$

3. Prediction Generation:

The prediction score $P(u, v_i)$ of a user from the set of users ($v_i \dots \dots N$). It shows how similar is that user to the current user (u) can be computed on the basis of linear combination of similarity scores of the four attributes mentioned above is given in equation 3:

$$P(u, v_i) = l(u_l, v_{l_i}) * A(u_a, A_i) * S(u_s, v_{s_i}) * O(u_o, v_{o_i}) \quad [3]$$

Following is the detailed algorithm: ALGORITHM

1. Start
2. Get the user data (u) containing Location, Age, Occupation and Sex
3. Finding similarity of the current user (u) with the set of users present in the system ($v_i \dots \dots N$) on the basis of four attributes which are locality, age, occupation, sex.
4. Multiplying the result of similarity scores of four attributes and storing it in the array for each user v_i present in the system.
5. Sorting the array in descending order.
6. Picking the first N users in the array. These are the first N similar users to the current user (u).
7. Classifying the top N similar user on the basis of rating (i.e. , rating > 3) given by other users as a recommendation set
8. Stop

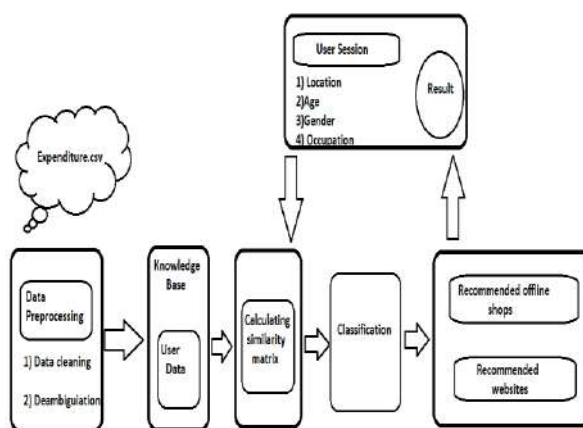


Figure 2: FRecS-FINANCIAL RECOMMENDER SYSTEM

II. Data collection and pre-processing

The data was collected via the google form because no such dataset was available online that could help in this particular case. The data was then cleaned for duplicates and for the sparsity. The clean data was then used to seed the two recommender techniques by allowing the computation of their similarity matrices.

1. Collecting user data

The data that is collected has 38 entries. You can get the data from [18]. In all of the 38 entries we used four entries for our recommender system which are:

The proposed recommender system uses four major attributes out of total 38. These are:

- **Occupation.** It shows the occupation of the user.
- **Age.** It shows the age of the user.
- **Gender.** It shows the gender of a user.
- **Location.** It shows the locality where the user lives.

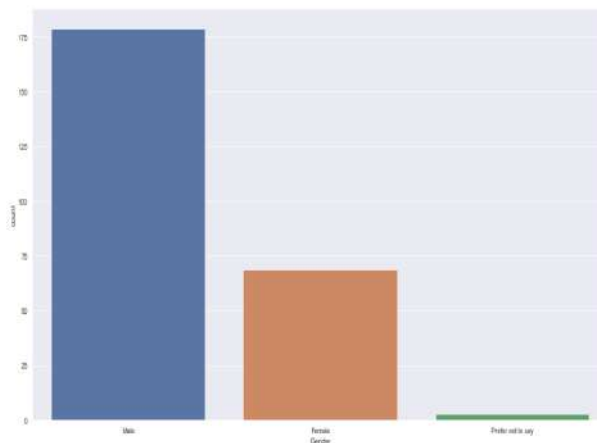
2. Data cleaning

The first step in preprocessing the data is the data cleaning, specifically removing data that is of use to the system (i.e. user having null entries in the attributes Age, Occupation, Gender, and Location). The next step in preprocessing is to apply normal NLP [16] in the entries that are to be shown as a result to another user requesting recommendations.

3. Data Visualization

Some data exploratory data visualization is given below:

- Figure 3 shows the count plot for the gender of the users in the database:



- Figure shows the count plot for the method user generally uses for shopping :

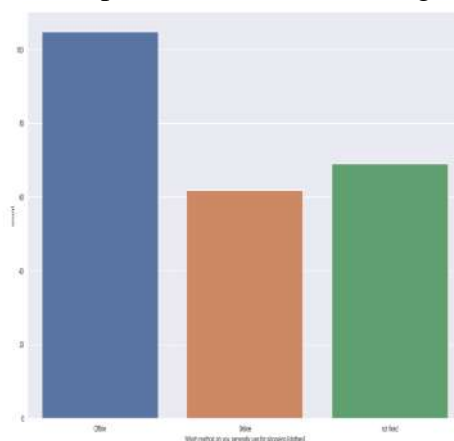


Figure 4: Shopping Trend

- Figure shows the count plot for the method user generally uses for shopping in hue of Gender:

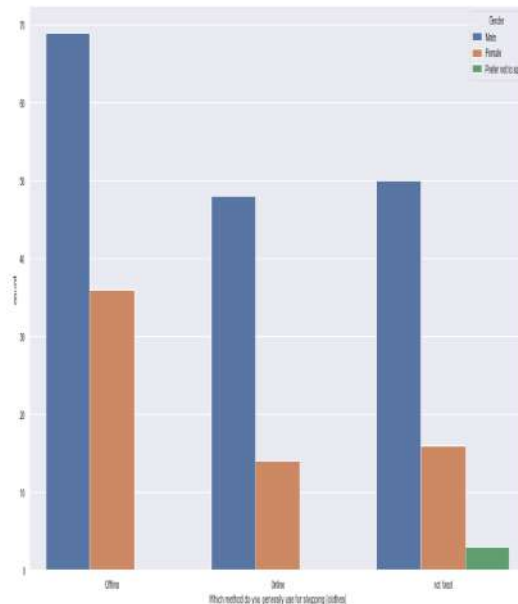


Figure 5: Shopping trend among males and females

III- Evaluation and Analysis

The results obtained by the recommender system will not always be in accordance with the user’s preferences. The results obtained are denoted as A and user’s original preferences are denoted as B with respect to the users. But it was not possible to accurately determine B. So, the principle of statistical estimation was opted which is based on the offline shops and websites that the user gave in the dataset. This way, the degree of utility of a particular recommendation to a user can be identified.

Accuracy metrics:

Accuracy metrics measure the quality of proximity to the truth or the true value achieved by the system. It is calculated from equation (4):

$$Accuracy = \frac{\text{Number of successful recommendations}}{\text{Number of recommendations}}$$

[4]

The accuracy shown by the FRecS is:

$$Accuracy = 68/83 = 0.819$$

$$Accuracy \text{ (in percentage)} = 81.9\%$$

Mean Absolute Error (MAE) metrics:

This metrics measure the average absolute deviation between each original preference B and the result X. the accuracy and MAE together add up to 1.

IV- Conclusion

This paper proposed a collaborative recommender algorithm which is able to recommend users websites and shops near there locality which provide great service and are highly rated by other users. To get these recommendations for a user, similarities are calculated between the current user and a set of users present in the system on the bases of four attributes and then similar users are marked. These marked users are then classified on the bases of ratings given by themselves. The final classified user’s data is shown as a result to the current user who requested for the

recommendations. Our System ensures that the user gets a relevant recommendation on the same time we give them very good privacy which must be very important.

Recommender systems involve an inherent trade-off between accuracy of recommendations and the extent to which users are willing to release information about their preferences. Using online applications users may share or upload their personal information but this information is shared within the specific scope. The privacy of the information means exposure of the information within a bounded scope

V- FUTURE WORK

It is observed that although a large amount of research and development has been done in the area of collaborative recommender systems, still a very small amount of research has been done related to the financial services. However, the recent increase in the number of publications and new findings point to a prosperous future regarding advanced and upgraded financial recommender systems that could revolutionize many applications such as money investment and expenditure analysis.

REFERENCES

- [1] H. Xu, R. Zhang, C. Lin, and W. Gan, “Construction of e-commerce recommendation system based on semantic annotation of ontology and user preference,” *Telkomnika Indonesian Journal of Electrical Engineering*, vol. 12, no. 3, 2013.
- [2] J. Lu, D. Wu, M. Mao, W. Wang and G. Zhang, “Recommender system application developments”, *Decis. Support Syst.*, vol. 74, pp. 12–32, 2015.
- [3] G. Linden, B. Smith, and J. York, " Amazon.com Recommendations: Item-to-Item Collaborative Filtering , *IEEE Internet Computing* , vol. 7, no. 1, pp. 76 80, 2003.
- [4] H. Tan and H. Ye, “A collaborative filtering recommendation algorithm based on item classification,” in *Proceedings of the 2009*.
- [5] Sanay Sharma, Aakriti Aharma, Yamini Sharma and Ms Manjot Bhatia, *Recommender System Using Hybrid Approach*, 2016.
- [6] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, “Item-based collaborative filtering recommendation algorithms,” in *Proceedings of the 10th International Conference on World Wide Web*. Hong Kong, China: ACM Press, May 2001, pp. 285–295.
- [7] Recommender_system page on Wikipedia.[Online]. Available: http://en.wikipedia.org/wiki/Recommender_system
- [8] A scalable collaborative filtering algorithm based on localized preference, vol. 1, 2008.
- [9] Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B. Kantor «Recommender Systems Handbook», 2011, 845p.
- [10] Dietmar Jannach, Markus Zanker, Alexander Felfernig, and Gerhard Friedrich, “Recommender Systems. An Introduction”, 2011
- [11] L. Iaquinta, M. D. Gemmis, P. Lops, G. Semeraro, M. Filannino, and P. Molino, “Introducing serendipity in a content-based recommender system,” in *International Conference on Hybrid Intelligent Systems*, 2008, pp. 168–173.
- [12] A. Hegde and S. K. Shetty, “Collaborative filtering recommender system,” *Esrta Publications*, vol. V4, no. 3, 2015.
- [13] M. Salehi and I. N. Kmalabadi, “A hybrid attribute-based recommender system for e-learning material recommendation,” *Ieri Procedia*, vol. 2, no. 4, pp. 565–570, 2012.
- [14] Company Blog E-Contenta. "Quality Metrics ranking".Access:<https://habrahabr.ru/company/econtenta/blog/303458/>
- [15] Francesco Ricci and Lior Rokach and Bracha Shapira, *Recommender Systems Handbook*, Springer, 2011.
- [16] Thomas Hornung, Cai-Nicolas Ziegler, Simon Franz, *Evaluating Hybrid Music Recommender System*, 2013
- [17] Shruthi Bhat, K.Aishwarya, *Item-Based Hybrid Recommender System For Newly Marketed Pharmaceutical Drugs*

E-Voting System Using Blockchain

Dr. Richa Sharam

Computer Science and Engineering
Dept, Bhagwan Parshuram Institute
Of Technology, Delhi/India
richasharma@bpitindia.com

Dr. Mugdha Sharma

Computer Science and Engineering
Dept, Bhagwan Parshuram Institute
Of Technology, Delhi/India
mugdhasharma@bpitindia.com

Akanksha Dhamija

Computer Science and Engineering
Dept, Bhagwan Parshuram Institute
Of Technology, Delhi/India
akankshadhamija12@gmail.com

Abstract: There's no doubt about the pervasive impact that digitalization has had on the lives of people, globally. However the electoral system still uses paper in its conventional implementation. The paper based centralized system (offline) has its own disadvantages like lack of transparency and a security threat. The general elections are organized by a centralized authority that has complete access of the database collected, which in turn can be tampered for considerable benefits.

The revolutionary concept of Blockchain tends to digitally solve the problem due to its decentralized nature. It embraces the decentralized system and not any single authority holds the database. The adoption of blockchain in the distribution of databases on e-voting systems can reduce one of the cheating sources of database manipulation. This research discusses the recording of voting result using blockchain algorithm from every place of election. The Ethereum platform is a smart contract that makes blockchain more reliable to be used for products of daily services. Smart contracts are meaningful pieces of codes, to be integrated in the blockchain and executed as scheduled in every step of blockchain updates. On the other hand, e-voting is a very volatile and crucial issue, therefore handling it with robust and secure concept of smart contracts turns out to be a viable solution to develop smarter, cheaper, secure, transparent and convenient electoral systems. The major benefit of using Ethereum is its consistency, widespread use and provision of smart contract logics. The system uses proof of work because of the hard to find solution to the problem property and once found, it can be easily verified.

Keywords-: Blockchain; e-voting; Ethereum; security.

I. INTRODUCTION

Blockchain is the dawn of new millennium's technologies that has a wide bandwidth of applications. The blockchain technology owes its success to the widely accepted and very first cryptocurrency, Bitcoin[1]. Studies suggest that the blockchain technology can be used not only for the monetary transactions but also in many other areas due to the high transparency of the system. As sited in Bitcoin, the wallets are in a distributed structure. Therefore, the total amount of coins and transactional volume in the world can be traced momentarily and clearly. This nullifies the need of a central authority or complete

the P2P-based system.

Because of this not only money transactions but structural values can also be evaluated and stored in this distributed chain, and can be maintained securely with the help of some cryptologic methods. The most important example for this is the Ethereum coin (Ether) revealing that this technology can produce structured systems with certain modifications [2] as described above. Smart Contracts [3] enforce software programs that are written into blockchain and are immutable. These programs cannot be manipulated once written. Hence, they continue to work autonomously, transparently and properly without involvement of any external stimuli [4]. The distributed nature of blockchain might address more issues than the digitalization purpose.

E-voting projects are being studied extensively, some still being used and implemented to find a reliable solution. There are many online polls and petitions being actively implemented into the regime of legislation yet we cannot say the same for elections as they define the democratic element of the constitution of nation and also form the basics of the administrative methodologies. Thus, the democratic society demands a more transparent, robust and secure approach to electoral process.

II. MOTIVATION

The solemn purpose of elections is to channel the sovereignty as a representative of the democracy. Each eligible voter, comes to the polling stations with government issued id proofs, mostly Voter ID cards, gets verified by the committee members and chooses a valid and legitimate option.

Many governmental and organizational election are held using sealed paper ballots, poll booths and EVM (Electronic Voting Machine). The notarized accounts are then counted and results are announced publically. The complete process of counting votes in conventional elections can take 3 to 7 working days depending on the speed of sending the sound to a higher level [5]. The reliability of notaries are at disposal of the committees involved. The most frequent problem in elections is the issue of data manipulation, security, and transparency.

The main motive of the project is to build a robust, reliable and secure e-voting system and show that blockchain is a viable solution to this as availability of an e-voting system will make this available for anyone who has a computer, or a mobile phone and people's opinion will be more public. When the opinion of public will be easily accessible by the politicians and managers, the society will be lead to the true direct democracy [6]. The e-voting system also promises to cut down the long term cost of conducting elections especially if the distribution is to be made at 1000s of locations involving millions of voters [7]. Also, it can resolve the problem of absenteeism while voting if the voter is unavailable in the region during the voting days due to any reason. E-voting system adds mobility to the voting regime.

III. RELATED WORK

Blockchain and Benefits of Blockchain

Blockchain is a decentralized technology that works on the principle of transparent Data distribution. It stores data record that grow continuously and are not controlled by any single authority. It is an immutable and transparent ledger [8]. It consists of a sequential series of several blocks which are related as hash from the previous block is used in making the next block. Therefore, attempt to change the information will be next to impossible as it will affect the trail [9].

The widely known Blockchain technology currently exists in the Bitcoin system which is the public ledger of all transactions. Bitcoin is a decentralized, peer-to-peer digital payments system based on the first public key cryptography proposed by Satoshi Nakamoto in 2008 [4]. Bitcoin uses a consensus protocol called PoW (Proof of Work) based on cryptocurrency to ensure only legitimate transactions are allowed within the system. Where each transaction is calculated its hash value and entered into a database called Blockchain as described in fig.1. To connect between one block with another block, the hash value of the previous block inserted into the next block then calculated its hash value. The hash value must meet certain requirements called difficulty in order to be considered a legitimate block. Searching for hash values that match those requirements is called Proof Of Work. Bitcoin stores all transaction information in a database called blockchain in the internet network. Blockchain consists of several blocks associated with each other and in sequence as shown in fig.1 The blocks are related because the hash values of the previous block are used in the next block creation process. The effort to change the information will be more difficult because it must change the next blocks.

The first block is called the genesis block.

In creating new blocks, miner required in the mining process using hash computing equipment. Miner compete against each other to create a new legitimate block in accordance with the specified difficulty. A new block is generally generated by a miner but there are times when more than one new block is generated by multiple miners that both meet the criteria even though the odds are small, making blockchain a fork. If this case occurs, then the voting process conducted by the miners

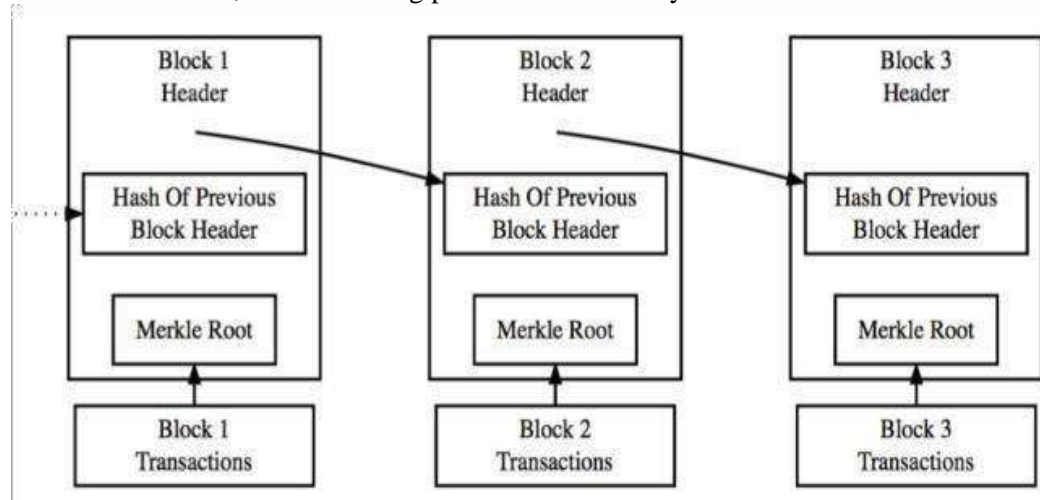


Figure 1. Blockchain Illustration Source :-www.blockchain.org

The voting process is done by way of the miner choosing one of several new blocks and then producing the discovery of a longer chain branch. Then the entire Bitcoin system uses the longest branch and deletes all other branches. Unused blocks are called block orphans and become invalid, also all transactions that have been recorded in the block orphan will be inserted into the new block. Blockchain comes with a variety of different types, but has several common elements :

- Blockchain is distributed digitally to a number of computers in almost real time.
- Blockchain is decentralized, the entire recording is available for all users and peer to peer network users. This eliminates the need for central authorities, such as banks, as well as trusted intermediaries.
- Blockchain uses many participants in the network to reach consensus.

Participants use their computers to authenticate and verify every new block. For example, to ensure that transactions not occur more than once, new blocks are only adopted by the network after the majority of its members agree that they are valid.

- Blockchain uses cryptography and digital signatures to prove identity.

Transactions can be traced back to the cryptographic identity, which is theoretically anonymous, but can be re-linked with real-life identity using reverse engineering techniques. • Blockchain has a difficult (but possibly) mechanism for altering stored records.

Although all data can be read and new data can be written, previously existing data on blockcahin can't be changed theoretically unless the rules embedded in the protocol allow such changes by requiring more than 50 percent of the network to approve the change.

- A Blockchain is time-stamped.

Transactions in blockchain are timed, so they are useful for tracking and verifying information

- Blockchain is programmable.

Instructions embedded in blocks, such as "if" this "then" do that "else do this, allow transactions or other actions to be performed only if certain conditions are met, and may be accompanied by additional digital data.

Blockchain has several advantages, which makes it a powerful and secure alternative to distributed databases [8]:

- High Availability: Distributed completely to all nodes and stored in the database completely.
- Verifiability and Integrity: Each block is verified and added to the blockchain. Therefore, it will be difficult to change the data in it because all the blocks have to be changed value. • Easy in determining a common starting point, where to store data - which is always added to the last block in the longest chain.

These advantages make the blockchain attractive for use in recording systems on e-voting.

B.Election and Blockchain Technology

E-voting currently widely used by some countries in the world, for example in Estonia. The country has been using the e-voting system since 2005 and in 2007 conducted online voting and was the first country in the world to conduct online voting [8]. Since then, a legally binding online voting system has been implemented in various other organizations and countries such as the Austrian Federation of Students, Switzerland, the Netherlands, Norway, and so on [9]. But it still has considerable security issues and the selection is often canceled [10]. Although getting a lot of attention, online voting system is still not widely done in various countries around the world. The traditional voting system has several problems encountered when managed by an organization that has full control over the system and database, therefore the organization can tamper with the database, and when the database changes the traces can be easily eliminated [11].

The solution is to make the database public, the database owned by many users, which is useful to compare if there are any discrepancies. The solution to the e-voting system is compatible with using blockchain technology. Blockchain technology allows in support of e-voting applications. Each voter's vote serves as a transaction that can be created into blockchain that can work to track voice counting. In this way, everyone can approve the final calculation because of the open blockchain audit trail, the vote count can be verified that no data is altered or deleted nor is there any unauthorized data entered in the blockchain.

IV. IMPLEMENTATION

Among many choices of different tech stacks in the study it is found that the best suited option for implementing the system is Ethereum blockchain network. It provides the bigger range of use cases with the power of smart contracts, through which many applications which require a web server can be run without it, on a peer to peer network, thus making them very hard if not impossible to mutate.

Another reason of selecting ethereum network is that the transactions happen in almost real time. The transaction on ethereum blockchain occurs in exchange of some ethers or gas which is used to cover the cost of operations, hardware and reward to the miners who verify these transactions. The contracts are written in solidity programming language which is used to develop smart contracts in ethereum.

To be able to hold elections following challenges which include transparency, authentication, verifiability and individuality shall be overcome. In order to make transparent and verifiable elections check signed and time stamped data is needed to be gathered and stored. Also, to preserve the individuality it is important to make sure that a person who is eligible to vote in particular election can only vote once and only on his behalf. No person must be able to vote on someone else's behalf.

These challenges can be solved by using blockchain by writing executable smart contracts according to the aforementioned challenges. Just like writing a code in any other programming language, one needs to define some rules, exceptions, error handling, objects, data models etc. After a smart contract is initialized it can be removed or discarded for the blockchain, and anyone on the network can verify the execution trail.

To deploy a contract on main ethereum network is costly and requires ether (A crypto currency) which is costly. So for testing purposes the contract has been deployed on a test network. There are many test networks available and one of them is rinkeby (<https://www.rinkeby.io/>). One can get some fake ethers on this network to deploy and perform transactions on this network. In order to use a test network users need to download a legit ethereum wallet to store and manage their account.

In the code given in the following figure - the Voter has been defined as a struct in the solidity. This structure contains a Boolean variable `isVoted` to check if the voter has voted before or not `hasRightToVote` to check if the person has the right to vote or not vote - an integer to select the index of the candidate the person to which the user has voted. `ID` - to keep the address of the voter.

There is another structure named `proposal` - which counts the number of votes a candidate or a proposal has got.

```
address chairPerson;  
struct Voter {  
    bool isVoted;  
    bool hasRightToVote;  
    uint8 vote;  
    address ID;  
}  
struct Proposal {  
    uint voteCount;  
}
```

In Fig. 2, giveRightToVote function

Owner of the contract, who was declared once during the construction of the contract is held in the chairperson variable. This function can only be executed by the owner of the contract. This property can be checked with a basic if statement. Then, the voting right to an eligible Voter's (wallet) address is given. Example as follows:

giveRightToVote(0xDF69B68b00A3a4e6F907eD

353467bA
C068aF0717);

The person, who has that Ethereum address, imposed by the chairperson, has the right to vote within this contract.

```
function giveRightToVote(address  
toVoter) public {  
    if (msg.sender !=  
        chairPerson) {  
        voters[toVoter].isVoted {  
            return;  
        }  
    }  
    else {  
        voters[toVoter].hasRightToVote  
        true;  
    }  
}
```

Fig. 3. Code block of the function that initialize voters.

The given code block can be executed by every voter, whenever they want to attend the voting (until the deadline). Voters just send the id of the proposal, which they want to vote as a parameter and their votes are hence recorded. This function firstly detects who currently are trying to execute that function of the contract. More, if the person has the right to vote, and casted his/hervote, thereafter the person is marked as already isVoted, and the vote count of the candidate (proposal) of his choice is incremented by one or by another number based on his/hervoter weight.

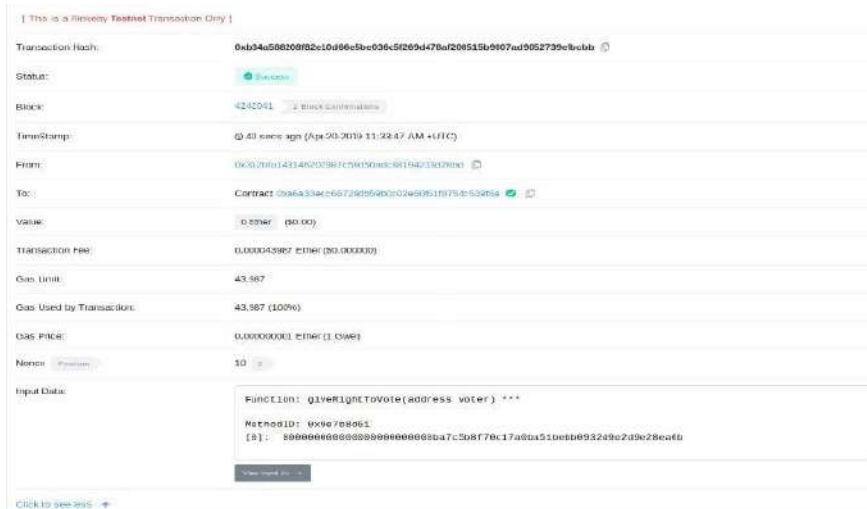
```
function vote(uint8 toProposal) public
{
    Voter storage sender =
    voters[msg.sender]; if
    (sender.isVoted || toProposal >=
    proposals.length &&
    !sender.hasRightToVote)
    return; sender.isVoted =
    true; sender.vote =
    toProposal;
    proposals[toProposal].voteCount += 1;
}
```

Fig. 4. The vote() function

```
function winningProposal() public
constant returns
(uint256 _winningProposal) {
    uint256 winningVoteCount = 2;
    _winningProposal = 0;
    for (uint8 prop = 0; prop
    < proposals.length; prop++)
    if
    (proposals[prop].voteCount > winningVoteCount)
    {
        winningVoteCount =
        proposals[prop].voteCount;
```

Fig. 5. Code block defining the vote casting process.

The winningProposal() function, presented in Fig. 4, returns the id of the winning candidate in the winningProposal variable. It doesn't finish the voting process itself, but it returns the winning proposal every



time it is executed. This function checks every proposal, counts the votes and then returns the one, who is the winner of the whole voting process as of the execution time, since it doesn't end the Election. Voter can verify the trail of the blockchain on the rinkeby network by going to the <https://rinkeby.etherscan.io> and entering the transaction hash received during the casting of the vote.



Fig. 6 and Fig. 7 show detailed records of the entries (blocks) regarding the vote creation and casting operations written in the blockchain. This information is publicly available to everyone tracking the network.

V. CONCLUSION

By building this proposed smart contract of ours, success has been achieved in moving e-voting to the blockchain platform and we addressed some of the fundamental issues that legacy e-voting systems have, by using the power of the Ethereum network and the blockchain structure. The trials have resulted in, the concept of blockchain and the security methodology which it uses, namely immutable hash chains, has become adaptable to polls and elections. This achievement may even pave the way for other blockchain applications that have impact on every aspect of human life. At this point, Ethereum and the smart contracts, which made one of the most revolutionary breakthroughs since the blockchain itself, helped to overturn the limited perception of blockchain as a cryptocurrency (coin), and turned it into a broader solution-base for many Internet-related issues of the modern world, and may enable the global use of blockchain.

E-voting is still a controversial topic within both political and scientific circles. Despite the existence of a few very good examples, most of which are still in use; many more attempts were either failed to provide the security and privacy features of a traditional election or have serious usability and scalability issues. On the contrary, blockchain-based e-voting solutions, including the one we have implemented using the smart contracts and the Ethereum network, address (or may address with relevant modifications) almost all of the security concerns, like privacy of voters, integrity, verification and non-repudiation of votes, and transparency of counting. Yet, there are also some properties that cannot be addressed solely using the blockchain, for example authentication of voters (on the personal level, not on the account level) requires additional mechanisms to be integrated, such as use of biometric factors [12].

The prominence of distributed systems stands out especially when considering the mitigation of the risk that storing the registrations at a central location (office). This can always somehow allow officials to have the opportunity to physically access to the vote records, which could lead to corruptions and cheatings by the authorities. Additionally, in today's connected world, with the concept of the Internet of Things (IoT), expectedly, many non-computer devices will gain access to the Internet. While we are still working on a mobile phone application as a supportive extension to our work to widen the usability; it is important to note that, apart from phones and tablets; air conditioning devices, cars, chairs, clothes, refrigerators, televisions, and many other everyday objects are/will be able to directly reach to the internet. In terms of blockchain, it won't be difficult to build such distributed systems when there is such a large network and a reserve processing power. Moreover, if all these devices work together as a grid to shorten the validation period of transactions in a blockchain, we will be able to do most of our online transactions securely, reliably, and effectively, not only in theory but also in practice.

ACKNOWLEDGMENT

This study is a part of a broader research related to evoting systems, conducted by the research and development group at BhagwanParshuram Institute of Technology, Computer Science and Engineering Department, and has been supported by Dr. DeepaliVirmani, Head of Department, Computer

Science and Engineering and Ms. AkankshaDhamija, Assistant Professor, Computer Science and Engineering Department.

REFERENCES

- [1] S. Nakamoto, “Bitcoin: a peer-to-peer electronic cash system”, [Online]. Available: <https://bitcoin.org/bitcoin.pdf> .
- [2] G. Wood, “Ethereum: a secure decentralised generalised transaction ledger”, Ethereum Project Yellow Paper, vol. 151, pp. 1-32, 2014.
- [3] C.D. Clack, V.A. Bakshi, and L. Braine, “Smart contract templates: foundations, design landscape and research directions”, Mar 2017, arXiv:1608.00771.
- [4] E. Maaten, “Towards remote e-voting: Estonian case”, *Electronic Voting in Europe-Technology, Law, Politics and Society*, vol. 47, pp. 83-100, 2004.
- [5] Christian, “Desain Dan Implementasi Visual Cryptography Pada Sistem E-Voting Untuk Meningkatkan Anonymity,” Institut Teknologi Bandung, 2017.
- [6] U.C. Çabuk, A. Çavdar, and E. Demir, "E-Demokrasi: Yeni Nesil Doğrudan Demokraside Türkiye'deki Uygulanabilirliği", [Online] Available: https://www.researchgate.net/profile/Umut_Cabuk/publication/308796230_E-Democracy_The_Next_Generation_Direct_Democracy_and_Applicability_in_Turkey/links/5818a6d408aee7cdc685b40b/E-Democracy-The-Next-Generation-Direct-Democracy-and-Applicability-in-Turkey.pdf.
- [7] "Final report: study on eGovernment and the reduction of administrative burden (SMART 2012/0061)", 2014, [Online]. Available: <https://ec.europa.eu/digital-single-market/en/news/final-report-study-egovernment-and-reduction-administrative-burden-smart-20120061>.
- [8] D. A. Wijaya, *Bitcoin Tingkat Lanjut*. 2016.
- [9] H. Watanabe, S. Fujimura, A. Nakadaira, Y. Miyazaki, A. Akutsu, and J. J. Kishigami, “Blockchain contract: A complete consensus using blockchain,” *2015 IEEE 4th Glob. Conf. Consum. Electron. GCCE 2015*, pp. 577–578, 2016.
- [10] T. Martens, “Verifiable Internet Voting in Estonia,” *October*, pp. 1–7, 2009
- [11] Follow My Vote, “Why Online Voting.” [Online]. Available: <https://followmyvote.com/>. [Accessed: 01-Jan-2017].
- [12] U.C. Çabuk, T. Üenocak, E. Demir, and A. Çavdar, “A Proposal on initial remote user enrollment for IVR-based voice authentication systems”, *Int. J. of Advanced Research in Computer and Communication Engineering*, vol 6, pp.118- 123, July 2017.

DICGM: A Deep Learning based Novel Image Caption Generator Model

Shweta Taneja¹, Bhawna Suri² and Savneet Kaur³

shwetataneja@bpitindia.com, bhawnasuri@bpitindia.com, savneetkaur@yahoo.com

1,2 Department of Computer Science & Engineering, Bhagwan Parshuram Institute of Technology
Guru Gobind Singh Indraprastha University, Delhi.

3 Department of Information Technology, Guru Teg Bahadur Institute of Technology, Delhi, India.

Abstract-Image caption generator is a task that involves computer vision and natural language processing concepts to recognize the context of an image and describe them in a natural language like English. It is used in a variety of applications like describing images to People with Visual Impairment (PVI's), web development, describing real time videos etc. With the advancement in Deep learning techniques, availability of huge datasets and computer power, we can build models that can generate captions for an image.

In this work, we propose a fully automated approach that uses deep neural networks to build image caption generator model – DICGM. The image features are extracted from Xception which is a CNN model trained on the ImageNet dataset and then fed to LSTM model, which is responsible for generating the image captions. The main goal here is to put CNN-RNN together to create an automated image captioning model that takes in an image as input and outputs a sequence of text that describes the image. DICGM is tested using Flickr8k Dataset. It contains a total of 8092 images in JPEG format with different shapes and sizes. Of which 6000 are used for training, 1000 for test and 1000 for development. DICGM works really well with a small number of photos. For production-level models, there is a need to train on datasets larger than 100,000 images which can produce better accuracy models. This proposal is a work in progress and we are convinced that with more extensive tuning and precise configuration the results will improve.

Keywords: Machine learning, Image caption generation, Image colorization, CNN, RNN, LSTM

I Introduction

For a long time, colorization has been a semi-automated process, relying on hints from the user. Levin et al.[1], for instance, proposed that neighboring pixels in space with similar intensities should have similar colors. Thus, in his work, hints are provided as rough inaccurate "scribbles" on a grayscale image, and the algorithm is able to generate Exploring Convolutional Neural Networks for Automatic Image. Several others have improved this algorithm further, including Huang et al [2] (by addressing color-bleeding issues) and Qu et al.[3] (by modifying the cost function to account for color continuity over similar textures in addition to similar intensities).

The objective of our research is to learn the concepts of a CNN and LSTM model and build a working model of Image caption generator by implementing them. In this work, we will be implementing the caption generator using *CNN (Convolutional Neural Networks)* and *LSTM (Long short term memory)*. The image features will be extracted from Xception which is a CNN model trained on the imagenet dataset and then we feed the features into the LSTM model which will be responsible for generating the image captions.

There are two main architectures of an image captioning model. The first one is an image-based model which extracts the features of the image, and the other is a language-based model which translates the features and objects given by our image-based model to a natural sentence. In this, we will be using a pretrained CNN network that is trained on the ImageNet dataset. The images are transformed into a standard resolution of 224 X 224 X 3. This will make the input constant for the model for any given image.

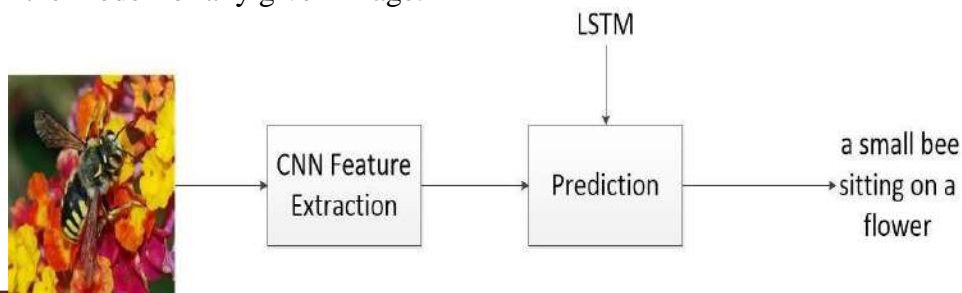


Figure 1 Image captioning model

The condensed feature vector is created from a convolutional neural network (CNN). In technical terms, this feature vector is called *embedding*, and the CNN model is referred to as an **encoder**. In the next stage, we will be using these embeddings from the CNN layer as input to the LSTM network, a **decoder**. In a sentence language model, LSTM is predicting the next word in a sentence. Given the initial embedding of the image, the LSTM is trained to predict the most probable next value of the sequence. Its just like showing a person a series of pictures and asking them to remember the details. And then later show them a new image which has similar content to the previous images and ask them to recall the content. This “recall” and “remember” job is done by our LSTM network. The symbols <start> and <stop> stoppers are used to signal the end of the caption. This way, the model learns from various instances of images and finally predicts the captions for unseen images. The main goal here is to put CNN-RNN together to create an automatic image captioning model that takes in an image as input and outputs a sequence of text that describes the image.

A captioning model relies on two main components, a CNN and an RNN. Captioning is all about merging the two to combine their most powerful attributes i.e. A **Convolutional Neural Network (ConvNet/CNN)** which is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. Convolutional Neural networks are specialized deep neural networks which can process the data that has input shape like a 2D matrix. Images are easily represented as a 2D matrix and CNN is very useful in working with images. **Long short-term memory (LSTM)** is a type of RNN (**recurrent neural network**) which is well suited for sequence prediction problems. Based on the previous text, the next word can be predicted. It has proven itself effective from the traditional RNN by overcoming the limitations of RNN which had short term memory. LSTM can carry out relevant information throughout the processing of inputs and with a forget gate, it discards non-relevant information. An LSTM consists of three main components: a forget gate, input gate, and output gate. Each of these gates is responsible for altering updates to the cell's memory state. The **long-term memory** is usually called the **cell state**. The looping arrows indicate the recursive nature of the cell. This allows information from previous intervals to be stored within the LSTM cell. Cell state is modified by the forget gate placed below the cell state and also adjusted by the input modulation gate. From the equation, the previous cell state forgets by multiplying with the forget gate and adds new information through the output of the input gates. The **remember vector** is usually called the **forget gate**. The output of the forget gate tells the cell state which information to forget by multiplying 0 to a position in the matrix. If the output of the forget gate is 1, the information is kept in the cell state. From equation, the sigmoid function is applied to the weighted input/observation and previous hidden state. The save vector is usually called the input gate. These gates determine which information should enter the cell state / long-term memory. The important parts are the activation functions for each gate. The input gate is a sigmoid function and have a range of [0,1]. The **focus vector** is usually called the **output gate**. Out of all the possible values from the matrix, which should be moving forward to the next hidden state. The **working memory** is usually called the **hidden state**. What information should I take to the next sequence? This is analogous to the hidden state in RNN and HMM.

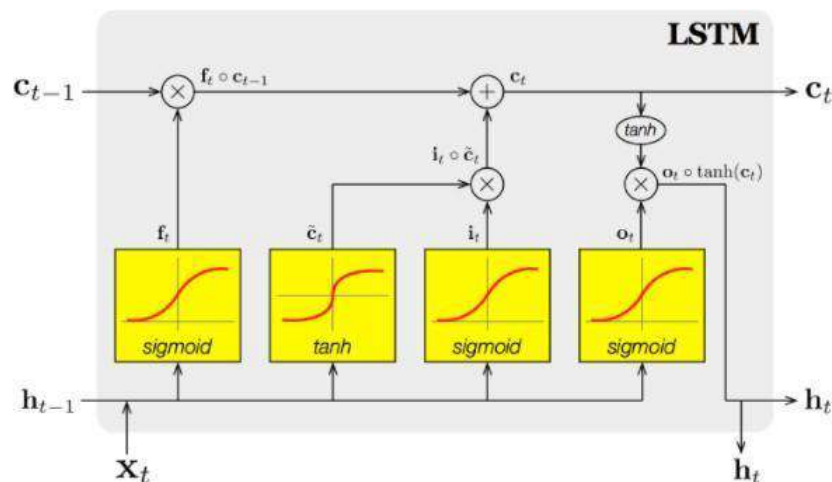


Figure 2 VIEW OF LSTM CELL

The first sigmoid activation function is the **forget gate**. Which information should be forgotten from the previous cell state (C_{t-1}). The second sigmoid and first tanh() activation function is our **input gate**. The last sigmoid is the **output gate** and highlights which information should be going to the next **hidden state**.

II State of Art

We review and describe the main categories of existing image captioning methods and they include template-based image captioning, retrieval-based image captioning, novel caption generation and Deep neural network-based image captioning.

2.1 Template-based image captioning

Template-based approaches have fixed templates with a number of blank slots to generate captions. In these approaches, different objects, attributes, actions are detected first and then the blank spaces in the templates are filled. For example, Farhadi et al. [4] use a triplet of scene elements to fill the template slots for generating image captions. Li et al. [5] extract the phrases related to detected objects, attributes and their relationships for this purpose. A Conditional Random Field (CRF) is adopted by Kulkarni et al. [6] to infer the objects, attributes, and prepositions before filling in the gaps. Template-based methods can generate grammatically correct captions. However, templates are predefined and cannot generate variable-length captions. Moreover, later on, parsing based language models have been introduced in image captioning [8,9,10,11] which are more powerful than fixed template-based methods.

2.2 Retrieval-based image captioning

One type of image captioning commonly used in early work is retrieval based. Here, the query image, retrieval-based methods produce a caption for it through retrieving one or a set of sentences from a pre-specified sentence pool. The generated caption can either be a sentence that has already existed, or a sentence composed from the retrieved ones. In [12], the $\langle \text{object, action, scene} \rangle$ meaning space to link images and sentences is established. Given a query image, they map it into the meaning space by solving a Markov Random Field, and use Lin similarity measure to determine the semantic distance between this image and each existing sentence. In [13], to caption an image global image descriptors are employed to retrieve a set of images from a web scale collection of captioned photographs. Then, they utilize semantic contents of the retrieved images to perform re-ranking and use the caption of the top image as the description of the query.

2.3 Novel caption generation.

Novel captions can be generated from both visual space and multimodal space. A general approach of this category is to analyze the visual content of the image first and then generate image captions from the visual content using a language model [9,10,11,12]. These methods can generate new captions for each image that are semantically more accurate than previous approaches. Most novel caption generation methods use deep machine learning based techniques. Therefore, deep learning based novel image caption generating methods are our main focus in this literature. Usually captions are generated for a whole scene in the image.

2.4 Deep neural network-based image captioning

Retrieval based and template-based image captioning methods are adopted mainly in early work. Due to great progress made in the field of deep learning [13] [14], recent work begins to rely on deep neural networks for automatic image captioning. Even though deep neural networks are now widely adopted for tackling the image captioning task, different methods may be based on different frameworks. Therefore, we classify deep neural network-based methods into subcategories on the basis of the main framework they use and discuss each subcategory respectively.

2.4.1 Encoder-Decoder Architecture-Based Image captioning

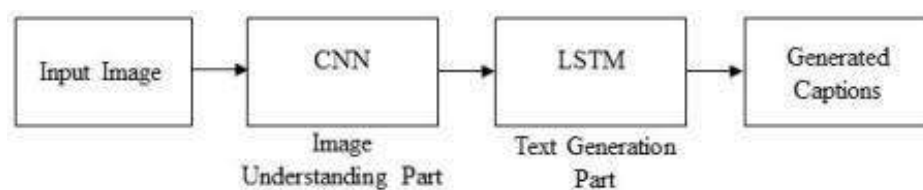


Figure 3: A block diagram of simple Encoder-Decoder architecture-based image captioning

The neural network-based image captioning methods work as just simple end to end manner. These methods are very similar to the encoder-decoder framework-based neural machine translation [15]. In this network, global image features are extracted from the hidden activations of CNN and then fed them into an LSTM to generate a sequence of words. A typical method of this category has the following general steps:

- (1) A vanilla CNN is used to obtain the scene type, to detect the objects and their relationships.
- (2) The output of Step 1 is used by a language model to convert them into words, combined phrases that produce an image caption. A simple block diagram of this category is given in Figure 3. Vinyals et al. [23] proposed a method called Neural Image Caption Generator (NIC). The method uses a CNN for image representations and an LSTM for generating image captions. This special CNN uses a novel method for batch normalization and the output of the last hidden layer of CNN is used as an input to the LSTM decoder. This LSTM is capable of keeping track of the objects that already have been described using text. NIC is trained based on maximum likelihood estimation. In generating image captions, image information is included to the initial state of an LSTM. The next words are generated based on the current time step and the previous hidden state. This process continues until it gets the end token of the sentence. Since image information is fed only at the beginning of the process, it may face vanishing gradient problems. The role of the words generated at the beginning is also becoming weaker and weaker. Therefore, LSTM is still facing challenges in generating long length sentences.

III Materials and Methods

3.1 Data set

We have used the following two datasets:

Flickr8k_Dataset: Contains a total of 8092 images in JPEG format with different shapes and sizes. Of which 6000 are used for training, 1000 for test and 1000 for development.

Flickr8k_text: Contains text files describing train_set ,test_set. Flickr8k.token.txt contains 5 captions for each image i.e. total 40460 captions.

This dataset contains 8000 images each with 5 captions (as we have already seen in the Introduction section that an image can have multiple captions, all being relevant simultaneously). The images do not contain any famous person or place so that the entire image can be learnt based on all the different objects in the image. It is small in size. So, the model can be trained easily on low-end laptops/desktops. Data is properly labelled, for each image 5 captions are provided and the dataset is available for free. Example: Each photo has 5 captions:



Figure4 sample data

```
[ 'A blond horse and a blond girl in a black sweatshirt be  
stare at a fire in a barrel.'  
'A girl and her horse stand by a fire.'  
'A girl holds a horse 's lead behind a fire.'  
'A man, and girl and two horse be near a contain fire.'  
'Two person and two horse watch a fire.'
```

These images are bifurcated as follows: Training Set — 6000 image and Test Set — 2000 images.

3.2 Proposed Approach

In our proposed approach, two main objectives are met- Colorizing Grayscale Videos And building an Image Caption Generator Model.

3.2.1 Method for Colorizing Grayscale Videos

- Phase one of the algorithm has five parts of action, as following:
- Accepting the colored source image.
- Converting that image to a grayscale image and extract the L (luminance) value out of given RGB image; using formula $Gray(L) = 0.299 \text{ Red} + 0.587 \text{ Green} + 0.114 \text{ Blue}$.
- Each pixel of the grayscale version is processed in which its luminance, average and the standard deviation of its 8-neighbours are recorded in a table called learning table. These three columns (the intensity of each pixel, the average of 8-neighborhood pixels, and the standard deviation of 8-neighborhood pixels) are applied as inputs to the 3 nodes neural network.
- The target data are prepared by recording the RGB values of each pixel of the colored version. So, there will be three inputs, and three targets.
- The whole dataset is divided to two subsets in a full random procedure, 90% for training set and 10 % for testing set.

Figure 5 Block diagram for the first phase.

Colorizing of the Target Grayscale Image



Once the neural network is trained it can be directly used to colorize the destination grayscale image. The following points explain the second phase of the technique .Accepting the gray scale image (destination image to be colorized).

- The luminance (L) of pixel values of the grayscale image will be determined and recorded as well as the average and standard deviation of its 8-neighborhood pixels.
- The variables (L) are applied as inputs to the trained ANN to produce three values which are the R, G, and B components of the colored version.
- The output RGB components are rearranged in three layered matrix to reconstruct a colored image.

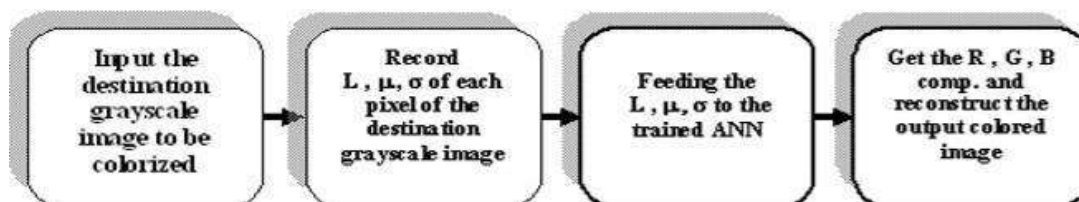


Figure 6 Block diagram for the second phase

3.2.2 Image Caption Generator Merged Model

So, to make our image caption generator model, we will be merging these architectures. It is also called the CNN-RNN model.

1. CNN is used for extracting features from the image. We will use the pre-trained model Exception.
2. LSTM will use the information from CNN to help generate a description of the image.

3.2.3 Proposed CNN-RNN Model

We'll be using a pre-trained network like VGG16 or Xception. Since we want a set of features that represents the spatial content in the image, **we're going to remove the final fully connected layer that classifies the image** and look at earlier layer that processes the spatial information in the image. So now CNN acts as a **feature extractor** that compresses the information in the original image into a smaller representation. Since it encodes the content of the image into a smaller feature vector hence, this **CNN is often called the encoder**.

When we process this feature vector and use it as an initial input to the following RNN, then it would be called decoder because RNN would decode the process feature vector and turn it into natural language.

Model - Image Caption Generator

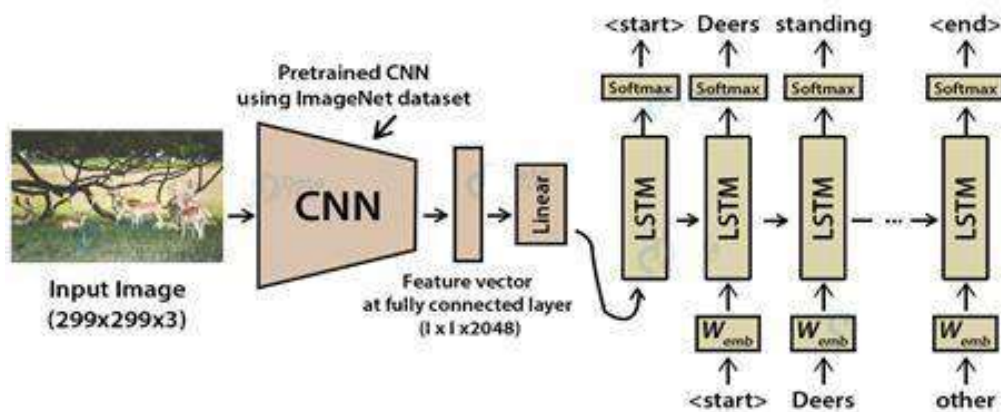


Figure 7 Combined Model-Image Caption Generator

Data pre-processing and cleaning is an important part of the whole model building process. Understanding the data helps us to build more accurate models.

Caption Preprocessing

Each image in the dataset is provided with 5 captions. For e.g image 1000268201_693b08cb0e.jpeg has captions

```
[ 'A child in a pink dress is climbing up a set of stairs in an entry way .',
  'A girl going into a wooden building .',
  'A little girl climbing into a wooden playhouse .',
  'A little girl climbing the stairs to her playhouse .',
  'A little girl in a pink dress going into a wooden cabin .']
```

Captions are read from Flickr8k.token.txt file and stored in dictionary k:v where k = image id and value = [list of caption]. Since there are 5 captions for each image and we have preprocessed and encoded them in below format

“startseq “ + caption + “ endseq”

The reason behind startseq and endseq is,

startseq : Will act as our first word when feature extracted image vector is fed to decoder. It will kick-start the caption generation process.

endseq : This will tell the decoder when to stop. We will stop predicting word as soon as endseq appears or we have predicted all words from train dictionary whichever comes first.

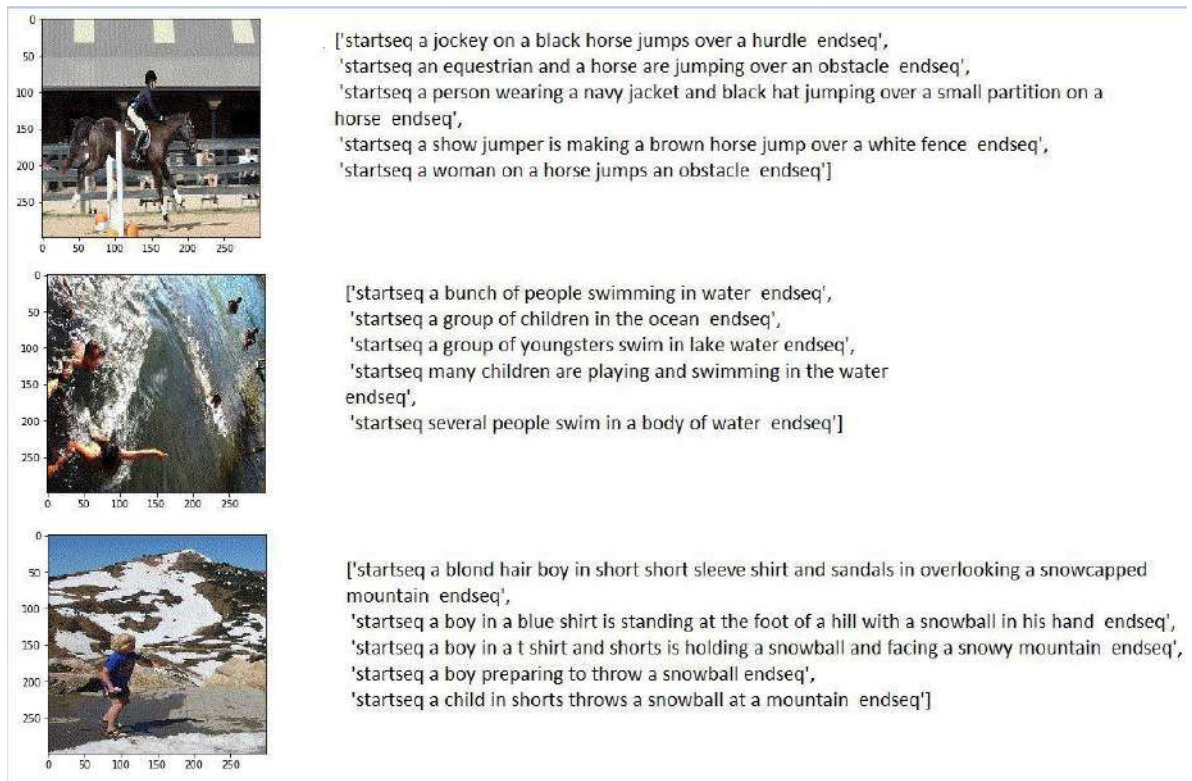


Figure 8 Captions after preprocessing

Tokenizing Captions

The RNN component of the captioning network is trained on the captions in the Flickr_8K dataset. We're aiming to train the RNN to predict the next word of a sentence based on previous words. For this we transform the caption associated with the image into a list of tokenized words. This tokenization turns any strings into a list of integers. First, we iterate through all of the training captions and create a dictionary that maps all unique words to a numerical index. So, every word we come across, will have a corresponding integer value that we can find in this dictionary. The words in these dictionaries are referred to as our vocabulary. This list of tokens (in a caption) is then turned into a list of integers which come from our dictionary that maps each distinct word in the vocabulary to an integer value. It transforms each word in a caption into a vector of a desired consistent shape. After this embedding step, we're finally ready to train an RNN that can predict the most likely next word in a sentence.

We will define the deep learning model and fit it on the training dataset. This section is divided into the following parts:

1. Loading Data.
2. Defining the Model.
3. Fitting the Model.
4. Evaluate the model.

Loading Data

First, we must load the prepared photo and text data so that we can use it to fit the model. We are going to train the data on all of the photos and captions in the training dataset. While training, we are going to monitor the performance of the model on the development dataset and use that performance to decide when to save models to file.

The train and development dataset have been predefined in the *Flickr_8k.trainImages.txt* and *Flickr_8k.devImages.txt* files respectively, that both contain lists of photo file names. From these file names, we can extract the photo identifiers and use these identifiers to filter photos and descriptions for each set.

The model we will develop will generate a caption given a photo, and the caption will be generated one word at a time. The sequence of previously generated words will be provided as input. Therefore, we will need a ‘*first word*’ to kick-off the generation process and a ‘*last word*’ to signal the end of the caption.

We will use the strings ‘*startseq*’ and ‘*endseq*’ for this purpose. These tokens are added to the loaded descriptions as they are loaded. It is important to do this now before we encode the text so that the tokens are also encoded correctly.

Defining the Model

We will define a deep learning based on the “*merge-model*” described by Marc Tanti, et al. in their 2017 papers. We will describe the model in three parts:

- **Photo Feature Extractor.** This is a 16-layer VGG model pre-trained on the ImageNet dataset. We have pre-processed the photos with the VGG model (without the output layer) and will use the extracted features predicted by this model as input.
- **Sequence Processor.** This is a word embedding layer for handling the text input, followed by a Long Short-Term Memory (LSTM) recurrent neural network layer.
- **Decoder** (for lack of a better name). Both the feature extractor and sequence processor output a fixed-length vector. These are merged together and processed by a Dense layer to make a final prediction.

The Photo Feature Extractor model expects input photo features to be a vector of 4,096 elements. These are processed by a Dense layer to produce a 256-element representation of the photo.

The Sequence Processor model expects input sequences with a pre-defined length (34 words) which are fed into an Embedding layer that uses a mask to ignore padded values. This is followed by an LSTM layer with 256 memory units. Both the input models produce a 256-element vector. Further, both input models use regularization in the form of 50% dropout. This is to reduce overfitting the training dataset, as this model configuration learns very fast.

The Decoder model merges the vectors from both input models using an addition operation. This is then fed to a Dense 256 neuron layer and then to a final output Dense layer that makes a SoftMax prediction over the entire output vocabulary for the next word in the sequence.

Fitting the Model

When the skill of the model on the development dataset improves at the end of an epoch, we will save the whole model to file. At the end of the run, we can then use the saved model with the best skill on the training dataset as our final model. We can do this by defining a *ModelCheckpoint* in Keras and specifying it to monitor the minimum loss on the validation dataset and save the model to a file that has both the training and validation loss in the filename.

Evaluate Model

Once the model is fit, we can evaluate the skill of its predictions on the holdout test dataset. We will evaluate a model by generating descriptions for all photos in the test dataset and evaluating those predictions with a standard cost function. First, we need to be able to generate a description for a photo using a trained model. This involves passing in the start description token ‘*startseq*’, generating one word, then calling the model recursively with generated words as input until the end of sequence token is reached ‘*endseq*’ or the maximum description length is reached.

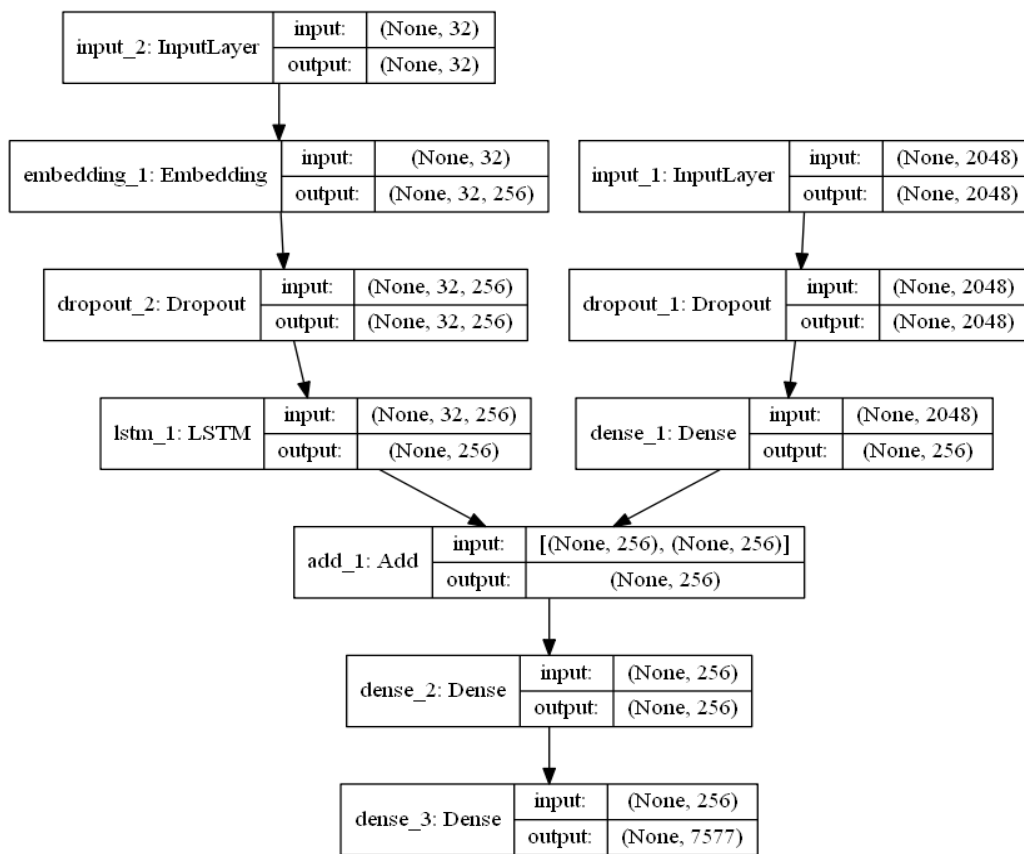


Figure 10 Model Summary

IV Results and Discussion

In this work, we have implemented a CNN-RNN model by building an image caption generator. Some key points to note are that our model depends on the data, so, it cannot predict the words that are out of its vocabulary. We used a small dataset consisting of 8000 images. For production-level models, we need to train on datasets larger than 100,000 images which can produce better accuracy models.

Caption : two girls are playing in the grass



Caption : man in yellow kayak is reflecting up river



Figure 11 Result for image captioning

Following colorization is performed on manually handpicked similar images from Unsplash dataset which is an ongoing project containing images of countryside, coasts, mountains, roads etc. One of these videos is converted into gray scale and other one is used to colorize first one. In this work, we figured out what is deep learning. we assembled and trained the CNN model to colorize black and white videos. We have tested that this model works really well with a small number of photos. For production-level models, we need to train on datasets larger than 100,000 images which can produce better accuracy models.

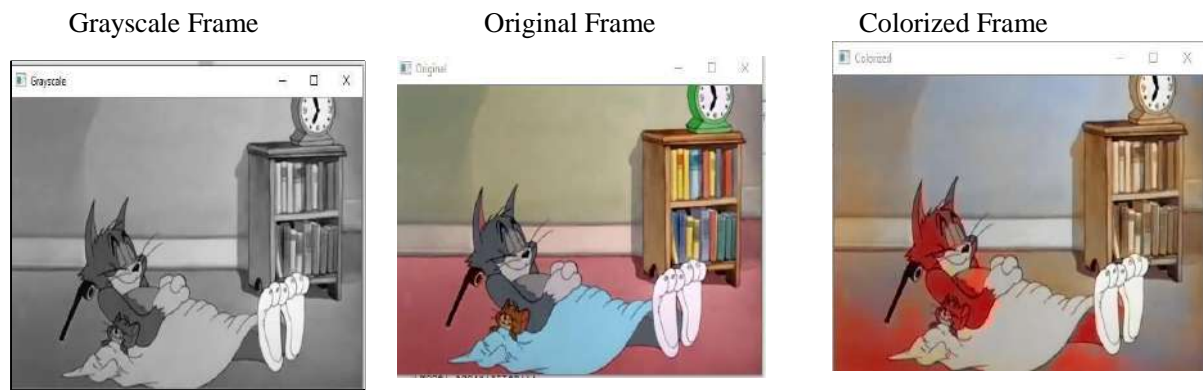


Figure 12 Results for Video Colorization

V CONCLUSION

In this overview, we have compiled all aspects of the image caption generation task, discussed the model framework proposed in recent years to solve the description task, focused on the algorithmic essence of different mechanisms. We summarize the large datasets and evaluation criteria commonly used in practice. Although image caption can be applied to image retrieval, video caption, and video movement and the variety of image caption systems are available today, experimental results show that this task still has better performance systems and improvement. It mainly faces the following three challenges: first, how to generate complete natural language sentences like a human being; second, how to make the generated sentence grammatically correct; and third, how to make the caption semantics as clear as possible and consistent with the given image content. In this work we design and develop a methodology, strongly based on neural networks, to colorize gray level videos. Our proposal consists on a series of steps to finally obtain a trained neural network which predicts the color of a gray level pixel. The methodology joins two main ideas from previous works: color reduction using vector quantization and using a group of neighboring pixels to predict the color of a single pixel. In our opinion the results are good and promising although the prediction scores obtained are below the 50%. To a human eye the images obtained show good results and have colors close to the intuition. As the prediction highly depends on the set of training images, the images more similar to the training set show better results than those more different.

References

1. Levin, A., Lischinski, D., & Weiss, Y. (2004). Colorization Using Optimization. *ACM Transactions on Graphics (Tog)*, 23, 689-694.
2. Huang, Yi-Chin, et al. "An adaptive edge detection-based colorization algorithm and its applications." *Proceedings of the 13th annual ACM international conference on Multimedia*. ACM, 2005
3. Qu, Yingge, Tien-Tsin Wong, and Pheng-Ann Heng. "Manga colorization." *ACM Transactions on Graphics (TOG)*. Vol.25. No. 3. ACM, 2006.
4. Ali Farhadi, Mohsen Hejrati, Mohammad Amin Sadeghi, Peter Young, Cyrus Rashtchian, Julia Hockenmaier, and David Forsyth. 2010. Every picture tells a story: Generating sentences from images. In *European conference on computer vision*. Springer, 15–29.
5. Siming Li, Girish Kulkarni, Tamara L Berg, Alexander C Berg, and Yejin Choi. 2011. Composing simple image descriptions using web-scale n-grams, 220–228.
6. Girish Kulkarni, Visruth Premraj, Sagnik Dhar, Siming Li, Yejin Choi, Alexander C Berg, and Tamara L Berg. 2011. Baby talk: Understanding and generating image descriptions. In *Proceedings of the 24th CVPR*. Citeseer.
7. J. Curran, S. Clark, J. Bos, Linguistically motivated large-scale nlp with cc and boxer, in: *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, pp. 33–36.

8. Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. 2015. Show, attend and tell: Neural image caption generation with visual attention. In International Conference on Machine Learning. 2048–2057.
9. V. Ordonez, G. Kulkarni, T. L. Berg., Im2text: Describing images using 1 million captioned photographs, in: Advances in Neural Information Processing Systems, 2011, pp. 1143–1151.
10. Ryan Kiros, Ruslan Salakhudinov, and Richard S Zemel. 2014. Unifying visual- semantic embeddings with multimodal neural language models. In Workshop on Neural Information Processing Systems (NIPS)).
11. Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. 2015. Show, attend and tell: Neural image caption generation with visual attention. In International
12. Conference on Machine Learning. 2048–2057.
13. Quanzeng You, Hailin Jin, Zhaowen Wang, Chen Fang, and Jiebo Luo. 2016. Image captioning with semantic attention. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 4651–4659.
14. Ting Yao, Yingwei Pan, Yehao Li, Zhaofan Qiu, and Tao Mei. 2017. Boosting image captioning with attributes. In IEEE International Conference on Computer Vision (ICCV). 4904–4912.
15. Y. Bengio, A. Courville, P. Vincent, Representation learning: A review and new perspectives, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (8) (2013) 1798–1828.

Ms. Preeti Arora
Assistant professor
Department of Computer Science
BPIT, New Delhi, India
preeti.arora@bpitindia.com

Ms. Tanisha Madan
Assistant professor
Department of Computer Science
BPIT, New Delhi, India
tanishamadan@bpitindia.com

Ms. Nikita Nijhawan
Assistant professor
Department of Computer Science
BPIT, New Delhi, India
nikitanijhawan@bpitindia.com

Abstract - Sarcasm is one of the sentiments that mean something positive with negative intentions. Nowadays Sarcasm is widely utilized on social platforms such as Twitter, FB, Instagram and comedy etc. Therefore we need a system to detect sarcasm using an effective model. Through an exhaustive study we conclude that the main issue of sarcasm detection is the general expression of sarcasm - “positive sentence address to a negative situation.” Our proposed model works effectively on the LSTM approach and provides 88% accuracy.

Keywords: LSTM, Sarcasm Detection, Sentiment analysis, SVM(Support Vector Machine), KNN, Naive Bayes, Accuracy

INTRODUCTION

Sentiment analysis is the natural language processing (NLP) and text analysis tools to associate and elicit the feelings of a writer from the text [1]. These emotions can be expressed as either positive, negative, or neutral. Sarcastic writing is common in many online texts[7]. Sarcasm can be identified when a person says the opposite of what they mean [3]. However, due to the restrictions of the unauthenticated language and characters used by News headlines, it is hard to understand the opinions of users. In addition, the appearance of sarcasm in any sentence is even more difficult. Therefore we need a system that Predict the emotions accurately thus, we proposed a model that uses Long short-term memory algorithm Our model can predict whether the sentence is sarcastic or non-sarcastic. We Used Dataset from Kaggle and News Paper headlines are used as Dataset. The main advantages of our model are as follows: (i) Sarcasm detection model can automatically differentiate a sarcastic text from a non-sarcastic one. (ii) We have compared our result with many existing algorithms such as SVM (Support Vector Machine), Naive Bayes, and KNN and Our proposed LSTM algorithm has the highest accuracy among all the pre existing algorithms and provides 88% accuracy.

II. LITERATURE STUDY

Long Short Term Memory (LSTM) is modified version of recurrent neural network and it is better than RNN in terms of memory, which makes task to remember past data in memory easy. LSTM has multiple hidden layers and information passes through every layer, the meaningful information is kept and all the non meaningful information is discarded in every cell. LSTM provide us with larger range of parameters such as learning rates and input/output. Therefore, no need for fine adjustments. The complexity to update each weight is reduced to $O(1)$ with it, Non Decaying Error back propagation.

III. PROPOSED FRAMEWORK

Here, We use an LSTM approach to our proposed model and framework as shown in Fig. 2. Our approach model provides efficient results and improves accuracy when compared with the existing algorithm. Each stage of the framework is explained as follows: In the first stage, we use the dataset of the news headlines from: <https://www.kaggle.com/datasets/rmisra/newsheadlines-dataset-for-sarcasm-Detection>. Dataset consists of 28617 headlines, with sarcastic headlines from the Onion and non-sarcastic headlines from HuffPost. The average of each headline is 12 words. The dataset consists of three columns, as shown in fig

1. And those fields are headline(headline of the news article), is_sarcastic(if the headline is sarcastic then the value is 1 otherwise, it is 0), and article link(link of the original news article).

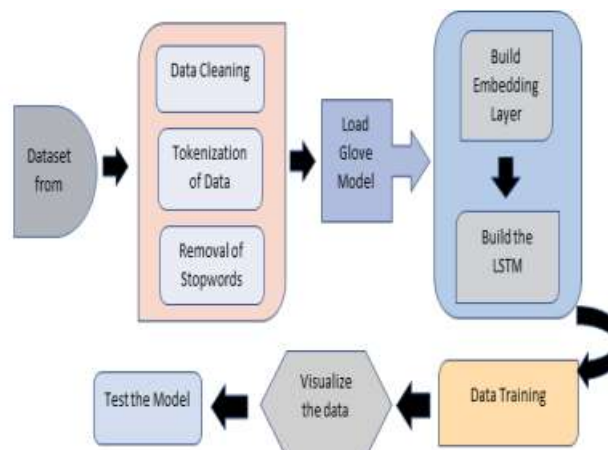


Fig 1.Proposed Framework

In the second stage, preprocessing of data is done after collecting the data. This is an essential phase of our model. Here we clean the data and remove all the unwanted information from the dataset. Such as special characters, urls, hashtags, and unnecessary symbols which do not contribute to the analysis purpose. Pre-processing represents the transformations applied to the data before feeding it to the learning algorithms. Before applying the data preprocessing task, nonEnglish words which are present in the headlines are removed. In figure 3, word cloud representations of the two datasets after applying preprocessing tasks are shown. It is clear from the figures that data preprocessing steps are very essential. Next, we convert the text data into a lower case such as “Post” and “post” are seen as two different words by the program. Therefore, it is important to normalize the case of the words so that every word is in the same case and the model does not process the same word as two different Tokens. After that stop word removal process is applied. Stop words are used more frequently than words in the language but they do not require adding more value to a sentence, hence it is safe to ignore them by removing them from the dataset. They are the most common words as “a”, “an”, “to”, “the”, and “is” in the English language. Thus after applying all the required data preprocessing tasks the total number of headlines in the dataset becomes 28617. In the third stage, the tokenizer is built based on the frequency of words occurring in the dataset. This feature converts each headline/sentence to a sequence of 25 words. Extra spaces after the 25 words would be padded. The number of unique tokens 28657 was gathered from the news headlines. The size of the vocab is 28658 words. Vocab size is taken one more than the unique token because if the word is not in the vocab, then it will give that extra space to that word. In the fourth stage, we generate the Embedding matrix. This is a vector value given to each of the unique words selected by the tokenizer in the previous step. The embedding matrix was constructed using pre-trained glove word embedding by Stanford. In the fifth stage, we sequence the embedded and tokenized words. In the sixth stage, the embedding layer is the first layer that is passed to the LSTM model. Then we add 23 more neural networks layer to it. The output is feeding to the LSTM model and then flattening all the LSTM sequences, feeding it. Our model works effectively on the following functions.

1. The sigmoid function helps to increase the accuracy of the model.
2. Binary cross-entropy is used for the loss and this loss function tells how good our

model predicts.

3. Adam optimizer is used which increases the performance of the model.
4. An Accuracy matrix is used for the performance check.

IV. PROPOSED ALGORITHM



Fig2. Proposed Algorithm

V. RESULT ANALYSIS

For result analysis, news headlines were collected from the KAGGLE. Python, machine learning, and deep learning are used to implement the algorithm. For the efficient results, 24 neural networks were used for each dataset, the algorithm is executed 'n' times, and the results are stored. Furthermore, Errors and accuracy are calculated on the basis of training and validating data. Our model uses dropout to reduce the overfitting by repeating the algorithm. We find minimum error from the test data to reduce the overfitting. The performance of the SVM(support vector machine), Naive Bayes, and Knn are not as good as our Proposed LSTM approach. LSTM model gave the best result in comparison to the other algorithm. The confusion matrix for each subset is calculated. It gives the performance measure of the classifier model, and how well the model predicts the correct and incorrect output values with the count. SVM gave an accuracy of 59% approx, Naive Bayes accuracy is 55% approx, and Knn accuracy is 61% approx. After training and testing the model, we see that our model predicts 88.35% accuracy and there is a loss of 0.2732. On plotting the graphs for Model Validated Accuracy and Validate loss, we get the following result for the prediction of the model. Fig 4 and fig 5 show the Model Training and the model validated

accuracy with each epoch of the data set.

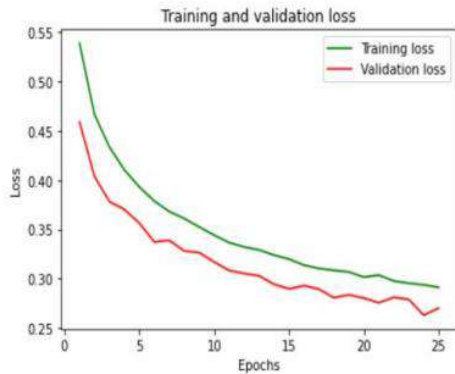


Fig4. Model Loss

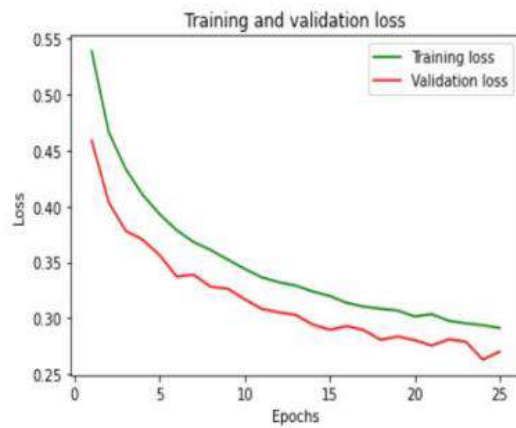


Fig4. Model Loss

VI. CONCLUSION

In this paper, LSTM approach is proposed. Proposed approach provides good accuracy and overcome the problem of Sarcasm Detection. For the adequate results, Binary entropy was used to find the loss and The sigmoid function helps to increase the accuracy of the model. Binary crossentropy is used for the loss and this loss function tells how good our model predicts. Adam optimizer is used which increases the performance of the model. An Accuracy matrix is used for the performance check. Model is trained on the basis of training & validating Data. Overall 88.35% of Accuracy is achieved by the proposed model. Our model is capable to detect sarcastic and nonsarcastic headlines with no context.

REFERENCES

- [1] B. Wagh, J. V Shinde, and P. A. Kale, “A Twitter Sentiment Analysis Using NLTK and Machine Learning Techniques,” *Int. J. Emerg. Res. Manag. Technol.*, vol. 6, no. 12, pp. 37–44, 2018
- [2] C. J. H. E. Gilbert, “Vader: A parsimonious rule-based model for sentiment analysis of social media text,” in *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*. Available at (20/04/16)
- [3] Shihab Elbagir and Jing Yang, “Twitter Sentiment Analysis Using Natural Language Toolkit and VADER Sentiment,” *Proceedings of the International MultiConference of Engineers and Computer Scientists, IMECS 2019, Hong Kong, March 13 15, 2019*.
- [4] S. B. Mane, Y. Sawant, S. Kazi, and V. Shinde, “Real-Time Sentiment Analysis of Twitter Data Using Hadoop,” *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 3, pp. 3098–3100, 2014.
- [5] M. Bouazizi and T. Ohtsuki, “A Pattern-Based Approach for MultiClass Sentiment Analysis in Twitter,” *IEEE Access*, vol 3536, no. c, pp. 1–21, 2017.
- [6] Wenhong Tian, Yong Zhao, in *Optimized Cloud Resource Management and Scheduling*, 2015
- [7] Medhat, Walaa, Ahmed Hassan, and Hoda Korashy. (2014) "Sentiment analysis algorithms and applications: A survey." *Ain Shams engineering journal* 5(4): 1093- 1113.

- [8] Pang, Bo, and Lillian Lee. (2008) "Opinion mining and sentiment analysis." *Foundations and Trends® in Information Retrieval* 2(1–2): 1-135.
- [9] Kolchyna, Olga, and et al. (2015) "Twitter sentiment analysis: Lexicon method, machine learning method and their combination." arXiv preprint arXiv: 1507.00955
- [10] Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. (2002) "Thumbs up? : sentiment classification using machine learning techniques." *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*. Association for Computational Linguistics, 10: 79-86.
- [11] Liu, Bing. (2012) "Sentiment analysis and opinion mining." *Synthesis lectures on human language technologies* 5(1): 1-167.
- [12] https://www.researchgate.net/publication/327097647_Scopus_Indexed_Journals.
- [13] https://www.scribd.com/document/483389700/ahuja_2018-pdf
- [14] Khattri, Anupam, and et al. (2015) "Your sentiment precedes you: Using an author’s historical tweets to predict sarcasm." *Proceedings of the 6th workshop on computational approaches to subjectivity, sentiment and social media analysis*. Lisboa, Portugal, 25-30.
- [15] Joshi, Aditya, et al. (2016) "Are word embedding based features useful for sarcasm detection?." *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, Texas, 1006–1011

A Study on Suicidal Rate Detection with social media using Deep Learning

Abhishek Singh¹, Akshay Khare¹, Anand Kumar Jha¹, Shikhar Kannoja¹, *Shimpy Goyal¹

Abstract - Social media platforms, such as Twitter and Reddit, can be used to identify individuals at risk of suicide with high accuracy using machine learning techniques. Researchers have developed systems that track words related to suicide and alert relevant organizations or individuals. The use of online data for suicide detection and prevention is an active area of research with the potential to save lives. Depression is also a major cause of suicide and some studies have focused on using machine learning to classify individuals at risk based on data from social media platforms.

Keywords: Algorithms, Artificial neural network (ANN), Support vector deep, Deep learning, Deep learning natural language processing, LSTM, CNN and Suicide detection.

INTRODUCTION

Suicide is a major public health concern and a leading cause of death worldwide. Online platforms, such as social media, have been studied as potential sources of data for detecting people with suicidal thoughts. These studies have shown that data from online platforms can be used to identify individuals at risk of suicide with high accuracy using machine learning techniques. Some researchers have developed systems that can track words related to suicide on social media platforms in order to identify individuals who may be at risk and alert relevant organizations or individuals. Depression is also a major cause of suicide and some studies have focused on using machine learning to classify individuals at risk based on data from social media platforms. In general, the use of online data for suicide detection and prevention is an active area of research with the potential to save lives.

LITERATURE REVIEW

According to Kasturi Dewi Varathan, Nur Hafizah Talib (2014) social networks such as Twitter are communication channels that allow users to disseminate information and activities and opinions through short texts. Suicide is a mental health problem which requires attention, controlling & preventing it is no easy task. It should be a weapon for quick detection. Twitter Suicide Detection is a system capable of identifying all words related to suicide on Twitter. The system also aims to detect suicides in time by listening to the coming tweets.

This system can be used by “NGOs” and psychologists to track suicidal individuals with the history of suicide attempts. This system is designed to help & prevent people from committing suicide.

Rohith Kumar Thiruvalluru, Manas Gaur, Krishnaprasad Thirunarayan say that suicide is 10th leading cause of death in US and the second leading cause of death among adolescents. Clinical and psychosocial factors contribute to suicide risk, but the documentation and self-report of these factors in EHRs and social networks varies. This study examines the degree of difference between EPA and social media. Using more than 13.8 million clinical records of 123,703 psychiatric patients, subjective analyzes of SRF, including “self-harm, bullying, impulsivity” and domestic violence/discord, were obtained. We grouped clinical notes using semantic integration according to a set of its SRFs. Similarly, we aggregated 2,180 suicidal users (~30,000 posts) on Suicide Watch for comparative analysis.

Mark Larsen, Jennifer Nicholas, Helen Christensen, (2016) suggest that the use of mobile health (mHealth) applications has increased rapidly to help, identify & support the people at risk of suicide. 123 suicide-related apps were identified and downloaded, and only 49 of them contained minimum one interactive feature of suicide prevention. 27 apps had focused on getting support from family & friends, 14 focused on safety planning, 13 were found to facilitate access to crisis assistance. Potentially dangerous content, such as inciting behavior during a crisis or listing lethal access to vehicles, has also been identified in the apps.

Xuening Wang et al. Says that depression is a usual mental illness which affects people of all ages and backgrounds. Depressed and depression-prone people often flock to online depression communities to meet the needs that otherwise could not meet. Such a community's impact is not

clear entirely due to the exposure of members to depression. Therefore, this study examined behavior when depressed people gather in the online community “Depression Super Topic” on Sina Weibo. rice field. Through website crawls, depression super topic posts were compared to the members normal timeline posts in terms of topics, published sentiments and number of likes & comments. The extracted posts topics were then coded to cover support, regulation, sharing emotions and life, and initiating discussion. A comparative analysis showed that posts in depression super-topic communities received more comments and revealed many emotions than regular timelines, and the members were more active now in the communities at night.

Scottye J. Cash, Jeffrey, A Bridge, Jeffrey A Bridge considers this review "a review of recent research on adolescent suicidal tendencies and suicidal behavior with a focus on epidemiological, psychiatric, psychological and environmental factors." Informed clinicians can make the most of this data to develop a more comprehensive understanding and assessment of suicide risk factors in adolescents, which can help guide targeted interventions that reduce the risk of deteriorating mental health and suicidal behavior. can be.

Marcel Adam Just¹, *Lisa Pan², Vladimir L. Cherkassky¹, Dana McMakin³, Kristin Cha,⁴ Matthew K. Knock⁵, and David Brent² suggest that clinical assessments of suicide risk may be substantially complemented by biological measures assessing changes in neural expression. said it would Concepts related to death and survival. from people with suicidal thoughts nearby. This study used machine learning algorithms (Gaussian Naïve Bayes) to identify these individuals (17 suicidal versus 17 controls) with high accuracy (91%), based on their fMRI-adjusted neural ratings of concepts related to life and death. The most discriminatory concepts are death, cruelty, trouble, recklessness, good, and praise. A similar classification correctly distinguishes (94%) 9 people with suicidal ideation and 8 people with no suicidal ideation. In addition, an important aspect of structural change is elicited emotion, where neural signatures serve as surrogate bases for accurate cluster classification (85%). The study established a biological and neurocognitive basis for altered conceptual representations in participants with suicidal ideation, allowing for group membership classification with high accuracy.

According to “S. Kumar, A.K Verma, S. Bhattacharya and S. Rathore” (2021),

Lucknow County Suicide Rates were collected for 5 years from January 2008 to October 2012 and disaggregated by sex and age group. The data show

that the rate in men is 56.61% and in women it is 43.38%. Suicides are concentrated among all ages of adolescents. The data also show that “the suicide tendency of men is higher than that of women of the same age. It is shown that suicide by poison is the most common method among men and women is suicide by poison”.

“Shaoxiong Ji” (2020) investigated “some factors such as prolonged exposure to negative emotions and life events that may lead to suicidal intentions and attempts. This study examines online social content for timely detection of suicidal ideation”. For the first time, it performed comprehensive content analysis to uncover knowledge of suicide texts and standard binary classifications of suicidal ideation, including the use of quote-based classifiers. feature export and deep neural networks. He considered the sentimental themes and signals contained in people's messages and proposed to establish a relationship between these factors and the message by using an attention relationship network to perform the detection. good idea of suicide. Finally, it tests the detection of suicidal ideation in another private chat situation. To address the challenge of sealed data in private chat rooms, he developed a knowledge transfer framework to shape a global knowledge sharing model using dispersion agent.

According to “Mitchell Weiss” et al. (2008) “Accurate information on preferred suicide methods is important for the development of suicide prevention strategies and programs. The aim of this study was to provide the first comprehensive look at international patterns of suicide methods”. An analysis of correspondence shows “a polarization between pesticide suicide and gun suicide, at the expense of traditional methods such as hanging from above and jumping between the two”. He shows “how suicide with drugs and guns is replacing traditional methods in many

countries”. They observed that “suicide patterns depended on the availability of the methods used. Available evidence indicates that limiting access to suicide drugs is more urgent and technically feasible than ever”.

According to “Benjamin Macharia et al. (2015)” “more than 400,000 people die by suicide each year. National, cultural, religious and social values play a role in suicide”. His 7-year retrospective study of all autopsies performed at “Moi Teaching and Referral Hospital in Eldoret, Kenya” was analyzed, and cases where identified as apparent suicide was investigated further. Data on age, sex, and method of suicide were collected and entered a data collection form. In summary, this study shows a marked difference in suicide patterns across the life cycle of the population of West Kenya. “Restricting access is means of suicide is an essential part of a comprehensive suicide prevention strategy”.

“Mojtaba Davoudi et al.” (2022) suggested that “natural disasters can affect the mental health of survivors and cause them to die by suicide”. A literature search was performed on some of the databases. Natural hazards include earthquakes, geomagnetic disturbances, droughts, dust storms, hurricanes, floods, and other natural disasters. Evidence supports increased suicide rates after disasters due to droughts, dust storms, hurricanes, and geomagnetic disturbances, but not floods. Based on these findings, they call for continued psychological research to pave the way for interventions to protect people's mental health and livelihoods following emergencies such as natural disasters.

According to “Manas Gaur et al.” (2019) “Mental illness, such as depression, is an important risk factor for suicidal thoughts, behavior, and intentions”. Their interdisciplinary study used Reddit as a separate data source to gather information about suicide and other related mental illnesses affecting users with depression. They provide detailed learning frameworks, including domainspecific knowledge to predict the severity of an individual's risk of suicide. They also used “language modeling, medical entity detection and normalization, and negativity detection to create a Redditor dataset of 2,181 people who discussed or suggested suicidal thoughts, behaviors, or attempts”. Due to the importance of knowledge, the benchmark “dataset of 500 Redditors (out of 2181)” had a pairwise annotation agreement of 0 as directed by the Columbia University, C-SSRS (Columbia Suicide Severity Rating Scale).

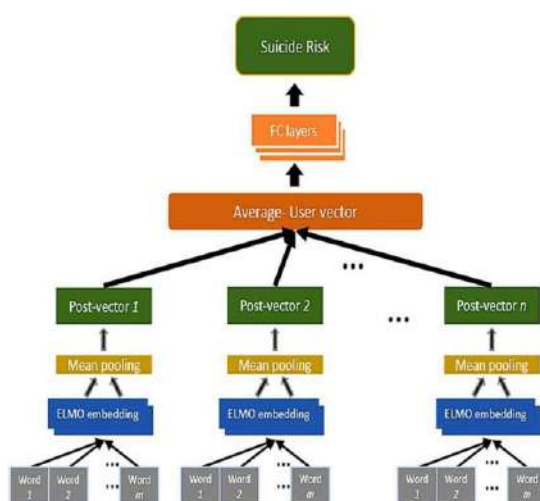


Fig.1: Suicidal Rate in Society Using DL-Deep Learning Flow.

According to “Shivangini Singh” (2021), suicide is a preventable tragedy and a major problem in developing countries like India, but it has been overlooked. The leading causes of suicide in India are hanging (53.6%), poisoning (25.8%), drowning (5.2%) and self-immolation (3.8%). A study by Jhansi concluded that the most common method of suicide among women was men's

self-immolation by running onto trains.

“Yujiro Kuroda et al.” (2021), “the suicide mortality rate in twelve cities Prefecture (Japan) designated as evacuation 2011 nuclear disaster the evolution of series model.” Whereas there in this study. First, some evacuees may have changed addresses after the disaster and the resulting data may make them untraceable, leading to an underestimation of suicide rates in affected areas. Second, due to the descriptive design of the published data, it was not possible to adjust for related factors such as presuicide economic and psychological variables. Third, because the study uses a two-year moving average model, the results of the age stratification analysis should be interpreted with caution.

Table 1: Details of the Techniques Used by Various Authors

S. No	Author Name	Title	Year	Database Used
1	Kasturi Dewi Varathan, Nur Hafizah Talib	Suicide detection system based on Twitter	2014	Twitter
2	Rohith Kumar Thiruvalluru, Manas Gaur, Krishnaprasad Thirunarayan	Comparing Suicide Risk Insights derived from Clinical and Social Media data	2021	Twitter, Reddit
3	Mark Larsen, Mark Larsen, Jennifer Nicholas, Helen Christensen	A Systematic Assessment of Smartphone Tools for Suicide Prevention	2016	123 Apps Downloaded and reviewed by two reviewers
4	Xuening Wang, Xianyun Tian, X. Pan, Dongxu Wei, Qi Qi	What Happens When People with Depression Gather Online	2021	Postings from Depression Super Topic community
5	Scott J. Cash, Jeffrey A. Bridge, Jeffrey A. Bridge	Epidemiology of Youth Suicide and Suicidal Behavior	2009	PubMed search for all English-language articles published between January 2007 and May 2009
6	S. Kumar, A. K. Verma, S. Bhattacharya, and S. Rathore	“Trends in Rates and Methods of Suicide in India,” Egyptian Journal of Forensic Sciences,	2021	Distribution of suicidal cases according to sex, age per year group, cases per year in Lucknow (2008–October 2012) NCRB Forensic Medicine & Toxicology Department, King George’s Medical University in the Lucknow

7	<u>Isabel De la Torre Díez</u> <u>Gema Castillo Jon</u> <u>Arambarri Basañez</u>	Mobile Apps for Suicide Prevention: Review of Virtual Stores and Literature	2017	Science Direct, Medline, PsycINFO, Embase, The Cochrane Library, IEEEXplore, and Google Scholar
8	<u>Shaoxiong Ji</u>	Suicidal Ideation Detection in Online Social Content	2020	Reddit, Twitter
9	Mitchell Weiss, Mariann Ring, <u>Urs Hepp</u>	Methods of suicide: International suicide patterns derived from the WHO mortality database	2008	WHO Mortality Database
10	<u>Benjamin Macharia, Iddah Maulid Ali, F. M. Ndiangui</u>	Pattern of Suicide: A Review of Autopsies Conducted at Moi Teaching and Referral Hospital in Eldoret Kenya	2015	MTRH autopsy record books
11	<u>Mojtaba Davoudi</u> , <u>Boroumand Alipour</u> , <u>Zahra Karimi Balouchi</u>	SUICIDES AFTER NATURAL DISASTERS: A REVIEW	2022	Web of Knowledge, PubMed and Scopus
12	<u>Manas Gaur</u> , <u>Amanuel Alambo</u> , <u>Joy Prakash Sain</u>	Knowledge-aware Assessment of Severity of Suicide Risk for Early Intervention	2019	Reddit
13	<u>Shivangini Singh</u>	Suicide in Uttar Pradesh: An Overview	2021	NCRB 2019
14	<u>Yujiro Kuroda</u> , <u>Masatsugu Orui</u> , <u>Arinobu Hori</u>	Trends in Suicide mortality in 10 years around The Great East Japan Earthquake: Analysis of evacuation and non-evacuation areas in Fukushima Prefecture	2021	Suicide rate numbers in Evacuation areas of Fukushima Prefecture (from March 2009 to December 2018)

III- CONCLUSION

In conclusion, the use of AI and machine learning in detecting and preventing suicides is a promising approach that has the potential to save lives. By analyzing data from social media platforms, AI can help identify individuals who may be at risk of suicide and provide them with appropriate support. For example, AI can be used to track words related to suicide on social media and alert relevant organizations or individuals. This can help identify individuals who may be at risk and provide them with the support they need before it is too late.

Additionally, AI can also help in detecting the causes of depression and other mental health issues, which can be used to prevent suicides from occurring. By analyzing data and patterns on social media, AI can identify common triggers or patterns that may lead to depression or suicidal thoughts. This information can be used to provide targeted interventions and support to individuals who may be at risk.

However, it is important to use AI and machine learning in this context with care and sensitivity. These technologies can have significant consequences if not used properly, and it is important to ensure that they are used ethically and in a way that respects the privacy and dignity of individuals. Additionally, it is important to recognize that AI and machine learning are not a panacea for preventing suicides, and that they should be used in conjunction with other interventions, such as education and awareness campaigns, to address this complex issue.

Overall, the use of AI and machine learning in the context of suicide detection and prevention is an important step towards addressing this global issue and providing support to those who may be at risk. By leveraging the power of these technologies, it is possible to identify individuals who may be at risk and provide them with the help and support they need, ultimately

References

1. K. D. Varathan and N. Talib, "Suicide detection system based on Twitter," 2014 Science and Information Conference, 2014, pp. 785-788, doi: 10.1109/SAI.2014.6918275.
2. Rohith Kumar, Thiruvalluru, Manas Gaur, Krishnaprasad Thirunarayan, "Comparing Suicide Risk Insights derived from Clinical and Social Media data" AMIA Jt Summits Transl Sci Proc. 2021: 364–373.
3. Larsen ME, Nicholas J, Christensen H (2016) A Systematic Assessment of Smartphone Tools for Suicide Prevention. PLoS ONE 11(4): e0152285. <https://doi.org/10.1371/journal.pone.0152285> .
4. Wang X, Tian X, Pan X, Wei D, Qi Q. What Happens When People with Depression Gather Online? Int J Environ Res Public Health. 2021 Aug 19;18(16):8762. doi: 10.3390/ijerph18168762. PMID: 34444519; PMCID: PMC8392513.
5. Scottye J. Cash Jeffrey, Bridge Jeffrey, Epidemiology of Youth Suicide and Suicidal Behavior
6. S. Kumar, A. K. Verma, S. Bhattacharya, and S. Rathore, "Trends in Rates and Methods of Suicide in India," Egyptian Journal of Forensic Sciences, vol. 3, no. 3, pp. 75–80, 2021.DOI: 10.1021/j.ejfs.2013.04.003
7. Isabel De la Tore DiexGema CastilloGema Castillo Jon Arambarri DasanezJon Arambarri Basanez, Mobile Apps for Suicide Prevention: Review of Virtual Stores and Literature JMIR Mhealth Uhealth. 2017 Oct; 5(10): e130.
8. Shaoxiong Ji, Suicidal Ideation Detection in Online Social Content, DOI:[10.13140/RG.2.2.19846.32328/1](https://doi.org/10.13140/RG.2.2.19846.32328/1)
9. Mitchell Weiss, Mariann Ring, Urs Hepp, Methods of suicide: International suicide patterns derived from the WHO mortality database, Bull World Health Organ.Bull World Health Organ. 2008 Sep; 86(9): 726– 732. doi: [10.2471/BLT.07.043489](https://doi.org/10.2471/BLT.07.043489)doi: [10.2471/BLT.07.043489](https://doi.org/10.2471/BLT.07.043489)
10. Benjamin Macharia, Iddah Maulid Ali, F. M. Ndiangui, Pattern of Suicide: A Review of Autopsies Conducted at Moi Teaching and Referral Hospital in Eldoret Kenya, June 2015, The Open Access Journal of Science and Technology, DOI:[10.11131/2015/101112](https://doi.org/10.11131/2015/101112)

11. Davoudi, Mojtaba & Alipour, Boroumand & Balouchi, Zahra & Saghi, Masoumeh & Lester, David. (2022). Suicide Studies, 2022, 3(4), 31-46 SUICIDES AFTER NATURAL DISASTERS: A REVIEW. 3. 31-46.
12. Manas Gaur, Amanuel Alambo, Joy Prakash Sain, “Knowledge-aware Assessment of Severity of Suicide Risk for Early Intervention” Companion Proceedings of The 2019 World Wide Web Conference, 2019, pages 514-525. Shivangini Singh, “Suicide in Uttar Pradesh : An Overview”, Indian Journal of Clinical Psychiatry(2021) <https://www.ipsupsb.com/Pdf/suicide-in-uttarpradesh.pdf>
13. Shivangini Singh, “Suicide in Uttar Pradesh : An Overview”, Indian Journal of Clinical Psychiatry(2021) <https://www.ipsupsb.com/Pdf/suicide-in-uttarpradesh.pdf>
14. Yujiro Kuroda, Masatsugu Orui, Arinobu Hori, “Trends in Suicide Mortality in 10 Years around the Great East Japan Earthquake: Analysis of Evacuation and Non-Evacuation Areas in Fukushima Prefecture”, June 2021, International Journal of Environmental Research and Public Health 18(11):6005, DOI: [10.3390/ijerph18116005](https://doi.org/10.3390/ijerph18116005)

Development of Student Interaction Platform

Suman Arora
CSE Department
Bhagwan Parshuram Institute of
Technology
Delhi, India
sumanarora@bpitindia.com

Siddhant Chadha
CSE Department
Bhagwan Parshuram Institute of
Technology
Noida, India
sid.01chadha@gmail.com

Tushar Jain
CSE Department
Bhagwan Parshuram Institute
of Technology
Delhi, India
tjain210@gmail.com

Abstract—Higher education, especially in the field of computer science and information technology, requires continuous learning, as there is always something new to learn. Web development is an excellent example of this. The students have experienced that sometimes it can be difficult to keep up with the pace. Learning new things can be burdensome, so this paper proposes an idea to help our college students cooperate and connect based on shared domain interests. Students usually want to learn something new that will help them progress in their careers. This paper proposes a solution to address the issues of current students by creating a platform where student communities would aid an institution in growing with their talent. Students can use the proposed platform, "TechBPIT" to learn, share their skills, collaborate, and design projects. Through collaborative learning, interaction with peers and teachers, students shape their perception of social media and mobile devices, and it has a significant impact on student's academic performance. A user can register or log in to their timeline and communicate with other users. Users can communicate, post, share, and update others on activities and events that take place online or in the real world. Users can also join a specific group (SIG)

based on their interests and learn about the latest technology. Mentors can also manage community events to develop students' professional skills.

Keywords—Chat, Multiple-user, NodeJS, Socket.io, Student Interaction, Engagement, Learning, Communities, Efficient

I. INTRODUCTION

In today's globalized era, skill development is the necessity for self-employment and entrepreneurship. Indeed, it is the need of the hour for the economic development of the 21st century. And, therefore, there is a strong need to provide skill-based learning in higher education to help students receive vocational training besides academics to improve their chances of better employment and a successful career. TechBPIT is a social network that connects students in our college and turns them into a group of friends. Users can get in touch with like-minded people based on their interests and hobbies. They also get to know about what their peers are learning/working on and can join them in group projects. Students can also join specific groups/communities according to your interests to learn about new tech. Mentors can also conduct community sessions to enhance the professional abilities of the students. The following sections first seek to place our research in the current landscape of theoretical and relevant research. Secondly, we present the design and implementation of our platform, and then conclude and provide the perspectives of future work aimed at inspiring designers seeking to design similar applications

II. LITERATURE REVIEW

[1] Suggested solution for managing alumni is a social networking site that enables users to easily share any type of information with other users. Using a registration form, they collect the users' fundamental data, after which login details were given. This management system offers a network that links Institute students in a single location, enabling them to communicate, exchange concepts, and access other data.

[2] To investigate the application and utility of social media and mobile devices in transmitting materials and interacting with academicians across the border in higher education institutions. It is based on a study of 360 students from an institution of higher learning in eastern India, which sought to understand how students felt about social media and mobile devices through group projects, peer interaction, and instructor interaction. A latent multivariate technique called variance-based structural equation model (SEM), which provides a concurrent estimate of structural and measurement models that do not meet parametric assumptions, was used to measure and evaluate the proposed model of the study. The confirmatory factor analysis (CFA) was carried out to determine whether the widely accepted criterion of discriminant and convergent validity was satisfied. It concludes that using social media for collaborative learning has a substantial positive correlation with online knowledge sharing behaviour, implying that university students who use social media more frequently for collaborative learning will share more knowledge with peers and colleagues.

[3] A P2P-based mobile app for collaborative learning in higher education was designed and put into use. They began by compiling the requirements for the mobile application and figuring out all of its features. Unified Modeling Language (UML) is employed for this step. Then, the Android mobile platform, Java and XML for programming languages, and Firebase for data administration are utilized in the development of the mobile app. The interfaces created by the mobile app's implementation were then displayed. 40 participants who were undergraduate students participated in the testing of the application. The findings of the study show that the majority of students are enthusiastic about peer education. Only a small percentage of students are not interested in peer learning, and the suggested mobile app can help students perform better academically and collaborate more effectively.

[4] A platform that makes it possible and guarantees better interaction between the facilities, the students, and the outside world, for which the students must be prepared. By putting students in touch with faculty members and business leaders directly, this platform aims to close the gap between expectations and reality. It is a web-based platform with tools for facilitating student discussion, enthusiastic question-asking, thought-expression, community building, fun, educational event planning, and off-course gatherings. It makes use of independent micro services so that if one of them fails, it won't bring down the entire platform. Each micro service has a model, a database, and a controller of its own. It attempts to manage and inform students about forthcoming events on their campus that they might attend, as well as a job site that aids in helping them locate pertinent employment, internships, and gigs.

[5] To increase campus-wide student connections and provide a platform that will make this possible. They develop a system that allows users to access message boards, Q&A forums, event alerts, and notes in order to encourage student cooperation and involvement. Users will have easy access to strategic materials, be able to participate in club activities online, create groups with their classmates, and feel prepared. Users will benefit from this when they have general inquiries about things like placement preparation, discussion of cutting-edge technology, exchange of business ideas, team building, etc.

[6] The chat app provides a better and more flexible chat system. Developed with the latest technology in the way of providing a reliable system. The main advantage of the system is instant messaging, real-world communication, added security, group chat, etc. This application may find the best demand in the market for most organizations that aim to have independent applications.

[7] Chat operation is a point or a program on the Internet to communicate directly among Internet users who are online or who are inversely using the internet. This allows users to communicate indeed though from a great distance. Thus, this application must be real-time and multiplatform to be used by numerous users. Application manufacture begins with the collection of applicable data that will be displayed in the web interpretation.

[8] The basic implementation for server – client communication for real time chat app. It shows that with the help of the chat application we can send the message to many users at a time, and we don't need to call every time to anyone for any type of queries or problem, we can just message them.

III. SYSTEM DESIGN

Social networking websites are frequently used by people. It provides a more efficient and useful method of communication. Our recommended Student Interaction System features a social networking platform that enables users to easily exchange any type of information with other users. This type of platform requires both client- and server-side development. The platform's client side will be built on Android, and its server side will use Node.js. The server uses Web Sockets for the full duplex paradigm, which is used to integrate real-time chat, and REST APIs for the requestresponse approach. All application data is stored in MongoDB, and passwords are encrypted using the Bcrypt library, protecting data from hackers and ensuring security. A onetime password (OTP) sent to the user's registered email will be used to verify their email address. A JSON web token stored on the user's device is delivered in the header of the request to authorize user requests. The Android application will be using Retrofit to make network calls.

A. METHODOLOGY

Rapid Application Development (RAD) is the software methodology that will be used in developing the platform. Various research papers were examined before finalizing technologies for implementation of various features. The aim is to reduce risks and costs by enlightening understanding of the proposed arrangements before further assets are submitted. All the steps necessary for the development will be completed correctly to attain a complete working system. A newly developed student interaction platform will be the outcome. Depending on the requirements, various features shall be added later.

B. ADVANTAGES

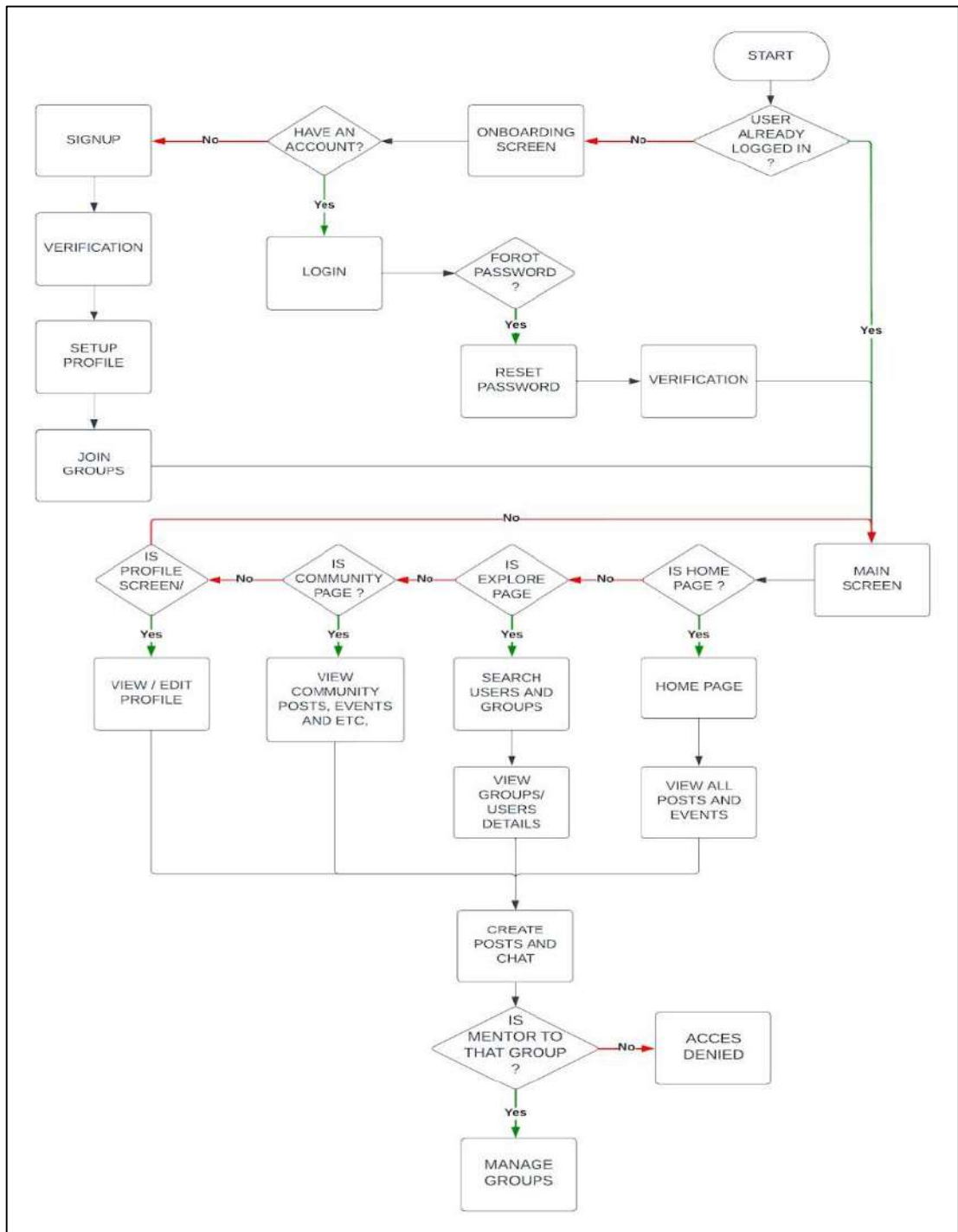


Figure 1: System Flowchart

C. LIST OF MODULES

1) Onboarding

This is the very first page visible when someone enters the platform. It contains the option to sign up for new users and login.

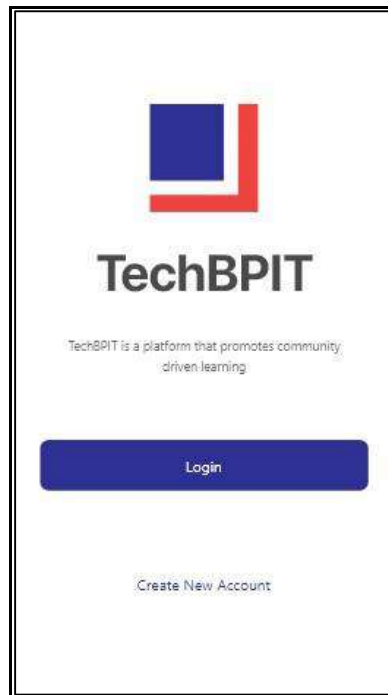


Figure 2: Login Page

2) Signup

The signup page allows the user to register on the platform by entering the required information. If all the entered information is correct, the user is registered and the information gets saved in the database else the error message is displayed. It then directs the user to the verification screen.

3) OTP Verification

OTP Authentication module provides a two-layer authentication feature and adds an extra layer of security. A One Time Password is received by the user on the entered email and after verifying it the user can fully access the platform.

4) Login

On the login page, the user is required to enter the email and password. The password is encrypted and hence the data is highly secure.

5) Forgot Password

If the user forgot the login credentials they can use this feature to recreate their password. The system will use the mail ID provided by the user to send the password reset link. The password is updated in the database.

6) Setup Profile

After registering, the user needs to set up his or her public profile by entering necessary details. This will help other users on the platform to discover and know about them.

7) Join groups

Users can join groups they are interested in which will make them part of the community and then they can chat with people in the community.

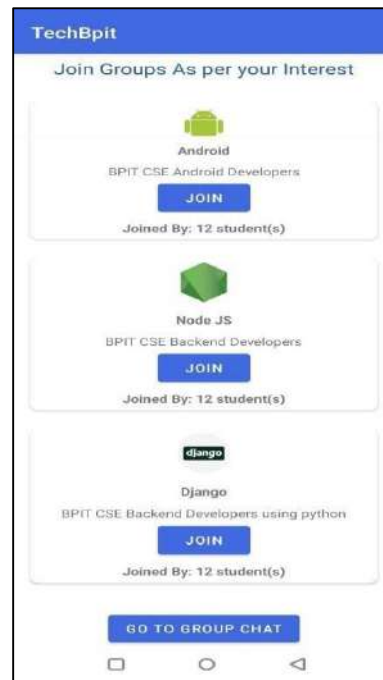


Figure 3: Join Groups Screen

8) Home Page

Home page is what the user sees when they login. It includes user feeds, the constantly updating list of posts from peers, groups you're in and events scheduled happening around the campus.

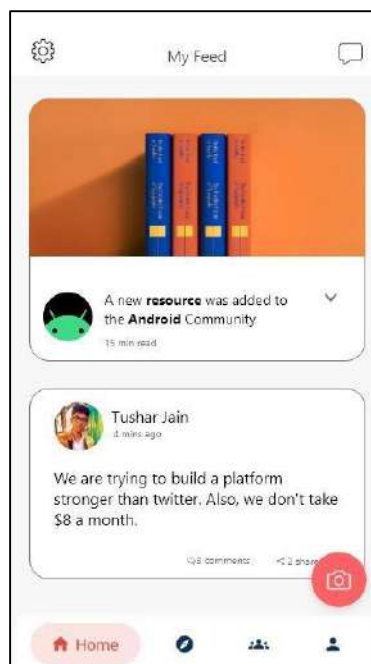


Figure 4: Home Screen

9) Create Post

This allows users to create and upload posts on the platform. While creating posts user can add images, text, links and tags to make posts more appealing.

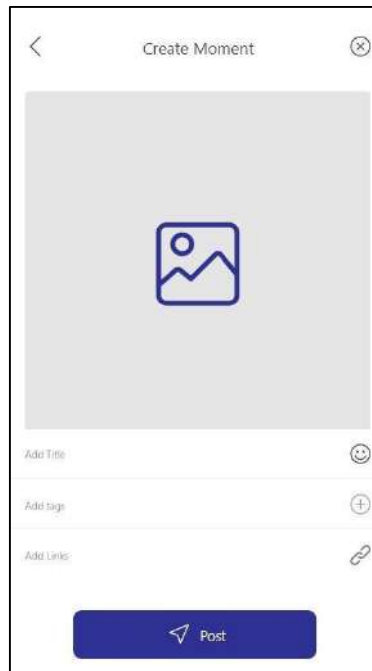


Figure 5: Create Post Page

10) Community Posts

Section specific to see updates, posts and modules posted by groups user follows.

11) Search Users, Groups and Posts

The user can search for other users and groups by typing their names in the search bar. Users can search any post with the help of keywords. The user will be displayed with multiple posts from the database.

12) Profile

User profile is a professional landing page for users to manage their personal brand. It's a great way for users to tell people who they are and what they do by displaying a general history of professional experiences and achievements. Users can also add their other social links on the platform to increase their social network.

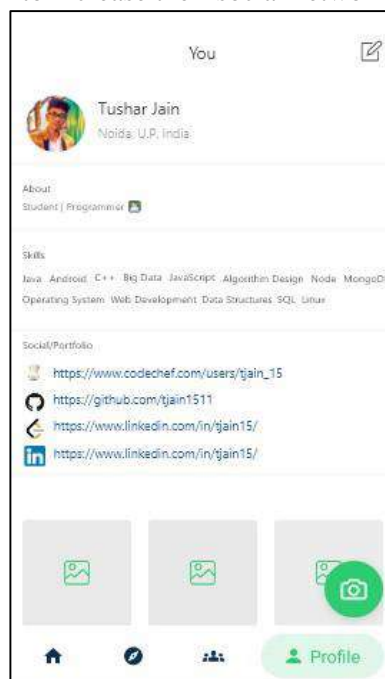


Figure 6: Profile Page

13) Manage groups and community

Users assigned as mentors to specific groups or communities can manage the functioning of the groups. They can post study materials, event posts and updates specific to the community they are mentoring.

14) Create Event Post

Mentor can create posts about an event or any session that is happening in the community to provide information about the events like when the event is happening, where it is happening, topic of event, who is the speaker and whether it is online or offline.

15) Personal Chat

Users registered on the platform can send direct messages to each other. Personal chat is a very important feature as this helps to bridge the gap between other websites and a social networking site.

16) Group Chat

Group chat is similar to chat features in that it gives you the ability to send instant messages to others in a community. More specifically, group chat is the ability to chat with multiple people. It structures teambased conversations without slowing the flow of information.

IV. FUTURE SCOPE

Future work may include adding the ability to save user posts for later use. Events can eventually be connected to the user's calendar to make reminders easier. The ability to customize and change the application's theme according to user preferences. System can be extended to connect alumni and students. It can be used to manage internships and career opportunities

v. CONCLUSION

Nowadays, students use various social networking sites and platforms. It provides a more efficient method for interaction. This paper focused on the design and implementation of a student interaction platform that improves student's interest in learning and promotes communication between students and mentors. It will help students gain the necessary skills required to boost their professional journeys and foster collaboration with other.

REFERENCES

- [1] Eileen Mona, Hari Ranjan M, Sowmya V, Mrs. S. Sivakumari, “Alumni Social Networking Site”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology.
- [2] Jamal Abdul Nasir Ansari, Nawab Ali Khan. “Exploring the role of social media in collaborative learning the new domain of learning”, Smart Learning Environment Journal
- [3] Asmae Blilat, Abdelali Ibriz, “Design and Implementation of P2P Based Mobile App for Collaborative Learning in Higher Education”, International Journal of Interactive Mobile Technologies
- [4] Jaydev Sharma, Mruthunjay Gurav, Himanshu Singh, Shaina Gupta, “Milo - A College Connect Platform”, International Journal Of Creative Research Thoughts
- [5] Siddharth Jadhav, Sachin Gunjal Sanket Jadhav, Sakshi Hole Prof. Nivedita Rawate. “COLLEGE COMMUNE’S - PLATFORM FOR STUDENT INTERACTION AND COLLABORATION”, Journal of Emerging Technologies and Innovative Research

- [6] Dr. Abhay Kasetwar, Ritik Gajbhiye, Gopal Papewar, Rohan Nikhare, Priya Warade. “Development of Chat Application”,International Journal for Research in Applied Science & Engineering Technology
- [7] Nita Thakare, Nitin Deshmukh, Anshul Vairagade, Ayush Nagarare, Himanshu Kamane, Rajat Mohod. “Real Time Chatting Web Application”,International Journal of Engineering Science and Computing
- [8] Rahul Khandelwal, Dishant Solanki, Teena Verma. “DESIGN & IMPLEMENTATION OF REAL TIME CHAT APPLICATION”,International Research Journal of Modernization in Engineering Technology and Science