

2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT 2021)

**Bhopal, India
18 – 19 June 2021**

Pages 1-447



**IEEE Catalog Number: CFP2118P-POD
ISBN: 978-1-6654-4624-2**

**Copyright © 2021 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP2118P-POD
ISBN (Print-On-Demand):	978-1-6654-4624-2
ISBN (Online):	978-1-6654-2306-9
ISSN:	2329-7182

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

2021 10th IEEE International Conference on Communication Systems and Network Technologies

CSNT 2021

Table of Contents

Welcome from IEEE CSNT 2021 General Chairs

Welcome from IEEE CSNT 2021 Program Chairs

Conference Committees

Track-A: Microwave Devices, Wave Propagation and MMIC

28 GHz inset feed circular shaped compact patch antenna array for 5G wireless communication.....	01
<i>Rohit Yadav, Leeladhar Malviya, Dhiraj Nitnaware and Ajay Parmar</i>	
A Compact Millimeter Wave 5G Patch Antenna For n-260 Band Frequency Range Applications	05
<i>Jai Mangal and Ajay Anil Kumar.</i>	
A Single Band E Shaped Microstrip Patch Antenna For WLAN Application using Coaxial Feed.....	10
<i>Arpita Vishwakarma and Devendra Kumar Meda.</i>	
Compact Dual-band Umbrella Shape slotted Patch Antenna for Wireless Applications.....	14
<i>Nidhi Sharma and Vikas Soni.</i>	
Compact Wideband Dual Polarized Circular Slot Antenna for ISM Band Applications.....	19
<i>Toshika Chand, Saurabh Singh and Sudhanshu Verma</i>	
Design, Testing, and Analysis of a Third Order Bandpass Filter For Ku-Band Applications.....	23
<i>Rashmita Mishra; Kailash Chandra Rout; Subhrajit Pradhan</i>	
MIMO-Array Antenna with Beamforming for 5G Applications	27
<i>Leevanshi Rao, Leeladhar Malviya, M. P. S. Chawla and Ajay Parmar</i>	
Parametric Study of Microwave Metamaterial Inspired High Sensitivity Sensor.....	33
<i>Divya Prakash and Nisha Gupta.</i>	
Polarization Controlled Ultrathin Quad Band Metamaterial Absorber for C and X-Band.....	38
<i>Gaurav Chaitanya, Abhay Suroliya, Jitendra Neema and Ankit Chandachoriya.</i>	
Single Slit Compact Patch Antenna with sectional Ground Plane for Bluetooth and WLAN 802.11b Application.....	45
<i>Akanksha Gupta; Nitesh Kahsyap; Ramesh kumar Sunkaria</i>	
A Compact and Wideband Filtenna Using Elliptical Patch and CSRR Structure for Wireless Application	50
<i>Bhawna Sharma, Anita Rani, Nitin Saluja, Geetanjali Singla, Dhawan Singh</i>	

Track-B: Computer Networking, Wireless Communication, Ad Hoc Networking and Vehicular Networking

An assessment of various clustering techniques for VANET	56
<i>Atul Barve and Pushpinder Singh Patheja</i>	
Attitude, Position and Velocity determination using Low-cost Inertial Measurement Unit for Global Navigation Satellite System outages.....	61
<i>Jyothirmai Vigrahala and B. Naresh Kumar Reddy.</i>	
HRSUD: Hybrid Roadside Unit Deployment System to Improve Vehicular Communication in VANET	66
<i>Prashant Kumar Shrivastava, Megha Kamble , Mayank Sharma,</i>	
<i>Priya Shrivastava, Ranjeet Singh Tomar</i>	
Impact of Varying Physical Links on QoS of Virtual Topology Management in WDM Network	74
<i>Pushpendra Singh Tomar, Sunil Phulre and Manish Shrivastava.</i>	
Modeling and Simulation of Vehicular Traffic at Intersection of Roads.....	81
<i>Hrishikesh Ugale, Pushpak Patil, Shubham Chauhan and Neeraj Rao.</i>	
A Geocast Based Routing for Cooperative Video Streaming Over Vehicular Networks.....	89
<i>K Shanoja, C. Ismayil Siyad and Renjith V Ravi.</i>	
Multi-Objective Emperor Penguin Optimizer for Tuning the Quality-of-Service in Cloud Computing	95
<i>Harsimran Kaur and Robin Singh Bhaduria</i>	
Quantized Directional Cosine measurements based Localization	99
<i>Balarami Reddy B N, Srinath Gunnery, Pardhasaradhi Bethi and Srihari Pathipati</i>	
Implementation of consensus with TCP/IP	105
<i>Parshal Chitrakar and Priya Ranjan</i>	

Track C: Data Mining, Big Data, Text & web mining and Social Networking

A Primitive Solution to Avoid Data Deduplication in Cloud and Overcome Security Threats....	113
<i>Y Bevish Jinila, Jithina Jose, J.S. Vimali, P. Ajitha, S. Gowri, Sivasangari</i>	
A Review on Automatic COVID-19 Lung Lesions Detection from Tomographical Image	118
<i>Jitendra Agrawal, Nivedita Rawat and Yogendra P S Maravi</i>	
A Statistical Analysis on Website Quality For Purchase Intention during Online Shopping	125
<i>Mohit Mittal; Shaifali Chauhan; Swadha Gupta; H. Singh</i>	
An Efficient method for Aspect based Sentiment Analysis Using SpaCy and vader	130
<i>Akhilesh Singh and Ananya Verma</i>	
COVID-19 Tweet Analysis using Hybrid Keyword Extraction Approach	136
<i>Rajni Jindal, Mansi Garg, Sakshi Vatsa and Surbhi Mathur.</i>	
Comprehensive Analysis of Feature Selection Techniques on Early Heart Stroke Prediction....	142
<i>Neeraj Sharma, Pratyush Sethi, Jasroop Singh Chandha and Praveen Lalwani</i>	
EEG Data Analysis for Stress Detection	148
<i>Lokesh Malviya, Sandip Mal and Praveen Lalwani.</i>	
Data Mining Techniques for Breast Cancer Detection: A Survey.....	153
<i>Surender Singh and Saruchi Kukkar</i>	
Evaluating the semantic property in an E-Learning ontology	159
<i>Amitkumar Bajpai, Rajiv Pandey and S.P. Tripathi</i>	

Medical Image Enhancement with Brightness Preserving Based on Local Contrast Stretching and Global Dynamic Histogram Equalization.....	164
<i>Vijay Panse and Rajendra Gupta</i>	
Naive and Neighbour Approach for Phishing Detection.....	171
<i>Naman Bhoj, Rakshika Bawari, Ashutosh Tripathi, and Neil Sahai.</i>	
Marker-Less Augmented Reality System for Home Interior and Designing.....	176
<i>Prashant Johri, Gaurav Dhuriya, Sudeept Singh Yadav and Swati Chauhan</i>	
Performance Analysis of Single Image Fog Expulsion Techniques.....	182
<i>Gaurav Saxena, Sarita Singh Bhadauria and Subodh Kumar Singh.</i>	
Optimal K-Means Clustering Algorithm For Weblog Mining.....	188
<i>Vipin Jain and Kanchan Lata Kashyap.</i>	
Sentiment analysis of Twitter data by making use of SVM, Random Forest and Decision Tree algorithm.....	193
<i>Jyotsna Singh, Pradeep Tripathi</i>	
Studying the Role of Data Quality on Statistical and Neural Machine.....	199
<i>Karunesh Kumar Arora, Geetam Singh Tomar, SS Agrawal</i>	
Text Mining Techniques for Knowledge of Defects in Power Equipment.....	205
<i>Shuaisong Yang, Jingdong Wang, Fanqi Meng and Shuqiang Guo.</i>	
Data Mining based Prediction of Diabetes using Firefly Optimized Neural Network.....	211
<i>Vivek Vaidya and L K Vishwamitra</i>	
Web User's Access Pattern Identification Using Clustering Algorithms: A Survey.....	220
<i>Preerna Kumari, Aneesh Mishra and Rahul Kumar</i>	
An Implementation of K-Means Clustering for Efficient Image Segmentation.....	224
<i>K. Venkatachalam, M. Amudhan, V. Purandhar Reddy, A. Raguraman, E. Mohan</i>	

Track D: Digital System Design, Fabrication and Robotics

Error Tolerant MBIST Design with Efficient High Throuput and Low Power Reversible Techniques.....	230
<i>S. Sambasiva Rao Dannina, U V Ratna Kumari</i>	
Design of Thermal-Aware and Power-Efficient LFSR on Different Nanometer Technology FPGA for Green Communication.....	236
<i>Keshav Kumar, Shivani Malhotra, Rubina Dutta and Amit Kumar.</i>	
Effect of Gate Metal Work Function on Leakage Current in Single Pocket FDSOI 28 nm Transistor.....	241
<i>Aryan Kannaujiya, Sandhya Kannaujiya and Rajeev Kumar Chauhan.</i>	
Efficient Realization of 1:2 Demultiplexer Using QCA.....	247
<i>Bhawna Sharma, Anita Rani, Nitin Saluja, Geetanjali Singla and Amandeep Kaur</i>	
Energy-Efficient Design Of RAM And It's Implementation On FPGA.....	252
<i>Bishwajeet Pandey, Amrit Ratnakar, Deepanshu Joshi, Amit Kumar Gupta</i>	
FPGA Implementation of Power Efficient Floating Pont Fused Multiply Add Unit.....	258
<i>Mounika Kajjam and Venkataramana P.</i>	
Noise margin enhancement of Conventional 6T SRAM Cell by aspect Ratio Optimization.....	264
<i>Ashish Sachdeva and V.K. Tomar.</i>	
VLSI Implemetation of Euler Number Computation and Stereo Vision Concept for CORDIC based Image Registration.....	269
<i>K Pradeep, Neenavath Veeraiah</i>	

Verilog based Implementation of Real Time Digital Video Stabilization.....	273
<i>Mohammed Ahmed and Laxmi Singh</i>	

Track E: Electrical Power System Design, Grid and Solar Devices

An Artificial Neural Network Based Forecasting Strategy for Estimating Weather Parameters: Application for Sizing Stand-alone Renewable Power System.....	279
<i>Naveen Kumar Dubey, M.P.S. Chawla and Leeladhar Malviya</i>	
Simulation Study of Five-Phase Dual Two Level Inverter System using Unified Space Vector PWM.....	285
<i>Shailesh Kumar Gupta, Md. Arif Khan, Omveer Singh and D.K. Chauhan</i>	
Contribution to Power Quality Improving the DFIG Control Using a Reduced Components of Multi-level Inverter.....	291
<i>Kalyan Chatterjee and Abhishek Kumar.</i>	
Detection of Insulators by Point Feature Matching Using Speeded Up Robust Features.....	295
<i>Ayushi Jadia and M.P.S Chawla.</i>	
Economic Load Dispatch Using Jellyfish Search Optimizer.....	301
<i>Himanshu Rai and H. K. Verma</i>	
Events Recognition and Power Quality Estimation in Distribution Network in the Presence of Solar PV Generation.....	305
<i>Rajkumar Kaushik, Om Prakash Mahela and Pramod Kumar Bhatt.</i>	
Load frequency control of a photovoltaic-pumped hydro power energy storage based micro-grid System	312
<i>Rebecca Lalparmawii, Subir Datta, Subhasish Deb and Sumana Das.</i>	
Predictive Maintenance of Machines and Industrial Equipment.....	318
<i>K. Purnachand, Md. Shabbeer, P.N.V. Syamala Rao M, Ch. Madhu Babu</i>	
S-LIGHT: Smart LED Lamppost using PWM-based Adaptive Light Controller.....	325
<i>Sara AlRuwais, Rawabi AlQahtani, Nuwayyir AlHajri, Batool AlHashim, Abul Bashar, Loay AlZubaidi</i>	

Track F: Image Processing, Signal Processing and Applications

A Hybrid Unsupervised Feature Selection Algorithm	332
<i>Rana Pratap Singh, Kuldeep Singh Jadon</i>	
A Review on Human Action Recognition Approaches.....	338
<i>Suresh Chand Gupta, Deepak Kumar and Vijay Anant Athavale.</i>	
A Survey on recent techniques in face recognition for various databases.....	345
<i>Vijaya Kumar H R, Mathivanan M.</i>	
An Ensemble Approach On Scientific Paper Reviews for Sentiment Analysis.....	350
<i>Divyank Ojha, Kuldeep Singh Jadon</i>	
An Improved Wrapper based Feature Selection using Hybrid BFO for Multimodal.....	358
<i>N Krishnaraj, Y. Baby Kalpana, Meenakshi Shunmugam</i>	
Automatic Diabetic Retinopathy Detection using Roberts Cross Edge Detection in DIP.....	363
<i>Anshu Sharma, Ruchi Dronawat and Arun Jhapate</i>	

COVID-19 Detection from Chest X-Rays and CT Scans using Dilated Convolutional Neural Networks.....	369
<i>Shailender Kumar, Kumar Abhishek, Kritika and Kashish Singh.</i>	
Designing and Evaluation of the Reduced Order IIR Filter Design for Signal De-Noising.....	375
<i>Nitin Jain, Shanti Rathore and S.K Singh.</i>	
Evaluation of Feature Extractors for Sign Language Recognition.....	381
<i>Nitish Kumar, Himanshu Shrivastava and Nitin Rakesh.</i>	
Optimized Color Correlogram based CBIR to beat the hole of Information and Interpretation.....	387
<i>Punit Soni, Vijay Lamba and Surender Kumar</i>	
Periocular Recognition using CNN based Feature Extraction and Classification.....	393
<i>Vedant Kayande, Pratyush Kaware, Aditya Khambete, Sukanya Kulkarni</i>	
Mycobacterium Tuberculosis using Directional Embossing and SVM from CT Scan Images.....	408
<i>Akanksha Soni and Avinash Rai. Satish Kumar Ahirwar</i>	
Study on Various Collaborative Filtering Techniques to Recommend Movies.....	414
<i>Megha Sahu, Khetwat Saritha</i>	
Speech Emotion Recognition System With Librosa.....	421
<i>P. Ashok Babu, V. Siva Nagaraju and Rajeev Ratna Vallabhuni</i>	
Face Mask Detection using CNN.....	426
<i>G. Saranya, K.Kumaran N.Ananthi, Dipshikha Sarkar, Sayan Ghosh, Lokenath Basu</i>	
Maritime Ship Detection using Convolutional Neural Networks from Satellite Images.....	432
<i>Jaafar Alghazo, Abul Bashar, Ghazanfar Latif, Mohammed Zikria</i>	
Paddy Blast Detection using Color Co-occurrence Features and Support Vector Machine Classifier.....	438
<i>Senthil Kumar K, S.Varalakshmi, Hemavathy N, E. Sivanantham. E.Mohan</i>	

Track -G: Internet of Things and Sensor Networks

Applied Internet of Things in Saudi Arabia Aquaculture System	442
<i>Ekhlas Alshelaly, Noha Alswaigh, Amani Samha and Ghalib Alshammri.</i>	
Architectural framework for Multi sensor Data fusion and validation in IoT Based system.....	448
<i>Sindhu S and Saravanan M.</i>	
Development of Cooperative Adaptive Cruise Control with multiple predecessors information and CarSim Validation	453
<i>Jaswandi Sawant and Uttam Chaskar.</i>	
Efficient Secret Key Generation for Internet of Things Communication Device Using Discrete Wavelet Packet Transform	459
<i>Nikhil Patankar, Parmalik Kumar, Rajneesh Karan and Animesh Dubey.</i>	
Analysis of SDN based IIoT networks Targeting Automation Processes in Smart Factory Environments	465
<i>Likhitha Emmadi, Raja VaraPrasad, Hrishikesh Venkataraman</i>	
IoT Based Real Time Electrical Parameter Monitoring and Control	473
<i>Jigyasa Singhai and Arun Parakh.</i>	

IOT based Soil Nutrition and Plant Disease detection system for smart agriculture	478
<i>Sushant Suhag, Nidhi Singh, Ayush Shukla, Prashant Johri, Sanskriti Jaduan, Nidhi Parashar</i>	
Prevent Data Counterfeiting among Internet of Things (IoT) devices using Ethereum Blockchain	484
<i>Chetan Bawankar, Sayali Kukkar, Vaishnavi Chaudhari, Suraj Rashinkar and Sanket Solanke.</i>	
Smart Health Monitoring During Pandemic Using Internet of Things	489
<i>Shardeep Kaur Sooch and Darpan Anand.</i>	
Swarm Technology based Self-Deployment with ACO for Wireless Sensor Network in the presence of Obstacles	494
<i>Uppalapati Srilakshmi.</i>	
Various Energy-Saving Approaches in Wireless Sensor Networks: An Overview	499
<i>Prasanta Pratim Bairagi and Mala Dutta.</i>	
IoT based Novel Hydration System for Smart Agriculture	505
<i>Sai Likhita Vunnava, Sri Chandana Yendluri, Sateeshkrishna Dhuli</i>	

Track -H: Machine Learning, Deep Learning and Soft Computing

A Deep Convolutional Neural Network For Detection of Malaria Parasite in Thin Blood Smear Images	510
<i>Milind Raj, Rohan Sharma and Deepak Sain</i>	
A Neural Network based Diabetes Prediction on Imbalanced Data	515
<i>Shivani Yadav, Yogendra P.S. Maravi, Jitendra Agrawal and Nishchol Mishra</i>	
A Stacked Long Short-Term Memory Neural Networks for Parking Occupancy Rate Prediction	522
<i>Anu Jose and Vidya Ubalanka.</i>	
An Adaptive Machine Learning-based Threat Detection Framework for Industrial Communication Networks	527
<i>Sujit Bebortha and Sumanta Kumar Singh.</i>	
An Analysis on Data Mining for Sentiment Analysis Using Different Classification Algorithms.....	533
<i>Divyank Ojha, Rana Pratap Singh, Kuldeep Singh Jadon</i>	
An Efficient Feature Selection Approach using Sensitivity Analysis for Machine Learning based Heart Disease Classification	539
<i>Saranya G and Pravin A</i>	
An Intelligent System to Forecast COVID-19 Pandemic using Hybrid Neural Network.....	543
<i>Supriya Vanahalli and Preethi Nanjundan.</i>	
Analyzing Computational Response and Performance of Deep Convolution Neural Network for Plant Disease Classification using Plant Leave Dataset	549
<i>Lalit Kumar and Dushyant Kumar Singh.</i>	
Classification of Diseased Potato Leaves Using Machine Learning	554
<i>Sakshi Sharma, Vatsala Anand and Swati Singh.</i>	
Content Based Image Retrieval using deep learning.....	560
<i>Manjula and Sanjay Kumar</i>	
Deep Learning based Forecasting of Consumption of Petroleum Products in India	565
<i>Kashish Wattal and Sandeep Kumar Singh.</i>	
Depression Detection and Analysis Using Deep Learning: Study and Comparative Analysis.....	570
<i>Siddharth Prabhudesai, Apurva Mhaske, Manvi Parmar and Sumedha Bhagwat.</i>	

Diagnosis of COVID-19 from Chest X-ray Images using Deep Learning Model	575
<i>Harsh Gupta and Shiv Pratap Singh Tomar</i>	
Early Prediction of Breast Cancer by using Artificial Neural Network and Machine Learning Techniques	582
<i>Krishna Mridha.</i>	
Efficient Design of Adaptable Deep Learning Accelerator	588
<i>Abhishek Choubey and Shruti Bhargava Choubey</i>	
Heart Stroke Risk Analysis: A Deep Learning Approach	593
<i>Neeraj Sharma, Jasroop Singh Chadha, Manas Kumar Mishra and Praveen Lalwani</i>	
Hue based Plant Leaves Disease Detection and Classification Using Machine Learning Approach	599
<i>Vijay Kumar Trivedi, Piyush Kumar Shukla and Anjana Pandey.</i>	
Intrusion Detection In Computer Networks Using Machine Learning Algorithms	605
<i>Gokul A, Sarath J N, Mohit Manoj, Niranjan M and Aswathy K N.</i>	
Machine Learning based Forecasting of Wind Power	612
<i>Nikhil Garg and Sandeep Kumar Singh.</i>	
Machine learning techniques for Morphologic and Clinical Features Extraction of Cerebral Aneurysm	617
<i>S.R.Sri Vidhya and AR Arunachalam.</i>	
Neural Dense Captioning with Visual Attention	622
<i>Akshat Jadhav, Aarti Kale, Ankita Joshi and Jyoti Kundale.</i>	
On the Use of SVR based Machine Learning Method for Nonlinearities Mitigation in Short Range Fronthaul Links	628
<i>Muhammad Usman Hadi and Isha Mittal.</i>	
Predicting Stock Price and Risk using Embaded Stacking Algorithm	632
<i>P.N.V. Shyamalarao M., N. Suresh Kumar</i>	
Real-Time Face Mask Detection in Deep Learning using Convolution Neural Network	639
<i>Syed Imran Ali, Muhammed Khurram, Shirin Shafiei Ebrahimi, Syed Imran Ahmed Qadri</i>	
Robust Malware Detection Using Machine Learning	643
<i>Naman Bhoj, Ritika Verma, Deeksha Chilkoti, Mayank Khari, Shalini Sahai.</i>	
Soft Computing Techniques for Dengue Prediction	648
<i>Dilip Kumar Choubey, Adweat Mishra, Sambeet Kumar Pradhan and Naman Anand.</i>	
Study and Analysis of Implementing A Smart Attendance Management System Based on Faced Recognition Technique Using OpenCV and Machine Learning	654
<i>Krishna Mridha and Nabhan Tawjih Yousef.</i>	
SVM based Supervised Machine Learning Framework for Glaucoma Classification using Retinal Fundus Images	660
<i>Deepak Parashar and Dheeraj Agrawal</i>	
Brain Tumor Detection Using Machine Learning in MR Images	664
<i>Roshan Jahan and Manish Madhav Tripathi</i>	
Fast Parallel SVM based Arrhythmia Detection on Multiple GPU Clusters	669
<i>Ghazanfar Latif; Jaafar M. Alghazo; Muhammad Mohsin Butt; Zafar Kazimi</i>	
A Real Time Driver Fatigue Detection System Using Multi-task cascaded Convolutional Neural Network	674
<i>Raja Mohana S.P., Manu Vidhya S., Reshma D.</i>	

A Novel Approach to Improve Employee Retention Using Machine Learning	680
<i>Abhay Charan Patro, Saad Aziz Zaidi, Aaradhyा Dixit and Manish Dixit</i>	
Application of Data Sampling for Crop Yield Prediction using Stochastic Gradient Descent Neural Networks	685
<i>Veena Patki, Parth Wazurkar</i>	
Advancing Deep Learning for Supply Chain Optimization of COVID-19 Vaccination in Rural Communities	690
<i>Miguel Barajas, Shilpa Bhatkande, Pireethi Baskaran, Hardik Gohel, Bishwajeet Pandey</i>	

Track -I: Cryptography, Security and BlockChain

A Literature Review of Various Techniques to Perform Encryption and Decryption of Data	696
<i>Ankita Laad and Khushboo Sawant.</i>	
A Novel Approach for Preservation of privacy by using location.....	700
<i>Hemantkumar Jadhav and Shila R. Pawar.</i>	
A Secure Audio Interface to Prevent Spying of Audio Data on Voice Controlled Devices	704
<i>Padmanabhan Rajanbabu and Pankaj Kumar Dubey</i>	
Achieving Effective Secrecy based on Blockchin and Data Sharing Scheme in Cloud Computing	709
<i>Anil Kesavareddy and Praveen Kumar Mannehalli</i>	
An Efficient Color Image Security Technique for IOT using Fast RSA Encryption Technique..	717
<i>Seema Vishwakarma and Neetesh Gupta</i>	
Android malware detection and build model using support vector machine	723
<i>Joseph Raymond V and Jeberson Retnaraj R.</i>	
Blockchain Implementation for Privacy preserving and securing the Healthcare data	729
<i>Mukesh Soni and Dileep Kumar Singh.</i>	
Decentralized Voting Application Using Blockchain	735
<i>Devansh Agarwal, Hrithwick Rao Rewatkar, Anmol Khandelwal and Subho Upadhyay.</i>	
Energy-efficient implementation of AES algorithm on 16nm FPGA	740
<i>Vaishnavi Bisht, Bishwajeet Pandey, Dil Muhammad Akbar Hussain, Mohsin Jamil and Mohammad Kamrul Hasan.</i>	
Exploring Blockchain enabled Smart Community with Electric Vehicles	745
<i>Athira Nair K, Chaitanya Kapoor, Nidhin Mahesh A, Sai Shibu N B and Balamurugan S.</i>	
Securing data in Decentralized Cloud Storage for Authorized Encrypted Search with Privacy - Preserving on Healthcare Databases	755
<i>Satish Pokharkar and Laxmikant Vishwamitra.</i>	
Security Challenges in Adopting Internet of Things for Smart Network	761
<i>Huma Gupta and Sanjeev Sharma.</i>	
The Hybrid Cryptography for Enhancing the Data Security in Fog Computing	766
<i>Sakshi Tripathi, Rakesh Kumar Tiwari, Rajesh Nigam, Neetesh Kumar Gupta and Bhupendra Verma.</i>	
Tree Based Classification of Firewall Logs to Dodge Intrusion	772
<i>Naman Bhoj, Sanchit Sayala, Tushar Sharma, Suraj Negi and Varun Mishra.</i>	
Trust Based Approach to Mitigate Wormhole Attacks in Mobile Adhoc Networks	776
<i>Mukul Shukla and Brijendra Kumar Joshi</i>	

Detection of DDoS Attacks in Software Defined Network using Decision Tree 783
Heena Kousar, Mohammed Moin Mulla, Pooja Shettar, Narayan D.G.

Trust Management in Social Internet of Things 789
Amandeep Kaur and Darpan Anand.

Track -J: Intelligent Systems, Software Testing and IT Applications

A Novel Approach in Requirement Engineering during Software Build-up 795
Devesh Kumar Srivastava and Aarushi Batta.

A Study of Bug Manifestation Process for Ensuring Software Quality 801
Raj Kumar and Singla Sanjay.

An Aid for Health monitoring during pregnancy 805
Khushboo Gairola and Dr. Ashish Bagwari.

Analysis of Diabetes Patients using Classification algorithms 810
Joe Abinas J, Radha Subramanyam, Hari Venkata Krishna Chandolu, Vempalli Murali Krishna and Nagabushanam Perattur.

Analysis of the Indian IT Sector—Investigating the Role of Technology for Software Process Improvement through Knowledge Management 815
Mitali Chugh and Robin Singh Bhadaria.

Application of AI in Video Games to Improve Game Building 821
Tanvi Rath and Preethi Nanjundan.

Automatic Parking Slot Occupancy Detection using Laplacian Operator and Morphological Kernel Dilation 825
Maysara Noor and Amit Srivastava.

Comparative Analysis of Research Status of Mobile Learning at China and Abroad (2011-2020) 832
Songbin Bao, Yujie Zheng and Fanqi Meng.

Saudi Vision 2030: Sustainable Economic Development through IoT 837
Faisal Yousif Al Anezi

Impact of tourist demand, from COVID-19, on the sustainable Inca infrastructure of Ollantaytambo, Urubamba – Cusco – Peru 842
Doris Esenarro, Jaqueline Santos, Christian Teodoro, Bishwajeet Pandey and Ciro Rodriguez.

New Approach for Online Examination Conduction System Using Smart Contract 848
Mayuri Kulkarni and Khalid Alfatmi.

Performance Evaluation of Execution Time Aware Bio-Inspired Virtual Machine Placement Strategy 853
Pradeep Singh Rawat and Robin Singh Bhadaria.

Recognition and Classification of Accompanying Audios of Kathak Dance 860
Anirban Saha and Indranil Sarkar.

Review of Classifiers Used for Novel Corona Virus Foot-Print Detection 865
Manish Assudani and Neeraj Sahu.

A Semantic Annotation Scheme for Mapping Provenance in Educational Ontology 872
Amit Kumar Bajpai, Rajiv Pandey, Surya Prakash Tripathi

Performance Analysis of Online Shopping for Customer Satisfaction Using PLS-SEM 876
Shaifali Chauhan, Richa Banerjee, Mohit Mittal, Hardeep Singh

Digital Healthcare System Vulnerability Analysis using Network Forensic Tool	881
<i>Ali Siddiqui, Oguntoyinbo Olufunmilayo, Hardik Gohel, Bishwajeet Pandey</i>	
The Research Status of Formulaic Language Identification	886
<i>Fanqi Meng, Yujie Zheng and Jingdong Wang</i>	
Intelligent Job Navigation System For Rural Development.....	892
<i>Ayushi Ghosh, Malvika Thakur, Roshini Praveen and M Saravanan.</i>	

Author Index

A Novel Approach to Improve Employee Retention Using Machine Learning

Abhay Charan Patro

Department of CSE
Birla Institute of
Technology Mesra
Ranchi, India

abhaycharanpatro@gmail.com

Saad Aziz Zaidi

Department of CSE
Birla Institute of
Technology Mesra
Ranchi, India

seanazarrin@gmail.com

Aaradhyaa Dixit

Department of ECE
Birla Institute of
Technology Mesra
Ranchi, India

aaradhyadixit@ieee.org

Manish Dixit

Department of CSE
Madhav Institute of
Technology and Science
Gwalior, India

manishdixit@ieee.org

Abstract—Employee attrition is one of the biggest challenges faced within business organizations. With the advancement in machine learning, various research papers brought into light the numerous reasons why an employee leaves, this paper puts forward a machine learning pipeline that not only predicts employee attrition but also suggests a minimum cost approach for the company so that the employee does not leave.

Keywords-Attrition; HumanResource; Machine Learning; Slack features; Willingness factor; H2O AutoML; Random Forests; Extremely Randomized Trees (XRT); Generalized Linear Models; XGBoost GBM; H2O GBM; Deep learning

I. INTRODUCTION

As defined by FurstPerson[1], attrition is a reduction of staff be it, voluntary or involuntary. Attrition is usually viewed negatively and as a burden for employers. According to FurstPerson, large enterprises, more often than not, don't calculate the turnover rate, downplaying it's costs. In a study conducted by furstperson[1], the financial toll employee turnover can take on an organization, in general, across several major industries, can cost anywhere from \$1,500 to \$16,650 per agent. According to salary.com's Compdata surveys [10], the total attrition in 2018 was 19.3% out of which 14.2% was voluntary in nature. The attrition in 2018 saw a 3.6% increase compared to 2014's total attrition of 15.7%.

Machine learning seems to be an ever progressing field that shows no signs of slowing down, with the monumental leaps in the way we recognize, process and find patterns in data, lost dreams seem within reach. Machine learning has taken care of predicting employee attrition.

Using these predictive models made and machine learning as our foundation, we tried to answer the question of what specific changes would an employee need to prevent them from leaving the company and thus reducing the attrition rate.

II. PROBLEM

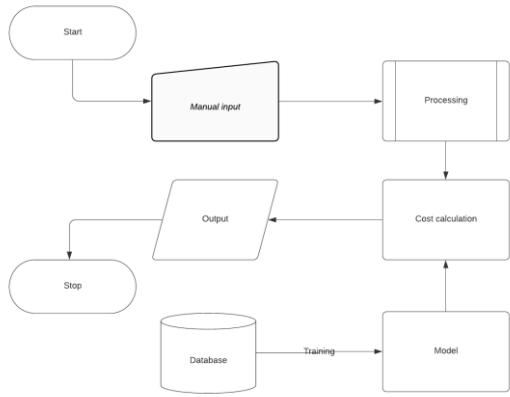
The Question of why an employee Quits [2] has been answered in quite an extensive manner, with various dependency graphs, correlation matrices and the like. The topic is quite important and thus multiple methods like decision trees[3], data-mining techniques [4][5] etc, have been employed for predictions, but all they could do was predict or try and show what factors had the most impact on the quitting employee. Here we try to build on our predecessors' incredibly insightful work to answer the pertinent question :-

What changes could have stopped the quitting employee?

III. CONCEPT

This paper has been inspired by Google's deep dream[9], wherein, using gradient ascent you try to change not the weights but the input itself to increase the activation of nodes, essentially flipping the neural net. Our concept for this pipeline, with its processes to change the input features for the desired outcome, was originally intended for a neural net, but owing to insufficient data, it was adapted for other models. This research seeks to help employers retain their employees, significantly reducing costs for the business organization in the long run.

We propose a pipeline that trains on the data of the past, predicting the future, while providing an appropriate minimum cost of, if you so wish, averting said future. In the context of our paper, this avertable future is Employee attrition. Here, we know the desired outcome, what we don't, is the change in our input that'll bring about the desired outcome to fruition



IV. METHODOLOGY

Figure 1. Flowchart of the proposed pipeline

A. Pre-Processing

The data used for this purpose was “IBM HR Analytics Employee Attrition & Performance” [6]. It consisted of 34 features and 1470 synthetic examples. The feature type and the number of unique values in them is given in Fig. 2.

Employee count, Over18 and StandardHours were removed from the feature list because they all had the same value in all the training examples. EmployeeNumber was also removed from the feature list because it was just a unique identifier for each employee. Moreover 'DailyRate', 'HourlyRate', 'MonthlyRate' were also removed because of their relation with MonthlyIncome. This brought the feature list down to 27 features.

Then the features 'BusinessTravel', 'Gender', 'OverTime', 'Marital Status' were label encoded and the features 'Department', 'EducationField', 'JobRole' were one-hot encoded.

The data was then randomly shuffled to a 80:20 split for training and testing respectively.

B. Training

Trying to find a model that could most effectively help us answer our question was tedious, we tried and tested multiple classification algorithms to train our model to accurately classify employee attrition, recreating results from [7] was the first step, we then moved on to change our pre-processing techniques and could obtain a much better AUC, precision, recall and F1_score in all our testing algorithms, but it was not enough, bear in mind the results by others had been astonishing as well given that the dataset was so imbalanced, an accurate representation would be, for every ‘YES’ there were almost 15 ‘NO’s. With odds like these an F1_score of ~0.5 was good.

But we decided to train various models using H2O’s Auto-ML. H2O’s Auto-ML is an algorithm for automatic machine

Feature	Feature Type	Unique values
Age	Numerical	43
BusinessTravel	Numerical	3
DailyRate	Numerical	886
Department	Categorical	3
DistanceFromHome	Numerical	29
Education	Numerical	5
EducationField	Categorical	6
EmployeeCount	Numerical	1
EmployeeNumber	Numerical	1470
EnvironmentSatisfaction	Numerical	4
Gender	Categorical	2
HourlyRate	Numerical	71
JobInvolvement	Numerical	4
JobLevel	Numerical	5
JobRole	Categorical	9
JobSatisfaction	Numerical	4
MaritalStatus	Categorical	3
MonthlyIncome	Numerical	1349
MonthlyRate	Numerical	1427
NumCompaniesWorked	Numerical	10
Over18	Categorical	1
OverTime	Categorical	2
PercentSalaryHike	Numerical	15
PerformanceRating	Numerical	2
RelationshipSatisfaction	Numerical	4
StandardHours	Numerical	1
StockOptionLevel	Numerical	4
TotalWorkingYears	Numerical	40
TrainingTimesLastYear	Numerical	7
WorkLifeBalance	Numerical	4
YearsAtCompany	Numerical	37
YearsInCurrentRole	Numerical	19
YearsSinceLastPromotion	Numerical	16
YearsWithCurrManager	Numerical	18

learning on tabular data, part of the H2O machine learning platform[8].

Figure 2. Type of Feature and its unique values

model_id	auc	logloss	aucpr	mean_per_class_error
XGBoost_grid_1_AutoML_20210319_095526_model_1	0.837357	0.319802	0.609676	0.238802
StackedEnsemble_BestOfFamily_AutoML_20210319_095526	0.83687	0.301567	0.652458	0.230209
StackedEnsemble_AllModels_AutoML_20210319_095526	0.83056	0.308226	0.636741	0.229186
XGBoost_grid_1_AutoML_20210319_095526_model_9	0.82957	0.324921	0.606539	0.256299
XGBoost_grid_1_AutoML_20210319_095526_model_5	0.827343	0.325788	0.609031	0.247814
XGBoost_grid_1_AutoML_20210319_095526_model_3	0.822915	0.328112	0.598294	0.228271
GLM_1_AutoML_20210319_095526	0.822277	0.330172	0.628944	0.269102
XGBoost_2_AutoML_20210319_095526	0.821109	0.328804	0.589495	0.272386
GBM_5_AutoML_20210319_095526	0.821055	0.337032	0.5835	0.284165
GBM_grid_1_AutoML_20210319_095526_model_1	0.820527	0.340547	0.550585	0.270857
XGBoost_grid_1_AutoML_20210319_095526_model_10	0.819685	0.335535	0.5676	0.261984
GBM_grid_1_AutoML_20210319_095526_model_2	0.818686	0.333837	0.583345	0.288095
GBM_grid_1_AutoML_20210319_095526_model_7	0.817499	0.339303	0.578365	0.272623
XGBoost_3_AutoML_20210319_095526	0.817243	0.343133	0.576583	0.239663
XGBoost_1_AutoML_20210319_095526	0.816952	0.33673	0.584532	0.245381
XGBoost_grid_1_AutoML_20210319_095526_model_11	0.815224	0.333925	0.585591	0.238145
XGBoost_grid_1_AutoML_20210319_095526_model_7	0.814791	0.343491	0.560935	0.281118
GBM_grid_1_AutoML_20210319_095526_model_5	0.813283	0.344926	0.573433	0.284531
GBM_grid_1_AutoML_20210319_095526_model_6	0.812812	0.363997	0.526772	0.274012
GBM_2_AutoML_20210319_095526	0.812484	0.344781	0.533989	0.270964
XGBoost_grid_1_AutoML_20210319_095526_model_4	0.809135	0.346493	0.562634	0.24158
DeepLearning_1_AutoML_20210319_095526	0.808508	0.353982	0.516117	0.267777
GBM_4_AutoML_20210319_095526	0.80513	0.353018	0.520182	0.261327

Figure 3. A few top models of the leaderboard of the various models generated using H2O's AutoML

Model ID	Importance
GLM_1_AutoML_20210319_095526	0.6782996382634335
XGBoost_grid_1_AutoML_20210319_095526_model_1	0.3964367864527333
XGBoost_2_AutoML_20210319_095526	0.13192081915394016
XGBoost_grid_1_AutoML_20210319_095526_model_5	0.10976832022050043
XGBoost_1_AutoML_20210319_095526	0.06323860700302551
XGBoost_grid_1_AutoML_20210319_095526_model_3	0.062410970916512085
XGBoost_3_AutoML_20210319_095526	0.04553848325053091
DeepLearning_1_AutoML_20210319_095526	0.02526754006536049
GBM_grid_1_AutoML_20210319_095526_model_5	0.023648664303890343
XGBoost_grid_1_AutoML_20210319_095526_model_2	0.022977488834341198
DeepLearning_grid_3_AutoML_20210319_095526_model_2	0.01894327135590139
DeepLearning_grid_2_AutoML_20210319_095526_model_2	0.014568108799226352
XGBoost_grid_1_AutoML_20210319_095526_model_6	0.008854624036856702

Figure 4. The importance of various models in the stacked ensemble model made using all the 40 models

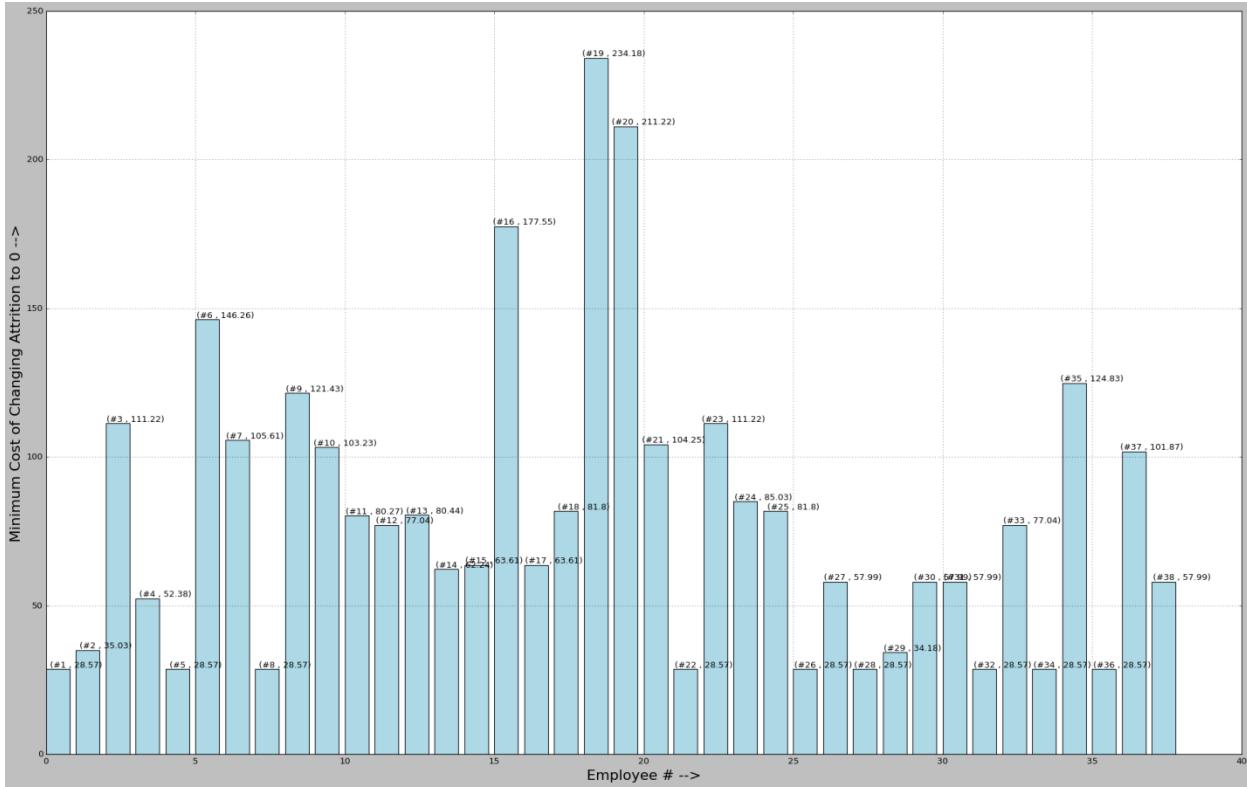


Figure 5. Minimum cost required to retain various employees

40 different models were trained using various algorithms such as Random Forest and Extremely Randomized Trees (XRT) models, Generalized Linear

minimum cost of changes that needed to be done on a set of features so that an employee, predicted to leave, wouldn't.

C. Cost calculation

The feature-set would be decided by the business organization, and the changes they are willing to make. We refer to these features as slack features.

Model	F1 Score
XGBoost_grid_1_AutoML_20210319_095526_model_1	0.4943820224719101
StackedEnsemble_BestOfFamily_AutoML_20210319_095526	0.6021505376344085
StackedEnsemble_AllModels_AutoML_20210319_095526	0.6086956521739131

Figure 6. F1 Score of the top 3 models of the leaderboard

Models, XGBoost, GBMs, and Deep Learning models. Apart from the 40 models, two more stacked ensemble models were created, one was created using the best model of the family and the other one using all 40 models.

From the leaderboard (Fig. 3) we calculated the F1 scores for the top 3 models (Fig. 6) and chose to use a stacked ensemble of all models as the main reference model as it gave the best F1 score. Then the attrition predictions were made on the test set. Now the main goal was to find the

Slack Feature	Willingness Factor
EnvironmentSatisfaction	0.21
JobLevel	0.14
StockOptionLevel	0.28
PercentSalaryHike	0.21
WorkLifeBalance	0.14

Figure 7. Slack Features and their corresponding willingness factors used for this paper

Each slack feature has a willingness factor ω associated with it. The willingness factors are calculated by taking input from the business organisation, willingness percentages of the slack features, these describe the “intent” of the

organisation, in making changes to those slack features. Then the willingness percentages are normalized so that they all add up to 1, this gives us the willingness factors. For the purpose of this paper we chose the features and willingness factors as shown in Fig. 7. Then we needed to solve the problem of choosing the best set of changes from the many possible and to do that we came up with the following minimization goal :

Cost required for employee e:

$$\min_{\text{for all changes}} \left(\sum_{x \in \text{slack features}} (1 - \omega_x) * \Delta x_e \right)$$

We used this minimization goal individually for each employee to find the best set of feature changes that would suit the employee.

V. FUTURE PROSPECTS

The cost we calculated was minimized over the cost of changes in the various possible solutions. This cost has no real relation to a monetary value that the company might incur to prevent attrition, albeit this is what was originally envisioned for the pipeline, but couldn't be applied due to the reluctance of companies, with regards to data and varied ways in which accounting is done for a particular company. For a company that has records of its employees and their CTCs, combining these datasets should be no problem, we would then be able to calculate the exact monetary cost required for preventing employee attrition, given past accounting data, this could then be taken further, using real-time comparison of incurring cost, with other viable options to decide on the change.

This could be addressed in successive works.

VI. CONCLUSION

In this paper a solution to increasing attrition rates through machine learning is discussed by answering the question, what would make the employee stay ? Through this paper, we have tried to introduce a Machine Learning Pipeline, which business organizations can utilise by choosing their own slack features and willingness factors, complemented with a large database, to drastically reduce employee attrition.

REFERENCES

- [1] S. Colberg, "What is Attrition & Attrition Rate, What Causes Attrition and Cost of Employee Turnover", Fursthperson.com, 2021. [Online]. Available: <https://www.fursthperson.com/blog/causes-employee-attrition-cost-employee-attrition>
- [2] Frye A, Boomhower C, Smith M, Vitovsky L, Fabricant S. Employee Attrition: What Makes an Employee Quit?. SMU Data Science Review. 2018;1(1):9.
- [3] Alao, D.A.B.A. and Adeyemo, A.B., 2013. Analyzing employee attrition using decision tree algorithms. Computing, Information Systems, Development Informatics and Allied Research Journal, 4(1), pp.17-28.
- [4] Shankar RS, Rajanikanth J, Sivaramaraju VV, Murthy KV. Prediction of employee attrition using datamining. In2018 ieee international conference on system, computation, automation and networking (icscan) 2018 Jul 6 (pp. 1-8). IEEE.
- [5] Yadav S, Jain A, Singh D. Early prediction of employee attrition using data mining techniques. In2018 IEEE 8th International Advance Computing Conference (IACC) 2018 Dec 14 (pp. 349-354). IEEE.
- [6] "IBM HR Analytics Employee Attrition & Performance", Kaggle.com, 2017. [Online]. Available: <https://www.kaggle.com/pavansubhasht/ibm-hr-analytics-attrition-dataset>
- [7] F. Fallucchi, M. Coladangelo, R. Giuliano, and E. William De Luca, "Predicting Employee Attrition Using Machine Learning Techniques," Computers, vol. 9, no. 4, p. 86, Nov. 2020
- [8] v.3.32.0.5 Erin LeDell and Sebastien Poirier. H2O AutoML: Scalable Automatic Machine Learning. 7th ICML Workshop on Automated Machine Learning (AutoML), July 2020. URL https://www.automl.org/wp-content/uploads/2020/07/AutoML_2020_paper_61.pdf
- [9] "Inceptionism: Going Deeper into Neural Networks", Google AI Blog, 2021. [Online]. Available: <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>.
- [10] "Compdata 2018 Turnover Report", Www2.salary.com, 2021. [Online]. Available: <https://www2.salary.com/turnover>