



AIST - 2023



# 5<sup>TH</sup> INTERNATIONAL **CONFERENCE ON Artificial Intelligence & Speech Technology**

## **AIST 2023**

26<sup>th</sup> - 27<sup>th</sup> December, 2023

**Supported By:**



विज्ञान एवं  
प्रौद्योगिकी मंत्रालय  
MINISTRY OF  
SCIENCE AND  
TECHNOLOGY

सत्यमेव जयते



**Centre of Excellence – Artificial Intelligence**

(Supported by Department of Science and Technology, GOI)

**Indira Gandhi Delhi Technical University for Women**

(Established by Govt of NCT of Delhi),

Kashmere Gate, Delhi-110006

[www.igdtuw.ac.in](http://www.igdtuw.ac.in)



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

26<sup>th</sup> - 27<sup>th</sup> December, 2023

### Guest of Honour



**Dr. Sachin Gulati**  
Director, Campus Hiring  
American Express

### Chief Guest



**Dr. (Smt.) N. Ranjana**  
Outstanding Scientist and Director,  
Directorate of Systems and Technology Analysis,  
DRDO

### Organizing Chair



**Prof. Arun Sharma**  
Dean Academics,  
IGDTUW

### Technical Program Chair



**Prof. S.S. Agrawal**  
Emeritus Scientist,  
CSIR

### General Chair



**Dr. (Mrs.) Amita Dev**  
Hon'ble Vice Chancellor  
IGDTUW, Delhi



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

26<sup>th</sup> - 27<sup>th</sup> December, 2023

### KEYNOTE SPEAKERS



**Prof. Yudong Zhang**  
School of Computing and Mathematical Sciences  
University of Leicester, UK



**Dr. Priyankoo Sarmah**  
IIT Guwahati, India



**Prof. Sardar M. N. Islam (Naz)**  
ISILC, Victoria, University,  
Melbourne, Australia



**Dr. Helmy Bin Abd Wahab,**  
University Tun Hussein Onn Malaysia



**Dr. Meena Jha,**  
Central Queensland University, Australia



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Message from the General Chair



It gives me immense pleasure to welcome all the participants, delegates, keynote speakers, and resource persons for Fifth International Conference on “Artificial Intelligence and Speech Technology” i.e. AIST-2023, to be held at Indira Gandhi Delhi Technical University for Women, Kashmere Gate, Delhi on 26<sup>th</sup> and 27<sup>th</sup> December 2023. The first four versions of the Conferences were a great success with the participation of experts from USA, Germany, Japan, Austria, Malaysia, and almost all corners of India.

The 5<sup>th</sup> version of AIST has scaled up with a greater number of quality papers from wider reach of academicians, professionals and researchers all over the world. The fifth International Conference on Artificial Intelligence and Speech Technology, 2023 has provided us the opportunity to interact with a great number of International Keynote Speakers and to learn from these experts.

The aim of the Conference is to serve as a forum for discussions on the state-of-the-art research, development, and implementations of Artificial Intelligence and Speech Technology. AIST-2023 is dedicated to cutting-edge research that addresses the scientific needs of academic researchers and industrial professionals to explore new horizons of knowledge related to Artificial Intelligence, Machine Learning, Deep Learning, Speech Synthesis, and Speech Recognition. Researchers from across the world are presenting their research revealing the latest and relevant research findings on almost all aspects of these domains.

As academicians, the responsibility to nurture complete professionals lies with us. This necessitates the knowledge of the latest trends in fast-changing technology. Conferences bring together people from all different geographical areas who share a common discipline or field and is found effective to extend one’s knowledge.

I, on behalf of the Steering Committee, would like to express my sincere thanks and appreciation to the world-renowned Professors and prominent Researchers for having agreed to deliver the keynote session and share their knowledge during the Conference. I am sure that this colloquy of researchers and experts from academia and industry would greatly benefit researchers, students and faculty. Young scientists and researchers will find the contents of the proceedings helpful to set roadmaps for their future endeavours.

I wish the conference great success.

**Dr. (Mrs.) Amita Dev**  
**Hon'ble Vice-Chancellor, IGDTUW**  
**General Chair, AIST-2023**



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Message from the TPC Chair



It is a matter of great pleasure and happiness to see that Indira Gandhi Delhi Technical University for Women, Delhi is organizing its Fifth International Conference on Artificial Intelligence and Speech Technology (AIST-2023). The objective of the conference is to provide a platform for a profound discussion and presentations on state-of-the-art research, development, innovations, and implementations of Artificial Intelligence and Speech Technology by researchers worldwide.

There has been tremendous advancement and innovations in Artificial Intelligence which is incomparable to what Artificial Intelligence emerged traditionally. We use Artificial Intelligence many times during the day-often, without even realizing it. Today Artificial Intelligence has greatly enhanced machine learning, Natural Language Processing (NLP), and Deep learning such that they are enabling new developments in Speech Technology like voice response user interactive systems, smartphones, and home appliances.

The demand of such appliances and systems is increasing tremendously with new features and the market for voice-enabled systems will increase by several billion dollars in the near future. India has a great opportunity scope of developing such systems in indigenous languages-22 official and 780 spoken languages. Looking to its huge hope and dimensions AIST-2023 brings together academics, industry experts and education leaders from all over the world to discuss an incredibly wide array of topics ranging from the Foundation of Artificial Intelligence and machine learning, data mining, and Cognitive science to Speech technology, to name a few.

I would like to express my sincere thanks and appreciation to the world-renowned Professors and prominent Researchers for having agreed to deliver the keynote talks and share their knowledge during the Conference. My warmest thanks go to the organizing committee colleagues including the co-chairs, the technical program committee members, the paper reviewers for their invaluable work in shaping the technical program, and not the least all the authors who kindly submitted their papers to AIST-2023. I am very happy to share that the accepted papers will be submitted to Springer for Publication, which is indexed by SCOPUS.

In summary, no doubt you all will appreciate the unique combination of the cutting-edge technical program, with the wonderful organization of the conference, Enjoy meetings with friends and colleagues as well as important discussions with eminent speakers. I look forward to seeing everyone in IGDTUW, Delhi India.

**Prof. S.S. Agrawal**  
**DG, KIIT Group of Colleges**  
**Ex. Emeritus Scientist, CSIR**  
**Technical Program Chair, AIST-2023**



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Message from the Organizing Chair



I take this opportunity to welcome you all to the 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology i.e., AIST-2023, to be held at Indira Gandhi Delhi Technical University for Women, Delhi during 26<sup>th</sup> -27<sup>th</sup> December, 2023. This conference will have an amalgam of researchers from the fields of Artificial Intelligence and Speech Technology.

The objective of the conference is to provide a forum for researchers worldwide to unveil their latest work in Artificial Intelligence and innovations in Speech Technology. Topics covered in this conference include fundamentals of AI, its tools and applications, Machine Learning, Deep Learning, Speech Analysis, Representation and Models, Spoken Language Recognition and Understanding, Affective Speech Recognition, various applications including NLP, Computer Vision, Robotics, Medical and others. The conference received more than 300 submissions from all over the globe, out of which the best 70 selected papers will be presented during these two days. The Conference Proceedings will be published by Springer publication.

All the paper submissions have gone through a careful anonymous review process (2 or more reviewers per submission) aided by Technical Program Committee members and Advisory Board. The AIST-2023 Conference includes prominent Keynote addresses by Prof. Yudong Zhang, School of Computing and Mathematical Sciences, University of Leicester, UK, Dr. Helmy Bin Abd Wahab, University Tun Hussein Onn Malaysia, Prof. Sardar M. N. Islam (Naz) ISILC, Victoria, University, Melbourne, Australia and Dr. Priyankoo Sarmah, IIT Guwahati. Also, the Technical Sessions will be chaired by eminent experts from AI and Speech Technologies including Dr. Ajay Kumar, Thapar University, Dr. Sunita Garhwal, Thapar University, Dr. Karunesh Arora, CDAC, Dr. Meena Jha and several others.

I would like to thank everyone who has given his or her time, energy and ideas to assist in organizing this event including all members of organizing committee, Technical Program Committee members and all reviewers and our distinguished keynote speakers who have agreed to address the conference attendees. I also wish to thank all of our sponsors and supporters especially American Express and DST who have made this event possible. It is through the collective efforts of these individuals and organizations that we are able to bring this conference a great event. Looking for the great success of the Conference

**Prof. Arun Sharma**  
**Organizing Chair - AIST 2023**  
**Professor & Head – CSE and Dean - Academic Affairs**



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Message from the Chief Guest



It is an immense honour to be part of 5th International Conference on Artificial Intelligence and Speech Technology being organised by IGDTUW. This conference is a testament to the remarkable strides being made in the field of technology and the boundless possibilities that lie ahead.

The convergence of Artificial Intelligence and Speech Technology represents a paradigm shift in operations, interactions and problem solving using the power of machines. Scientists, researchers, and innovators, are at the forefront of shaping a future where intelligent systems will seamlessly collaborate with human capabilities to enable better discoveries and improve efficiencies and make innovations possible.

Artificial Intelligence is not just about creating smarter machines; but it will augment human capability by solving complex challenges that were once thought insurmountable. Speech technology, in particular, opens up avenues for more natural and intuitive human-machine interactions, providing a bridge between the digital and physical worlds. AI and Speech Technology will have a transformative impact on various domains, from national defence to healthcare, education, and beyond.

Applications of Artificial intelligence brings in plethora of avenues in advancing defence technologies, where multiple developments are happening on not only functionalities but also on its safety aspects.

As the conference navigates through the presentations, discussions, exchanges of ideas over the course of present work, results and achievements, it is important to keep a focus on the possibilities of tomorrow. Meanwhile ethical implications of AI are also to be carefully considered, to channelise innovations towards the betterment of humanity.

I encourage all participants to discuss new ideas, debate the best way forward, forge new collaborations, and foster an environment where products can take concrete shape.

May the ideas that are exchanged sow new seeds of knowledge during the deliberations. I look forward to the insights that will emerge from this conference and the impact they will have on shaping the trajectory of endeavours in AI and speech technology.

Best wishes to the participants and organisers for a successful conference.

**Dr. (Smt.) N. Ranjana**  
**Outstanding Scientist and Director,**  
**Directorate of Systems and Technology Analysis, DRDO**



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Program Schedule

Time	Schedule
10:10 AM – 10:15 AM	Kulgeet of the University and Lighting of Lamp
10:15 AM – 10:20 AM	Welcome of Guests
10:20 AM – 10:25 AM	Inaugural Address by Prof. Arun Sharma, Dean Academics Affairs and Organizing Chair, AIST 2023
10:25 AM – 10:30 AM	Highlights of Technical Program by Dr. S. S. Agrawal, DG KIIT, Emeritus Scientist, CSIR and Technical Program Chair, AIST 2023
10:30 AM – 10:35 AM	Presidential Keynote Address by Prof. Amita Dev, Hon'ble VC, IGDTUW and General Chair, AIST2023
10:35 AM – 10:40 AM	Address by Guest of Honor – Dr. Sachin Gulati, Director – Campus Hiring, American Express
10:40 AM – 10:50 AM	Address by Chief Guest – Dr. Ranjana N. , Scientist H and Director, Directorate of Systems and Technology Analysis (DSTA), DRDO
10:50 AM – 10:52 AM	Release of Conference Souvenir
10:52 AM – 10:55 AM	Presentation of Mementoes to the Guest
10:55 AM – 11:00 AM	Vote of Thanks and National Anthem
11:00 AM – 11:45 AM	High Tea
11:45 AM – 12:30 PM	Panel Discussion (American Express) Session Chair: Ms. Karuna Kadian, IGDTUW Co Chair: Ms. Priyamvada Mathur, IGDTUW



# 5<sup>th</sup> International Conference AIST 2023



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology, AIST 2023

Date: 26<sup>th</sup> December, 2023

<b>Online Mode</b> Time: 12:00 PM – 12:30 PM Venue: Seminar Hall	<b>Keynote Address 1:</b> The Transformative Power of Artificial Intelligence in Medical Image Analysis	<b>Speaker: Prof. Yudong Zhang</b> School of Computing and Mathematical Sciences, University of Leicester, UK Session Chair: Prof. Ajay Kumar, Thapar University Session Co-Chair: Dr. Ritu Rani, COE – AI, IGDTUW
<b>Offline Mode</b> Time: 12:30 PM -01:45 PM Venue: Conference Hall	<b>Technical Paper Presentation 1:</b> Trends and Applications in Speech Processing	Session Chair: Dr. Karunesh Arora, CDAC Session Co Chair: Ms. Sukhandeep Kaur Shergill, BML Munjal University, Dr. Vijay Kumar Yadav, IGDTUW Student Coordinator:
<b>Offline Mode</b> Time: 12:30 PM - 01:45 PM Venue: Seminar Hall	<b>Technical Paper Presentation 2:</b> Recent Trends in Speech and NLP	Session Chair: Prof. Girish Nath Jha, JNU Session Co Chair: Dr Goldie Gabrani, MRIU Faridabad/Ms. Bhawna Jain, IGDTUW
<b>Offline Mode</b> Time: 02:15 – 15:00 PM Venue: Seminar Hall	<b>Keynote Address 2:</b>	<b>Speaker: Dr. Priyankoo Sarmah, IIT Guwahati</b> Session Chair: Prof. S.S. Agrawal, DG KIIT, Emeritus Scientist, CSIR, TPC Chair Session Co-Chair: Dr Himanshu Mittal, IGDTUW
<b>Online Mode (in parallel)</b> Time: 02:15 – 15:00 PM Venue: Conference Hall	<b>Keynote Address 3:</b> <b>Quantum Machine Learning</b>	<b>Speaker: Prof. Sardar M. N. Islam (Naz)</b> ISILC, Victoria, University, Melbourne, Australia Session Co Chair: Prof. A.K. Mohapatra, IGDTUW
<b>Offline Mode</b> Time: 15:00 – 16:30 PM Venue: Seminar Hall	<b>Technical Paper Presentation 3:</b> Recent Trends in Machine Learning and Deep Learning	Session Chair: Dr. Sunita Garhwal, Thapar University Session Co Chair: Ms. Karuna Kadian, IGDTUW Dr. Garima Jaiswal, Bennett University
<b>Online Mode</b> Time: 15:00 – 16:30 PM Venue: Conference Hall	<b>Technical Paper Presentation 4:</b> Emerging trends in Speech Processing	Session Chair : Dr. Devendra Kumar, NSUT Session Co Chair: Dr. Deepak Kumar Sharma, IGDTUW
Date: 27 <sup>th</sup> December, 2023		
<b>Online Mode</b> Time: 10:00 AM – 10:45 AM Venue: Seminar Hall	<b>Keynote Address 4:</b> <b>Smart Application for Community</b>	<b>Speaker: Dr. Helmy Bin Abd Wahab</b> University Tun Hussein Onn Malaysia Session Chair: Prof. Arun Sharma, IGDTUW Session Co-Chair: Dr. Sanjib Kumar Sahu, IGDTUW
<b>Online Mode</b> Time: 10:45 AM - 12:15 PM Venue: Seminar Hall	<b>Technical Paper Presentation 5:</b> Analysis using Hybrid technologies with Artificial Intelligence	Session Chair: Prof. Savita Gupta, UIET Punjab University, Chandigarh Session Co Chair: Bhawana Narwal, IGDTUW
<b>Online Mode</b> Time: 10:45 AM - 12:15 PM Venue: Conference Hall	<b>Technical Paper Presentation 6:</b> Advances in Computational Linguistics and NLP	Session Chair: Prof. Brijesh Kumar, IGDTUW Session Co Chair: Mr. Rahul Sachdeva, Anveshan Foundation
<b>Online Mode</b> Time: 12:45 PM – 01:30 PM Venue: Seminar Hall	<b>Keynote Address 5:</b>	<b>Speaker: Dr. Meena Jha,</b> Central Queensland University, Australia Session Chair: Dr. Ritu Rani, IGDTUW Session Co Chair: Ms. Shreshtha Singh, IGDTUW
<b>Online Mode</b> Time: 02:00 – 03:30 PM Venue: Seminar Hall	<b>Technical Paper Presentation 7:</b> Exploring New Horizons in Computer Vision Research	Session Chair: Prof. Shakir Khan, Al Imam University Session Co Chair: Dr. Kamal Kumar, IGDTUW
<b>Online Mode</b> Time: 02:00 – 03:30 PM Venue: Conference Hall	<b>Technical Paper Presentation 8:</b> Applications of Machine Learning and Deep Learning in Healthcare	Session Chair: Prof. A.K. Mohapatra, IGDTUW Session Co Chair: Dr. Gaurav Indra, IGDTUW



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Technical Session: 01

### Trends and Applications in Speech Processing

Date: 26<sup>th</sup> December 2023 (Offline Mode) Time: 12:30 PM - 01:45 PM

S No	Paper ID	Authors	Paper Title
1	228	Manisha Gupta, Amita Dev and Poonam Bansal	Recent Trends in text to Speech Synthesis in Context with Indian Languages
2	68	Anunya Sharma , Kiran Malik and Poonam Bansal	Speech Emotion Recognition using Convolutional Neural Networks
3	177	Gokul Krishnan, Dhruthi Rao, Anshu Sandur, Ananya Jandhyala, and Pooja Agarwal	A Comprehensive Review of Instructional Tools and Applications for Dyslexic Learners
4	108	Pooja Gambhir, Amita Dev, Poonam Bansal	Text-Independent Voiceprints Identification using Feed- Forward Back-propagation with layered strategies
5	202	Vartika Tyagi, Amita Dev and Poonam Bansal	Speech Recognition using Adaptation of Whisper Models
6	205	Surbhi Bharti, Prerna Jha, Sakshi Sahu, Sheba Alam, and Ashwini Kumar	Voice Stress Analysis using Machine Learning
7	221	Suhani, Amita Dev and Poonam Bansal	A Comprehensive Analysis on Kaldi-Based Speech Recognition for Low Resource Indian Languages
8	180	Divya Taneja, Dr. Shobha Bhatt	Deep Learning-based Speech Recognition Models: A Review
9	173	Nidhi Bhatt, Shobha Bhatt, Geetanjali Garg	Hindi Speech Recognition using Deep Learning: A Review



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Technical Session: 02

### Recent Trends in Speech and NLP

Date: 26<sup>th</sup> December 2023 (Offline Mode) Time: 12:30 PM - 01:45 PM

S No	Paper ID	Authors	Paper Title
1	120	Deepanshi Jalan, Anandita Tuli, Shradha Jain, Ashwani Kumar and Vaishali Kikan	Email Bot- Voice Based Email System for Blind
2	144	Mayank Singhal, Komal, Mohammad Zeeshan, Ishika Saini and Preeti Nagrath	Hate Speech Detection Using Glove and BERT
3	209	R. C. Barik, Tanuj Kumar, Anunay Sinha and Vishwajeet Kumar	A Novel and Intelligent approach for Indian Locale Based Text-to-Speech Model by hybridizing Wave Net and Wave Glow with Mel-Spectrogram Analysis
4	216	Khushi Anand, Bhawna Jain, Ananya Verma, Anushka Gupta, Niharika Chhabra	A Comparative Analysis of Deep Learning Architecture for Accurate Gender Classification using Vocal Data
5	48	Apoorva Tiwari, Karuna Kadian, Nikita Garg, Sukhandeep Kaur, and Vimal Dwivedi	Scaling Language Boundaries: A Comparative Analysis of Multilingual Question-Answering Capabilities in Large Language Models
6	199	Sapna Yadav, Syed Bilal Abbas Rizvi and Syed Afzal Murtaza Rizvi	A Comprehensive Approach to AI-Enabled Diagnosis of Lung Diseases: Utilizing Pretrained Models for Fibrosis, Tuberculosis, and Covid-19 Detection
7	193	Ankit Kumar, Richa Sharma, and Punam Bedi	Enhancing Named Entity Recognition with DistilBERT and Attention Ensemble Fusuion
8	125	Abdul Moid and Narendra Sharma	Investigation on Existing Blockchain based Architecture, AI/ML driven for Boosting IoT Security & Privacy
9	260	Pooja Bhati	A Comprehensive Review on Impact of Computer Vision in Healthcare Domain



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Technical Session: 03

### Recent Trends in Machine Learning and Deep Learning

Date: 26<sup>th</sup> December 2023 (Offline Mode) Time: 03:00 PM - 04:30 PM

S No	Paper ID	Author	Paper Title
1	76	Tanya Kumari, Ritika Kumari and Poonam Bansal	Diagnosis of Alzheimer Disease using Machine Learning Algorithms
2	62	Astha Sharma and Ashwni Kumar	Efficient Detection and Multi-Level Classification of Tomato Plant Leaves using fused deep and hand-crafted features and LSTM-ANFIS
3	222	Niranjana Sowpari, Poonam Bansal, and Nandini Sethi	Comparative Analysis Of Deep Learning Models For Text Summarization On Hindi Corpus
4	198	Surbhi Bharti, Divya Verma, Anamika Kumari and Ashwni Kumar	Analysing Social Media Data for emotion detection and sentiment analysis
5	200	Surbhi Bharti, Medha Arora, and Ashwni Kumar	Campus Exploration: Figma-Driven UI Design for a Campus Tour App
6	206	M Deekshitha Reddy, Geetika Vadali, Ritu Rani, Arun Sharma, and Amita Dev	Predicting Optimal Startup Exit Strategies with Ensemble Learning for Venture Capitalists
7	33	Akash Shah, Sapna Varshney and Monica Mehrotra	GRUbd-SM: Gated Recurrent Unit based Bot Detection on Social Media
8	143	Saumya Mehra and Bhawna Jain	Closing the Data Divide in COVID-19 X-ray Datasets: Advancing Diagnosis through Stable Diffusion-based Synthetic Image Generation
8	190	Kajal Tyagi, Anil Ahlawat and Himanshi Chaudhary	Machine Learning in IoT: An In-depth Dataset Analysis based on Attack Detection
9	225	Subhranshu Behura, Arham Alam, Nishtha Phutela, Atul Mishra, and Goldie Gabrani	Tune into Your Feelings: NLP-Powered Emotion-Driven Music Recommender Systems



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Technical Session: 04

### Emerging trends in Speech Processing

Date: 26<sup>th</sup> December 2023 (Online Mode) Time: 03:00 PM - 04:30 PM

S No	Paper ID	Author	Paper Title
1	63	A.S. Sri Saila, A.T. Venkata Subramani, M.D. Harsha Prada, and G. Madhu Priya	Deep Learning Based Multilingual Voice Recognition System and Analytics for Organization Surveys
2	82	Gitesh Kambli, Jay Oza, and Amit Maity	MelSpectroNet: Enhancing Voice Authentication Security with AI-based Siamese Model and Noise Reduction for Seamless User Experience
3	183	Puneet Mittal, Sukhwinder Sharma and Khyati Marwah	Analysis of Acoustic Features for Gender Identification using Punjabi Speech Dataset
4	134	Shivani Trivedi, Sugandh Rastogi, Sneha Agrawal, and Ravish Sharma	Empowering Hate Speech Detection: A Comparative Exploration of Deep Learning Models
5	73	Pavel Posokhov, Stepan Skrylnikov, Olesia Makhnytina and Yuri Matveev	Hybrid Approach to The Personification of Dialog Agents
6	166	Sandip Jana , Joyanta Basu ,Tapan Kumar Basu and Amiya Karmakar	Santali Vowel Recognition: An Under-Resourced Tribal Language
7	195	Jesin James, Balamurali B. T., Binu Abeysinghe, and Junchen Liu	Explaining Spectrograms in Machine Learning: A Study on Neural Networks for Speech Classification
8	250	Rizwana Kallooravi Thandil, Mohamed Basheer K.P and V.K Muneer	Enhancing Emotion Classification in Malayalam Accented Speech: An In-Depth Clustering Approach



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Technical Session: 05

#### Analysis using Hybrid technologies with Artificial Intelligence

Date: 27<sup>th</sup> December 2023 (Online Mode) Time: 10:45 PM - 12:15 PM

S No	Paper ID	Author	Paper Title
1	135	Shivani Trivedi, Ankit Yadav, Suraj Kumar Gupta, and Shivam Gupta	FarmScan: Smart Rotten Fruit and Vegetable Detector
2	28	Ms. Prachi Sasankar and Dr. Gopal Sakarkar	Fault Predictions based on base learnings and mean weighted score using machine learning techniques
3	53	Mr. Omprakash A. Jaisinghani, Dr. P. L Ramteke Department and Mr. Bharat S. Dhak	Ensuring Trustworthy Elections using IoT-Enabled Blockchain EVM Voting Mechanism with Aadhaar Card-Based Face Verification
4	197	Panimalar Kathirolu, Vijayalakshmi.V, Piyush Gopal, Sivananda Panda	Investigation of satellite data for monitoring air quality over remote sensing technology
5	17	Adria Binte Habib, Faisal Bin Ashraf Muhammad Iqbal Hossain and Golam Rabiul Alam	N-BEATS & Temporal Fusion Transformer based Surface Temperature Prediction and Forecasting for Realizing Global Warming Trends
6	249	Muneeb V.K, Mohamed Basheer K.P, Rizwana Kallooravi Thandil and Muhamed Ilyas P	A Hybrid Travel Recommender Model Based on Deep level Autoencoder and Machine learning algorithms
7	35	R.Revathi	A Journey through Multifaceted Data in Machine Learning Predictions on Financial Viability
8	49	Ankita Chopra, Madan Lal Saini	Neural Network Based Employability Index for IT Graduates
9	253	Garima Jaiswal, Mala Saraswat, Drishti Rai, Shristi Agarwal	Future-Proofing Careers: The AI Era and the Imperative of Reskilling
10	109	Kesana Vanisri, K Srujan Raju, Bagam Laxmaiah	An Image processing-based Tomato Leaf Disease prediction using Deep CNN



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Technical Session: 06

### Advances in Computational Linguistics and NLP

Date: 27<sup>th</sup> December 2023 (Online Mode) Time: 10:45 PM - 12:15 PM

S No	Paper ID	Author	Paper Title
1	21	Dhananjay Bhagat, Kalyani Sawarkar, Dewanshu Barde, Rashi Channawar ,Mayur Talmale ,Prathamesh Kale,Rahul Agrawal	Text Scribe: Unveiling New Dimensions in Text Summarization
2	191	Kalpna B. Khandale and C. Namrata Mahender	Context Based Anaphora Resolution of English Discourses using Rule Based Approach
3	58	Arrun Sivasubramanian, Sheel Shah, Akash Narayanaswamy, Rindhya C, and Barathi Ganesh HB	Performing Text Segmentation to Improve OCR on Multi Scene Text
4	212	Kata Venkata Sai Krishna, Goli Abhilash, Devanjali Relan, Kiran Khatter	Optimization of Indian Sign Language Detection using Data Generators
5	210	Dr. Vinisha Sumra, Dr. Naveen Malik, Santosh Kumar	An Inclusive Analysis on Deep Learning hinged Malware Detection Techniques
6	220	Aditi Chauhan, Harsh Vishwakarma, Yashika Kukkar, Abhinav Maheshwari	Advances in Text Summarization Techniques: A Comprehensive Review and Future Prospects
7	167	Shakir Khan, Mohammad Serajuddin, Zafarul Hasan, Syed Ahad Murtaza Alvi, Rashid Ayub, Arun Sharma	Natural Language Generation (NLG) with Reinforcement Learning (RL)
8	169	Nimish Goel, Shushil Paudel, Srushti Jagtap, and Ishan G. S. Kakodkar	Revolutionizing Writing: Personalized Neural Classifier for Handwritten Text
9	18	Deepa. R, Dr. T. Sree Sharmila and Dr. R. Niruban	An Efficient Deep Learning based Seq2Seq Model for Abstractive Text summarization
10	208	Raghuveera T, Akshayalakshmi V K, Nisha B A, Easwarakumar K S	Efficient Real-Time Indian Sign Language Fingerspelling Recognition in Natural Settings using Heuristics
11	98	Nachiket Rathod and Prabhakar Ramteke	Detecting Image-based Fake News with Neural Sleuths



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Technical Session: 07

### Exploring New Horizons in Computer Vision Research

Date: 27<sup>th</sup> December 2023 (Online Mode) Time: 02:00 PM - 03:30 PM

S No	Paper ID	Author	Paper Title
1	102	Shashank Pandey, Sakshi Dwivedi, Vijay Bhan Singh, Neetu Verma, and Ranvijay5	Video Summarization using Firefly Algorithm
2	37	V. Manikandan, P. Dinesh, M. Chantilyan, C. Gomathi	Enhancing Mint Plant Disease Detection Accuracy through Deep Reinforcement Learning with YOLO Algorithm
3	229	Rashmi Gandhi, Yash Pasrija, Amisha Chauhan, Ritu Rani	Multiple Object Detection and Tracking Using DeepSORT
4	100	Manisha, Avadhesh Kumar, Dileep Kumar Yadav	Background Subtraction Model with Advance GMM in Dynamic Background
6	83	Lakshmish Nayak, Bharathi Ramudu, Malay Kumar Nema and Rituraj Kumar	A framework for Information Retrieval using Domain Specific Dictionary: Illustration through enhancing the Intelligence Cycle
7	108	Pooja Gambhir, Amita Dev, Poonam Bansal	Text Independent voice prints identification using feed forward back propagation with layered strategies
8	133	Shivani Trivedi, Jahanvi Gupta, Himanshu Singh	Machine Learning- Based Estimation of Caloric Expenditure





## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### Technical Session: 08

#### Applications of Machine Learning and Deep Learning in Healthcare

Date: 27<sup>th</sup> December 2023 (Online Mode) Time: 02:00 PM - 03:30 PM

S No	Paper ID	Author	Paper Title
1	19	Apeksha Koul, Rajesh K. Bawa, Yogesh Kumar	Tuberculosis Detection and Classification in Chest X-ray Images Using Optimized CNN Architecture
2	188	Soumya Jindal and Meemansa Jindal	i-BMD: AI-based Opportunistic Screening for Osteoporosis on Abdominal CT using Deep Learning
3	04	Snehal Moghe and Dr. Harsh Pratap Singh	Machine Learning Prediction for Knee Condition: Osteoarthritis and TKA, using Image Processing Among Medical Patients
4	09	Laxmi Ahuja	Machine Learning Techniques to Study: Medical and Biomedical Signal Analysis
5	145	Sakshi Shete, Suvarna Udgire, Anuradha Joshi, Shravani Desai, Ronit Murpani	Machine Learning Algorithms for Predicting Cardiac Arrhythmias
6	57	Kanchan A. Khedikar and Dr. Piyush Kumar Pareek	Data Mining Based Heart Disease Prediction using Hybrid Optimization technique of derived features
7	66	Mohit M Rao and Virendra Kumar Shrivastava	Unveiling the Enigma: Sudden Cardiac Arrest among Youth and Uncovering Underlying Causes Using Machine Learning Models
8	72	Himani Agarwal and Shweta Singhal	Using Encoder Models To Predict Ratings Using Flipkart Reviews



## 5<sup>th</sup> International Conference on Artificial Intelligence and Speech Technology (AIST 2023)

### ABSTRACTS

#### Technical Session: 01 Trends and Applications in Speech Processing

### Recent Trends in text to Speech Synthesis in Context with Indian Languages

Manisha Gupta, Amita Dev and Poonam Bansal

Abstract. The most important type of communication in daily life is speech. However, for those who are physically or visually challenged or illiterate, using computers is confusing due to the dependence of human computer interaction on visuals as well as printed work. For novices, it can be challenging to start researching the creation of text-to-speech synthesis (TTS) platforms for Indian languages, even though there are numerous speech synthesis models available for a range of languages. This is because there is currently a lack of data on the latest developments in TTS for the Indian languages. This paper describes the latest developments in the area of speech synthesis in Indian languages, as reported by a number of scholars. Different types of Speech synthesis tools and services are also described here. Afterwards, this study mentions the speech datasets that are currently available for the Hindi language.

Keywords: Speech synthesis, Indian Languages, Neural Speech synthesis.

### Speech Emotion Recognition using Convolutional Neural Networks

Anunya Sharma, Kiran Malik and Poonam Bansal

Abstract. An efficient conversation requires understanding of speech along with accurate interpretation of the tone and emotion of the speaker. Misinterpretation of emotions is one of the biggest communication barriers. It affects personal communication, business dealings, customer handling, interviews, etc. It occurs because of the simple fact that people think, speak and interpret differently. Speech Emotion Recognition, often abbreviated as SER, is based on the observation that emotions are frequently conveyed through voice, in the form of tone and pitch variations. With the rise of service industry, exact interpretation and correctly gauging a message has become a basic requirement. Multiple industries can leverage this capability to provide a range of services. For instance, marketing companies can recommend products based on a person's emotions, call center representatives can interact more effectively by adapting to the customer's mood, and the auto-motive sector can detect a person's emotions to adjust autonomous vehicle speed for collision avoidance. Consequently, this type of application holds significant potential in the world, offering benefits to companies and enhancing consumer safety. The motivation behind this paper was to build a machine learning model capable of discerning emotions conveyed through speech, assessing and comprehending an individual's emotional state, and subsequently providing personalized recommendations based on their mood in the future. The proposed Spectrogram-MFCC fusion model's impressive accuracy of 97% in detecting emotions represents a significant advancement, particularly in applications such as chatbots and social robots, where the ability to perceive concealed emotions in speech is vital for fostering meaningful interactions.

Keywords: Speech Emotion Recognition (SER), Convolutional Neural Networks (CNNs), Machine Learning (ML), Deep Learning (DL).

### A Comprehensive Review of Instructional Tools and Applications for Dyslexic Learners

Gokul Krishnan, Dhruthi Rao, Anshu Sandur, Ananya Jandhyala, and Pooja Agarwal

**Abstract.** Dyslexia is a neurological disorder which can affect the reading, writing and speaking skills of an individual. It is often characterized by difficulties in processing phonological information. Dyslexic students often struggle with tasks such as decoding words, recognizing spelling patterns and comprehending text. This can pose problems with respect to their teaching-learning experience. Numerous instructional aids have been created over time to help dyslexic children overcome their challenges and reach their full learning potential. Within the realm of dyslexic education, there is a growing emphasis on personalized interventions, the integration of technology, and the development of multi-sensory teaching methods. This review paper offers a thorough summary and analysis of the existing research on instructional materials created especially for dyslexic learners.

**Keywords:** Dyslexia · Assistive Technologies · Interactive Learning · Audio-Visual · Literacy Aids

---

## **Text-Independent Voiceprints Identification using FeedForward Back-propagation with layered strategies**

**Pooja Gambhir, Amita Dev and Poonam Bansal**

**Abstract.** The task of identifying speakers forms a formidable field in the area of Speaker Recognition which compares and matches the set of utterances spoken by an unknown or a known speaker with the available trained database of reference templates. And Deep Learning has proven to deliver better prediction and computational efficiency in many application domains of Speaker Recognition. Deep Neural Networks provide the best results as a backbone of deep learning due to their portability, versatility and computational efficiency. This work presents a Fast Fourier Transform (FFT)- based extraction of MFCC and GFCC features from a sets of text-independent audios from the different number of speakers containing high ground noise. A Feed-Forward Back- propagation Neural Network (FFBNN) is used to categorise the voices of the selected speakers into tensor labels during the learning phase. These tensor labels are further tested with sample voice sets to identify the correct speaker. A comparative predictive outcome showed that FFBNN worked efficiently for 50 speakers, generating 85 % of accuracy with 64-32 combinations of neurons at the hidden layer. The performance of GFCC features which simulates the physiological behaviour of human ear and reduces the noise effects is better in terms of accuracy than that given by MFCC features in FFBNN.

**Keywords:** Mel-frequency Cepstral Coefficient (MFCC), Feed-Forward Backpropagation Neural Network (FFBNN), Text-Independent, Speaker Identification, Gammatone Frequency Cepstral Coefficients (GFCC).

---

## **Speech Recognition using Adaptation of Whisper Models**

**Vartika Tyagi , Amita Dev and Poonam Bansal**

**Abstract.** Remarkable advancements have been made in automatic speech recognition (ASR) by the implementation of various deep-learning techniques. In this paper, the authors transcribed recordings of OOD-Speech. For Bengali automatic speech recognition (ASR) with respect to the training set using the Whisper model and the Multilingual Automatic Speech Recognition approach, the first out-of-distribution (OOD) benchmarking dataset is available. The authors implemented an English normalizer to standardize the transcription and computed the Word Error Rate (WER): 4.27%; which is a very low error rate and indicates highly accurate Bengali speech recognition. This work is part of a code competition of Bengali.AI Speech Recognition on Kaggle to recognize Bengali speech from out-of-distribution audio recordings

**Keywords:** Automatic Speech Recognition · Bengali · Whisper · Transcription · Word Error Rate

---

## **Voice Stress Analysis using Machine Learning**

**Surbhi Bharti, Prerna Jha, Sakshi Sahu, Sheba Alam , and Ashwni Kumar**

**Abstract.** Voice stress analysis (VSA) has received a lot of attention as a non-invasive tool for detecting deceit and emotional states by analysing speech patterns. Machine Learning (ML) has emerged as a potent technique for improving the precision and efficiency of VSA approaches. This paper provides an in-depth examination of the applications and improvements in voice stress analysis using ML approaches. This study explores various ML algorithms and feature extraction approaches typically used in VSA, highlighting their strengths and limitations. Furthermore, this approach discusses the problems and prospects in the field of VSA, emphasising the potential for incorporating cutting-edge ML methods to improve accuracy, dependability, and real-time applications.

**Keywords:** Machine learning · Pitch Energy · Sentimental analysis · Speech sentiment · Voice stress.

## A Comprehensive Analysis on Kaldi-Based Speech Recognition for Low Resource Indian Languages

Suhani, Amita Dev and Poonam Bansal

**Abstract.** The accessibility of open-source software is greatly aiding the popularity of speech recognition and deep learning. As an illustration, Kaldi has become a well-established framework for creating cutting-edge speech recognition systems in current times. This paper provides summary of the various ASR systems based on Kaldi for Indian Languages. The results acquired using different models and classifiers on multiple datasets are strictly observed and compared. Experiments on various Indian datasets and objectives demonstrate that Kaldi can be successfully utilized to create advanced speech recognizers. It offers a speech recognition system that utilizes finite-state transducers, accompanied by comprehensive documentation and scripts for constructing full-fledged recognition systems. The phonetic context modelling capability of Kaldi, which encoded in C++, is enabled by its core library. Along with standard Gaussian Mixture Models, it also makes acoustic model-ling using Subspace Gaussian Mixture Models (SGMM) easier. Additionally, a large variety of popular linear and affine transformations are supported, including support for a wide range of them.

**Keywords:** Automatic Speech Recognition · Kaldi · Acoustic Modeling · SGMM · WFST · OpenFST

---

## Deep Learning-based Speech Recognition Models: Review

Divya Taneja, Dr. Shobha Bhatt

**Abstract.** Models for speech recognition using deep learning are reviewed for a deeper understanding of the speech recognition process. A key focus of this survey is the recent proliferation of deep learning techniques in speech recognition. Extensive study of deep learning models is presented with a comparative analysis. Different speech recognition tools were presented in paper. Difference evaluation metrics for deep learning-based speech recognition has been illustrated. Comparative analysis for different speech recognition methods was presented. The research finding show that when different combined deep learning model showed reduce word error rate. The research work's primary contribution is through theoretical and technical details of speech recognition using deep learning. This is accomplished by providing details information about suitable models, tools and feature that can be used to enhance the system. Research finding can be applied to improve deep leaning based speech recognition system.

**Keywords:** Speech Recognition, CNNs, RNNs, LSTM, BLSTM, LAS, CTC

---

## Hindi Speech Recognition using Deep Learning: A Review

Nidhi Bhatt, Shobha Bhatt, Geetanjali Garg

**Abstract.** Speech recognition is a technology that enables humans and machines to communicate by converting spoken words into text. Hindi is spoken extensively in various regions of India and is also used in various global locations. This paper comprehensively reviews deep learning methods for Hindi speech recognition. Challenges associated with speech recognition are presented. Various speech recognition tools and databases are discussed. A comparative analysis of different deep learning-based Hindi speech recognition systems is presented. Other performance parameters are studied. The findings demonstrate that the convolution neural network (CNN) is mainly used for speech recognition. The main contribution of this research work is providing a comprehensive background and technical details of the deep learning-based Hindi speech recognition system. This is achieved by presenting details of appropriate models, tools, and features that can be utilized to improve the system. The work can be extended by going deeper into deep learning-based systems.

**Keywords:** Hindi, Speech recognition, CNN, Deep Learning

---

## Technical Session: 02 Recent Trends in Speech and NLP

## Email Bot- Voice Based Email System for Blind

Deepanshi Jalan, Anandita Tuli, Shradha Jain, Ashwani Kumar and Vaishali Kikan

**Abstract.** Email is the most important medium of communication in the present world for sharing professional and time sensitive information. It has been estimated that globally around 2.2 million people are blind. Though there have been many advancements in order to allow these people to use the email system, the existing systems cannot help them use the email systems much effectively as they require the person to use keyboard or mouse. Thus, this paper aims to provide a solution- Email Bot, which is an interactive system for visually impaired persons to use the email system with just voice commands. It is built using speech recognition library of python and can be used by any untrained person owing to the interactive voice commands at every step. It is useful not only for blind persons but also to people who have trouble in reading and writing. The system functions entirely on voice commands from the user's perspective and the user does not need to type/ press any key to utilize any option of the email system.

**Keywords:** Voice Interactive Email System, Visually Challenged, Email Bot, text to speech translation, speech recognition.

---

## Hate Speech Detection Using Glove and BERT

Mayank Singhal, Komal, Mohammad Zeeshan, Ishika Saini<sup>4</sup> and Preeti Nagrath

**Abstract.** The increasing prevalence of hate speech, on platforms poses a threat to inclusive communication. In our research we have utilized word embeddings to acknowledge the challenge of identifying Hate speech. For enhancing the accuracy and contextual understanding of our hate speech detection models we have incorporated both BERT (Bidirectional Encoder Representations, from Transformers) and GloVe (Global Vectors for Word Representation) embeddings. By preprocessing input and training machine learning models using these embeddings we have developed a methodology that effectively recognizes hate speech. Our experiments demonstrate the effectiveness of combining GloVe and BERT in detecting instances of hate speech. This proposed approach not outperforms techniques but also captures the intricate context in which hate speech manifests. Through this study we contribute to endeavors aimed at mitigating the impact of Hate speech on platforms.

**Keywords:** Hate speech, Naive Bayes, Machine Learning, Decision Tree, Logistic Regression, CountVectorizer, TF-IDF, Natural Language Processing.

---

## A Novel and Intelligent approach for Indian Locale Based Text-to-Speech Model by hybridizing Wave Net and Wave Glow with Mel-Spectrogram Analysis

R. C. Barik, Tanuj Kumar, Anunay Sinha and Vishwajeet Kumar

**Abstract.** This paper introduces a innovative architecture for text-to-speech synthesis, comprising a recurrent speaker encoder, a sequence-to-sequence synthesizer, and various vocoder options like Griffin-Lim, WaveNet, and WaveGlow. The speaker encoder extracts a fixed-dimensional vector from speech signals, leveraging a text-independent speaker verification model trained on reference speech samples from the target speaker, resulting in the generation of a speaker embedding vector. The training data includes text transcripts paired with target audio, and transfer learning is facilitated through a pre-trained speaker encoder. The vocoder component plays a critical role in converting synthesized mel-spectrograms into time-domain waveforms, exploring various vocoder options such as Griffin-Lim, WaveNet, and WaveGlow. This paper aims to evaluate the performance of two vocoder variants, including the hybridization of Waveglow and Wavenet. The research seeks to advance the field of multispeaker text-to-speech synthesis by enhancing the quality of generated speech and enabling speaker adaptation. The paper explores various vocoder options, including Griffin-Lim, which iteratively estimates missing phase information, while WaveNet employs up-sampling layers to achieve high-resolution audio. WaveGlow efficiently synthesizes high-quality audio without the need for autoregression by combining insights from Glow and WaveNet. The research that we have performed focuses on the accent-based speech within the context of a novel text-to-speech synthesis. These findings hold significant promise for applications in speech synthesis, voice cloning, and natural language processing, contributing to the development of more versatile and high-quality text-to-speech systems.

**Keywords:** WaveNet, WaveGlow, Text-To-Speech, Mel-Frequency Cepstral Coefficients (MFCC), Speech Synthesis

## A Comparative Analysis of Deep Learning Architecture for Accurate Gender Classification using Vocal Data

Khushi Anand, Bhawna Jain , Ananya Verma, Anushka Gupta, Niharika Chhabra

**Abstract** Gender categorization from vocal data is the method of detecting a speaker's gender using speech or voice data. Speech recognition, speaker identification, and emotion detection are just a few of the practical uses of this job. Researchers have tried to create an automated system to adequately discern the speaker's gender based on their speech without prior knowledge of their gender. The primary issue in this subject is identifying gender-specific traits in voice data and developing models that can successfully capture these qualities. Convolutional Neural Networks (CNNs) and other Machine Learning (ML) and Deep Learning (DL) technologies have shown considerable potential in handling this challenge. The current focus of the study is on enhancing the accuracy and generalization of these models. This study investigates the efficacies of different CNN models for gender categorization from voice data utilizing spectrograms as input. A publicly available dataset containing various vocal features was selected for this study. It comprises voice recordings of male and female speakers and includes 3,168 voice samples, each with 20 acoustic characteristics. Three CNN models were tested on this dataset, and the three-layer CNN model fared the best, with an accuracy of 98.21% and a loss of 5.64%. This research improves the performance of deep learning models in voice detection and conducts a comparative analysis among various models, ultimately determining the most effective model.

Keywords: Convolution Neural Network (CNN), Vocal Data, Speech Recognition, Gender Identification, Deep Learning, Comparative Analysis

---

## Scaling Language Boundaries: A Comparative Analysis of Multilingual Question-Answering Capabilities in Large Language Models

Apoorva Tiwari<sup>1</sup>, Karuna Kadian, Nikita Garg, Sukhandeep Kaur, and Vimal Dwivedi

**Abstract.** Large Language Models (LLMs), which are endowed with the capacity to understand and produce text similar to human language, have emerged as transformative entities within the field of natural language processing (NLP). Particularly, multilingual LLMs show enormous potential for bridging linguistic divides and expanding technical access to a variety of linguistic communities. In light of this, our research compares and contrasts GPT and Bloom, two well-known LLMs, across the range of regional Indian languages. We carefully evaluate the models' ability to produce language-specific replies with a focus on English, Hindi, and Punjabi prompts. To fully comprehend the strengths and weaknesses of the models, this study includes measures for linguistic alignment, accuracy assessment, and human review of model-generated results. The accuracy analysis revealed that while BLOOM displayed a range of performance levels, GPT consistently demonstrated competency in all of the languages tested, suggesting its versatility across a wide range of languages. This comparative study has consequences that go beyond the field of AI studies. We enable informed decision-making when deploying these models for multilingual communication, content creation, and cross-cultural interactions by highlighting the strengths and shortcomings of these models. In the end, by bridging the gap between robots and other linguistic communities, this work advances both technological aspirations and societal inclusion.

Keywords: Multilingual Large language model · GPT · Bloom

---

## A Comprehensive Approach to AI-Enabled Diagnosis of Lung Diseases: Utilizing Pretrained Models for Fibrosis, Tuberculosis, and Covid-19 Detection

Sapna Yadav, Syed Bilal Abbas Rizvi and Syed Afzal Murtaza Rizvi

**Abstract.** Chest radiographs are extensively employed for the detection and diagnosis of several pulmonary conditions, including fibrosis, tuberculosis, and COVID-19. The manual processing of chest X-ray images is a time-consuming and error-prone task, yet its analysis is crucial for timely disease identification and diagnosis. This study utilizes Convolutional Neural Networks and five pretrained models viz. ResNet50, MobileNetV2, InceptionNetV3, Xception, and VGG16 to detect lung infectious diseases using chest X-ray. The main purpose of this work is to assess the performance of deep learning models in diagnosing the lung diseases namely Covid-19, Fibrosis, and Tuberculosis. These models are trained on publicly available datasets and assessed for the performance parameters such as accuracy, recall, precision, F1-score, AUC-score, and false positive, false negative counts. The performance of CNN, and five pretrained models are evaluated and compared with respect to these performance metrics. It has been observed from the results that ResNet50, MobileNetV2, and VGG16 show incredibly high performance for all

performance parameters used in this work. To provide early diagnosis and patients outcomes, these automated diagnosis methods can be used by the medical professionals.

Keywords: Pretrained Model, Convolutional Neural Network, Fibrosis, Tuberculosis, Covid-19.

---

## **Enhancing Named Entity Recognition with DistilBERT and Attention Ensemble Fusion**

**Ankit Kumar, Richa Sharma, and Punam Bedi**

**Abstract.** This paper presents Attention Ensemble, a model based on a combination of multi-head self-attention for long-term context retention and local attention for short-term context, to significantly enhance contextual embeddings for Named Entity Recognition (NER) task. NER, an integral natural language processing (NLP) task, involves determining entities in text, such as names of persons, locations, and organizations. We have used tokenizer and the contextual embeddings obtained from the pre-trained Distil BERT model as a foundation for capturing linguistic structures and context. Our Attention Ensemble model refines these embeddings with the fusion approach for improved NER performance when fine-tuned over CoNLL-2003 and HiNER datasets. The experiments conducted on these datasets demonstrated substantial improvements in comparison to five baseline models taken into this study. Notably, our approach achieved an F1 score of 0.9757 on CoNLL 2003 and 0.9495 on HiNER datasets surpassing the scores achieved by the baseline models ACE+document-context (0.946 on CoNLL 2003) and XLM-Roberta (Large) (0.8878 on HiNER) datasets respectively. The substantial improvement exhibited by our model underscores its distinct contribution towards balancing long-term and short-term contexts in English as well as non-English languages, thus establishing its effectiveness over the reputed models available in the literature.

Keywords: Named entity recognition, multi-head self-attention, local attention, attention ensemble, Distil BERT, fine-tuning

---

## **Investigation on Existing Blockchain based Architecture for Boosting IoT Security & Privacy**

**Abdul Moid and Narendra Sharma**

**Abstract.** In the social world, the Internet of Things (IoT) can assimilate the physical as well as computational procedures as services. The quantity of administrations at the IoT edge is rising quickly because of the predominant purposes of smart gadgets and cyber-physical systems. One of the difficulties in exploring the encouraging applications of IoT services is making the services interoperable in a decentralized environment. It has been demonstrated that the Blockchain technology (BCT) is a promising option for calling for executions and establishing data trust. It could theoretically be utilized to aid in service interoperability. BCT records a verification of info or a procedure as a operation in a distributed ledger. Like the geography to IoT, considering BCT at the edges of the organization shows the dispersed trademark. Be that as it may, presently, BCT is as yet confronting the moves for interoperability because of various factors, for example, agreement conventions, block sizes, and time frame. The current BC-related security systems for the IoT environment are discussed in this paper. This paper further makes sense of the meaning of the provision of a safeguarded communication architecture for IoT devices utilizing a BC with Software defined network (SDN) empowered network. Further, this paper depicts the significance of the AI/ML driven techniques in tightening the security and confidentiality level of IoT frameworks.

Keywords: Internet of Things (IoT), Blockchain (BC) Technology, smart frameworks, Artificial Intelligence, SDN, Security.

---

## **Tune into Your Feelings : NLP-Powered Emotion-Driven Music Recommender System**

**Subhranshu Behura, Arham Alam, Nishtha Phutela, Atul Mishra, and Goldie Gabrani**

**Abstract.** Our research delves into enhancing user experiences by introducing mood-based music recommendations without the need for intrusive wearable devices. Through natural language processing (NLP), we successfully classified user-entered text into five emotional categories with a commendable 74.22% accuracy which is later used to recommend appropriate songs to the user. This methodology not only revolutionizes personalized music curation but also extends its applications to broader realms, such as assessing overall emotional well-being. Our non-intrusive and easily implementable approach offers a scalable solution, providing a seamless and meaningful way for applications to understand and respond to users' emotions in real-time, transcending the limitations of current solutions.

Keywords: Music, Emotion, NLP, Machine Learning, Classification

---

## Technical Session: 03 Recent Trends in Machine Learning and Deep Learning

### Diagnosis of Alzheimer Disease using Machine Learning Algorithms

Tanya Kumari, Ritika Kumari and Poonam Bansal

**Abstract.** Alzheimer's disease (AD) is a prominent nervous ailment. Even though its symptoms are mild at first, they get worse with time. To foretell the onset of AD utilizing psychological criteria such as age, gender, memory, Clinical Dementia Rating (CDR), and Mini-Mental State Examination (MMSE) scores, this study uses machine learning algorithms (MLAs): Naïve Bayes (NB), Decision Tree (DT), Random Forest (RF), Support vector machine (SVM), and K closest neighbor (KNN) to identify AD in an Alzheimer's clinical dataset from the Kaggle Repository. We use five performance metrics for a comparative study: Accuracy, Precision, Recall, F1 score, and AUC.

Keywords: Alzheimer's Disease, dementia, machine learning, performance metrics.

---

### Efficient Detection and Multi-level Classification of Tomato Plant Leaves using fused deep and hand-crafted features and LSTM-ANFIS

Astha Sharma and Ashwini Kumar

**Abstract:** Tomato leaf diseases affects the superiority and amount of the crops to a large extent. Consequently, it is vital to notice the crop diseases at an initial phase to maintain a social, economic and ecological balance. In this direction this paper presents a framework which integrates both ML and DL-based technology for recognition and sorting of tomato leaf bugs. The proposed framework leverages fusion of deep-features and hand-crafted features for efficient feature representation. Further, the fused features obtained are utilized by LSTM-ANFIS based classification network. This proposed framework efficiently classifies the tomato plant leaf diseases by utilizing ANFIS-LSTM based classification module. To validate the effectiveness of proposed model extensive experiments are conducted on PlantVillage dataset. Furthermore, the proposed framework provided an accuracy of 98.83%.

Keywords: Machine Learning (ML), Deep-Learning (DL), Tomato disease detection, Vegetal bug detection, Modified LBP, Inception-V3, LSTM, ANFIS

---

### Comparative Analysis Of Deep Learning Models For Text Summarization On Hindi Corpus

Niranjana Sowpari, Poonam Bansal, and Nandini Sethi

**Abstract.** We often come across large texts for example, stories, news articles etc. Sometimes it is required to have a quick summary of the context to save our time in reading the long paragraphs. In literature we often know this process by the name of Text Summarization. Text summarization comes under a field in Data Science called as Natural Language Processing. Over the years, a lot of work towards Text Summarization has been presented on English Corpus, but not much work has been done for our Indian Languages like Hindi. Our goal for this research work is to find a model which is able to generate good Hindi summary out of Hindi texts and is efficient in terms of time and resources. To be able to achieve our goal, we have performed a comparative analysis of the Deep Learning Models for text summarization so that we can compare the models on various parameters and their results. In our paper, we have trained the models BART and T5 on the 'Hindi Text Short Summarization' Corpus where T5 performed better than BART and with a rouge-L (f-measure) score of 23.6. For experimentation, we have used the 'Hindi Text Short Summarization' Corpus available from Kaggle as not much work has been performed until now on this dataset and we wanted to learn about the essential data transformations or data pre-processing that can be done on a Hindi dataset so that our



model yields us good results for the Hindi Corpus as most of times the industrial data are not very clean and understanding the data and performing data-preprocessing becomes a very crucial step.

Keywords: Deep Learning · IndicBART · BART · T5 · mT5 · Natural Language Processing · BERT · Transformers · Encoder-Decoder · Neural Networks

## **Analysing Social Media Data for emotion detection and sentiment analysis**

**Surbhi Bharti, Divya Verma , Anamika Kumari , and Ashwni Kumar**

**Abstract.** In the digital era, social media has become a rich source of insights into individuals' emotional state and mental well-being. This paper introduces a cutting-edge framework for sentiment analysis that transcends conventional positive, negative, or neutral annotation of social media text. Sentiment analysis, commonly referred to as opinion mining, is an area of Natural Language Processing (NLP) that has received a lot of attention since it focuses on identifying the emotional tone or sentiment portrayed in textual data. Through rigorous training and fine-tuning of advanced neural networks on a diverse and extensive dataset, our framework achieves an unprecedented level of accuracy in discerning nuanced emotional states and mental health concerns. Our research underscores the potential of leveraging social media as a tool for early intervention and support for individuals navigating mental health challenges. Employing state-of-the-art deep learning models, we were successfully able to predict the sentiment with an accuracy of 78%. In order to comprehend human emotions through textual data, we study several sentiment analysis approaches and applications. We also talk about the difficulties, developments, and potential future directions in this area. By shedding light on the language - emotion relationship in the digital sphere, we contribute to a more empathetic and informed digital landscape.

Keywords: Bidirectional Encoder Representations from Transformers (BERT) · Cutting - edge framework · Deep Learning (DL) · Long Shortterm Memory (LSTM) · Sentiment Analysis · Text Classification

---

## **Campus Exploration: Figma-Driven UI Design for a Campus Tour App**

**Surbhi Bharti, Medha Arora, and Ashwni Kumar**

**Abstract.** Virtual campus tours have become increasingly popular as a useful resource for prospective students, parents, and other stakeholders to remotely tour educational institutions. In order to develop a user-friendly and engaging interface for virtual campus tours, this research study investigates the convergence of Human-Computer Interaction (HCI) concepts and User Interface (UI) design. The goal is to enhance user interaction and engagement throughout the virtual tour, which will ultimately help students make better decisions when choosing their academic communities. In order to improve the campus experience for both prospective students and visitors, this research study examines the design, development, and implementation of a campus tour app using the Figma design platform. The detailed process of UI design, the tenets of user-centered design, and the seamless integration of technology to produce an interesting and educational app are all covered in this article. It also emphasizes the possibility for future improvement by incorporating AR technology. The article emphasizes the transformative effect of UI design on the educational landscape and explores the triumphs, problems, and lessons learned throughout the app's development. The introduction of augmented reality (AR) is expected to elevate the campus tour experience as the educational technology landscape continues to develop, offering an immersive and interactive way to explore educational facilities. This study lays the groundwork for new advancements in educational technology, with an emphasis on creative UI and AR integration to improve the campus tour experience.

Keywords: Artificial intelligence · Augmented Reality (AR) Integration · Figma · User Interface (UI) Design · Virtual Campus Tours.

---

## **Predicting Startup Exit Strategies with Ensemble Machine Learning**

**M Deekshitha Reddy, Geetika Vadali, Ritu Rani , Arun Sharma, and Amita Dev**

**Abstract.** In the current landscape, Venture capitalists are opting for more data-driven decision-making to achieve high returns while mitigating risks. All VCs seek a home run in their portfolio, for which it is essential to select the optimal exit. The process can be smoothed with the prediction of the type of exit- IPO, M&A, or the next funding round. Further, it is also to gauge which type of exit yields maximum profit for the particular climate. In this paper, through the utilisation of six machine learning models - Elastic Net, SVR, Kernel Ridge, Gradient Boost, XGBoost, and Random Forest, trained on factual data sourced from Crunchbase. The machine learning models are evaluated on the basis of 2 metrics- RMSLE and R-squared Score. We propose an ensemble learning model with Root Mean Square Logarithmic Error (RMSLE) score of 0.29 and R-squared score of 0.76.

This weight-averaged ensemble has been validated to predict the optimal exit for a startup under six broad labels - IPO, Acquisition, Leveraged Buyout, Merger, Acquire, and Management Buyout.

Keywords: Startup Exit Prediction · Ensemble Learning · Machine Learning.

---

## GRUBBD-SM: Gated Recurrent Unit based Bot Detection on Social Media

**Akash Shah, Sapna Varshney and Monica Mehrotra**

**Abstract.** Social media platforms have emerged as powerful tools for fostering instantaneous connections and networking across geographical boundaries. However, it is imperative to recognize that these platforms also harbor potential detrimental outcomes. The proliferation of bots contributes to the propagation of unsolicited messages, fraudulent schemes, and a spectrum of cyber threats, thereby compromising users' data security and privacy. Previous research predominantly centered on crafting bot detection models tailored to specific social media platforms. This has led to a dearth of holistic endeavors aimed at creating a unified bot detection technique effective across diverse social media networks. This paper introduces a generalized approach, GRUBBD, designed to detect bot accounts on social media. The GRUBBD model leverages user profile metadata-derived attributes and behavioral patterns for bot detection. Integrating a Gated Recurrent Unit (GRU) and an attention mechanism, the GRUBBD model is engineered to bolster its performance. To streamline computational efficiency, a dimensionality reduction technique, specifically linear discriminant analysis, is applied during data preprocessing phase. The effectiveness of GRUBBD is demonstrated by empirical evaluations performed leveraging datasets sourced from Twitter and Instagram. The GRUBBD model has a noteworthy level of accuracy, achieving 99% and 92% accuracy for the Twitter and Instagram datasets, respectively. This research contributes to the advancement of bot detection strategies while offering broader applicability across diverse social media platforms.

Keywords: Social media, Bots, Gated recurrent unit, Attention mechanism, Linear discriminant analysis.

---

## Machine Learning- Based Estimation of Caloric Expenditure

**Shivani Trivedi , Jahanvi Gupta , Himanshu Singh**

**Abstract.** In today's fast-paced technological age, people tend to neglect the health and opt for unhealthy fast food due to busy schedules. This leads to consuming more calories, contributing to obesity. This paper gives a thorough analysis of five different machine learning algorithms—Linear Regression, Random Forest, Support Vector Machine (SVM), Decision Tree, and XGBoost—to predict the number of human calories burned during physical activity. This study aims to evaluate the predictive Mean Absolute Error of different algorithms and pinpoint the most precise human calorie burn calculation model. To train and test the models on a diverse dataset including details on age, gender, amount of activity, heart rate, and other pertinent variables. Following feature engineering and preprocessing, we assessed the algorithms' capacity to forecast calorie burn. The findings of this study have practical ramifications for a variety of applications, such as fitness monitoring, individualized health advice, and the creation of wearable technology. People can improve their general health, fitness, and well-being by making educated judgments about their physical activities and diet by precisely forecasting their calorie burn. Additionally, this study adds to the body of knowledge in machine learning for health-related predictions and emphasizes the significance of selecting appropriate algorithms for specific predictive tasks.

Keywords: Machine learning, Kaggle, XGBoost, Decision tree regression, Support Vector Machine, Linear regression, Random Forest.

---

## Closing the Data Divide in COVID-19 X-ray Datasets: Advancing Diagnosis through Stable Diffusion-based Synthetic Image Generation

**Saumya Mehra and Bhawna Jain**

diagnostic tools, this research harnesses Stable Diffusion, a robust generative modelling technique. Deep neural networks are employed to learn the underlying data distribution and generate synthetic X-ray images. By employing the Stable Diffusion model, the study aims to overcome the scarcity of annotated COVID-19 patient X-ray datasets. Through a process of finetuning with a Kaggle dataset, synthetic X-ray images closely resembling authentic patient X-rays are created. This synthetic data generation technique effectively addresses the data scarcity challenge, facilitating more robust model training and improving the accuracy of COVID-19 detection and categorization. The methodology for utilizing Stable Diffusion for data augmentation involves several crucial steps and processes. Initially, the model is fine-tuned to capture essential characteristics and patterns

present in real patient X-rays. Following fine-tuning, the model generates synthetic X-ray images that closely mimic genuine patient X-rays by drawing from the learned data distribution. These synthetic images are then seamlessly integrated into the original dataset, thereby expanding and diversifying the collection of X-ray images. In the final stage, machine learning models, such as deep neural networks, are trained using the augmented dataset. This approach leverages information from both real and synthetic images, significantly enhancing the accuracy and robustness of COVID-19 case detection and classification.

Keywords: Stable Diffusion ,COVID 19, Generative AI

---

## Machine Learning in IoT: An In-depth Dataset Analysis based on Attack Detection

Kajal Tyagi, Anil Ahlawat and Himanshi Chaudhary

**Abstract.** Internet of Things (IoT) has transformed how people interact with the physical environment, generating vast amounts of data from interconnected devices and sensors. This study offers a comprehensive survey of IoT network traffic datasets. The efficient analysis of these massive datasets has become a challenging task, leading researchers and practitioners to harness the power of machine learning algorithms. This paper delves into the exciting field of IoT dataset analysis using machine learning techniques. We explore the different IoT datasets on which ML is implemented, offering insights into the future prospects of this rapidly evolving domain. Overall, we examine seventeen datasets that we found through the analysis of various scientific publications, which we have covered in this study. Additionally, we examined which ML approach is more effective at classifying attacks and how we can implement Feature Engineering or Data Preprocessing before applying ML techniques on these datasets. Many researchers get some limitations while applying the ML techniques on different datasets which we have discussed in our analysis, also they highlights emerging trends and future directions on these datasets. Atlast, we have discussed five well known datasets which has been created and given labels corresponding to their different attack categories.

Keywords: IoT, Machine Learning(ML), Datasets, Data Preprocessing, Attack samples.

---

## Technical Session: 04 Emerging trends in Speech Processing

### Deep Learning Based Multilingual Voice Recognition System and Analytics for Organization Surveys

A.S. Sri Saila, A.T. Venkata Subramani, M.D. Harsha Prada, and G.Madhu Priya

**Abstract.** Most organizations frequently undervalue the importance of employee surveys, which results in lost opportunities for both reputational and financial improvement. In order to support organizational development, this study recommends a thorough survey approach that incorporates both written and oral feedback responses and takes into account a variety of linguistic preferences. A novel intelligent system that makes use of cutting-edge voice recognition technology is suggested for the smooth collection of feedback. The system that Survey Spark has proposed makes use of state-of-the-art voice recognition technology, allows oral feedback from respondents in their native tongues, and allows text-based feedback input according to user preferences. The sentiment analysis model based on deep learning is an essential element that reveals emotional subtleties in the gathered data. A word cloud visualization shows the terms that respondents use most frequently. This work enhances multilingual speech recognition, supporting 60 global languages, including 5–6 indian languages. Survey Spark achieves 97% accuracy with voice recognition system and 99% accuracy in sentiment analysis. By merging linguistic diversity and advanced analytics, this system offers insights into the collective employee mindset, enhancing decision-making and organizational efficiency.

**Keywords:** Survey System, Voice recognition system, Deep Learning, Sentiment analyzer, Textual recognition, Multilingual voice transcription

---

## MelSpectroNet: Enhancing Voice Authentication Security with AI-based Siamese Model and Noise Reduction for Seamless User Experience

**Gitesh Kambli , Jay Oza and Amit Maity**

Abstract Voice authentication has become critical for secure access control while achieving usability. Background noise and increased security requirements, however, continue to be problems. This paper presents MelSpectroNet, an innovative voice authentication system using a Siamese neural network trained on over one million samples. It leverages mel spectrograms for efficient feature extraction and employs noise reduction, enhancing reliability. The model achieves 96.62% test accuracy, demonstrating efficacy. Our methodology involves audio denoising, meticulous spectrogram preprocessing, a tailored Siamese architecture, and rigorous training. Testing demonstrates MelSpectroNet's exceptional performance and ability to generalize. However, enhancing longitudinal accuracy by accounting for natural voice variations over time still needs exploration. Overall, MelSpectroNet pioneers highly accurate and usable voice authentication with enhanced security. It balances user convenience and stringent authentication needs. This research motivates further work to optimize these systems for diverse conditions while advancing security and inclusiveness.

Keywords: voice-authentication, siamese-neural-network, mel-spectrogram

---

## **Analysis of Acoustic Features for Gender Identification using Punjabi Speech Dataset**

**Puneet Mittal, Sukhwinder Sharma and Khyati Marwah**

Abstract. Recognition accuracy of Automatic Speech Recognition (ASR) System primarily depends upon training and testing of underlying acoustic and language models. The corpus used for training and testing usually consists of speech recordings from different speakers, which are generally taken without considering their gender. Voice of males and females is quite different. Male voice has low pitch and high pitch period and for female voice it is vice versa. While collecting speech and data developing the ASR model these variabilities need to be considered else model may result into biased model for one group. This paper aims to evaluate and analyze the dissimilarity between voices of male and female speakers based on their acoustic features for an ASR system. Exclusive model for male and female voices have been developed and cross testing has been done. The results show that model developed for one gender is not efficient in recognizing speech from other gender. The comparison of gender-based recordings results in certain speech characteristics which need to be considered while training and testing ASR systems.

Keywords: ASR, Female voice, HMM, Male voice, MFCC, Pitch.

---

## **Empowering Hate Speech Detection: A Comparative Exploration of Deep Learning Models**

**Shivani Trivedi, Sugandh Rastogi, Sneha Agrawal, and Ravish Sharma**

Abstract Hate Speech Detection, employing Natural Language Processing (NLP) and Machine Learning techniques, serves as a crucial tool for automatically identifying and flagging discriminatory language, hatred, or other protected characteristics within the vast landscape of social media. The escalating volume of social media posts, generated every second, has led to a surge in hateful and disrespectful comments. The imperative to automatically detect such content arises due to the time and resource intensity involved in creating high-quality human-labeled datasets. This study meticulously selects a dataset ensuring its inclusion of a substantial amount of hate speech instances across diverse languages, capturing the nuanced nature of online discourse. To enhance model performance, each model, including BERT, GPT, and LSTM, undergoes fine-tuning on this dataset, incorporating personalized adjustments based on hyperparameters and preprocessing strategies. The comparative analysis yields insightful findings on the effectiveness of BERT, GPT, and LSTM models in Hate Speech Detection. Notably, this research advances beyond the previous abstract, acknowledging the diverse linguistic landscape by utilizing data generated from social media in various languages. The study contributes valuable insights to intellectuals and researchers, emphasizing the deployment of state-of-the-art techniques in real-world applications across linguistic boundaries. This paper underscores the importance of a nuanced approach in model selection for different Hate Speech Detection tasks, recognizing the challenges posed by diverse languages. Moreover, it highlights the necessity of research-centric innovation in addressing the critical issue of hate speech across multilingual social media platforms.

Keywords: Bidirectional encoder representations from transformers (BERT), Generative pre-trained transformer (GPT), long-short-term memory (LSTM), natural language processing (NLP), machine learning.

---

## **Hybrid Approach to The Personification of dialogue Agents**

**Pavel Posokhov, Stepan Skrylnikov, Olesia Makhnytina, and Yuri Matveev**

**Abstract.** This paper describes the application of knowledge grounded generation for the personification of dialogue agents in Russian language. To implement own model for Russian language the authors propose two different architectures for personified systems: BERT&GPT and T5-multitask. The former consists of two independent components ranking module using an encoder-only BERT-style model and a generative module using a GPT model. The latter employs a unified multitask encoder-decoder model for both ranking and generation tasks. The research is based on the crowdsourced dataset Toloka Persona Chat Rus pre-processed to match the personification task. The authors used full golden knowledge set during training and hyperparameter tuning to avoid noise and improve quality of the knowledge-generation task. Man-ual evaluation of the proposed models' metrics demonstrate a significant improvement in SSA scores for the Russian language, outperforming base models by 51%.

**Keywords:** Personified dialogue systems · Knowledge-grounded generation · Hybrid models.

---

## **Santali Vowel Recognition: An Under-Resourced Tribal Language**

**Sandip Jana, Joyanta Basu, Tapan Kumar Basu and Amiya Karmaka**

**Abstract.** Santali is a tribal language spoken in various regions of India. It challenges speech recognition due to limited resources and linguistic diversity. This paper offers a comprehensive investigation into Santali vowel recognition through the integration of acoustic phonetics and machine learning techniques. The primary objective is to explore the recognition of Santali vowels from spoken utterances. In order to tackle the difficulties arising from the limited resources available for the language, various approaches encompassing feature extraction methods, model architectures, and training strategies are investigated. In this study, the acoustic phonetics approach incorporates formant frequencies, Short-time energy (STE), and short-time zero crossing rate (ST-ZCR). Meanwhile, the machine learning-based approach involves feature extraction techniques like MFCC, Chroma, and Mel-spectrogram, coupled with various classifiers such as SVM, Random Forest, Gradient Boosting Machines, and XGBoost. The results of this study showcases promising outcomes, underscoring the potential for advancements in Santali speech processing. Additionally, this research contributes to the preservation and analysis of the under-resourced tribal language, supporting efforts to bridge the gap in Santali language technology and promoting its sustainable development.

**Keywords:** vowel duration, formant frequencies, acoustic characteristic, underresourced tribal language

---

## **Explaining Spectrograms in Machine Learning: A Study on Neural Networks for Speech Classification**

**Jesin James, Balamurali B. T., Binu Abeysinghe, and Junchen Liu**

**Abstract.** This study investigates discriminative patterns learned by neural networks for accurate speech classification, with a specific focus on vowel classification tasks. By examining the activations and features of neural networks for vowel classification, we gain insights into what the networks “see” in spectrograms. Through the use of class activation mapping, we identify the frequencies that contribute to vowel classification and compare these findings with linguistic knowledge. Experiments on a American English dataset of vowels showcases the explainability of neural networks and provides valuable insights into the causes of misclassifications and their characteristics when differentiating them from unvoiced speech. This study not only enhances our understanding of the underlying acoustic cues in vowel classification but also offers opportunities for improving speech recognition by bridging the gap between abstract representations in neural networks and established linguistic knowledge.

**Keywords:** Spectrograms, Linguistics, Explainable Speech Recognition, Interpretable, Activation Maps, Vowels

---

## **Enhancing Emotion Classification in Malayalam Accented Speech: An In-Depth Clustering Approach**

**Rizwana Kallooravi Thandi, Mohamed Basheer K.P and V.K Muneer**

**Abstract.** Accurate emotion classification in accented speech for the Malayalam language poses a unique challenge in the realm of speech recognition. In this study, we explore the application of various clustering algorithms to this specific dataset, evaluating their effectiveness using the Silhouette Score as a measure of cluster quality. Our findings reveal significant insights into the performance of these algorithms. Among the clustering methods, Affinity Propagation emerged as the frontrunner, achieving the

highest Silhouette Score of 0.5255. This result indicates a superior cluster quality characterized by well-defined and distinct groups. OPTICS and Mean Shift Clustering also demonstrated strong performance with scores of 0.4029 and 0.2511, respectively, indicating the presence of relatively distinct and well-formed clusters. In addition, we introduced Ensemble Clustering (Majority Voting), which achieved a score of 0.2399, indicating moderate cluster distinction. These findings provide a valuable perspective on the potential advantages of ensemble methods in this context. Our experiment results shed light on the effectiveness of various clustering methods in the context of emotion classification in accented Malayalam speech. This study contributes to the advancement of speech recognition technology and lays the groundwork for further research in this area.

Keywords: Speech Emotion Recognition (SER), Accented Speech Recognition, Clustering Techniques, Speech Feature Engineering

---

## **Technical Session: 05 Analysis using Hybrid technologies with Artificial Intelligence**

### **FarmScan: Smart Rotten Fruit and Vegetable Detector**

**Shivani Trivedi , Ankit Yadav , Suraj Kumar Gupta , and Shivam Gupta**

**Abstract.** Quality evaluation of fruits and vegetables is a vital factor in maintaining qualities of food and ensures consumer satisfaction. Freshness is an essential asset of fruit and vegetable quality that impacts the physical health and motivation of consumers. This paper presents a transformative approach to revolutionizing quality evaluation in fruits and vegetables by applying image processing and machine learning algorithms, specifically utilizing the GrabCut algorithm for background subtraction. The traditional subjective methods of assessing product quality are replaced with objective and data-driven techniques. Through case studies, the efficacy of this approach is demonstrated, showcasing its potential to enhance quality assessment precision and efficiency. This innovative fusion of image processing and machine learning presented a promising avenue to redefine industry standards, ensuring higher consumer satisfaction and reduced wastage through consistent and reliable quality evaluation.

Keywords: Image processing, image segmentation, background subtraction, grab cut algorithm.

---

## **Fault Predictions based on base learnings and mean weighted score using Machine Learning Techniques**

**Ms. Prachi Sasankar and Dr. Gopal Sakarkar**

**Abstract.** Software Fault Prediction (SFP) processes find utility across various phases of the software development lifecycle, with a particularly valuable role in early identification of flawed or buggy components. Machine learning (ML) techniques are commonly harnessed in this domain. Predictive models rely on a set of 23 object-oriented metrics drawn from the CK and Martin metric collections, serving as input features. To evaluate the efficacy of these prediction models, experiments are conducted using ten publicly available datasets sourced from the PROMISE repository. Performance metrics, including Receiver Operating Characteristic (ROC) curve, accuracy(Acc), precision(Pre), F1-score(Fs), recall(Rc), and mean weighted error, are employed to gauge the predictive capabilities of the models. In this study, nine machine learning (ML) classifiers are employed to predict the accuracy rates of software fault proneness. We utilize ten of the most popular datasets in their latest versions from the PROMISE repository to evaluate the performance of algorithms, including Logistic Regression-LR, Decision Tree-DT, Support Vector Machine-SVM, Naïve Bayes-NB, Random Forest-RF, K-Nearest Neighbor-Knn, AdaBoost-AB, Gradient Boosting-GB, and Extreme Boosting-XB. The results highlight the superiority of the Random Forest over other examined algorithms in prediction accuracy. When using the RF classifiers, the achieved benchmarks include an Accuracy of 90%, Precision of 89%, Recall of 89.50%, F1-Score of 88%, and AUCROC of 75%. The comparative analysis indicates that the Random Forest delivers outstanding accuracy and reduced error rates. This research underscores the potential of the Random Forest, alongside other deep learning techniques, in advancing the field of software fault prediction precision.

---

## **Ensuring Trustworthy Elections using IoT-Enabled Blockchain EVM Voting Mechanism with Aadhaar Card Based Face Verification**

**Mr. Omprakash A. Jaisinghani , Dr. P. L Ramteke Department and Mr. Bharat S. Dhak**

**Abstract:** India is the biggest democracy country in the world, so sustaining the integrity of its electoral process is a challenge. To prevent voting fraud, booth capturing, and electronic voting machine (EVM) tampering in this situation, secure electronic voting mechanisms have become important. This presented work suggests a unique approach to improve voting mechanism security, transparency, and efficiency. The goal of this paper is to propose a robust EVM based system by combining blockchain technology, the Internet of Things (IoT), and face recognition based on Aadhaar cards. By verifying voters through facial feature matching with photos from their Aadhaar cards, this innovative method will significantly lower the number of duplicate and bogus ballots. Future electoral processes in India may be completely reinvented once this biometric identity verification system is implemented with seamlessness. In this paper propose a system design which can improve accuracy while retaining voter anonymity when EVM integrating IoT devices with blockchain technology. The cryptographic algorithms that support our technology also highlight how crucial data security and privacy are. By utilizing this cutting-edge technology, we are able to enhance the voter identification process also reducing the risk of issues like voting fraud, booth capture, and modification of electronic voting machines (EVMs) which are the common threats for voting process.

**Keywords:** EVM, Aadhaar, Face detection and recognition, Blockchain, UIDAI

---

## **Investigation of satellite data for monitoring air quality over remote sensing technology**

**Panimalar Kathirolu, Vijayalakshmi.V, Piyush Gopal, Sivananda Panda**

**Abstract:** This paper presents a comprehensive Investigation of Satellite data for Monitoring Air quality through Remote sensing Technology tool (ISMART tool). Specifically, it utilizes data from the Sentinel-5 precursor (Sentinel-5p) Tropospheric Monitoring Instrument (TROPOMI) datasets to monitor the columnar surface of co gas across India, with a focus on high-pollution regions. After rigorous data processing, essential attributes were extracted to monitor the columnar CO level. Haversine formula is employed for specific area monitoring, Basemap is used to visualize the pollution levels and machine learning models (ARIMA, LSTM) for precise monitoring and forecasting. This research culminated in a user-friendly website, built with the Django framework, where users can easily identify pollution hotspots by inputting specific parameters. Ultimately, this paper underscores the cost-effectiveness and vast coverage advantages of remote sensing over traditional ground-level instruments, urging governments to harness this technology for timely interventions against pollution.

**Keywords:** Tropomi; Sentinel-5p; Air pollution; Long Short-Term Memory; Auto Regressive Integrated Moving Average; Carbon Monoxide

---

## **N-BEATS & Temporal Fusion Transformer based Surface Temperature Prediction and Forecasting for Realizing Global Warming Trends**

**Adria Binte Habib, Faisal Bin Ashraf, Muhammad Iqbal Hossain and Golam Rabiul Alam**

**Abstract:** At the pinnacle of civilization, where the impacts of climate change have been increasingly felt, weather prediction plays a critical role in mitigating the potential disasters that may arise. Moreover, with the gradual change on climate, surface temperature of the earth is increasing. This increasing rate of the surface temperature causing global warming which is a matter of intimidation. To leave off this global warming, weather forecasting can be used as an arsenal. Selecting the appropriate tools and models for weather prediction is a crucial step in ensuring accurate forecasts. In this research paper, the focus was on studying the versatility of three specific architectures for weather prediction: LSTM, Temporal Fusion Transformer, and N-BEATS. To assess these architectures' performance, we conducted a number of experiments. With the lowest Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) of the three, NBEATS stood out. This shows that when compared to the other models, the NBEATS architecture had greater prediction accuracy. It's vital to remember, too, that the trials also showed that the Temporal Fusion Transformer and LSTM performed well. The only distinction was that these models required larger sizes in terms of parameters and computational complexity to achieve their performance levels. Consequently, considering both performance and model size, the researchers determined that N-BEATS was the most optimal and versatile architecture for weather prediction. Its ability to achieve excellent results with a smaller model size makes it a favourable choice for practical applications.

**Keywords:** Weather Prediction, Time Series Analysis, LSTM, Temporal Fusion Transformer, N-BEATS.

---

## **A Hybrid Travel Recommender Model Based on Deep level Autoencoder and Machine learning algorithms**

**Muneer V.K, Mohamed Basheer K.P, Rizwana Kalllooravi Thandil and Muhamed Ilyas P**

**Abstract.** This research investigates the application of autoencoders in processing travelogues written in the Malayalam language on Facebook. The main objective is to harness the capabilities of autoencoders to learn a compressed representation of the input data and employ it to train various machine learning models for enhanced accuracy and efficiency. The major challenge of unavailability of a benchmark dataset in the Malayalam language for the travel domain was overcome by employing NLP techniques on the unstructured, lengthy, imbalanced travelogues, applying some additional filtering methods, and the creation of an exclusive POT Tagger along with lookup dictionaries. The study follows a two-step approach. Initially, an autoencoder neural network architecture is utilized to encode the travelogues into a lower-dimensional latent space representation. The compressed representation obtained from the encoder fed into the decoder, which reconstructs the original travelogues. Subsequently, the encoded model is employed to train diverse machine learning models, including logistic regression, decision tree classifier, SVM, RFC, KNN, SGD, and MLP. The model attained an impressive validation accuracy of 95.84%.

**Keywords:** Autoencoders, self-supervised learning, Recommender model, Natural language processing, Malayalam.

---

## **A Journey through Multifaceted Data in Machine Learning Predictions on Financial Viability**

**R. Revathi**

**Abstract.** This research explores the application of various machine learning models for credit risk assessment in the context of a highly imbalanced dataset. The study encompasses logistic regression, random forest, gradient boosting algorithms (LightGBM and XGBoost), neural networks, and ensemble methods. An extensive data processing pipeline, including categorical encoding, missing value imputation, and feature scaling, is implemented to enhance model performance. Evaluation metrics, primarily focusing on AUC-ROC scores, are used to assess model accuracy. The findings reveal that gradient boosting algorithms, specifically LightGBM, outperform other models, achieving the highest AUC-ROC score of 0.80123. Deep learning models, such as TabNet, demonstrate potential but require further development for competitive performance. Logistic regression and random forest models yield respectable results, showcasing the significance of feature engineering in enhancing their predictive capabilities. Ensemble methods, including stacking and weighted averaging, are explored to leverage the strengths of individual models. The study emphasizes the importance of model selection and parameter tuning in achieving optimal results. Logistic regression, despite its simplicity, performs well due to the incorporation of a feature-engineered pipeline. In conclusion, the research suggests that, while current deep learning models for tabular data show promise, gradient boosting algorithms remain superior for credit risk assessment. The study highlights the need for further exploration of advanced techniques and emphasizes the potential of regression algorithms for a comprehensive credit assessment score, complementing the classifier algorithms discussed.

---

## **Neural Network Based Employability Index for IT Graduates**

**Ankita Chopra and Dr. Madan Lal Saini**

**Abstract.** With the rapid development of information technology and the internet, employment prospects for IT graduates are evolving. While demand for software developers and engineers remains strong, entrepreneurship is an increasingly appealing option for IT graduates. This paper proposes an optimization model for employment and entrepreneurship guidance for university IT graduates using credible neural network technology. A survey of IT graduates was conducted to understand needs and challenges around employment and entrepreneurship. Based on the survey results, a credible neural network model was developed to provide personalized guidance on employment and entrepreneurship. The model utilizes deep learning techniques to analyze career preferences, skills, interests and other traits of graduates. It then matches graduates to suitable job openings, entrepreneurial opportunities, training programs and networking events. The network continuously learns from user interactions and feedback to improve recommendations over time. Experimental results demonstrate that the optimized guidance model achieves 63% user satisfaction compared to 51% for a baseline model without optimization. The proposed approach enables universities to better support IT graduates in navigating the evolving employment landscape and showing the impact of neural networks in understanding the employability index for IT graduates.

**Keywords:** IT graduates, credible neural network technology, baseline model

---

## **Future-Proofing Careers: The AI Era and the Imperative of Reskilling**

**Garima Jaiswal, Mala Saraswat, Drishti Rai, Shristi Agarwal**



**Abstract.** Foundational disciplines of the wide, interdisciplinary study of Artificial Intelligence (AI) include statistical inference and data, software application, rational analysis, information theory neurology, linguistics, and information theory. The way AI is implemented is a crucial factor in how it will change the workforce. It follows that AI systems can grow. Unlike earlier systems that typically involved human labor, these systems are intelligent and capable of learning beyond pre-defined structures and logic. This creates the opportunity for AI systems to take on formerly reserved for humans, experience, and understanding tasks. Organizations' prospective use of cutting-edge AI technologies raises serious questions concerning both security and the welfare of their workers. In the future, technology will probably replace a significant portion of the workforce, which, if not effectively managed, might result in shockingly high unemployment rates throughout society. The skills required today might not be significant in the future as much. Our research focuses on the future of work and how artificial intelligence (AI) will change it, particularly in terms of job displacement, reskilling, and upskilling. Everyone agrees that AI is profoundly changing the workplace and is poised to redefine what it means to be a worker. Artificial intelligence (AI) has the potential to boost production and efficiency, but there is also a concern that it may lead to the replacement of human labor as computers take over tasks that they were previously performing.

**Keywords:** Artificial Intelligence, Automation, Job Displacement, Reskilling, Upskilling, Workforce Development

---

## **An Image Processing-Based Tomato Leaf Disease Prediction Using Deep CNN**

**Vanisri , K Srujan Raju , Bagam Laxmaiah**

**Abstract:** Pest outbreaks have a significant impact on plants and crops, important to a reduction in national agricultural productivity. Conventionally, farmers and specialists rely on physical observation, and it can be time-consuming, more expensive, and prone to errors, to classify diseases in plants. Though, by employing leading-edge image processing techniques, an instinctive detection system can be established, providing rapid and precise results. This paper aims to produce a Tomato Leaf Disease identification model with different categorizing leaf images and utilizing deep neural networks. The development of specific plant protection has the potential for substantial growth and improvement, with computer vision applications in the precision agriculture market also showing promising growth. This discussion about outlines all the critical stages for implementing the disease recognition model, initial from the image database group, assessment by agricultural specialists, and applying a deep learning framework to conduct the training of deep CNNs. By using this CNN, we will apply Max pooling layer, Multiple implementations are applied and finally we will get more accuracy.

**Keywords:** Diseases, Deep CNN, Dataset, neural network.

---

## **Technical Session: 06 Advances in Computational Linguistics and NLP**

### **Text Scribe: Unveiling New Dimensions in Text Summarization**

**Dhananjay Bhagat, Kalyani Sawarkar, Dewanshu Barde, Rashi Channawar, Mayur Talmale, Prathamesh Kale, Rahul Agrawal**

**Abstract.** The massive quantities of facts available have made it hard to retrieve and understand vital information efficiently. This has brought the many improved techniques of text summarization over the past few years which includes shortening of long texts and maintaining the essence of the information present in the text. Summarizing the text includes records, and making it simpler for humans to apprehend. The cause of this venture is to create a precise set of rules a good way to concisely summarize textual content yet nevertheless consist of essential pieces of records. The machine might need to be in a position to extract applicable data, distinguish essential issues, and remove incoherent or repetitive factors from the content. This project will involve the construction of a scalable, dependable and unique textual content summarization device able to be applied to lots of fields consisting of articles, scientific papers, files, and social media posts. A modern-day text summarizer using cutting-edge (NLP) strategies, which includes deep learning and neural networks, will be provided for imparting honest precis. There, it will use the text because the input and generate a summary that encapsulates the main concepts and important facts from the original textual content. In order to assess the overall performance of the machine, precis of nice contained therein accuracy, coherence, and clarity. The end result of this assignment is an effective and dependable machine for text summarization that can be used in different fields for enhancing accessibility and performance in data provision to people. In addition, individuals who continuously have interaction with large quantities of texts can preserve some time, recognize informative texts better, and improve their productiveness.

**Keywords:** Deep Learning, nlp, Machine Learning, TFIDF, LSA

---

## **Context Based Anaphora Resolution of English Discourses using Rule Based Approach**

**Kalpana B. Khandale and C. Namrata Mahender**

**Abstract.** Anaphora is a common phenomenon in discourses and a significant area for research in the use of natural language processing. In this paper we have concentrated on context within the discourse and trying to resolve anaphora with the rule based approach. In applications of NLP like machine translation, summarization, question-answering systems, sentiment analysis, and others, anaphora resolution is essential. In this article, we make the case that many common adverbial phrases that are typically thought to indicate a discourse relation between syntactically related units within discourse structure really add relational meaning anaphorically. So many work has been done on anaphora resolution for simple sentences but very less amount of work has been done on complex English discourses. We have developed manual dataset of 100 complex English discourses to resolution of an anaphora from discourses with appropriate rules and algorithm which we have developed on the basis of grammar. The system achieved 94% accuracy.

**Keywords:** Anaphora Resolution, English Discourse, NLP

---

## **Performing Text Segmentation to Improve OCR on Multi Scene Text**

**Arrun Sivasubramanian<sup>1</sup>, Sheel Shah , Akash Narayanaswamy, Rindhya C , and Barathi Ganesh HB**

**Abstract.** Optical Character Recognition (OCR) is a revolutionary technique that aids machines in retrieving textual content from images to perform further analysis. However, OCR has its limitations, especially when dealing with degraded or low-quality images, which can impact the overall reliability of the text recognition process. Thus, the system's accuracy is contingent upon the quality of the input (digital or handwritten documents). Efforts to modify the text detection and text recognition modules in existing OCRs fail to work in complex dynamic environments due to the complexity of the background information of the input data. Thus, a new first-of-its-kind annotated dataset called OCR-SBT for digital text segmentation is proposed in this work, along with a novel preprocessing pipeline using deep learning that performs text retrieval from images having varying and complex backgrounds using binary semantic segmentation. With quantitative metrics such as the DICE coefficient as high as 99.56%, the qualitative performance improvement of OCR has also been validated on real-world test samples containing varying contextual information to validate the model's efficacy. Ablation experiments are also performed to determine the importance of super-resolution of input images using Stable Diffusion and ESRGAN. This work will help the research community to improve OCR for several real-world applications by alleviating the problems related to background contextual information obfuscating the text recognition module. The dataset and the codes will be made publically available at the Github Link: <https://github.com/argon125/OCR-SBT-Performing-Text-Segmentation-to-Improve-OCR>

**Keywords:** Optical Character Recognition · Foreground Extraction · Text Segmentation · Language Understanding · Super Resolution

---

## **Optimization of Indian Sign Language Detection using Data Generators**

**Kata Venkata Sai Krishna , Goli Abhilash , Devanjali Relan , Kiran Khatter**

**Abstract.** Communication via Sign Language is important for people who can't speak and hear. However, ordinary people find it difficult to communicate through sign language with hard-of-hearing people. In this paper, we propose to develop an Indian Sign Language recognition system that converts sign language into regular text. The study uses Convolutional Neural Networks for Indian sign language detection and data generators for optimizing memory and CPU utilization. The research uses 42000 images of 35 signs, including A to Z alphabets and 1 to 9 numbers. With this collected dataset of images, we have experimented with CNN architecture by changing the different hyperparameters in terms of filters and optimizers. The proposed model performance is evaluated using a confusion matrix, precision, recall, and F- score. A model trained with an SGD optimizer achieved an accuracy of 99.99%. Further, a GUI-based tool is developed to recognize Indian Sign Language hand gestures in real-time and convert the sign language to normal readable language.

---

## **An Inclusive Analysis on Deep Learning hinged Malware Detection Techniques**

**Dr. Vinisha Sumra, Dr.Naveen Malik, Santosh Kumar**

Abstract. Malwares are touching the sky in connection with their numbers and this accretion is contemporaneous with the amelioration in usage of Android smart phones. Nearly everything is a blessing and curse at the same time. This applies to innumerable applications available in play stores too. Malware developers earmark this platform to consummate their thrust. These malicious programs can torment the devices in disparate ways such as absconding with the privileged information, superintending the device, wreaking havoc on the battery and maneuvering the data stockpiled in the device. Malware has become proficient in coping with various customary detection techniques. Static detection methods are powerless to seize the malware that utilizes dynamic features such as network traffic features whereas dynamic detection procedures discover it strenuous to detect malware that lodges static features like permissions, images and intents etc. Therefore, the exigency surfaces to maneuver some advanced methodologies in this direction. Deep learning coupled with numerous techniques fulfill this objective. This paper analyzes numerous deep learning techniques which are worthwhile for the detection of these stubborn malwares. The analysis throws light on various deep learning models also.

Keywords: Malwares, deep laerning, machine learning, static analysis, dynamic analysis, hybrid analysis, Android, neural network.

---

## **Advances in Text Summarization Techniques: A Comprehensive Review and Future Prospects**

**Harsh, Yashika Kukkar, Aditi Chauhan, Abhinav Maheshwari, Dharmender Saini, Preeti Nagrath**

Abstract. This research work provides a concise comparative analysis of prominent content summarization models, namely Textrank, T5, Pegasus, and Bart. The models are evaluated for their specific strengths: Textrank excels in sentence extraction, T5 showcases versatility, Pegasus demonstrates superior abstractive summarization, and Bart proves robust across diverse content types. Evaluation metrics, including ROUGE scores and human assessment, offer insights into each model's performance. Beyond performance, the paper considers computational efficiency. The study contributes to content summarization by incorporating machine learning and LSTM techniques, advancing automated text processing. This concise reference aids professionals in selecting optimal summarization models for specific use cases and resource constraints. Evaluation metrics, including ROUGE scores, highlight Pegasus as the top performer out of the implemented models. Pegasus demonstrates exceptional capability in generating summaries aligned with references.

Keywords: Long Short-Term Memory (LSTM), Encoder, Decoder.

---

## **Natural Language Generation (NLG) with Reinforcement Learning (RL)**

**Shakir Khan, Mohammad Serajuddin, Zafarul Hasan, Syed Ahad Murtaza Alvi, Rashid Ayub, Arun Sharma**

Abstract. Finally, utilizing an interpretative framework as well as a deductive method this study studied the use of reinforcement learning strategies in the development of natural languages. Although secondary data collection provided a solid foundation, the descriptive method allowed for a complete study. The technical solutions used included the incorporation of RL computer programs, advanced text processing, linguistic analysis, contextual adaptability, and moral concerns. As the consequence of our research, we now have a better understanding of RL-NLG systems, which opens the way to more sophisticated and environmentally conscious text production. This research not only enhances the field of NLG, although it also emphasizes the importance of ethical as well as objective applications for artificial intelligence. The insights have application that encompass intelligent machines to the generation of customized content.

Keywords Reinforcement Learning, Natural Language generating, RL-NLG systems, mechanism, disinformation penalized, Python modules, RL engine

---

## **Revolutionizing Writing: Personalized Neural Classifier for Handwritten Text**

**Nimish Goel, Shushil Paudel, Srushti Jagtap, and Ishan G. S. Kakodkar**

**Abstract.** This research presents an innovative method to enhance the effectiveness of a Convolutional Neural Network (CNN) classifier when dealing with offline handwritten text. By adapting to the writing style of individual users, this method achieves lightweight and privacy-preserving classification without requiring specialized deep-learning frameworks or hardware on the user's end. The training phase, which takes place on the server side, involves training the base CNN and defining the writing styles of different characters through clustering in the feature space learned by the CNN. In the adjustment phase, implemented on the user side, the system determines the user's writing style based on their writing and corrections of recognitions offered by the base CNN. This information is integrated into an ensemble of alternative k-NN classifiers with a lightweight Bayesian selection mechanism that learns when to correct the base CNN. Essentially, this ensemble construction mechanism conducts automated error-space analysis of the base CNN to improve its decisionmaking. To evaluate the proposed method, we tested it on two widely-used datasets - NIST Special Database 19 and ETH Zurich Deepwriting dataset - achieving up to 2.7 percent improvement in classification accuracy. This resulted in state-of-the-art classification results on these datasets.

**Keywords:** Convolutional Neural Network (CNN), Privacy-Preserving Classification, User Adaptation, Ensemble Classifier, k-NN classifier

---

## **An Efficient Deep Learning based Seq2Seq Model for Abstractive Text summarization**

**Deepa. R ,Dr. T. Sree Sharmila and Dr. R. Niruban**

**Abstract.** The stacks of literal verse contributed has increased drastically in recent millennia, which creates a opulence of particulars for exploration and the extortion of knowledge. The consumption and processing of the ever-growing volume and complexity of text data produced daily, including Web broadcast, Editorial pieces, E-messages, and messaging are rendered challenging by people due to the overwhelming nature of the information. However, this also means that it can be prolonged and rigorous to manually sift over generous textual quantity to discern pertinent insights. Text summarization overcomes these hurdles by instinctively condensing text into a shorter, more manageable form, allowing users to rapidly acquire the Critical insights. Amid the prevailing data-centric conditions, this emerged as a vital tool that is applied in a variety of context like Analysis of insights, personalized content endorsements, and knowledge recuperation. This article proposes an Abstractive Text Summarization (ATS) framework which utilizes a combination of Seq2Seq model with Long ShortTerm Memory (LSTM) and Convolutional Neural Network (CNN), bidirectional LSTM and Hybrid LSTM with attention to construct new sentences. This text summarization framework, referred as Deep Learning based Abstractive Text Summarization (ATSDL) which yields summarization from origin sentences using deep learning and is designed to explore more fine-grained fragments of text, such as semantic phrases, to create summaries. A massive corpus of text data called the News summary dataset is frequently used to train algorithms for multi-sentence summarizing tasks. To assess the quality of text produced by the deep learning model, two assessment metrics that are often employed in the natural language processing discipline are ROUGE and BLEU. Experimental results on news summary and CNN/DailyMail dataset revealed that ATSDL method achieves recent contemporary models in aspects of Grammatical structures and contextual relations, as well as having challenging outcomes in manual textual performance evaluations.

**Keywords:** Text summarization, LSTM, Bidirectional LSTM, Abstractive Text Summarization

---

## **Efficient Real-Time Indian Sign Language Fingerspelling Recognition in Natural Settings using Heuristics**

**Raghuveera T , Akshayalakshmi V K, Nisha B A, Easwarakumar K S**

**Abstract.** Sign language is the primary mode of communication for Hearing and Speech Impaired (HSI) people. However, the complexity and intricate nature of Indian Sign Language, which includes a majority of double-handed signs, poses a challenge for HSI people to communicate effectively with others. Moreover, the expanding vocabulary of sign language makes it difficult for those without access to updates to communicate effectively. Fingerspelling is most widely used by the HSI for general and easy day-to-day communication. A real-time and efficient fingerspelling system is thus crucial to facilitate communication for HSI people in a natural setting. However, existing real-time recognition systems are cumbersome and inefficient as they employ complex deep-learning architectures and primarily use RGB image and video data that are sensitive to lighting and background conditions and therefore are more error prone and moreover do not perform well under natural settings. This study proposes a simple and efficient real-time fingerspelling system for recognizing static fingerspelling gestures using Leap Motion Controller. The study employs a random forest classifier with translation-independent features to recognize signs, while achieving comparable accuracy, making the overall system lightweight. We achieve a real-time validation accuracy of 71% while also predicting the sample instantaneously with an average response time of 3.02 milliseconds. Since fingerspelling can introduce

spurious signs during transitions and can be ambiguous when recognizing similar signs, our system also includes a word fine-tuning phase that uses a dictionary-based approach to simplify the recognition process, making our system well-suited for real-time deployment in natural settings.

Keywords: Indian Sign Language, Fingerspelling, Natural Setting, Real-time, Dictionary approach, Heuristics.

---

## **Detecting Image-based Fake News with Neural Sleuths**

**Mr. Nachiket Rathod and Dr. Prabhakar Ramteke**

**Abstract:** In the period of extensive social media use, unconventional formats for news sharing often lack reliability, with misleading information disseminated through graphics emerging as a unique tactic. While the verification of the authenticity of news is a broad research domain, scholars have mainly focused on textual data. The misleading information detection model described in this paper uses neural sleuth-based networks to capture image data from the freely accessible Fakeddit dataset. The results are scrutinized using a confusion matrix, and a comprehensive evaluation of the model's performance metrics is conducted across six categories: "true," "satire," "false connection," "imposter content," "manipulated content," and "misleading content." This work contributes an inclusive analysis of the model's effectiveness as a problem posed by misleading information on social media.

Keywords: Deep Network, Image Fake News, Convolution, Neural Network.

---

## **Technical Session: 07 Exploring New Horizons in Computer Vision Research Video Summarization using Firefly Algorithm**

**Shashank Pandey , Sakshi Dwivedi , Vijay Bhan Singh, Neetu Verma , and Ranvijay**

**Abstract.** In this research, Firefly algorithm to generate summary of a given video is proposed. The Firefly Algorithm is an optimization algorithm that draws inspiration from the behaviour of fireflies. There are many types of existing video summarization techniques. Here in this implementation, we focus on summarizing the video based on Object of Interest. Object of interest-based video summarization is a technique for identifying and choosing key portions of a video based on the existence of specified items or regions of interest within the frames. This method involves identifying and tracking relevant things in video frames, and then picking keyframes or short video clips containing those objects or regions of interest. In a surveillance video, for example, the item of interest could be a person or a vehicle entering a restricted area, and the summarization algorithm would select the frames or clips in which the object is present to provide a summary of the video's relevant events.

Keywords: Object Of Interest, Firefly Algorithm, Video Summarization.

---

## **Enhancing Mint Plant Disease Detection Accuracy through Deep Reinforcement Learning with YOLO Algorithm**

**V.Manikandan, P. Dinesh, M. Chantilyan , C. Gomathi**

**Abstract-** This paper proposes a unique way for mint plant disease diagnosis that improves accuracy by employing the You Only Look Once (YOLO) algorithm with deep reinforcement learning. The YOLO method is used to extract differentiating traits from mint plant photos, portraying both healthy and damaged plant portions properly. This data is then analysed by a deep reinforcement learning system to identify diseases precisely. Extensive tests on a large collection of mint plant photos with annotated disease conditions proved the usefulness of the suggested technique. The results demonstrated a significant improvement in illness diagnosis accuracy when compared to conventional approaches. The YOLO method was successful in recognizing substantial trends, and the reinforcement learning component improved the detection process even further. This study has the potential to change mint plant disease diagnosis in the agricultural industry by offering a dependable approach for early disease identification. This enables swift action to limit disease spread and increase agricultural productivity. The research lays the path for comparable applications in other domains of agriculture.

Keywords: agriculture, mint, plant disease detection, deep reinforcement learning, you only look once algorithm, diagnosis, accuracy, swift action.

---

## Multiple Object Detection and Tracking Using DeepSORT

Rashmi Gandhi , Yash Pasrija , Amisha Chauhan , Ritu Rani

Abstract- The traffic surveillance and control have always been a major concern. Detecting and tracking vehicles or automobiles through satellites and surveillance cameras has made the task of tracking the defaulters and managing the vehicles on the roads a lot easier. Only detecting a moving object is not enough rather keeping track of it could possibly help in better management of traffic. So, the main objective is to detect and track the automobiles, person on the roads and provide an effective model for the same. The main challenge is to increase the detection count of the objects so that the tracking could be done efficiently and for that purpose we need to use proper detection algorithm. The dataset with around 4680 drone images has been used for the model training. After training the detection models; YOLOv5, YOLOv7, YOLOv8 the results are fed to the tracking algorithm i.e. DeepSORT. The results of the experiment reflects that the proposed object detection method has higher accuracy for faster real-time detection than the other models and the number of objects detected are also higher which helped in efficient tracking. Among all the detection algorithms YOLOv8 produced the best results with higher accuracy and maximum detections.

Keywords- Object Detection, Object Tracking, YOLOv8, DeepSORT, RCNN

---

## Background Subtraction Model with Advance GMM in Dynamic Background

Manisha, Avadhesh Kumar, Dileep Kumar Yadav

Abstract. The identify objects in video is a crucial task that faces numerous challenging issues due to the real-time environment challenges. These challenges are illumination variation, cluttered background, rain, fog, dust, mist, etc. Such works are very helpful for security surveillance, transportation, agriculture, manufacturing, medical, healthcare, robotics, and multimedia applications. The BGS approach is an effective way of separating the foreground region from the background scene. It's often used in applications that operate in real-time for surveillance systems, motion detection, and recognition, or gestures-based pixel classification BGS that works by removing the static or dynamic background from the current frame to detect the foreground moving objects. So, this work focuses on video datasets to detect objects and remove the background scene by applying a background-subtraction approach. This work also focuses on the lack of a detailed description of the considered dataset. Here, a deep learning technique and BGS method are applied for object detection. This work enhances the existing GMM method and also compares it with advance GMM. The proposed advance GMM based method provides the better outcome as compared with existing method.

Keywords: BGS model, GMM, Deep learning Model, intelligent video analytics

---

## A Comprehensive Review on Impact of Computer Vision in Healthcare Domain

Aiman Zakir, Jeevisha Shukla, Manjusha Iyer and Pooja Bhati

Abstract: In today's era, Healthcare is evolving due to the current technology boom and changing patient needs. This review highlights the effect of Computer Vision (CV) in the healthcare domain, mainly focusing on the progress in medical imaging. This work enumerates the impact of CV in various medical and healthcare areas such as cardiology, pathology, dermatology, ophthalmology, neurology, psychotherapy, nephrology, and surgery. This integration of CV learning techniques has also played a role in healthcare progress during the recent pandemic, including using advanced surveillance tools like drones and cameras to monitor mask-wearing and social distancing. This paper also highlights the challenges of applying CV in real-world clinical settings.

Keywords: CV, Healthcare, Medical Imaging

---

## A framework for Information Retrieval using Domain Specific Dictionary: Illustration through enhancing the Intelligence Cycle

Lakshmish Nayak, Bharathi Ramudu, Malay Kumar Nema and Rituraj Kumar

**Abstract.** The present work aims at relevant information retrieval in context of a chosen domain. The paper proposes a framework which employs dedicated Domain Dictionary for this purpose. It is anticipated that the proposed framework will make the Intelligence (INT) cycle efficient. The details on the framework modules have been provided. The proposed approach to realize the identified components has been elaborated. In a military context, the module has been evaluated and observed to be useful in Situation (SIT) Assessment. In a nutshell, leveraging the proposed framework and automating the process of retrieving relevant information specific to a given domain will lead to an enhanced situational assessment with INT cycle by making it significantly faster.

**Keywords:** Domain Dictionary, Relevancy Tagging, Situation Picture Generation.

---

## **Technical Session: 08 Applications of Machine Learning and Deep Learning in Healthcare**

### **Tuberculosis Detection and Classification in Chest X-ray Images Using Optimized CNN Architecture**

**Apeksha Koul, Rajesh K. Bawa, Yogesh Kumar**

**Abstract.** Tuberculosis (TB) is a global threat to public health, and pathology is one of the most important diagnostic tools to treat it in clinical practice. However, due to the small size and the large number of bacilli bacteria, its treatment is time-consuming and requires an experienced pathologist. Thusly, in this age of AI, the development of modern computer techniques has escalated the process of diagnosing tuberculosis. In this article, Convolutional Neural Network (CNN) model has been developed to detect as well as classify tuberculosis in chest X-ray images. The model is trained using 700 TB and 2800 Normal images (publicly available images) from the Tuberculosis chest x-ray (CXR) dataset. Initially, the images are pre-processed by reducing their size and increasing their contrast to reduce noisy signals. Later, several parameters and layers are added to the CNN model for improving the prediction rate of tuberculosis detection. The applied proposed model achieved a 91% prediction rate with a minimal loss value, which shows a better prediction than the existing techniques applied on the CXR dataset.

**Keywords:** Chest X-Ray, Artificial Intelligence, Tuberculosis, Learning Model, CNN

---

## **i-BMD: AI-based Opportunistic Screening for Osteoporosis on Abdominal CT using Deep Learning**

**Soumya Jindal , and Meemansa Jindal**

**Abstract.** Osteoporosis is a common curable ailment, more commonly seen in geriatric population, particularly prevalent in post-menopausal women. It often presents with generalized bone pain, and fragility fractures, most common in the distal radius, vertebrae, and pelvic bones. When it comes to diagnosing osteoporosis, the best and widely accepted golden standard is the DEXA scan, often interpreted in the form of a T-score and a Z-score of less than -2.5. Though, even in the modern era of evolved imaging, DEXA machines continue to be scarce and are not so profitable to be installed, leading to a massive population of the elderly having untreated and under-diagnosed osteoporosis. CT scan is one of the most common imaging techniques utilized in India for a variety of indications. There has been a lot of buzz regarding the discussions over – “What if one could utilize HU values for diagnosing osteoporosis?” “What if one could predict the risk of fragility fractures without letting one happen?”. A lot of research lately has been focused on employing ML algorithms for achieving this, by using texture analysis of the vertebral body, cortex and trabecular pattern, HU values, and also by using Genant classification of vertebral body fractures. This article aims to review the current situation of opportunistic screening of osteoporosis worldwide and the implications of ML in achieving the same.

**Keywords:** Osteoporosis · Deep learning · CT · DEXA

---

## **Machine Learning Usage for Knee Condition : Osteoarthritis and TKA, using Image Processing Among Medical Patients**

**Snehal Moghe and Dr. Harsh Pratap Singh**

**Abstract.** Artificial Intelligence is gaining grounds in not only technical areas of applications like IT Industry, Management areas like Logistics Management, Human Resources, Fintech, but also, Biomedical, Healthcare and Hospitality services. To the extent that the patients are getting automated phone calls for doctor appointments, medicine refilling, body check-up routine, etc. Also, huge amounts of data which is generated in forms of X-Ray images, CT Scans, MRIs, Sonography, Blood Test reports, Urine Reports, etc are all used for Healthcare analytics and predicting future behaviour of individual patients visiting the Hospital concerned. Moreover, AI has also entered the Operation Theatre (OT) to assist doctors, best example is Total Knee Arthroplasty, where robotic arm is programmed to perform robotic surgery on knee and do knee replacement. In this paper we intend to review systems where we have AI based model (on human body parameters) compared with doctor's analysis on the same parameters. We then formulate the results and come up with concrete solution for the problem. Literature review has clearly brought down that there are different aspects of Healthcare and AI which is a very vast domain in itself, but narrowing it to Knee Operations and Osteoarthritis, use of AI as power of Prediction, it is quite a game changer worth the efforts as it has gained much importance technically both in medical as well as computer science field.

**Keywords:** Artificial Intelligence, Medical Data, Healthcare Analytics, Medical Data, Reports, Doctors, Big Data, Data Science, Knee Operations, Image Processing, Sensors, Robotics, Robotic Arm

---

## **Machine Learning Techniques to Study: Medical and Biomedical Signal Analysis**

**Laxmi Ahuja and Ayush Thakur**

**Abstract.** This paper presents a study of machine learning techniques for medical and biomedical signal analysis. Measurements of physiological processes that reveal the state of health of living things are known as medical and biomedical signals. These signals are difficult to analyse using traditional signal processing techniques because they are frequently influenced by noise, non-stationarity, complexity, and high dimensionality. Algorithms that use machine learning may learn from data and provide hypotheses or predictions based on the patterns they spot. This paper examines several common electrocardiograms and electroencephalograms, among other biological and medical signals, and demonstrates how machine learning techniques may be used to analyse them for a variety of purposes, including diagnosis, treatment, monitoring, and research.

**Keywords:** Medical signals · Biomedical signals · Machine learning · Signal processing · Healthcare applications

---

## **Machine-Learning Algorithms for Predicting Cardiac Arrhythmias**

**Sakshi Shete, Suvarna Udgire, Anuradha Joshi, Shravani Desai, Ronit Murpani**

**Abstract.** Cardiac arrhythmia, a frequently encountered cardiac disorder characterized by irregular heart rhythms, poses a substantial risk to a patient's well-being. This research explored various algorithms for forecasting cardiac arrhythmias derived from ECG datasets available on Kaggle. Employing machine learning techniques, the study achieved significant accuracy rates: Logistic Regression (71.43%), Random Forest Classifier (58.24%), and Decision Tree Classifier (65.93%). Extracting the different characteristics of cardiac electrograms (ECG), focusing on R peak, QRS duration, and T interval, forms the foundation for arrhythmia categorization. The results are employed to transform arrhythmia management, emphasizing accuracy and key features, with Logistic Regression leading at 73.28%, Random Forest Classifier at 76.82%, and Decision Tree Classifier at 61.22% in AUC-ROC scores. This research contributes to advancing data-driven methodologies in cardiac care, enhancing patient outcomes by facilitating early detection and personalized treatment plans.

**Keywords:** Electrocardiography (ECG), R-R Interval (Peak-to-Peak Interval), Heart Rate (Beats per Minute), P-Wave to QRS Peak Interval, Convolutional Neural-Networks (CNN), Artificial-Neural- Network (ANN).

---

## **Data Mining Based Heart Disease Prediction using Hybrid Optimization technique of derived features**

**Kanchan A. Khedikar and Dr. Piyush Kumar**

**Abstract:** Data mining difficulties include improving forecasts, medical domain data from clinical data, and research observations. This section analyses and discusses medical data mining applications and research. Many medical treatment issues are investigated using data mining. One of the biggest challenges for medical service providers is providing high-quality treatment at low cost (such as hospitals and medical centres). Excellent service involves accurate patient evaluation and therapy



management. Clinical choice is crucial to diagnosing a patient's health. The hospital management must do many clinical tests to reduce treatment costs. Data mining uses clinical data and scientific observations to improve medical data and projections. This article highlights several of the most critical issues facing this business for massive data collections and streams. The suggested study examines the dataset size, cost sensitivity, location, and heart disease prediction factor relevance. In sequence, the proposed activity incorporates dataset collection, feature extraction, feature selection, feature fusion, and data classification. Feature Selection is the most important phase in data mining using Hybrid Grey Wolf and Particle Swarm Optimization (HGWPSO). Classifying Cardiovascular Disease using an Enhanced Probabilistic Neural Network (IPNN). Because of this, we aim to highlight sensitivity vs. specificity in the graph despite the data's high accuracy. Iterating over the standard classifiers yields the following results: the area under the ROC curve should be a maximum of 0.9063 and the accuracy should be 96.667%.

Keywords: Hybrid Grey Wolf with Particle Swarm Optimization, Improved Probabilistic Neural Network, Data Mining

---

## **Unveiling the Enigma: Sudden Cardiac Arrest among Youth and Uncovering Underlying Causes: A Data Analysis Study**

**Mohit M Rao, and Virendra Kumar Shrivastava**

**Abstract.** This study focuses on analysing electrocardiogram (ECG) records related to sudden cardiac death in youth and analyse the factors influencing a sudden cardiac arrest. Sudden cardiac arrest poses a severe threat to life and can result in death if not addressed promptly. Recent American research indicates a 13% increase in sudden cardiac arrest cases among individuals aged between 35- 45. Furthermore, the Indian Heart Association has reported that heart disease tends to affect Indians at a younger age compared to other demographic groups, often without prior warning. Studies reveal that Indians tend to experience heart diseases approximately a decade earlier than their Western counterparts. This work aims to conduct data collection and surveys to collect ECG records and other relevant information that is crucial for training the machine learning models to achieve the study outcomes. The study's scope encompasses evaluating the feasibility of the task, identifying relevant datasets, conducting data preprocessing, and assessing the performance using K-Nearest Neighbours (KNN), Support Vector Machine (SVM), and random forest algorithm. Among these, both SVM and Random Forest demonstrated superior performance, with the Random Forest algorithm further fine-tuned to align with the dataset's specific requirements. The refined model exhibited 93% accuracy. **Keywords:** Sudden cardiac arrest, sca, ecg dataset, heart failure, image processing, chi-square test, machine learning, artificial intelligence.

---

## **Using Encoder Models to predict Ratings Using Flipkart Reviews**

**Himani Agarwal and Shweta Singhal**

**Abstract.** Humans are highly subjective beings and present their opinions on products and services in the form of reviews, which need to be analyzed to understand the consumer base. However, deciphering the consumer sentiment from online reviews is a massive task due to the diversity in the reviews and their un- structured nature. To tackle this challenge, in this study we employ the BERT (Bidirectional Encoder Representations from Transformers) model to predict numerical ratings from textual reviews from a Flipkart product review dataset sourced from Kaggle. Our findings demonstrate BERT's suitability for analyzing diverse and complex textual data, with an impressive accuracy of 73%, due to its ability to employ contextual analysis. This study showcases the potential of utilizing advanced NLP techniques to derive insights from otherwise huge and un- decipherable data, which can be used to make well informed decisions by business looking to maximize gains and customers looking for make better purchasing choices alike.

Keywords: BERT, Predicting Ratings, Natural Language Processing, Textual Data, Machine Learning

---



# Indira Gandhi Delhi Technical University for Women

(Established by Govt. of Delhi vide Act 9 of 2012 : ISO 9001:2015 Certified University)

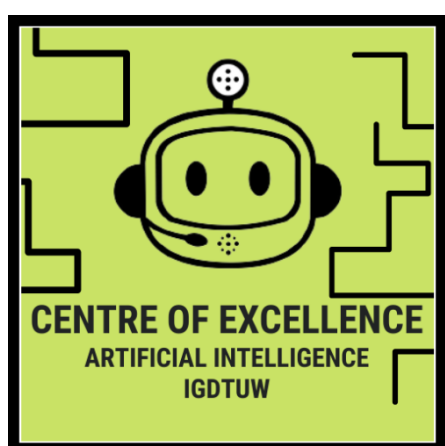
Kashmere Gate, Delhi -110006

**Women Education | Women Enlightenment | Women Empowerment**

## ABOUT IGDTUW

Indira Gandhi Delhi Technical University for Women (IGDTUW) was established by the Govt. of Delhi in May 2013, as a non-affiliating teaching and Research University at Delhi. The aim of the University is to facilitate and promote studies, Research, Technology, Innovation, Incubation and extension work in emerging areas of professional education among women, with focus on Engineering, Technology, Applied Sciences, Architecture and Management with the objective to achieve excellence in these and related fields. In last few years, the University has achieved a number of milestones, to name a few recent ones are:

- Ranked 40th in overall top 100 Universities globally and 8th in Special Category of Entrepreneur Spirit by World University Ranking Impact (WURI) 2023.
- IGDTUW received the highest 'AAA' Rating for NPTEL & Swayam Courses and was ranked among top 10 institutions in India for December, 2022.
- Awarded with Grade "A" by NAAC with score 3.24
- IGDTUW has been recognized as Zero Waste and ZERO Plastic Campus by Municipal Corporation of Delhi (MCD) on 15th March, 2023.
- IGDTUW has been conferred with New Code of Education Award by AICTE for 2022 India Today Rankings 2022.



## CENTRE OF EXCELLENCE ARTIFICIAL INTELLIGENCE

**SUPPORTED BY  
DEPARTMENT OF SCIENCE AND TECHNOLOGY, GOI**



सत्यमेव जयते

Department of Science & Technology  
Govt. of India

IGDTUW has established Centre of Excellence in Artificial Intelligence with the support of the Department of Science and Technology, GOI. Under the magnanimous leadership of Dr. (Mrs.) Amita Dev, Honourable Vice Chancellor IGDTUW, avant-garde and state of the art infrastructure with computing facilities/Research Centers with modern AI-enabled infrastructure facilities to enhance functional efficiency for Teaching, Research, and Innovation purposes have been created under COE - AI. The Centre of Excellence (CoE) in Artificial Intelligence (AI) at IGDTUW, caters to the requirements of Undergraduate, Post-graduate, and Doctorate programs in the domains of AI, Machine Learning, and Deep Learning and various applications including Robotics, Drones, NLP and others. The Centre serves as the perfect platform with the necessary hardware and software infrastructure to serve as a playground for the creative minds that solve real data driven problems at hand.



# STATE OF THE ART WORLD RENOWNED NVIDIA DGXA100 SERVER



## INFRASTRUCTURE SETUP

- \*Advanced PCs with i7 processor
- \*Interactive panels
- \*Data recording Software
  - PRAAT
  - WAVESURFER
  - MIC/SPEAKER,
  - FFMPEG pre-processing tool
  - KALDI toolkit, ESPNET
- \*Research Lounge
- \*Mini Conference room
- \*Conference Room with all amenities and facilities
- \*Counselling rooms and discussion areas
- \*CCTV surveillance
- Other Software available:
  - Python
  - MATLAB
  - PTC Creo
  - FluidSim

## AREAS OF RESEARCH

Use of Machine Learning and Deep Learning in:

- Computer Vision
- Speech Technology
- Natural Language Processing
- Large Language Models
- Medical Imaging
- Agricultural Innovations

## NVIDIA DGX A100 SERVER

GPUs	8x NVIDIA A100 40GB Tensor Core GPUs	
GPU Memory	320 GB total	
Performance	5 petaFLOPS AI 10 petaOPS INT8	
NVIDIA NVSwitches	6	
System Power Usage	6.5 kW max	
CPU	Dual AMD Rome 7742, 128 cores total, 2.25 GHz (base), 3.4 GHz (max boost)	
System Memory	1TB	
Networking	Up to 8x Single-Port NVIDIA ConnectX-7 200 Gb/s InfiniBand 1 Dual-Port ConnectX-7 VPI 10/25/50/100/200 Gb/s Ethernet	Up to 8x Single-Port NVIDIA ConnectX-6 VPI 200 Gb/s InfiniBand 1 Dual-Port ConnectX-6 VPI 10/25/50/100/200 Gb/s Ethernet

Interested Researchers, UG and PG students, Research Scholars, Faculty and Academicians from all over India are encouraged to utilize the computational power of NVIDIA DGX A100, to facilitate collaborative quality research developments.

For more information please contact: [coei@igdtuw.ac.in](mailto:coei@igdtuw.ac.in)



# IGDTUW Anveshan Foundation

An Incubation Center Promoted by  
**Indira Gandhi Delhi Technical University for Women**

Women Entrepreneurship | Women Empowerment | Women Enlightenment

## Convert Your IDEA to a STARTUP

& Start your Entrepreneurial Journey in support with 'ANVESHAN'

And get benefits like,

<b>Office Space</b> Fully Furnished Air-Conditioned Modular Cubical	<b>ICT Facility</b> Proper LAN Connection, Wi-Fi, Phone lines	<b>Growth Hacking</b> Pre-Sales, Sales & Post-Sales	<b>Seed-Money</b> Seed Fund of up to INR 20 lakhs
<b>Shared Space</b> Labs, Conference/ Seminar Halls	<b>Acceleration</b> Various Area Specific Programs	<b>Mentorship</b> Mentor Pool with 250+ Experts	<b>Investors Connect</b> Regular Investors Connect



For More Information

**Email**

[anveshan@igdtuw.ac.in](mailto:anveshan@igdtuw.ac.in)

**Logon**

<https://anveshanfoundation.org/>

Or, Visit : - IGDTUW Anveshan Foundation, IGDTUW Campus, Kashmere Gate, DELHI - 110006