3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY

(AIST – 2021)

(www.aistconference.com)

12th - 13th November, 2021

SOUVENIR

Publication Partner

Supported by

Organized By

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

Publication Partner

Supported by

(Centre of Excellence – AI, IGDTUW)

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

Dates: 12th – 13th November, 2021

(www.aistconference.com)

Chief Guest

Shri Parimal Kumar
Scientific Advisor to Chief of Integrated Defence Staff

General Chair

Dr. Amita Dev
Vice- Chancellor, IGDTUW

Guests of Honor

Inauguration Function
Ms. Bindu Narayan
Associate Director (Data Science and AI), Ernest & Young Global Limited

Valedictory Function
Dr. Anupam Shukla
Director, IIIT Pune

Keynote Speakers

Dr. Satoshi Nakamura,
NAIST, Japan

Dr. Tanja Shultz
University of Bremen, Germany

Dr. Laurent Besacier
Naver Labs Europe France

Dr. S. Sakti,
JAIST, JAPAN

Prof. S. Umesh
IIT Madras

Dr. Rajkumar Buyya
University of Melbourne Australia

Technical Program Chair

Dr. S.S. Agrawal
Emeritus Scientist, CSIR

Conference Convener

Dr. Arun Sharma
HoD - AI and Data Sciences, IGDTUW

For further details, please visit the conference website: **www.aistconference.com** or write at aist2021@igdtuw.ac.in.

Message from the General Chair

It gives me immense pleasure to welcome all the participants, delegates, keynote speakers, resource persons for Third International Conference on "Artificial Intelligence and Speech Technology" i.e. AIST-2021, to be held at Indira Gandhi Delhi Technical University for Women, Kashmere Gate, Delhi on 12th and 13th November 2021 in an online mode. The first two versions of the Conferences were a great success with the participation of experts from Japan, Hungary, Czech Republic, Myanmar and almost all corner of India.

The 3rd 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. Due to the pandemic situation all around the world, the Conference is being organized in an online mode which has provided us the opportunity to interact with a greater 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-2021 is dedicated to cutting edge research that addresses 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 latest and relevant research findings on almost all the aspects of these domains.

As academicians, the responsibility to nurture complete professionals lies with us. This necessitates the knowledge of 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 endeavors.

I wish the conference a great success.

Dr. (Mrs.) Amita Dev
Vice-Chancellor, IGDTUW
General Chair, AIST-2021

Message from the TPC Chair

It is the matter of great pleasure and happiness to see that Indira Gandhi Delhi Technical University for Women, Delhi is organizing its Third International Conference on Artificial Intelligence and Speech Technology (AIST-2021). 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 the researchers world-wide.

There has been a tremendous advancement and innovations in Artificial Intelligence which is incomparable to what Artificial Intelligence emerged traditionally. We use Artificial Intelligence many times during a 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, processing of bio-signals emitted during speech production, low-resource and multi-lingual speech processing, speech to speech translation systems etc. Looking to its huge hope and dimensions AIST-2021 brings together academics, industrial experts and education leaders from all over the world to discuss an incredibly wide array of topics ranging from Foundation of Artificial Intelligence and machine learning, data mining, 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 session and share their knowledge during the Conference. My warmest thanks go to the organizing committee colleagues including the tutorial program 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-2021.

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

Prof. S.S. Agrawal
Emeritus Scientist, CSIR
Technical Program Chair, AIST-2021

Message from the Conference Convener

I take this opportunity to welcome you all to the 3rd International Conference on Artificial Intelligence and Speech Technology i.e., AIST-2021, to be held at Indira Gandhi Delhi Technical University for Women, Delhi during 12-13th November 2021. 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, Soft Computing and Applications, Speech Analysis, Representation and Models, Spoken Language Recognition and Understanding, Affective Speech Recognition, Interpretation and Synthesis, Speech Interface Design and Human Factors Engineering, Speech Emotion Recognition Technologies, Audio-Visual Speech Processing, IoT Security. The conference received more than 180 submissions from all over the globe, out of which the best 55 selected papers will be presented during these two days. The Conference Proceedings will be published by Springer in its CCIS series.

AIST-2021 is an effort of IGDTUW to share knowledge and current research on Artificial Intelligence and its innovation in Speech technology. 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-2021 Conference includes prominent Keynote addresses by Prof. Satoshi Nakamura (NAIST, Japan), Prof. S. Umesh, Indian Institute of Technology – Madras, Prof. S.S. Agrawal (Emeritus Scientist, CSIR), Dr. Rajkumar Buyya, University of Melbourne, Australia, Dr. S. Sakti, JAIST, JAPAN, Dr. Laurent Besacier, Naver Labs Europe, France, Dr. Tanja Shultz, University of Bremen, Germany. Also, the Technical Sessions will be chaired by eminent experts from AI and Speech Technologies including Prof. Samudra Vijay, IIT Guwahati, Dr. Karunesh Arora, CDAC, Prof. Poonam Bansal, MSIT, Dr. Monica Mehrotra, Jamia Millia Islamia, Prof. D.K. Vishwakarma, DTU, Dr. Mani Madhukar, IBM Research, Prof. Anurag Jain, GGSIPU, Prof D P Mohapatra, NIT Rourkela and Prof P K Behera, Utkal University.

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 DST (Curie Grant) 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
Conference Convener, AIST-2021
Prof. and Head of Dept (AI and DS)
IGDTUW

CHIEF GUEST

Shri PARIMAL KUMAR
Scientific Advisor to Chief of Integrated Defence Staff

Mr. Parimal Kumar, Scientist-G, took over as Scientific Advisor to Chief of Integrated Defence Staff (SA to CISC), HQ_IDS, Ministry of Defence w.e.f. 14th Oct 2019. He is a postgraduate in Science, Engineering as well as in Business Administration. He obtained his MSc (Masters of Science) in Computer Science from University of Allahabad, MTech (Master of Technology) in Computer Science & Engineering from IIT Delhi and MBA (Master of Business Administration) from IGNOU. He joined DRDO as Scientist- 'B' in 1992 and has served in various Scientific & Techno-Managerial positions that include his responsibilities at RAC (Recruitment & Assessment Centre), SAG (Scientific Analysis Group) and DRDO Headquarters. During his tenure at RAC from 07th May 1992 to 30th June 2005, he was responsible for automation of Recruitment & Assessment Processes. He was also instrumental in formulation & implementation of various Recruitment & Assessment related policies during his tenure at RAC. During his first tenure at SAG from 01st June 2001 to 18th April 2006, he contributed in two very important projects namely PATCLAN and PICKPACKET. He is also founder member of the Side Channel Attack Group.

During his tenure at At DRDO Headquarters from 19th April 2006 to 28th March 2017, he served in several Techno-Managerial positions that included TSO to CC R&D(SI), Additional Director/Officiating Director, Dte of Industry Interface & Technology Management (DIITM) and Technical Adviser to Chief Controller R&D(PC&SI) (now DGR&D(PC&SI)). He was instrumental in promulgation of 'DRDO Guidelines for Transfer of Technology', which was first such document publicly available in the country. He also contributed towards formulation of various Policy issues in DPP. He was also member of various Technical Offset Evaluation Committees (TOECs).

During his second tenure at SAG from 28th March 2017 to 10th Oct 2019, he was member of the various Laboratory Level decision making bodies e.g., Technology Council, Management Council and HR-Council. He was also heading the Project Management Cell, Materials Management Group (MMG), Works, Quality-Reliability-Safety and Raj-Bhasha. He introduced systematic improvements to synergize the efforts of Project-Groups, MMG & Works resulting in Clearing Procurement related backlogs in the Projects, Installation of equipment's and Creating Laboratory Infrastructure for Projects. He has been associated with other Government Departments like DSIR (Department of Scientific & Industrial Research) and National Informatics Centre (NIC) in various capacities. He has presented research papers at several seminars & conferences and his work has been published in various journals. His current areas of interest are Big-Data Analytics, Artificial Intelligence, Quantum Computing and Technopreneurship. He is a Life Member of AIMA (All India Management Association) and CENJOWS (Centre for Joint Warfare Studies). He is an avid speaker at various Industry Forums & Academic Institutions and has received letters of appreciation from CII and DG MSME. In his spare time, he loves to read, travel and interact with students to motivate them for making India Technologically Self-Reliant and Economically Prosperous.

Keynote Speakers/Session Chairs

Prof. Satoshi Nakamura, Nara Institute of Science and Technology, Japan

Dr. Satoshi Nakamura is a Professor of Nara Institute of Science and Technology (NAIST), Team Leader of Riken AIP, Honorary professor of Karlsruhe Institute of Technology, Germany. He received his B.S. from Kyoto Institute of Technology in 1981 and Ph.D. from Kyoto University in 1992. He was an Associate Professor of the Graduate School of Information Science at NAIST in 1994-2000. He was the Department head and Director of ATR Spoken Language Communication Research Laboratories in 2000-2004, and 2005-2008, respectively, and Vice president of ATR in 2007-2008. He was Director-General of Keihanna Research Laboratories and the Executive Director of Knowledge-Creating Communication Research Center, National Institute of Information and Communications Technology, Japan in 2009-2010. He is currently Director of the Data Science Center and a full professor of Information Science Division, Graduate School of Science and Technology of NAIST, and Team Leader of the Tourism Information Analytics Team at Center for Advanced Intelligence Project (AIP), RIKEN, Japan. His research interests include modeling and systems of speech processing, speech-to-speech translation, speech recognition, spoken dialog systems, natural language processing, and big data analytics. He is one of the world leaders of speech-to-speech translation research and has been serving various speech-to-speech translation research projects. He was a committee member of IEEE SLTC in 2016-2018, Elected Board Member of International Speech Communication Association, ISCA in 2011-2019. He received Antonio Zampolli Prize in 2012 and retained the title of ATR Fellow, IPSJ Fellow, ISCA Fellow, and IEEE Fellow.

Title: Recent Advances in Speech-to-speech Translation

Abstract: "Speech translation" is a technology to translate a source language speech to a target language speech, which includes speech recognition, machine translation, and speech synthesis. The research and development of this technology was launched in 1980's, and the great effort and hard work of many researchers brought the "speech translation" to the current stage of practical use. However, what people expect to the real-time speech translation is the "interpretation". The human interpreters simultaneously and incrementally listen, understand, and produce interpreted utterances considering word order differences between source and target languages. In this talk, I will introduce the recent advances in speech-to-speech translation towards automatic speech interpretation at our team including incremental speech recognition, machine translation, and speech synthesis for distant language pairs like English and Japanese.

Prof. S Umesh, IIT-Madras

Dr. Umesh is a professor in Department of Electronics and Electrical Engineering, Indian Institute of Technology Madras, Tamil Nadu. Having specialisation in Speaker & Speech Recognition, Spectral Analysis, Detection & Estimation Theory, Digital Signal Processing. His research Interests lies in Low-Resource Automatic Speech Recognition, Acoustic Modelling, Speaker Normalization & Adaptation, Speaker Recognition and Diarisation, Spectral Analysis, and Digital Signal Processing. He is having a vast teaching and research experience of approx 28 years. Prof. Umesh worked as a Germany Visiting Researcher from May 2004 to June 2005 through Alexander von Humboldt Research Fellowship and from June 2003 to April 2004 as a UK Visiting Researcher in Machine Intelligence Lab-Cambridge University.

Title: Multilingual Speech Recognition of Indian Languages

Abstract: In this talk, I will talk about our efforts to build speech recognition systems in Indian languages. In particular, we have recently developed the idea of common label set for Indian languages which helps pool speech data and address the problems of resource scarcity. I will also talk about our recent efforts to make technical lectures available in Indian languages by automatic transcription and translation. This is part of the National Language translation Mission.

Prof. K Samudravijaya, IIT Guwahati

Prof. Samudravijaya is currently visiting faculty at Centre for Linguistic Science and Technology, IIT Guwahati. He has worked as Scientific Officer at TIFR Mumbai, Project Leader at Speaker verification lab, Aum Systems Inc., USA, and Visiting Scientist at Carnegie Mellon University, Pittsburgh, USA. His core area of research is Speech Technology. He has been awarded several prestigious awards including Prof. Rais Ahmed Memorial Lecture Award (2008), Sir C V Raman Award, Acoustic Society of India (2003), Indian Phonetic Society Lecture Award (2002), UNDP Fellowship for research at CMU, Pittsburgh (1988), Best Ph.D. Thesis Award (1986) and several others. He has worked extensively on various projects from government agencies including DeitY, DIT, UNDP and others. He has published approx. 100 papers in reputed Journals and conferences and travelled extensively in India and abroad to deliver Keynote addresses and expert talks.

Prof. S Sakti, JAIST, JAPAN

Sakriani Sakti is currently an associate professor at Japan Advanced Institute of Science and Technology (JAIST) Japan, adjunct associate professor at Nara Institute of Science and Technology (NAIST) Japan, visiting research scientist at RIKEN Center for Advanced Intelligent Project (RIKEN AIP) Japan, and adjunct professor at the University of Indonesia. She received her B.E. degree in Informatics (cum laude) from Bandung Institute of Technology, Indonesia, in 1999. In 2000, she received DAAD-Siemens Program Asia 21st Century Award to study in Communication Technology, University of Ulm, Germany, and received her MSc degree in 2002. During her thesis work, she worked with Speech Understanding Department, DaimlerChrysler Research Centre, Ulm, Germany. Between 2003–2009, she worked as a researcher at ATR SLC Labs, Japan, and during 2006–2011, she worked as an expert researcher at NICT SLC Groups, Japan. While working with ATR-NICT, Japan, she continued her study (2005–2008) with Dialog Systems Group University of Ulm, Germany, and received her Ph.D. degree in 2008. She actively involved in collaboration activities such as Asian Pacific Tele community Project (2003–2007), A-STAR, and U-STAR (2006–2011). In 2009–2011, she served as a visiting professor of Computer Science Department, University of Indonesia (UI), Indonesia. In 2011–2017, she was an assistant professor at the Augmented Human Communication Laboratory, NAIST, Japan. She also served as a visiting scientific researcher of INRIA Paris-Rocquencourt, France, in 2015–2016, under JSPS Strategic Young Researcher Overseas Visits Program for Accelerating Brain Circulation. In 2018–2021, she was a research associate professor at NAIST and a research scientist at RIKEN, Center for Advanced Intelligent Project AIP, Japan. Currently, she is an associate professor at JAIST, adjunct associate professor at NAIST, visiting research scientist at RIKEN AIP, and adjunct professor at the University of Indonesia. She is a member of JNS, SFN, ASJ, ISCA, IEICE, and IEEE. She is also a committee member of IEEE SLTC (2021-2023) and an associate editor of the IEEE/ACM Transactions on Audio, Speech, and Language Processing (2020-2023). Furthermore, she is the chair of ELRA/ISCA Special Interest Group on Under-resourced Languages (SIGUL) and a Board Member of Spoken Language Technologies for Under-Resourced Languages (SLTU). Her research interests include statistical pattern recognition, graphical modeling framework, deep learning, multilingual speech recognition and synthesis, spoken language translation, affective dialog system, and cognitive-communication.

Title: Listening while Speaking and Visualizing: A Semi-supervised Approach with Multimodal Machine Speech Chain

Abstract: The development of advanced spoken language technologies, such as automatic speech recognition (ASR) and text-to-speech synthesis (TTS), has enabled computers to either learn how to listen or speak. Many applications and services are now available, but the construction is commonly done based on a supervised fashion where a large amount of paired speech and corresponding transcription is required. In this talk, we will introduce a semi-supervised learning mechanism based on a machine speech chain framework. First, we describe the primary machine speech chain architecture that learns not only to listen or speak but also to listen while speaking. The framework enables ASR and TTS to teach each other given unpaired data. Then, we describe the recent multimodal machine chain framework that mimics overall human communication to listen while speaking and visualizing. With the support of image captioning and production models, the framework further reduces the need for a large amount of unpaired data. It enables ASR and TTS to improve their performance using an image-only dataset.

Prof. Laurent Besacier, Naver Labs Europe

Laurent Besacier is a principal scientist and Natural Language Processing (NLP) research group lead at Naver Labs Europe. He became a professor at the University Grenoble Alpes (UGA) in 2009 where he led the GETALP group (natural language and speech processing). Laurent is still affiliated with UGA. His main research expertise and interests lie in the field of natural language processing, automatic speech recognition, machine translation, under resourced languages, machine-assisted language documentation and the evaluation of NLP systems.

Title: Self-Supervised Learning for Low Resource Speech Tasks

Abstract: Self-supervised learning using huge unlabeled data has been successfully explored for image processing and natural language processing. Since 2019, recent works also investigated self-supervised representation learning from speech. They were notably successful to improve performance on downstream tasks such as speech recognition. These recent works suggest that it is possible to reduce dependence on labeled data for building speech systems through acoustic representation learning. In this talk I will present an overview of these recent approaches to self-supervised learning from speech and show my own investigations to use them in spoken language processing tasks for which size of training data is limited.

Prof. Tanja Schultz, University of Bremen, Germany

Tanja Schultz is Professor for Cognitive Systems of the Faculty of Mathematics & Computer Science at the University of Bremen, Germany and adjunct Research Professor of the Language Technologies Institute at Carnegie Mellon, PA USA. She received the diploma and doctoral degrees in Informatics from University of Karlsruhe and a Master degree in Mathematics and Sport Sciences from Heidelberg University, both in Germany. In 2007, she founded the Cognitive Systems Lab (CSL) and serves as Director since then. She is the spokesperson of the University Bremen high-profile area “Minds, Media, Machines” and helped establish the Leibniz Science Campus on Digital Public Health in 2019 for which she serves on the board of directors. Professor Schultz is a recognized scholar in the field of multilingual speech recognition and cognitive technical systems, where she combines machine learning methods with innovations in bio signal processing to create technologies such as in "Silent Speech Communication & quot; and "Brain-to-Speech". She is a Fellow of the IEEE, elected in 2020 “for contributions to multilingual speech recognition and biosignal processing”; a Fellow of the International Speech Communication Associ ation, elected in 2016 “for contributions to multilingual speech recognition and biosignal processing for human-machine interaction”; a Fellow of the European Academy of Science and Arts (2017), and a Fellow of the Asian-Pacific Artificial Intelligence Association (2021). Her recent awards include the Google Faculty Research Award (2020 and 2013), the ISCA/EURASIP Best Journal Paper Award (2015 and 2001), the Otto Haxel Award (2013), and the Research Award for Technical Communication from the Alcatel-Lucent Award (2012) “for her overall scientific work in the interaction of human and technology in communication systems”.

Title: Beyond Acoustic Speech Communication

Abstract:
Speech is a complex process emitting a wide range of biosignals, including, but not limited to, acoustics. These biosignals – stemming from the articulators, the articulator muscle activities, the neural pathways, and the brain itself – can be used to circumvent limitations of conventional speech processing in particular, and to gain insights into the process of speech production in general. In my talk I will present ongoing research at the Cognitive Systems Lab (CSL), where we investigate a range of speech-related biosignals beyond acoustics, such as muscle and brain activities to establish alternative communication interfaces. Based on machine learning methods we implement biosignal-based speech processing systems and devices for communication applications in everyday situations and for speech rehabilitation. Several applications will be described such as Silent Speech Interfaces that rely on articulatory muscle movement captured by electromyography to recognize and synthesize silently produced speech, Brain-to-text interfaces that use brain activity captured by electrocorticography to recognize speech, and Brain-to-Speech interfaces that synthesize speech with low latency to close the interaction loop.

Dr. RajKumar Buyya, University of Melbourne, Australia

Dr. Raj kumar Buyya a Redmond Barry Distinguished Professor and Director of Cloud Computing and Distributed Systems (CLOUDS) Laboratory at the University of Melbourne, Australia and is also serving as the founding CEO of Manjrasoft, a spin-off company of the University, commercializing its innovations in Cloud Computing. He has authored over 725 publications and four text books including "Mastering Cloud Computing" published by McGraw Hill, China Machine Press, and Elsevier/Morgan Kaufmann for Indian, Chinese and international markets respectively. He is one of the highly cited authors in computer science and software engineering worldwide. He led the establishment and development of key community activities, including serving as foundation Chair of the IEEE Technical Committee on Scalable Computing and five IEEE/ACM conferences.

Dr. Karunesh Kumar Arora, Sr. Director, CDAC, Noida

Dr. Karunesh Kumar Arora is working as Senior Director & Group Coordinator, Speech & Natural Language Processing Group of Centre for Development of Advanced Computing (CDAC), Noida. Having vast experience of research & development in the field and have completed / handled various major projects including Consortium projects - Development of Indian Languages to Indian Languages Machine Translation Systems, Development of Cross Lingual Information Access for Indian languages, English-Indian Language Machine Translation System based on Angla Bharati approach, all sponsored by Technology Development in Indian Languages (TDIL), Ministry of Electronics & Information Technology (MeitY), Govt. of India, U-STAR – Universal Speech to Speech Translation (International Consortium project) led by NICT, Japan. Looking after developments pertaining to Hindi language along with several others.

He has received renowned IMC Information Technology Excellence Award, Manthan Award & Skoch Project of Merit award for various projects. He has also worked as Invited Researcher at Advanced Technical Research Institute International (ATR now NICT), Japan. He has published more than 25 research papers in International, national conferences & journals. He has contributed chapters to International and national books. He has been Co-editor of Proceedings of International Conference on Natural Language Processing. He is the member of various national and international committees, and Project Review & Steering Groups. He has delivered various invited talks, invited expert for interview boards and session chair in various conferences.

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY

(AIST-2021)

12th-13th November, 2021

Inauguration Schedule

Time	Activity
10:00 AM – 10:05 AM	Welcome of the Guests
10:05 AM – 10:10 AM	Saraswati Vandana and Lighting of Lamp
10:10 AM – 10:15 AM	Introduction to Conference by Prof. Arun Sharma, Conference Secretary, AIST2021
10:15 AM – 10:20 AM	Address by Dr S. S. Agrawal, Technical Program Chair, AIST2021
10:20 AM – 10:30 AM	Address by Dr Amita Dev, General Chair-AIST2021 and Hon'ble VC, IGDTUW
10:30 AM – 10:35 AM	Address by Dr. Satoshi Nakamura, Honorary Chair, AIST 2021
10:35 AM – 10:45 AM	Address by Guest of Honor- Dr. Bindu Narayan, Associate Director (Data Science and AI), Ernst & Young Global Limited.
10:45 AM – 10:55 AM	Address by Chief Guest- Shri. Parimal Kumar, Scientific Advisor to Chief of Integrated Defence Staff, Ministry of Defence, GOI
10:55 AM – 10:57AM	Release of Conference Souvenir
10:57 AM – 11:00 AM	Vote of Thanks- Prof Jasdeep Kaur, IGDTUW
11:00 AM – 11:10 AM	University at a Glance

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY

(AIST-2021)

12th-13th November, 2021

Master Program Schedule

Day 1 (12th November, 2021)

Time	Activity
9:00 AM - 10:00 AM	Registration
10:00 AM - 11:00 AM	Inauguration
11:00 AM - 11:15 AM	Break
11:15 AM - 12:15 PM	Technical Session 1
12:15 PM – 01:00 PM	Key Note Address by Dr. Rajkumar Buyya, Professor, University of Melbourne, Australia
01:00 PM - 01:30 PM	Lunch Break
01:30 PM - 02:15 PM	Key Note Address by Dr. Satoshi Nakamura, NAIST, Japan
02:30 PM - 03:15 PM	Keynote Address by Dr. Tanja Shultz, Professor, University of Bremen, Germany
03:15 PM -04:30 PM	Technical Session - II Technical Session - III
04:30 PM – 5:15 PM	Keynote Address by Dr. Laurent Besacier, Principal Scientist, Naver Labs Europe, France

Day 2 (13th November, 2021)

Time	Activity
10:15 AM - 11:00 AM	Key Note Address by Dr. S Sakti, Professor, JAIST/NAIST, Japan
11:00 AM - 11:45 AM	Key Note Address by Prof. S. Umesh, Professor, IIT Madras
11:45 AM - 01:00 PM	Technical Session-IV
01:00 PM – 02:00 PM	Lunch Break
02:00 PM - 03:15 PM	Technical Session-V Technical Session-VI
03:15 PM - 04:30 PM	Technical Session-VII
04:30 PM - 05:00 PM	Valedictory Function

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

12th-13th November, 2021

Program Schedule for Valedictory Function

Date and Time: 13/11/2021, 4:30 PM

Time	Activity
04:30 PM – 04:35 PM	Welcome of the Guests
04:35 PM – 04:40 PM	Conference Report by Dr S S Agrawal, TPC Chair, AIST-2021
04:40PM – 04:50 PM	Address by Dr Amita Dev, General Chair, AIST-2021 and Hon'ble VC, IGDTUW
04:50PM – 05:00 PM	Address by Guest of Honor, Dr Anupam Shukla, Director, IIIT Pune
05:00 PM – 05:02 PM	Declaration of Best Paper Awards
05:02 PM – 05:05 PM	Vote of Thanks

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

12th-13th November, 2021

Technical Sessions Details

Track 1: Trends and Applications in Speech Processing Date: 12th November, 2021 Time: 11:15 Noon - 12:15 PM	Session Chair: Dr. Samudra Vijay, IIT G/Prof Anurag Jain, IP University Coordinator – Dr Ritu Rani, IGDTUW and Ms. Nandini Sethi, IGDTUW
Track 2: NLP and its Applications Date: 12th November, 2021 Time: 03:15 PM -04:30 PM	Session Chair: Prof Poonam Bansal, MSIT/Dr. Karunesh Arora, CDAC Coordinator: Dr Ritu Rani, IGDTUW and Ms. Nandini Sethi, IGDTUW
Track 3: Applications of Machine Learning and Deep Learning in Healthcare Date: 12th November, 2021 Time: 03:15 PM -04:30 PM	Session Chair: Prof D K Vishwakarma, DTU/Prof. Prafulla Kumar Behera, Utkal University Coordinator: Dr. Swati Sharma, IGDTUW and Ms Dimple Sethi, IGDTUW
Track 4: Recent Trends in Machine Learning and Deep Learning Date: 13th November, 2021 Time: 11:45 AM - 01:00 PM	Session Chair: Prof Monica Mehrotra, Jamia Millia Islamia/Prof Nidhi Goel, IGDTUW Coordinator: Dr. Swati Sharma, IGDTUW and Ms Dimple Sethi, IGDTUW
Track 5: Analysis using Hybrid Technologies with Artificial Intelligence Date: 13th November, 2021 Time: 02:00 PM - 03:15 PM	Technical Session Chair: Dr Mani Madhukar, IBM Research/ Prof Arun Sharma, IGDTUW Coordinator: Dr Ritu Rani, IGDTUW and Ms. Nandini Sethi, IGDTUW
Track 6: Image Analysis using Machine Learning and Deep Learning Date: 13th November, 2021 Time: 02:00 PM - 03:15 PM	Technical Session Chair: Prof. Laxmi Ahuja, Amity University, Noida/ Prof A K Mohapatra IGDTUW Coordinator: Dr. Swati Sharma, IGDTUW and Ms Dimple Sethi, IGDTUW
Track 7: Recent Advancements Using Hybrid Technology Date: 13th November, 2021 Time: 03:15 PM - 04:30 PM	Technical Session Chair: Prof Durga Prasad Mohapatra, NIT Rourkela/Prof. Jasdeeep Kaur, IGDTUW Coordinator: Ms. Garima Jaiswal, IGDTUW and Ms Pooja Gambhir, IGDTUW

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

12th-13th November, 2021

Technical Session: 01 **Date and Time: 12/11/2021, 11:15 AM – 12:15 PM**

Trends and Applications in Speech Processing

S. No	Paper ID	Authors	Title
1	11	Kiran Suryawanshi and Suvarnsing Bhable	A Survey on Automatic Visual Speech Recognition System
2	27	Rizwana Kallooravi Thandil	Speaker Independent Accent Based Speech Recognition for Malayalam Isolated Words: An LSTM-RNN Approach
3	42	Ruchika Kumari, Amita Dev and Ashwni Kumar	A Review on Speech Synthesis based on Machine Learning
4	43	Shobha Bhatt, Amita Dev and Anurag Jain	Hindi Phoneme Recognition- A Review
5	89	Aishwarya Suresh, Anushka Jain, Kriti Mathur and Pooja Gambhir	Comparison of modelling ASR system with different features extraction methods using Sequential model
6	103	Amritpreet Kaur, Rohit Sachdeva and Amitoj Singh	Recent Advances in Deep Learning for Automatic Speech Recognition System
7	114	Alisha Goel, Arun Sharma, Sparsh Sharma and Advikaa Kapil	Deep Learning Approaches for Speech Analysis: A Critical Insight
8	76	Nandini Sethi and Amita Dev	Survey on Automatic Speech Recognition Systems for Indic Languages

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

12th-13th November, 2021

Technical Session: 02 **Date and Time: 12/11/2021, 03:15 PM – 04:30 PM**

NLP and its Applications

S No	Paper ID	Authors	Title
1	21	Arushi Garg, Soumya Vats, Garima Jaiswal and Arun Sharma	Analytical Approach for Sentiment Analysis of Movie Reviews Using CNN and LSTM
2	28	Janhavi Jain, Debadrita Dey, Bhavika Kelkar and Khyati Ahlawat	Analysis of Indian News with Corona Headlines Classification
3	55	Neha Vaish, Nidhi Goel and Gaurav Gupta	Feature extraction and sentiment analysis using machine learning
4	64	Dr. Gaurav Dubey and Pinki Sharma	A Neural Network based approach for Text-Level Sentiment Analysis using Sentiment Lexicons
5	90	Amita Dev, Shweta Bansal and Shyam S. Agrawal	Cross Linguistic Acoustic Phonetic Study of Punjabi and Indian English
6	96	Priya Sharma, Sparsh Sharma and Pooja Gambhir	Prashn: University Voice Assistant
7	137	Kavita Sharma and S.R.N. Reddy	Spectrogram Analysis and Text Conversion of Sound Signal for Query Generation to Give Input to Audio Input Device
8	144	Mukund K Roy, Sunita Arora, Karunesh Arora and Shyam S Agrawal	Building Speech Corpus in rapid manner to adapt general purpose ASR System to specific domain.

3rd INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

12th-13th November, 2021

Technical Session: 03 **Date and Time: 12/11/2021, 03:15 PM – 04:30 PM**

Applications of Machine Learning and Deep Learning in Healthcare

S.No	Paper ID	Authors	Title
1	20	Harsimar Kandhari, Sagar Deep, Garima Jaiswal and Arun Sharma	Critical Insights on Cancer Detection using Deep Learning
2	52	Ruchika Bala, Arun Sharma and Nidhi Goel	A lightweight deep learning approach for diabetic retinopathy classification
3	54	Arun Gopi and Sajini Thattankandy	Auscultation of respiratory sounds for diagnostics using Deep Learning approaches – Machine hearing a cognitive service for aiding clinical diagnosis
4	60	Shivani Dhiman, Anjali Thukral and Punam Bedi	OHF: An Ontology based Framework for Healthcare
5	65	Namrata Singh, Meenakshi Srivastava and Geetika Srivastava	Enhancing the Deep Learning-Based Breast Tumor Classification Using Multiple Imaging Modalities: A Conceptual Model
6	94	Nishtha Phutela, Devanjali Relan, Goldie Gabrani and Ponnurangam Kumaraguru	EEG based stress classification in response to stress stimulus
7	143	Dimple Sethi, Chandra Prakash and Sourabh Bharti	Latest Trends in Gait Analysis using Deep Learning Techniques: A Systematic Review

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

12th-13th November, 2021

Technical Session: 04 **Date and Time: 13/11/2021, 11:45AM – 1:00 PM**

Recent Trends in Machine Learning and Deep Learning

S.No	Paper ID	Authors	Title
1	41	Sumit Bhardwaj, Ayush Somani and Khushi Gupta	Detection of Skin Lesion Disease using deep learning algorithm
2	10	Nidhi Agarwal, Amita Jain, Ayush Gupta and Devendra Tayal	Applying XGBoost Machine Learning model to succor astronomers detect exoplanets in distant galaxies
3	30	Sapna Arora, Manisha Agarwal, Shweta Mongia and Ruchi Kawatra	PSRE Self-Assessment Approach for Predicting the Educators' Performance using Classification Techniques
4	37	Richa Sharma, Nishita Malhotra, Nisha Rathee and Sakshi Singhal	Comparative Analysis of Traditional and Optimization Algorithms for Feature Selection
5	56	Punam Bedi, Purnima Khurana and Ravish Sharma	Session based Recommendations using CNN-LSTM with Fuzzy Time Series
6	61	Garima Singh, Nidhi Goel and Neeru Rathee	Feature Extraction and Classification for Emotion Recognition using Discrete Cosine Transform
7	63	Nisha Rathee, Neeraj Rani, Tanya Sharma and Aakriti Mittal	Sarcasm Detection in social media using Hybrid Deep Learning and Machine Learning Approaches
8	112	Singhal Syed Umar, Nilesh Gole, Gemechu Boche Beshan, Endal Kachew Mosisa, Tadele Debisa Deressa and Tariku Birhanu Yadesa	An Effective Machine Learning Approach for Clustering Categorical Data with High Dimensions

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

12th-13th November, 2021

Technical Session: 05 **Date and Time: 13/11/2021, 2:00 PM – 3:15 PM**

Analysis using Hybrid Technologies with Artificial Intelligence

S. No	Paper ID	Authors	Title
1	12	Kanchan Naithani, Y. P. Raiwani and Rajeshwari Sisodia	Text Based Analysis of COVID-19 Comments using Natural Language Processing
2	35	Prashant Verma, Vijay Kumar and Bharat Gupta	Indian languages requirements for String search/comparison on Web
3	47	Sunita Singhal	Identification of Disease Resistant Plant Genes Using Artificial Neural Network
4	50	S Rama Sree, A Vanathi, Ravi Kishore Veluri and S.N.S.V.S.C Ramesh	A Comparative Study on a Disease Prediction System using Machine Learning Algorithm
5	88	Ruchi Verma and Varun Singh	Leaf Disease Identification using DenseNet
6	95	Shyamala Devi Munisamy, Manivannan D, Manikandan N.K, Ankita Budhia, Sagar Srivastava and Manshi Rohella	Dictionary Vectorized Hashing of Emotional Recognition of Text in Mutual Conversation
7	116	Chandra Prakash and Kartik Kharbanda	A Pilot study on FoG prediction using Machine Learning for Rehabilitation
8	145	Pooja Gambhir, Amita Dev and Shyam Sunder Agrawal	A Contrastive View of Vowel Phoneme Assessment of Hindi, Indian English and American English Speakers

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

12th-13th November, 2021

Technical Session: 06 **Date and Time: 13/11/2021, 2:00 PM – 3:15 PM**

Image Analysis using Machine Learning and Deep Learning

S.No.	Paper ID	Authors	Title
1	17	Sumit Bhardwaj and Sagar Cv	Comparing the accuracy and the efficiency in detection of Coronavirus in CT scans and X Ray images
2	22	Mohit Kumar and Pooja Dehraj	An Analysis on Image Compression Using Neural Network
3	32	Vanshika Jain, Kshitiz Shah and Meghansh Goel	Deep learning on small tabular datasets: using image conversion and transfer learning
4	68	Tanishka Dixit, Namrata Singh, Geetika Srivastava and Meenakshi Srivastava	Implementation of a method using Image Sequentialization, Patch Embedding and Vision Transformer Encoder to detect the Breast Cancer on RGBA Images and Binary Masks
5	87	Dhivyashree M, Sarumathi K R and Vishnu Durai R S	An Ensemble model for face mask detection using Faster RCNN with ResNet50
6	102	Sumita Gupta, Rana Majumdar and Shivam Deswal	A Novel Approach to Detect Face Mask in Real Time
7	115	Rishi Kakkar and Srinivasa Murthy Y V	A Novel Approach for Detecting Facial Key Points using Convolution Neural Networks
8	120	Sushmita Gautam and Kailash Chandra Tiwari	Effective Hyperspectral Image Classification using Learning Models

3RD INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND SPEECH TECHNOLOGY (AIST-2021)

12th-13th November, 2021

Technical Session: 07 Date and Time: 13/11/2021, 3:15 PM – 4:30 PM

Recent Advancements Using Hybrid Technology

S.No	Paper ID	Authors	Title
1	67	Barkha Shree and Dr. Parneeta Dhaliwal	Reinforcing Digital Forensics through Intelligent Behavioural Evidence Analysis: Social media hate speech profiling
2	75	Muneer V.K and Dr. Mohamed Basheer Kp	Collaborative Travel Recommender System based on Malayalam Travel Reviews
3	99	Hemlata Goyal, Amar Sharma, Ranu Sewada, Devansh Arora and Sunita Singhal	Comparative Analysis of NLP Text Embedding Techniques with Neural Network Layered Architecture on Online Movie Reviews
4	121	Surbhi Khurana and Amita Dev	Current state of Speech Emotion Dataset-National and International level
5	125	Anushka Gupta, Bhumika Bhatia, Diksha Chugh, Gadde Satya Sai Naga Himabindu, Divyashikha Sethia, Ekansh Agarwal, Depanshu Sani and Saurabh Garg	Context-Aware Emoji Prediction Using Deep Learning
6	126	Thiyam Susma Devi and Pradip K. Das	Development of ManiTo: A Manipuri Tonal Contrast Dataset
7	142	Shweta Sinha, SS Agrawal	Deep Neural Networks for Spoken Language Identification in Short Utterances
8	146	Inderdeep Kaur, Sahil Dubra, Nidhi Goel, Arun Sharma	Hybrid Framework Model for Group Recommendation

A Critical Insight into Automatic Visual Speech Recognition System

Kiran Suryavanshi[1,] Suvarnsing Bhable[2] and [3]Charansing Kayte
[12] Department of Computer Science & IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, MH[2]
Government Institute of Forensic Science, Aurangabad, MH, INDIA
[1]suryawanshikiran@gmail.com, [2]suvarnsingbhable@gmail.com, Charankayte@gmail.com

This research paper investigated the robustness of the Automatic Visual Speech Recognition System (AVSR), for acoustic models that are based on GMM and DNNs. Most of the recent survey literature is surpassed in this article. Which shows how, over the last 30 years, analysis and product growth on AVSR robustness in noisy acoustic conditions has progressed? There are various categories of languages covered, with coverage, development processes, Corpus, and granularity varying. The key advantage of deep-learning tools, including a deeply convoluted neural network, a bi-directional long-term memory grid, a 3D convolutional neural network, and others, is that they are relatively easy to solve such problems, which are indissolubly linked to feature extraction and complex audio-visual fusion. Its objective is to act as an AVSR representative.

Keywords: Speech Recognition, AVSR, MFCC, HMM, CNN, DNN

Speaker Independent Accent Based Speech Recognition for Malayalam Isolated Words: An LSTM-RNN Approach

Rizwana Kallooravi Thandil[1[0000-0001-9305-681X]] and Muhamed Basheer K.P[2[0000-0003-3847-4124]]
[1,2] Sullamussalam Science College, Areekode, Kerala, India
ktrizwana@gmail.com

Automatic speech recognition (ASR) has been a very active area of research for the past few decades. Though there are great advancements in ASR in many languages accent-based speech recognition is an area that is yet to be explored in many languages. Speech recognition by humans is an intuitive process and so is a tough process to make the computers automatically recognize human speech. Although speech recognition has achieved promising achievements for many languages; speech recognition for the Malayalam language is still in infancy. The scarcity of the datasets makes it researchers difficult to do the experiments. Here in this paper, we have experimented with Long Short-Term Memory (LSTM) a Recurrent Neural Network (RNN), for recognizing the accent-based isolated words in Malayalam. The datasets we used here have been constructed manually under a natural recording environment. We used Mel Frequency Cepstral Coefficient (MFCC) methods to extract the features from the audio signals. LSTM with RNN is used to train and build the model since this technology significantly outperforms all other feed-forward deep neural networks and other statistical methodologies.

Keywords: Malayalam Speech Recognition, Accent-Based ASR, LSTM, RNN, MFCC.

.A Review on Speech Synthesis based on Machine Learning

Ruchika Kumari[1], Amita Dev[2] and Ashwani Kumar[2]
[1] Research Scholar, Indira Gandhi Delhi Technical University for Women, Delhi, India

[2]Indira Gandhi Delhi Technical University for women, New Delhi, India.
[1]ruchika.kumari33@gmail.com,[2]amita_dev@hotmail.com, [2]drashwnikumar@gmail.com

Recently, Speech synthesis is one of the growing techniques in the research domain that takes input as text and provides output as acoustical form. The speech synthesis system is more advantageous to physically impaired people. In execution process, there arise some complications by surrounding noises and communication style. To neglect such unnecessary noises various machine learning techniques are employed. In this paper, we described various techniques adopted to improve the naturalness and quality of synthesized speech. The main contribution of this paper is to elaborate and compare the characteristics of techniques utilized in speech synthesis for different languages. The techniques such as support vector machine, Artificial Neural Network, Gaussian mixture modeling, Generative adversarial network, Deep Neural Network and Hidden Markov Model are employed in this work to enhance the speech naturalness and quality of synthesized speech signals.

Keywords: Text to speech, Syllables, Database, Artificial Neural Network, Gaussian Mixture Modeling, Deep Neural Network, Hidden Markov Model

---

Hindi Phoneme Recognition- A Review

Shobha Bhatt[1], Amita Dev[2,] and Anurag Jain[1]
[1]University School of Information and Communication Technology, GGSIP University, New Delhi, India
[2]Indira Gandhi Delhi Technical University for women, New Delhi, India.
[1]bhattsho@gmail.com, [2]amita_dev@gmail.com, [3]anurag@ipu.ac.in;

Abstract. A review for Hindi phoneme recognition is presented to address Hindi speech recognition. Different issues related to Hindi phonemes such as Hindi speech characteristics, features used in phoneme recognition, and classification method highlighted. Related work was also presented to highlight issues concerned with feature extraction, classification, and distinct features. Earlier reviews mostly addressed speech recognition technologies. This work is an early research study presented for Hindi phoneme recognition. A phoneme-based system is used to overcome the constraint of the requirement of large training samples for word-based models. As phoneme-based systems are widely used for large vocabulary speech recognition, different issues related to consonants and vowels were also included. The comparative analysis is presented for different feature extraction and classification techniques with a recognition score. The research helps by presenting issues related to phoneme recognition. Research findings may be used to improve speech recognition by choosing appropriate methods for phoneme recognition.

Keywords: Review, Hindi, phoneme, features, consonants, vowels

---

Comparison of modelling ASR system with different features extraction methods using Sequential model

Aishwarya Suresh[1], Anushka Jain[1], Kriti Mathur[1], Pooja Gambhir[1]
[1]Indira Gandhi Delhi Technical University for Women, Delhi
aishwarya100btit19@igdtuw.ac.in

Speech recognition refers to a device's ability to respond to spoken instructions. Speech recognition facilitates hands-free use of various gadgets and appliances (a godsend for many incapacitated persons), as well as supplying input for automatic translation and ready-to-print dictation. Many industries, including healthcare, military telecommunications, and personal computing, use speech recognition programmes. In our paper, we are including the comparison between the different feature extraction methods (BFCC, GFCC, MFCC, MFCC Delta, MFCC Double Delta, LFCC and NGCC) using neural networks.

Keywords: Feature extraction method, Speech Recognition, Neural Network, FFT

Recent Advances in Deep Learning for Automatic Speech Recognition System

Amritpreet Kaur[1], Rohit Sachdeva [2], Amitoj Singh[3]
[1] Research Scholar, Doctor of Philosophy in Computer Science, Punjabi University, Patiala
[2]Assistant Professor, Department of Computer Science, M.M Modi College, Patiala
[3]Assistant Professor, Department of Computational Science, MRS PTU, Bathinda
amritkaur251990@gmail.com

In the field of Computer Learning and Intelligent Systems research, Deep Learning is one of the latest development projects. It's also one of the trendiest areas of study right now. Computational vision and pattern recognition have benefited greatly from the dramatic advances made possible by deep learning techniques. New deep learning approaches are already being suggested, offering performance that outperforms current state-of-the-art methods and even surpasses them. There has been much significant advancement in this area in the last few years. Deep learning is developing at an accelerated rate, making it difficult for new investigators to keep pace of its many kinds. We will quickly cover current developments in Deep Learning in the last several years in this article.

Keywords: Deep Learning, Machine Learning, Neural Networks.

Development of ManiTo: A Manipuri Tonal Contrast Dataset

Thiyam Susma Devi [0000−0002−6774−3760], Pradip K. Das [0000−0002−2363−2671]
Indian Institute of Technology Guwahati, Assam, India
thiyam.devi@iitg.ac.in, pkdas@iitg.ac.in

Tone recognition plays a vital role in understanding speech for tonal languages. Integrating tonal information from a robust tone recognition system can improve the performance of Automatic Speech Recognition (ASR) for such languages. The tonal recognition approaches adopted so far have focused on Asian, African and Indo-European lan

guages. In India, there are very few works on tonal languages, especially those spoken in its North-Eastern part, from which the Manipuri lan guage is largely unexplored. This paper presents the development of a Tonal Contrast dataset for Manipuri, a low resource language. It also presents an initial analysis of the recorded data.

Keywords: Tonal Contrast · Speech Recognition · Low Resource Languages.

Analysis of Indian News with Corona Headlines Classification

Janhavi Jain*, Debadrita Dey, Bhavika Kelkar, Khyati Ahlawat
Indira Gandhi Delhi Technical University for Women
Kashmere Gate, Delhi -110006, India
janhavi.j.98@gmail.com, debadrita3456@gmail.com,
bhavika12.1999@gmail.com, khyatiahlawat@igdtuw.ac.in

With the advent of the world wide web, the world has seen an explosion in the amount of information that is available online. People stay informed about the national and international affairs through online news which is readily available and portable allowing ease of access. These news pieces tend to shape people's thoughts and provoke emotions, which may be positive, neutral or negative, without them realizing their effect. The objective of this work is to create a hybrid model that can analyze the overall effect of digital news content in India. The hybrid approach of sentiment analysis encompasses lexicon and machine learning algorithms as well as a self-created scored corpus of corona related words to classify all sorts of headlines. The labelled dataset is used to train decision tree and random forest algorithms. They are evaluated based on their accuracy scores, classification reports and confusion matrices. The results prove that both the algorithms perform well on the dataset and that the Indian media highlighted neutral news the most. This finding can be very useful for the Indian news agencies since they can alter their reporting strategies to create an impact of their choice on the readers' minds and thus, increase the readership.

Keywords: Sentiment Analysis, Lexicon, corona corpus, Machine learning, Random Forest, Decision trees.

Feature Extraction and Sentiment Analysis Using Machine Learning

Neha Vaish[1], Nidhi Goel[1], Gaurav Gupta[2]
[1]Department of Electronics and Communication, IGDTUW, Delhi
[2] College of Science and Technology, Wenzhou-Kean University, Wenzhou, China
nehavaish6@gmail.com

The role of social networks has bought a tremendous change in the analysis of the opinions. Understanding people sentiments or opinion helps the business or organization to better understand their customers. There are several platforms where people can easily post their views about a service or products, these can be facebook, twitter e.t.c. Feature extraction or aspect extraction becomes important since one needs to know the qualities a product or a service have. In this research, we have analyzed hotel reviews by applying n-gram for feature. As the dataset is always noisy so basic preprocessing steps are applied before extraction. The features extracted are trained and tested by basic machine learning classifiers. Various machine learning algorithms like KNN, SVM, and random forest are used for the analysis of the performance. The evaluation measures are calculated at the end to validate the results. K-fold cross validation scheme is also applied on the dataset to improve the overall accuracy of the results.

Keywords: Feature, Sentiment, polarity, classifier, machine learning.

A Neural Network based approach for Text-Level Sentiment Analysis using Sentiment Lexicons

[1]Gaurav Dubey, [2]Pinki Sharma
[1,2] ABES Engineering College, Ghaziabad, Uttar Pradesh, India

There have been many discussions on forums, e-commerce sites, sites for reviewing products, social media which helps in exchanging opinions, thoughts through free expression of users. Internet as well as web 2.0 is overflowing with the data generated by users which provides a good source for various sentiments, reviews, and evaluations. Opinion mining more popularly known as sentiment analysis classifies the text document based on a positive or negative sentiment that it holds. This is an open research domain and this particular research paper puts forth a model called Artificial Neural Network Based Model i.e., ANNBM. The model is trained and tested through Information Gain as well as three other popular lexicons to extract the sentiments. It's a new approach that best utilizes the ANNBM model and the subjectivity knowledge which is available in sentiment lexicons. Experiments were conducted on the mobile phone review as well as car review to derive that this approach was successful in finding best output for sentiment-based classification of text and simultaneously reduces dimensionality.

Keywords: PSO, PSO-TVAC, Parameter tuning

Prashn: University Voice Assistant

Priya Sharma1, Sparsh Sharma and Pooja Gambhir1
1 Indira Gandhi Delhi Technical University for Women, Delhi, India
Delhi Technological University, New Delhi
priya003btcsai20@igdtuw.ac.in, sparsh1706@gmail.com

University websites are the best source to get information about a university. But each university has a different methodology for the implementation of its website and there is no common layout for finding specific information. This can be difficult for users, especially potential students who are trying to compare different universities to decide where to get admitted in. Natural Language Human-Computer Interaction (HCI) is trending these days due to its ease of use.[1] Thus, implementing a chatbot on university websites can reduce a lot of time students spend looking for information they need like office timings, address, admission procedure, accommodation information, etc. Voice Assistants are the future of Natural language HCI. Prashn is a web application that allows users to get all the necessary information about Indira Gandhi Delhi Technical University for Women, Delhi either through a chat interface or through speech recognition.

Keywords: Voice Assistant, University Chatbot, CNN, Flask.

Context-Aware Emoji Prediction Using Deep Learning

Anushka Gupta1, Bhumika Bhatia1, Diksha Chugh1, Gadde Satya Sai Naga Himabindu1,
Divyashikha Sethia1, Ekansh Agarwal2, Depanshu Sani2, and Saurabh Garg2
1 Delhi Technological University, Delhi - 110042, India
2 Samsung Research Institute, Noida, India

Emojis are a succinct and visual way to express feelings, emotions, and thoughts during text conversations. Owing to the increase in the use of social media, the usage of emojis has increased drastically. There are various techniques for automating emoji prediction, which use contextual information, temporal information, and user-based features. However, the problem of personalised and dynamic recommendations of emojis persists. This paper proposes personalised emoji recommen-dations using the time and location parameters. It presents a new an-notated conversational dataset and investigates the impact of time and location for emoji prediction. The methodology comprises a hybrid model that uses neural networks and score-based metrics: semantic and cosine similarity. Our approach differs from existing studies and improves the accuracy of emoji prediction up to 73.32% using BERT.

Keywords: NLP · Emoji Prediction · BERT.

Spectrogram Analysis and Text Conversion of Sound Signal for Query Generation to Give Input to Audio Input Device

Kavita Sharma[1] and S.R.N. Reddy[2]
[1,2] Indira Gandhi Delhi Technical University for Women, Delhi. kavita004phd20@igdtuw.ac.in

The world is being reshaped by Natural Language Processing. Audio inputs are used in modern electronics. Different types of people supply input to the system in their native language. The system accepts the person's speech, processes it, and responds accordingly. Cooking is a huge problem for a variety of people, including the elderly, those who are confined to their beds, and those who have a specific sort of handicap, such as those who are unable to use their hands and require assistance at all times. To help these people reach their full potential, an audio input device for giving cooking instructions to a cooking system has been proposed in this paper. The gadget takes the user's spoken English language as input, converts it to text using deep learning algorithms, and generates instructions with the help of context aware words extracted from the recorded audios to send the instruction to the cooking device. To analyses the audio signal for user authentication is a challenging task due to gaps and pauses between spoken characters, and existing noise in the environment. As a result, the audio input device developed for kitchen systems must analyse the audio input signal in order to create a more secure environment for authenticated users. As a result, the objective of this paper is to analyse the audio input signal captured in real time and process the accepted signal to convert into text in order to generate instructions for a larger system. The sound signals captured in real environment is analysed with Mel spectrogram, MFCC spectrogram and PRAAT software. The sound signal is processed with the help of natural language toolkit to generate instructions.

Keywords: Mel-spectrogram, audio signal, MFCC, tokenization.

Detection of Skin Lesion Disease using deep learning algorithm

Sumit Bhardwaj[1], Ayush Somani[2] and Khushi Gupta[3] [1,2,3]ECE Dept, ASET,
AMITY UNIVERSITY NOIDA, Sumitb460@gmail.com

Skin lesions are a part of the skin that has an unusual development or appearance contrasted with the skin around it. They may be something you are born with or something you acquire over your lifetime. They can be classified into two types: benign(non-cancerous) or malignant(cancerous). Some studies have been conducted on the computerised

detection of malignancy in images. However, due to various problematic aspects such as reflections of light from the skin's surface, the difference of colour lighting, and varying forms and sizes of the lesions, analysing these images is extremely difficult. As a result, evidence-based automatic skin cancer detection can help pathologists improve their accuracy and competency in the primitive stages of ailment. Our proposed method is to detect the early onset of skin lesions using python as a tool to detect benign(non-cancerous) or severe(cancerous) lesions using a machine learning approach. The dataset consists of nine different classes of skin lesion diseases: Melanoma (MEL), Melanocytic nevus (NV), Basal cell carcinoma (BCC), Actinic keratosis (AK), Benign keratosis (BKL), Dermatofibroma (DF), Vascular lesion (VASC), Squamous cell carcinoma (SCC), None of the above (UNK). In our proposed work, a DCNN model is created for classifying cancerous and non-cancerous skin lesions. We use techniques such as filtering, feature extraction for better categorization which will enhance the final analysis value. From our proposed model we have achieved a training accuracy of 90.7%.

Keywords: Skin lesions, deep convolutional neural network, benign, malignant

---

A Lightweight Deep Learning Approach for Diabetic Retinopathy Classification

Ruchika Bala1, Arun Sharma2, Nidhi Goel3
[1, 2, 3] Indira Gandhi Delhi Technical University for Women, Delhi, India
ruchikabala@gmail.com, arunsharma2303@gmail.com, nidhi.iitr1@gmail.com

In the present time, chances of suffering from diabetes have drastically increased due to the genetic probability, lack of physical ac tivities, high blood pressure and modern lifestyle related problems. Dia betic Retinopathy (DR) is an intense problem which affects blood vessels in the eye retina. Early detection of DR can avoid severe eye damage. Several machine-learning and deep-learning based techniques have been used for DR detection and classification. However, these techniques are complex, time consuming, and take millions of parameters in training and deploying the DR classifier. In this paper, a lightweight dual-branch based CNN architecture is proposed for DR classification. The proposed architecture involves 84,645 (0.084M) parameters for training and de ploying the model. APTOS dataset has been used for analysis.

Keywords: Diabetic Retinopathy, Classification, Grading, Fundus Im ages, Deep Learning, Transfer Learning, CNN, Convolution Neural Net work

---

OHF: An Ontology based Framework for Healthcare

Shivani Dhiman[1], Anjali Thukral[2], Punam Bedi[1],
[1] Department of Computer Science, University of Delhi, India, [2] Keshav Mahavidyalaya, University of Delhi, India
{pbedi, shivani} @cs.du.ac.in[1], athukral@keshav.du.ac.in[2]

Timely and holistic recommendations in disease diagnosis process may prove a great assistance to a medical practitioner. However, the amount of structured or unstructured data generated in medicinal domain are voluminous and thus imposes challenges. In order to make machines understand the data semantically and infer useful insights from patient's information, the data needs to be represented in a semantically organized knowledgebase, an ontology. Therefore, in this paper we explore various ontologies existing in the medicinal domain. The paper proposes an ontology-based framework for Healthcare (OHF) using these existing ontologies, and also proposes Healthcare Ontology (HO) which is a semantic representation of knowledgebase of patients' healthcare information available in the form of Electronic Health Record (EHR). The OHF consisting of systematically generated and exhaustive ontologies may be utilized for predicting semantic inferences related to a patient's medical condition. A case study is being used to explain working of the framework in disease diagnosis.

Keywords: Electronic Health Record, Ontologies, Clinical Diagnosis, Reasoning.

Enhancing the Deep Learning-Based Breast Tumor Classification Using Multiple Imaging Modalities: A Conceptual Model

Namrata Singh[1], Dr. Meenakshi Srivastava[2] and Dr. Geetika Srivastava[3]
[1,2]Amity Institute of Information Technology, [1,2]Amity University, Lucknow, India
[3]Department of Physics & Electronics, [3]Dr. Ram Manohar Lohia Avadh University, Ayodhya, India
[1]nam2817120@gmail.com, [2]msrivastava@lko.amity.edu, [3]geetika_gkp@rediffmail.comu

Breast tumors' preliminary and unambiguous prognosis is critical for early detection and diagnosis. A specific study has established automated techniques that use only science imaging modalities to speculate on breast tumor development. Several types of research, however, have suggested rephrasing the current literature's breast tumor classifications. This study reviewed various imaging modalities for breast tumors and discussed breast tumor segmentation and classification using preprocessing, machine learning, and deep neural network techniques. This research aims to classify malignant and benign breast tumors using appropriate medical image modalities and advanced neural network techniques. It is critical to improving strategic decision analysis on various fronts, including imaging modalities, datasets, preprocessing techniques, deep neural network techniques, and performance metrics for classification. They used preprocessing

techniques such as augmentation, scaling, and image normalization in the respective investigation to minimize the irregularities associated with medical imaging modalities. In addition, we discussed various architectures for deep neural networks. A convolutional neural network is frequently used to classify breast tumors based on medical images to create an efficient classification paradigm. It could be an existing network or one that has been developed from scratch. The accuracy, area-under-the-curves, precision, and F-measures metrics of the developed classification paradigm will be used to evaluate its performance.

Keywords: Classification of breast cancer, deep learning, medical imaging techniques, convolutional neural network

---

EEG Based Stress Classification in response to Stress Stimulus

Nishtha Phutela1, Devanjali Relan1, Goldie Gabrani2, and Ponnurangam Kumaraguru3
1 BML Munjal University, Gurugram, Haryana nishtha.phutela@bmu.edu.in, devanjali.relan@bmu.edu.in
2 Vivekananda Institute of Professional Studies, Pitampura, New Delhi, India. ggabrani@gmail.com
3IIIT Hyderabad, Telangana, India

Stress, either physical or mental, is experienced by almost every person at some point in his lifetime. Stress is one of the leading causes of various diseases and burdens society globally. Stress badly affects an individual's well-being. Thus, stress-related study is an emerging field, and in the past decade, a lot of attention has been given to the detection and classification of stress. The estimation of stress in the individual helps in stress management before it invades the human mind and body. In this paper, we proposed a system for the detection and classification of stress. We compared the various machine learning algorithms for stress classification using EEG signal recordings. Interaxon Muse de vice having four dry electrodes has been used for data collection. We have collected the EEG data from 20 subjects. The stress was induced in these volunteers by showing stressful videos to them, and the EEG signal was then acquired. The frequency-domain features such as absolute band powers were extracted from EEG signals. The data were then classified into stress and non-stressed using different machine learning methods - Random Forest, Support Vector Machine, Logistic Regression, Naive Bayes, K-Nearest Neighbours, and Gradient Boosting. We performed 10- fold cross-validation, and the average classification accuracy of 95.65% was obtained using the gradient boosting method.

Keywords: Stress classification, Machine learning, MUSE hadband, EEG signal.

A Pilot study on FoG prediction using Machine Learning for Rehabilitation

Kartik Kharbanda1 and Chandra Prakash1
National Institute of Technology Delhi, 110040, India
191220027@nitdelhi.ac.in, cprakash@nitdelhi.ac.in

Walking has a significant impact on one's quality of life. Freezing of Gait (FoG) is a typical symptom of Parkinson's disease (PD). FoG is characterized by quick and abrupt transient falls, as a result of which the patient's mobility is limited and their independence is lost. Thus, early detection of FoG in PD patients is necessary for diagnosis and rehabilitation. The present strategies for early detection of FoG are ineffective and have a low success rate. This study illustrates the comparative analysis of ML techniques (K Nearest Neighbors (KNN), Decision Trees, Random Forest, Support Vector Classifier (SVC), and Ada Boost Classifier), using time and statistical features to perform detection and prediction tasks on the publicly available DaphNet database. FoG prediction is highly patient dependent and achieved a peak F1 - score of 80% for one of the patients. The paper also presents a combined analysis of all the patients which may aid in designing wearable sensors for detection. This system detects FoG with a precision value of about 81%.

Keywords: Freezing of Gait (FoG), Machine learning techniques, Prediction, Detection, Parkinson's Disease

---

Latest Trends in Gait Analysis using Deep Learning Techniques: A Systematic Review

Dimple Sethi1, Chandra Prakash2, and Sourabh Bharti1
1Indira Gandhi Delhi Technical University, Delhi, India
2 National Institute of Technology. Delhi, India
dimple9203@gmail.com, cprakash@nitdelhi.ac.in, sourabhbharti@igdtuw.ac.in

Marker-less analysis of human gait has made considerable progress in recent years. However, developing a gait analysis system ca pable of extracting reliable and precise kinematic data in a standard and unobtrusive manner remains an open challenge. This narrative review considers the transformation of methods for extracting gait extremity information from videos or images, perceived how analysis methods have improved from arduous manual procedures to semi-objective and ob jective marker-based systems and then marker-less systems. The gait analysis systems widely used restrict the analysis process with the use of markers, inhibited environmental conditions, and long processing du ration. Such limitations can impede the use of a gait analysis system in multiple applications. Advancement in marker-less pose estimation and Q-learning-based techniques are opening the possibility of adopting pro ductive methods for estimating precise poses of humans and information of movement from video frames. Vision-Based gait analysis techniques are capable of providing a cost-effective, unobtrusive solution for estima tion of stick images and thus the analysis of the gait. This work provides a comprehensive review of marker-less computer vision and deep neural network-based gait analysis, parameters, design specifications, and the latest trends. This survey provides a birds-eye

view of the domain. This review aims to introduce the latest trends in gait analysis using computer vision methods thus provide a single platform to learn various marker less methods for the analysis of the gait that is likely to have a future impact in bio-mechanics while considering the challenges with accuracy and robustness that are yet to be addressed.

Keywords: Human Gait Analysis · Marker-less Techniques · Pose Estimation · Stick Images · Abnormality Detection.

Applying XGBoost Machine Learning model to succor astronomers detect exoplanets in distant galaxies

Nidhi Agarwal[1, 4][0000-0002-9044-8989] Dr. Amita Jain[2][0000-0003-0891-3675], Ayush Gupta[3] and Dr. Devendra Kumar Tayal[4][0000-0001-7388-5880]
[1] Kiet Group of Institutions, INDIA, [2] NSUT Delhi, INDIA, University of Turku, Finland, IGDTUW, Delhi, INDIA

The time when TRAPPIST-1 news became official on 22.02.2017, detection of planets beyond Milky Way Galaxy or planets orbiting around their own sun-like stars became one of the burning topics unlike prior times. There are seven famous exoplanets in TRAPPIST-1 system which are just forty light-years distant, and are available to be explored by our planets and other spacial telescopes. But several thousand other exoplanets are known to astronomers whose habitability is misleading as there is no evidence about contrasting effects take place between these bright stars and their suspected exoplanets. Since majority of the exoplanets are found using transit principal method, so in this research paper a new tool using XG Boost supervised Machine Learning Model is proposed to detect their presence. The results show that the prediction accuracy, precision and F1-score of this model is very high as compared to the other methods used in literature till now. This work is novel as till now no research work implements XGBoost based model of Machine Learning with highly accurate predictive power. None of the previous work has taken care of all the steps of data pre-processing and handling imbalanced data.

Keywords: Exoplanets, Machine Learning, Transit Principle, XG Boost Model, Supervised Machine Learning, SMOTE.

PSRE Self-Assessment Approach for Predicting the Educators' Performance using Classification Techniques

Sapna Arora[1, *], Manisha Agarwal[2], Shweta Mongia[3] and Ruchi Kawatra[4]
[1,2] Banasthali Vidyapith, P.O. Banasthali Vidyapith, Rajasthan, 304022, India
[3] University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India
[4] Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
*sapnaarora023@gmail.com

With the growing interest in and significance of Educational Data Mining to educators' performance, there is a vital need to comprehend the full scope of job performance that can substantially impact teaching quality. However, a few educational institutions are attempting to improve educator effectiveness to improve student outcomes. Furthermore, for reasons of confidentiality, most institutions do not share their data. As a result, an assessment of a self-assessment strategy is required to improve educators' performance. The professional, research, social, and emotional characteristics that contribute to the results and behavior of self and students taught are all part of an educators' overall job performance. With four input parameters and five classifiers (Logistics Regression, Naive Bayes, K-nearest Neighbor, Support Vector Machine- Linear, and Radial Basis Function), the proposed PSRE (Professional, Social, Research, and Emotional behavior) self-assessment approach is modeled to predict the overall performance of educators working in various Higher Educational Institutions. Overall, K-nearest neighbor has a high accuracy of 95.43 percent, which may help determine educators' progress and assist them in reaching new professional heights.

Keywords: Behavior, Classification, Educational Data Mining, Job Performance, Quality Teaching, Self-Assessment.

Hybrid Framework Model for Group Recommendation

Inderdeep Kaur, Sahil Dubra, Nidhi Goel, Arun Sharma
Indira Gandhi Delhi Technical University for Women, Delhi
nidhigoel@igdtuw.ac.in, arunsharma@igdtuw.ac.in

Abstract Group-based recommender systems are known for web-based applications to satisfy the preferences of every user in the group equally. The recommender system aims to identify user-preferred items, such as movies, music, books, and restaurants that tend to satisfy each individual and their collective needs in a group. They have a proclivity to solve the problem of information overload by probing through large volumes of dynamically generated information and provide its users with significant content and services. The recommendations for a group of users is pertinent, as there is a huge diversity in tastes and preferences among various members of the group residing at different locations. This paper presents the results computed from traditional filtering approaches and the Hybrid filtering mode. These approaches are evaluated using various offline metrics on Movie-Lens dataset and the proposed model is based on the working efficiency and the quality of recommendations generated. This Hybrid Filtering Model proposed as Hyflix is explored for generating recommendations to a group of users and compared with the existing traditional filtering approaches in a social network.

Keywords: Recommender System, Collaborative Filtering, Content-based Filtering, Hybrid Filtering, Hyflix, Movie recommender, Movie-Lens

Comparative Analysis of Traditional and Optimization Algorithms for Feature Selection

Sakshi Singhal, Richa Sharma, Nishita Malhotra, Nisha Rathee
Information Technology Department
Indira Gandhi Delhi Technical University for Women, Delhi
nishrathee@igdtuw.ac.in

Machine learning enables the automation of the system to generate results without direct assistance from the environment once the machine is trained for all possible scenarios. This is achieved by a series of processes such as collecting relevant data in raw format, exploratory data analysis, selection and implementation of required models, evaluation of those models, and so forth. The initial stage of the entire pipeline involves the necessary task of feature selection. The feature selection process includes extracting more informative features from the pool of input attributes to enhance the predictions made by machine learning models. The proposed approach focuses on the traditional feature selection algorithms and bio-inspired modified Ant Colony Optimization (ACO) algorithm to remove redundant and irrelevant features. In addition, the proposed methodology provides a comparative analysis of their performances. The results show that the modified ACO computed fewer error percentages in the Linear Regression Model of the dataset. In contrast, the traditional methods used outperformed the modified ACO in the SVR model.

Keywords: Feature Selection, ACO, Linear Regression, SVM, SVR.

Session based Recommendations using CNN-LSTM with Fuzzy Time Series

Punam Bedi1, Purnima Khurana2, Ravish Sharma3
1,2Department of Computer Science, University of Delhi, Delhi, India
3P.G.D.A.V. College, University of Delhi , Delhi, India
{pbedi@cs.du.ac.in, pk0403@gmail.com, ravish.sharma@pgdav.du.ac.in}

Recommender systems are a class of algorithms that help us to alleviate the problem of information overload. Over the years, recommender systems have evolved from generating recommendations based on user's long-term preferences to short term interests. Session based Recommender systems consider change in preferences by focusing on user's short-term interests that may change over a period of time. This paper proposes FS-CNN-LSTM-SR, a hybrid technique that uses CNN (Convolutional Neural Networks) and LSTM (Long Short-Term Memory) deep learning techniques with fuzzy time series to recommend products to user based on his activities performed in a session.

The advantage of our proposed method is that it combines the benefits of both CNN and LSTM. CNNs are capable of extracting complex local features and LSTM learn long term dependencies from the sequential session data. The Input time series data is preprocessed to obtain fuzzy time series data and is fed into CNN-LSTM model for prediction. A ranked list of items is recommended to the user based on predicted items, their co-occurring items and popularity of items. The performance of FS-CNN-LSTM-SR is evaluated on YOOCHOOSE dataset from RecSys Challenge 2015 and is compared with three variations viz. LSTM-SR, CNN-LSTM-SR and FS-LSTM-SR. We observed that our proposed approach performed better than the other three variations. The proposed technique is applicable on any E-commerce dataset where user purchasing choices need to be predicted.

Keywords: Session based Recommender System, Deep Learning, Fuzzy Time Series.

Sarcasm Detection in social media using Hybrid Deep Learning and Machine Learning Approaches

Tanya Sharma, Neeraj Rani, Aakriti Mittal and Nisha Rathee
nishrathee@igdtuw.ac.in
Department of Information Technology, Indira Gandhi Delhi Technical University for Women, Delhi

Sarcasm refers to the use of ironic language to convey the message. It is mainly used in social sites like Reddit, Twitter, etc. The identification of sarcasm improves sentiment analysis efficiency, which refers to analyzing people's behavior towards a particular topic or scenario. Our proposed methodology has used a hybrid supervised learning approach to detect the sarcastic patterns for classification. The supervised machine learning approaches include Logistic Regression, Naïve Bayes, Random Forest, and hybrid deep learning models like CNN and RNN. Before implementing the models, the dataset has been preprocessed. The data in the dataset is usually not fit for extracting features as it contains usernames, empty spaces, special characters, stop words, emoticons, abbreviations, hashtags, time stamps, URLs. Hence null values, stop words, punctuation marks, etc., are removed, and lemmatization is also done. After preprocessing, the proposed methodology has been implemented using various supervised machine learning models, hybrid neural network models, ensemble hybrid models, and models implementation by using word embeddings. The models have been implemented on two datasets. The outcome revealed that the hybrid neural network model RNN worked the best for both datasets and got the highest accuracy compared to other models.

Keywords: Sarcasm Detection, CNN, RNN, Naïve Bayes, Logistic Regression, Random Forest, Deep Learning Models.

Collaborative Travel Recommender System Based on Malayalam Travel Reviews

Muneer V.K and Dr. Mohamed Basheer K.P
Research Department of Computer Science, SS College, Malappuram, Kerala, India
vkmuneer@gmail.com

Recommender System is an unsupervised machine learning technique used for users to make decisions on their own choices and interests. Travel and Tourism is one of the important domains in this area. Rapid growth of redundant, heterogeneous data and web has created the importance for Recommender system. Travel Recommender System helps to users or travelers to decide on their choices and preferences on their travel. The work proposes a machine learning approach for a personalized travel recommender system in Malayalam Language, one of the prominent languages used in Southern part of India. The system developed using Collaborative Filtering technique, Malayalam Text processing with the help of Cosine similarity and TF-IDF methods. Data has been taken from the largest Malayalam Travel group in Facebook named "Sanchari". A customized scraping algorithm also developed to collect relevant and enough information from social media. The RS could suggest most suitable destinations to each of users with the accuracy of 93%, which is fair result as compared with other algorithms on same domain.

Keywords: Recommender System [RS], Collaborative Filtering, Natural Language Processing, Malayalam, Cosine similarity.

An Effective Machine Learning Approach for Clustering Categorical Data with High Dimensions

Dr. Syed Umar *, Tadele Debisa Deressa*, Tariku Birhanu Yadesa*, Gemechu Boche Beshan*,
Endal Kachew Mosisa*, Prof. Nilesh Gole*
Dept. of Computer Science, Wollega University, Nekemte, Ethiopi
Dept. of Information Technology, Wollega University, Nekemte, Ethiopia
Dept. of Mechanical Engineering, Wollega University, Nekemte, Ethiopia
Dept. of Computer Science & Engineering, Nagpur, India

Many modern real-world databases include redundant quantities of categorical data that contribute in data processing and efficient decision-making with their advances in database technology. However, for the reasons that they are identical to measurements the clustering algorithms are only devised for numerical results. An immense amount of work is being performed on the clustering of categorical data using a specifically defined similarity measure over categorical data. Thereby, the dynamic issue with real-world domain, which does not clearly take the predictive form, is the inner function. The function is based on both unseen and transonic perspective. This then offers a detailed and inventive collaboration with categorical results. The paper describes a stratified, immune-based approach with a new similarity metric, in order to reduce distance function, for clustering CAIS categorical data. For successful exploration of clusters over categorical results, CAIS adopts an immunology focused approach. It also selects subsistent nomadic characteristics as a representative entity and organize them into clusters that quantify affinity. To minimize database throughput, CAIS is segmented into several attributes. The analytical findings show that the proposed solution yields greater mining performance on different categorical datasets and outperforms EM on categorical datasets.

Keywords— High Dimensional Data, Categorical Data, Immune based Clustering, Maximization.

Text Based Analysis of COVID-19 Comments using Natural Language Processing

Kanchan Naithani, Y. P. Raiwani, Rajeshwari Sissodia
Department of Computer Science and Engineering,
HNB Garhwal University, Srinagar Garhwal, Uttarakhand, India
kanchannaithani696@gmail.com, yp_raiwani@gmail.com, raj.sissodia@gmail.com

In dialectology, Natural Language Processing is the process of recognizing the various ontologies of words generated in human language. Various techniques are used for analyzing the corpus from naturally generated content by users on various platforms. The analysis of these textual contents collected during the, COVID-19 has become a goldmine for marketing experts as well as for researchers, thus making social media comments available on various platforms like Facebook, Twitter, YouTube, etc., a popular area of applied artificial intelligence. Text-Based Analysis is measured as one of the exasperating responsibilities in Natural Language Processing (NLP). The chief objective of this paper is to work on a corpus that generates relevant information from web-based statements during COVID-19. The findings of the work may give useful insights to researchers working on Text analytics, and authorities concerning to current pandemic. To achieve this, NLP is discussed which extracts relevant information and comparatively computes the morphology on publicly available data thus concluding knowledge behind the corpus.

Keywords: NLP: Natural Language Processing, NLU: Natural Language Understanding, TBA: Text Based Analysis, Knowledge Representation, WSD: Word Sense Disambiguation.

Indian languages requirements for String search/comparison on Web

PrashantVerma*, Vijay Kumar*, Bharat Gupta*
*WSI, MeitY, New Delhi, India, vermaprashant1@gmail.com
* MeitY, New Delhi, India, {vkumar, bharatg}@meity.gov.in

Abstract.The document formats and protocols that based on character data is mainly prepared for the web. These protocols and formats can be access as resources that contain the various text files that cover syntactic content and natural language content in some structural markup language. In order to process these types of data , it requires various string based operations such as searching, indexing, sorting, regular expressions etc. These documents inspect the text variations of different types and preferences of the user for string processing on the web. For this purpose, W3C has developed two documents Character Model: String Matching and searching that act as building blocks related two these problems on the web and defining rules for string manipulation i.e. string matching and searching on the web. These documents also focus on the different types of text variations in which same orthographic text uses different character sequences and encodings. The rules defined in these documents act as a reference for the authors, developers etc. for consistent string manipulation on the web. The paper covers different types of text variations seen in Indian languages by taking Hindi as initial language and it is important that these types of variations should reflect in these documents for proper and consistent Indian languages string manipulations on the web.

Keywords: W3C, CSS, Unicode, html, head, style, body, Doctype

---

Identification of Disease Resistant Plant Genes Using Artificial Neural Network

Tanmay Thareja, Kashish Goel, Sunita Singhal, Manipal University Jaipur, Jaipur 303007, India
tanmay.189301081@muj.manipal.edu, kashish.189301023@muj.manipal.edu,
sunita.singhal@jaipur.manipal.edu

Much like animals have their defenses against disease-causing pathogens, plants have their own mechanisms to identify and defend against pathogenic microorganisms. Much of this mechanism depends upon disease-resistant genes, also known as 'R' genes. Early identification of these R genes is essential in any crop improvement program, especially in a time when plant diseases are one of the biggest causes of crop failure worldwide. Existing methods operate on domain dependence which have several drawbacks and can cause new or low similarity sequences to go unrecognized. In this paper, a Machine Learning method, employing a domain-independent approach, was developed and evaluated which improves upon or eliminate the drawbacks of existing methods. Data sets were obtained from publicly accessible repositories, and feature extraction generated 10,049 number of features. Batch Normalization was used on the models, and we were able to achieve a 97% accuracy on the test dataset which is greater than anything else in the literature that uses the same approach.

Keywords: Disease Resistant Genes, Plants, Artificial Neural Network, Deep Learning

---

A Comparative Study on A Disease Prediction System Using Machine Learning Algorithms

S.Rama Sree*, A.Vanathi *,Ravi Kishore Veluri*, S.N.S.V.S.C Ramesh*
Professor, Aditya Engineering College,Surampalem,India
Associate Professor, Aditya Engineering College,Surampalem,India
ramasree_s@aec.edu.in, vanathi.andiran@aec.edu.in,ravikishore1985@aec.edu.in, ramesh.snsvsc@acet.ac.in

One of the most essential and fundamental factors that motivates people to seek assistance is their physical well-being. A vast range of ailments affect people nowadays, making them extremely vulnerable. Thus, disease prediction at an early stage has now become increasingly relevant as a result of these developments. Machine Learning is a relatively new technique that can aid in the prediction and diagnosis of diseases. To diagnose respiratory difficulties, heart attacks, and liver disorders, this study employs machine learning in conjunction with symptoms. All of these diseases are the focus of our investigation because they are extremely prevalent, incredibly expensive to treat, and impact a significant number of people at the same time. A variety of supervised machine learning methods, including Naive Bayes, Decision Trees, Logistic Regression, & Random Forests, are used to forecast the disease based on the provided dataset. The discussion of learning categorization based on correctness comes to a close with this conclusion. Flask is also used to construct a platform that allows visitors to forecast whether they will contract a specific illness and take appropriate precautions if they do contract the illness. Among the most important aspects of healthcare informatics is the prediction of chronic diseases. It is critical to diagnose the condition at the earliest possible opportunity. Using feature extraction and classification methods for classification and prognosis of chronic diseases, this study gives a summary of the current state of the art. The selection of elements that are appropriate for a classification system is critical in improving its accuracy. The decrease of dimensionality aids in the improvement of the overall performance of the machine learning system. The use of classification algorithms on disease datasets gives promising results in the development of adaptive, automated, and smart medical diagnostics for chronic diseases, according to the researchers. Using parallel classification systems, it is possible to speed up the process while also increasing the computing efficiency of the final findings. This paper provides a complete analysis of several feature selection strategies, as well as the advantages and disadvantages of each method.

Keywords: Asthma Prediction, Heartstroke Prediction, Liver disease prediction, Naive Bayes, Logistic regression.

Reinforcing Digital Forensics through Intelligent Behavioural Evidence Analysis: Social media hate speech profiling

Barkhashree, Dr. Parneeta Dhaliwal
Manav Rachna University, Haryana, India
barkhashree55@gmail.com, parneeta.cst@mru.edu.in

Cumulative boom in the use of social media for crime incidents creates a need to achieve improvement in overall efficiency by developing tools and techniques. The present manuscript discusses this issue by proposing a model which offer a feasible solution to escalate the criminal investigation by performing digital criminal profiling using social media hate speeches content. The proposed expert system will consider the hate speech content analysis along with other digital footprints of the suspects. The achieved analysis along with the proofs collected manually will be feed into the knowledge base of expert system. The system will automatically process the dataset to lower down the list of suspects using intelligent ML algorithms. Hate speech content analysis achieved using latest intelligent mechanisms, can perform in the boundaries of BEA-S model to create criminal profiling which again act as a base for future investigations. The whole concept and the model are discussed in the paper which surely will upgrade the current investigation process to an innovative stature of digital forensics.

Keywords: criminal profiling, expert system, hate speech, digital forensics, intelligent behavioural evidence analysis, BEA.

Leaf Disease Identification using DenseNet

Ruchi Verma, Varun Singh
Dept. of Computer Science and Engineering Jaypee University of Information Technology, Solan, India
ruchi.verma@juit.ac.in1, varunsingh00712@gmail.com2

To maintain a promising status of global food security, it is imperative to strike a congruous balance between the estimated alarming growth in the global population and the expected agricultural yield to cater to their needs appropriately. An agreeable balance has not been acquired in this respect which could be the cause of the origin of food crisis across the world. Therefore, it is crucial to prevent any direct or indirect factors causing this. Proper growth of plants and protection against diseases is a very instrumental factor towards meeting the qual- ity and quantity of food requirements globally.
Deep learning Methods have gained successful results in the spheres of image processing and pat- tern recognition. We have made an effort in implementing the methods of deep learning for analyzing leaves of plants for prediction and detection of any diseases. Here, we have considered two majorly grown crops in Himachal Pradesh i.e., tomato and potato, for performing various experiments. In our result analysis, we have achieved an accuracy of 96.24% while identifying the diseases in the leaves.

Keywords: Leaf Disease · Deep Learning · DenseNet · Algorithm ·Plants · Detection

---

Dictionary Vectorized Hashing of Emotional Recognition of Text in Mutual Conversation

M. Shyamala Devi, D. Manivannan, N. K. Manikandan, Ankita Budhia, Sagar Srivastava, and Manshi Rohella
Computer Science & Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute ofScience and Technology, Chennai, Tamilnadu, india, shyamaladevim@veltech.edu.in

Emotion detection is a subset of sentiment classification that interacts with emotion processing and analysis. The condition of just being emotional is frequently associated with making sensible qualitative stimulation of feelings or with environmental influence. With increase in the social media usage, people tend to have frequent conversation through several applications. Even police department tend to analyze the victim of any suicidal cases through the personal chat conversation. Machine learning could be used to analyze emotional detection of the person through text processing of their personal conversation. The text conversation dataset with 7480 conversations from KAGGLE warehouse and is used in the execution analysis to detect the emotional analysis. The text conversation dataset is preprocessed by removing the stop words. The tokens are extracted from the text using NGram method. The emotional labels are assigned for the tokens and the machine is trained to identify the emotions during testing. The emotional labels are converted into features to form corpus text for classifying the emotions in the conversation. The corpus is splitted to form training and testing dataset and is fitted to Dictionary Vectorizer, Feature Hashing, Count Vectorizer and Hash Vectorizer to extract the important features from the text conversation. The extracted features from the text conversation are the subjected to all the classifiers to analyze the performance of the emotion prediction. The scripting is written in Python and implemented with Spyder in Anaconda Navigator IDE, and the experimental results shows that the random forest classifier with dictionary vectorizer is exhibiting 99.8% of accuracy towards predicting the emotions from the personal conversations.

Keywords: Machine Learning, Corpus, Vectorizer, Accuracy.

Comparing the accuracy and the efficiency in detection of Coronavirus in CT scans and X Ray images

Sagar C V, Sumit Bhardwaj
Department of Electronics and Communication Engineering, Amity University, Noida
cvsagar304@gmail.com, sumitb460@gmail.com

The Coronavirus pandemic, also known as the Covid pandemic, is a global disease (Coronavirus) pandemic caused by SARS Covid 2019 that causes severe respiratory illness (SARS-CoV-2). Side effects differ incredibly in seriousness, going from subtle to perilous. Individuals who are old or have basic clinical issues are more inclined to foster serious infection. Coronavirus is spread by means of the air when beads and small airborne particles dirty it. In this project we would be analysing the data set images of Chest CT scans and Chest X Rays for the Detection of Corona Virus using the different kind of deep learning algorithms and checking the efficiency of both of them as to which is more accurate and beneficial for detection of the corona virus pandemics so that this study can be used for future detection of COVID in the patients.

Keywords: Coronavirus, Xray, CT scans, Machine Learning

Deep learning on small tabular dataset: using Transfer Learning and Image Classification

Vanshika Jain[1], Kshitiz Shah[2], Meghansh Goel[3]
[1,2,3] Dept. of CSE MAIT New Delhi, India
j.vanshika11@gmail.com, Kshitizshah.work@gmail.com, Meghansh.work@gmail.com

Deep Learning is a subset of machine learning inspired by the human brain. It uses multiple layers of representation to extract specific knowledge from raw input. It is best suited for large image or sound-based datasets. Deep learning methods are generally avoided for small datasets because they tend to overfit them. Transfer learning can be one approach used to solve this problem. However, in the case of tabular datasets, their heterogeneous nature makes transfer learning algorithms inapplicable. This paper discusses a few approaches using a literature review to convert tabular data into images to overcome such limitations. The paper provides a 2-part study wherein we first give a brief overview of transfer learning, enhancing the efficiency of deep learning algorithms and drastically reducing the training time for small datasets. Secondly, we provide a detailed study of different techniques available to convert tabular data into images for image classification, such as SuperTML, IGTD, and REFINED approach. Furthermore, we propose a novel approach inspired by IGTD to create a blocked image representation of the tabular data on which we apply transfer learning to demonstrate the application of deep learning methods on small tabular datasets (with less than 1000 data points).

Keywords: Transfer Learning, Classification, Dataset.

Implementation of a method using Image Sequentialization, Patch Embedding and ViT Encoder to detect the Breast Cancer on RGBA Images and Binary Masks

Ms. Tanishka Dixit+,Ms. Namrata Singh*, Dr. Geetika Srivastava*, Dr. Meenakshi Srivastava!
+Department of Computer Science and Engineering,
NarainaVidhyapeeth Engineering & Management Institute, Panki, Kanpur, India
*Department of Physics & Electronics, [3]Dr. Ram Manohar Lohia Avadh University, Ayodhya, India
! Amity Institute of Information Technology, Amity University, Lucknow, India
tanishka1125@gmail.com, nam2817120@gmail.com, geetika_gkp@rediffmail.com, msrivastava@lko.amity.edu

This paper uses an approach where we will be training the unlabelled and labelled datasets into the system and these datasets comprises of: RGBA images and Binary masks. This will be a benefit in order to get a much better result and more accurate also as pre-training method always helps in getting excellent output. We will use a Transformer Model in this paper where we will apply an image sequentialization technique having 51 million parameters which will help us in attaining the smoother images without any noise. At times, this approach could be time consuming due to used datasets and its sizes, however it shows more accurate and efficient results. Further, we will try to figure out the pixel wise label map using patch embedding technique. Once these techniques are applied then CNN- Transformer Hybrid will come into the role which will encode and decode the images to high-level feature extractions and full spatial resolution respectively. This way of doing encoding and decoding is also known as forward pass and back propagation. Also, it will involve the cascaded upsampler where we will try to use self-attention processes into the design of encoder using the transformers. This entire mechanism will involve few of the best evaluation metrics and those are: Pixel accuracy, IoU, Mean-IoU and Recall/Precision/F1 Score giving the effective and best results.

Keywords: Deep learning, Vision Transformer (ViT), Convolutional Neural Networks (CNNs), Mammography, Image Sequentialization, Up sampler, Breast Cancer

---

An Ensemble model for face mask detection using Faster RCNN with ResNet50

DhivyaShree M, Sarumathi K R, Vishnu Durai R S
Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India.

The computer vision which is an important aspect of Artificial Intelligence. The object detection is the most researchable area with deep learning algorithms. Now in the current COVID – 19 pandemics, the social distancing is a mandatory factor to prevent this transmission of this deadly virus. The government is struggling to handle the persons without wearing masks in public places. Our work concentrates on the object detection of face masks using the state-of-the-art methodologies like YOLO, SSD, RCNN, Fast RCNN and Faster RCNN with different backbone architectures like ResNet, MobileNet, etc. This paper brings out various ensemble methods by combining the state of art methodologies and compare those combinations to identify the best performance, in choice of the dataset of the application. We have obtained the highest performance benchmark with the usage of Faster RCNN – ResNet50 among the other ensemble methods. All the performance evaluation metrics are compared with one other with the same face mask detection image dataset. In this paper, we present a balancing collation of the ensemble methods of object detection algorithms.

Keywords: Face mask, Neural networks, Object detection, Faster RCNN, Resnet

---

A Novel Approach to Detect Face Mask in Real Time

Sumita Gupta*, Rana Majumdar+, Shivam Deswal*
*Department of CSE, Amity School of Engineering & Technology, Amity University, Noida
+Department of Computer Science & Engineering, Swami Vivekananda University, Kolkata

COVID-19, a deadly virus outbreak in the entire world, infected countless number of people and leads to death of millions of people. Economy of countries halted, people are stuck up, the situation is becoming worse day by day that no one expected. The COVID-19 can be spread through airborne droplets, aerosols and other carriers. So, the guidelines are released which promotes 3 key points for its prevention: (a) maintain social distancing, (b) sanitization, and (c) most important, wearing of mask in public places. But unfortunately, people are avoiding these measures leading to spread of the disease. So, there is need of some sort of security guards on ground that ensures people would strictly follow the guidelines. But, it's too risky for the lives of security guards. Fortunately, we live in 21st century. There are so powerful technologies like Machine Learning, Artificial Intelligence, Deep Learning and many more. So, it is possible to develop a way where machines will help us to ensure the guidelines to be followed. There is no need for a physical person to watch over crowd. This research paper proposed a work to implement machine-learning algorithm to ensure people wear the proper face mask or not. It can be used in public places like airports, railway stations or at main gate of societies to ensure that no one without mask may enter the society and similarly can be used at stores.

Keywords: COVID-19, Face Detection, Mask Detection, Machine Learning, SVM, Viola-Jones Algorithm

---

A Novel Approach for Detecting Facial Key Points using Convolution Neural Networks

Rishi Kakkar, Y.V. Srinivasa Murthy
Department of CSE, VIT University, Vellore, Tamil Nadu, India - 632 014.
rishi.2018@vitstudent.ac.in, vishnu.murthy@vit.ac.in

The task of face recognition is having many real-time applications in which the process of facial keypoint detection is considered to be an intermediate and crucial step. The amount of keypoints that are using for face recognition decides the computational requirements of the algorithm. In this paper, an effort has been made to detect the useful 15 facial key points using convolutional neural networks and compared with the state-of-the-art system with 30 facial key points. We made an effort to identify the 15 facial key points (6 points from eye + 4 points from eyebrows + 4 points from lips + 1 point from the nose) by using the proper hyperparameters for convolutional neural network. It is found that the performance of the proposed system is quite similar when compared to the system with 30 facial key points.

Keywords: Batch size · Convolutional neural networks · Dropout Layer· Facial keypoint detection · Face recognition.

Effective Hyperspectral Image Classification using Learning Models

Sushmita Gautam, Kailash Chandra Tiwari
Multidisciplinary centre for Geoinformatics, Delhi Technological University, Delhi, India
sushg019@gmail.com

Recently, machine learning has produced appreciable performance results on various visual computing related studies, including the classification of common hyperspectral images. This study aims to compare the results of different machine learning models for the classification of a hyperspectral image dataset. The hyperspectral data captured from AVIRIS sensor covering scene over the Indian Pines test site in North-western Indiana and consists 224 spectral

reflectance bands. The ground truth has sixteen classes including vegetation crops, built structures, etc. Accuracy assessments and confusion matrices were used to evaluate classification performance. The study includes classification results of mainly three learning models including dimensionally reduced data via PCA for SVM classification, CNN and k-NN. The overall accuracy in PCA-SVM results was 72.38%, CNN was 85% and k-NN was 66.21% concluding the better efficiency of CNN classification for the hyperspectral dataset. Hence CNN classification technique succeeded in the hyperspectral image classification.

Keywords: Hyperspectral, classification, accuracy, Support vector machines, CNN, k-NN.

Survey on Automatic Speech Recognition Systems for Indic Languages

Nandini Sethi, Amita Dev
Indira Gandhi Delhi Technical University for Women, Delhi, INDIA-110006
nandini053phd20@igdtuw.ac.in, vc@igdtuw.ac.in

For the past few decades, Automatic Speech Recognition (ASR) has gained a wide range of interest among researchers. From just identifying the digits for a single speaker to authenticating the speaker has a long history of improvisations and experiments. Human's Speech Recognition has been fascinating problem amongst speech and natural language processing researchers. Speech is the utmost vital and indispensable way of transferring information amongst the human beings. Numerous research works have been equipped in the field of speech processing and recognition in the last few decades. Accordingly, a review of various speech recognition approaches and techniques suitable for text identification from speech is conversed in this survey. The chief inspiration of this review is to discover the prevailing speech recognition approaches and techniques in such a way that the researchers of this field can incorporate entirely the essential parameters in their speech recognition system which helps in overcoming the limitations of existing systems. In this review, various challenges involved in speech recognition process are discussed and what can be the future directives for the researchers of this field is also discussed. The typical speech recognition trials were considered to determine which metrics should be involved in the system and which can be disregarded.

Keywords: Speech Recognition, Acoustic modelling, Hidden Markov model, Dynamic Time Wrapping, Mel frequency Cepstrum Coefficient.

Feature Extraction and Classification for Emotion Recognition using Discrete Cosine Transform

Garima#*1, Nidhi Goel#2, Neeru Rathee*3
#Indira Gandhi Delhi Technical University for Women, Delhi, India
*Maharaja Surajmal Institute of Technology, Delhi, India
11592garima@gmail.com, 2nidhi.iitr1@gmail.com, 3neeru_rathee@msit.in

In recent years, the rigorous development in tools and techniques for biomedical signal acquisition and processing has drawn interest of researchers towards EEG signal processing. Human emotion recognition using Electroencephalography (EEG) signals has proved to be a viable alternative as it cannot be easily imitated like the facial expressions or speech signals. In this research, authors have explored EEG signals for behavior analysis using Discrete Cosine Transform and classifying the signals using K-Nearest Neighbors. The algorithm is then evaluated on publically available DEAP Dataset. Experimental results are expressed in terms of F1 score, accuracy, precision and recall. The performance metrics evaluation for the classification of the emotional labels of DEAP dataset has further confirmed the effectiveness of the research. Comparison evaluation with the recent state-of-the-art methods further confirms the efficacy of the proposed work.

Keywords: Discrete Cosine Transform, EEG, KNN, classification

---

Cross Linguistic Acoustic Phonetic Study of Punjabi and Indian English

Amita Dev, Shweta A. Bansal and Shyam S. Agrawal
1.3Indira Gandhi Delhi Technical University for Women, Delhi, 2K.R Mangalam University, Gurugram
1amita_dev@hotmail.com., 2s.bansal6281@gmail.com, 3ss_agrawal@hotmail.com

Punjabi and English are the languages that does not belong to the similar family. For instance, English belongs to the West German languages while Punjabi is a part of the Indo Aryan family. Regional languages have an impact on English in this borrowing and code-mixing process because borrowed words go through a make-up caused by phonetic features of the regional languages. In India, English is the medium of media, science, and technology & its influence on the native languages of the country is significant. This study shows how the regional dialects of the country have a significant influence on the way English is pronounced. It is an effort to define the phonemic changes in Punjabi and Indian English. The purpose of this study was to see if the sound pattern of Indian English varies depending on the speakers' native languages or if it is the same irrespective of the speakers' native languages and also to identify words that differ in pronunciation from Standard English and are clearly marked by an influence of the first language of the native Punjabi speakers of India.

Keywords: Acoustic Phonetics, Indian English, Punjabi

Comparative Analysis of NLP Text Embedding Techniques with Neural Network Layered Architecture on Online Movie Reviews

Hemlata Goyal[1] Amar Sharma[2] Ranu Sewada[3] Devansh Arora[4] Sunita Singhal[5]
[12345]Manipal University Jaipur

In NLP world, there is a need to convert the text data into numerical form in a smart way of text embedding with the machine learning architecture. In this research, the comparative text embedding methods of Binary Term Fre quency, Count Vector, Term Frequency - Inverse Document Frequency, and Word2Vec is used for converting text to meaningful representations of vectors, containing numerical values. In order to analyze the performance of the various text embedding techniques, Neural Network Layered Architecture is designed for movie review's polarity classification to include input layer, dense layers followed by the ReLU (Rectified Linear Unit) activation layers and Sigmoid activation function to make classifications on
the basis of training-testing performance. Word2Vec text embedding scored the highest training and testing accuracy among all the text embedding techniques of Binary Term Frequency, Count Vector, Term Frequency - Inverse Document Frequency, and Word2Vec with 89.75% and 86.94% respectively with $\pm 1.0\ error$ for the online movie reviews.

Keywords: Natural Language Processing, Text Embedding, Count Vector, Binary Term Frequency, TF-IDF, Word2Vec

---

Current state of Speech Emotion Dataset- National and International level

Surbhi Khurana Amita Dev 1Surbhi001phd20@igdtuw.ac.in 1vc@igdtuw.ac.in
1Indira Gandhi Delhi Technical University for Women, Delhi, INDIA

Research on emotion extraction from human speech is transitioning from a phase of exploratory study to one with the potential for significant applications, particularly in human–computer interaction. The establishment of relevant databases is critical to progress in this area. Scope, naturalness, context, and descriptors are the four primary challenges that must be considered while constructing databases of emotional speech. The current state of the art is examined for status of available datasets for internally spoken languages and Indian languages.

Keywords: Speech emotion recognition, Datasets, Context, Scope, Naturalness, International Datasets, Indian datasets.

Deep Learning Approaches for Speech Analysis: A Critical Insight

[1]Alisha Goyal, [2]Advikaa Kapil, [3]Sparsh Sharma, [4]Arun Sharma
[1,4]Indira Gandhi Delhi Technical University for Women, Delhi, India.
[2]Sanskriti School, Chanakyapuri, New Delhi, INDIA
[3]Delhi Technological University, New Delhi
[1]alishagoyal132@gmail.com, [2]kadvikaa6739@gmail.com, sparsh1706@gmail.com, [4]arunsharma@igdtuw.ac.in

The main objective of speaker recognition is to identify the voice of an authenticated and authorized individual by extracting features from their voices. The number of published techniques for speaker recognition algorithms is text dependent. On the other hand, text-independent speech recognition appears to be more advantageous since the user can freely interact with the system. Several scholars have suggested a variety of strategies for detecting speakers, although these systems were difficult and inaccurate. Relying on WOA and Bi-LSTM, this research suggested a text-independent speaker identification algorithm. In presence of various degradation and voice effects, the sample signals were obtained from an available dataset. Following that, MFCC features are extracted from these signals, but only the most important characteristics are chosen from the available features by utilizing WOA to build a single feature set. The Bi-LSTM network receives this feature set and uses it for training and testing. In the MATLAB simulation software, the proposed model's performance is assessed and compared to that of the standard model. Various dependent factors, like accuracy, sensitivity, specificity, precision, recall, and Fscore, were used to calculate the simulated outputs. The findings showed that the suggested model is more efficient and precise at recognizing speaker voices.

Keywords: Speaker recognition system, Artificial intelligence, Whale optimization algorithm, Recurrent neural network, etc.

Building speech corpus in rapid manner to adapt a general-purpose ASR system to specific domain

Mukund K Roy1, Sunita Arora2, Karunesh K Arora3, Shyam S Agarwal4
1,2,3 SNLP Lab, CDAC Noida, UP, India, 4 KIIT, Gurugram, Haryana, India
mukundkumarroy@cdac.in, sunitaarora@cdac.in, karuneshaarora@cdac.in, karuneshaarora@cdac.in

The situation prevalent due to Covid-19 has affected the traditional speech database collection process by reaching out persons in one-to-one man ner. In this paper, we describe an alternate approach adopted for faster speech dataset construction in Hindi language for building domain adapted Automatic Speech Recognition (ASR) for agriculture domain. We resorted to two methods – one utilizing App for speech samples collection and second through domain specific YouTube videos. In this paper we outline building of App and several filtering (signature music, advertisements and cross talks) and post-processing steps for speech database collected through on-line videos. The

paper also de scribe's novel idea of making speech segments suitable for training an end-to end ASR system. The process of annotating included combination of utilizing existing ASR systems and manual post correction to save time. Our experiment resulted in collection of speech data from 236 speakers through App and 106 hours of speech data through on-line videos. The experiment of re-training ASR with enhanced data reveals that exercise results in adapting it for a particular domain in a rapid manner.

Keywords: Speech corpus, Data collection, Speech recognition.

A contrastive view of vowel phoneme assessment of hindi, indian english and american englishspeakers

Pooja Gambhir[1] Amita Dev[2] S.S. Agrawal[3]
[123]Indira Gandhi Delhi Technical University for Women (IGDTUW), Delhi India

Acoustic, Phonetic and accented variations play an important role in pronunciation assessment of different languages uttered by speakers. In this paper, we consider the pronunciation assessment of vowels of North Indian Hindi, Indian English and American English uttered by varying subjects. Indian English is a language spoken in the India as a second language. It is highly influenced by variety of Native Indian languages as well as usage, cultural, regional, social and educational background. It is observed that Indian speakers tends to speak English phonemes close to the articulation of their native language rather than that of the American English. This paper contributes towards the contrastive study of vowel triangles and the appearance of the closeness of vowel cardinal space (ɑ, i, u) of Hindi, Indian English and American English using distance formula.

Keywords: Vowels, Pronunciation, Hindi, Indian English, American English, Articulation.

A Critical Insight into Automatic Visual Speech Recognition System

Kiran Suryavanshi1, Suvarnsing Bhable2 and 3Charansing Kayte
12 Department of Computer Science & IT,
Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, MH, INDIA
2 Government Institute of Forensic Science, Aurangabad, MH, INDIA
1suryawanshikiran@gmail.com, 2suvarnsingbhable@gmail.com, 3Charankayte@gmail.com

This research paper investigated the robustness of the Automatic Visual Speech Recognition System (AVSR), for acoustic models that are based on GMM and DNNs. Most of the recent survey literature is surpassed in this article. Which shows how, over the last 30 years, analysis and product growth on AVSR robustness in noisy acoustic conditions has progressed? There are various categories of languages covered, with coverage, development processes, Corpus, and granularity varying. The key advantage of deep-learning tools, including a deeply convoluted neural network, a bi-directional long-term memory grid, a 3D convolutional neural network, and others, is that they are relatively easy to solve such problems, which are indissolubly linked to feature extraction and complex audiovisual fusion. Its objective is to act as an AVSR representative.

Keywords: Speech Recognition, AVSR, MFCC, HMM, CNN, DNN

Analytical Approach for Sentiment Analysis of Movie Reviews Using CNN and LSTM

Arushi Garg1, Soumya Vats2, Garima Jaiswal3, Arun Sharma4
Indira Gandhi Delhi Technical University for Women, Delhi
gargarushi2002@gmail.com1, vatssoumya03@gmail.com2, garima121@gmail.com3, arunsharma@igdtuw.ac.in4

With the rapid growth of technology and easier access to the internet, several forums like Twitter, Facebook, Instagram, etc., have come up, providing people with a space to express their opinions and reviews about anything and everything happening in the world. Movies are widely appreciated and criticized art forms. They are a significant source of entertainment and lead to web forums like IMDB and amazon reviews for users to give their feedback about the movies and web series. These reviews and feedback draw incredible consideration from scientists and researchers to capture the vital information from the data. Although this information is unstructured, it is very crucial. Deep learning and machine learning have grown as powerful tools examining the polarity of the sentiments communicated in the review, known as 'opinion mining' or 'sentiment classification'. Sentiment analysis has become the most dynamic exploration in NLP (natural language processing) as text frequently conveys rich semantics helpful for analyzing. With ongoing advancement in deep learning, the capacity to analyze this content has enhanced significantly. Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) are primarily implemented as powerful deep learning techniques in Natural Language Processing tasks. This study covers an exhaustive study of sentiment analysis of movie reviews using CNN and LSTM by elaborating the approaches, datasets, results, and limitations.

Keywords: CNN, LSTM, Movie Reviews, Sentiment Analysis

Auscultation of respiratory sounds for diagnostics using Deep Learning approaches – Machine hearing a cognitive service for aiding clinical diagnosis

Arun Gopi and Sajini T
CDAC, Thiruvananthapuram, Kerala, India

Auscultation is being used for screening and monitoring respiratory diseases and is performed using a stethoscope. Auscultation the detection of abnormal respiratory sounds requires skilled medical professionals or clinical experts for diagnosis and an early diagnosis is always recommended for getting a higher probability of both curing and recovery. Respiratory disease being the most common with high morbidity the biggest challenge faced is the scarcity of clinical experts, non-availability of experts in rural and geographically challenged regions. Auscultation is an essential part of the physical examination, real-time and very informative, but based on the auditory perception of lung sounds. This requires clinicians' considerable expertise and the perception variability may lead to misidentification of respiratory sounds. In this work, we are proposing an objective evaluation approach using deep learning techniques to address the limitations of the existing approach. A machine hearing technique to aid clinical decisions. In this work, breath sounds are used for analysis. The wheeze and crackles are the indicators of underlying ailments like namely Pneumonia, Bronchiectasis, Chronic Obstructive Pulmonary Disease (COPD), Upper Respiratory Tract Infection (URTI), Lower Respiratory Tract Infection (LRTI), Bronchiolitis, Asthma, and healthy. These sounds were analyzed for classifying the 8 categories of pulmonary diseases. CNN and RNN architectures were used for the classification of respiratory diseases. Features like Mel Frequency Cepstral Coefficients are extracted from the breath sounds. These coefficients were used as the feature for training the CNN and RNN architecture. Data Augmentation techniques like time stretching and shifting were applied to handle the imbalance in the data set. The CNN architecture gave a better accuracy of 0.89 and RNN with a slightly low of 0.833. The proposed approach proves to be a successful solution for the classification. The accuracy can be further improved with real-time data. In future, this can be extended to develop a machine hearing as a decision support system for the clinical experts.

Keywords: deep learning, pulmonary sound classification, wheeze, crackle, MFCC, auscultation.

An Analysis of Image Compression Using Neural Network

[1]Mohit, [2]Pooja Dehraj
1,2Department of Computer Science Engineering, Noida Institute of Engineering and Technology, India
1motilalnehroo@gmail.com, 2drpooja.cse@niet.co.in

Image compression belongs to the area of data compression because the image is itself made up of data and the task of compressing images has become vital in our current life. Because the scenario is that images are required to build more attractive contents, also in today's world smartphones cover large fraction of internet traffic and are having low data bandwidth on average. Due to these factors and restrictions on bandwidth and other computing capabilities it has necessary for developers of websites/applications to reduce either size or resolution or both of an image to improve responsiveness of your websites/apps. For this purpose, image compression is divided into two categories these are lossless image compression and lossy image compression. The requirement for lossless image compression is that during the decompression process the image data must be recovered without/with negligible loss in image quality while in lossy image compression certain amount or level of error is allowed in image data to achieve better compression ratios and performance [1]. Neural Networks because of their good performance have been used to implement the task of image compression and there are multiple modified neural networks that are proposed to perform image compression tasks; however, the consequent models are big in size, require high computational power and also best suited for fixed size compression rate [2]. and some of them are covered in this survey report.

Keywords: Quantization, PSNR, loss function, Image scaling, convolution layer

---

Deep Neural Networks for Spoken Language Identification in Short Utterances

Shweta Sinhar1and S S Agrawal2
1 Amity School of Engineering and Technology, Amity University Haryana, India
2 KIIT College of Engineering, Gurugram, Haryana
meshweta 7@rediffmail.com

Abstract. This work addresses the problem of language identification (LID) in short-time utterances. This paper presents a feed-forward deep neural network (FF-DNN) for language identification using acoustic features of short-time utter ances. Two network topologies for DNN have been checked for their performance in LID task. The results are compared to a state-of-the-art i-vector system using MFCC-SDC as acoustic feature and support vector machine (SVM) as backend classifier. Both the systems have been implemented for the identification of two Indian languages (Hindi and Punjabi). The speech utterances are divided into short segments of 5 sec, 10 Sec, 20 Sec and 35 Sec. The performance of the system is measured in EER (%) and for short time segments, a relative improvement of 3% is achieved by the DNN system, whereas the average error rate overall the utterances was decreased by 2% using DNN.

Keywords: LID · DNN · i-vectors · Support Vector Machine · MFCC-SDC

Indira Gandhi Delhi Technical University for Women (IGDTUW)

(Established by the Government of NCT of Delhi vide Delhi Act 9 of 2012)
Education - Enlightenment- Empowerment

Indira Gandhi Delhi Technical University for Women (IGDTUW), Delhi was established by Delhi Government in May 2013 vide Delhi State Legislature Act 9,2012 in order to play a major role in the capacity building and imparting high quality technical education to the Women of our country.

The University offers B. Tech. in (CSE (AI), CSE, IT, ECE and MAE), Integrated B.Tech. (ECE)+MBA, Integrated BTech (MAE)+MBA, B. Arch., M.Tech. in (CSE-AI, IT-ISM, ECE –VLSI and Robotics & Automation), M. Plan (Urban Planning), BBA, MBA, MCA and Ph.D. courses. In last seven years, the University has achieved a number of milestones and has been awarded the 2nd fastest growing university by India Today group. University has also achieved ISO 9001:2015 QMS Certification from STQC Certification Services, Government of India. The recent achievements of the University are as under:

- Ranked **2nd** (under Women Colleges and Universities) and in Band A (under Overall Ranking) in **Atal Ranking of Institutions on Innovation Achievements** (ARIIA)-2020 **Awarded by Hon'ble Vice President of India** on 18th August, 2020.
- Conferred with **FICCI University of the Year-2021** Award for its **continuous commitment towards Quality Education**.
- Awarded the 2nd **fastest growing University** by India Today group-2021.
- Ranked in the band of 101-200 by prestigious '***Times Higher Education World Impact Rankings 2021'*** for Sustainable Development Goal **(SDG-4 & SDG-5) i.e., Quality Education & Gender Equality**
- IGDTUW is the only Government University in India, selected by **QS I QUAGE for awards of E LEAD INSTITUTE** for E-Learning Excellence for Academic Digitization.
- Awarded with a **Diamond Rated University** through rigorous analysis against performance metrics by **QS I-Gauge** (Indian College & University Rating)
- Ranked **16th for Entrepreneurial Spirit** and **64th in Innovative University** in **World University Ranking Impact (WURI) 2021Research and Industrial Collaborations**

University has signed several MoU with leading industries including NASSCOM Foundation, IBM, Cyberpeace Foundation, Fluor Daniel, EATON and several others for technical upgradation of students. IGDTUW has also received research and project grants from various Government and Private agencies including of Rs.363 Lakhs from Department of Science and Technology for consolidation of University Research and Excellence and Rs. 70 Lakhs for establishing Centre of Excellence in Artificial Intelligence. University is having strong and active linkage with various Government bodies including DSSSB, MCD, Delhi Police and others for developing and customizing their software and to provide trainings to the staff on emerging trends like Cyber Security, Forensics and others.

Prestigious Collaborations: IGDTUW has partnered with several industries for knowledge updation & bridging the gap between academia and industry. Few of these collaborations worth mentioning are: **NASSCOM & CISCO** for setting up of **thingQbator** Innovation Lab & **M/s Flour Daniel** for setting up Turning Lab (both under their CSR activities), **EATON Foundation** for Internship placement and scholarship to students, **IBM India Pvt Ltd** for academic growth and Technological advancement, **Amazon AWS**, **Ennoble IP, GiZ Germany,** etc.

Placement

Retaining and revitalizing the past glory, the highly energetic and dynamic team of students and faculty ensure that the students have an incredibly great platform to prove their skills and merit. Many of our students have been placed in renowned National and Multinational companies. We are honored to have corporate giants like Google, Microsoft, Uber, Intuit, Goldman Sachs, American Express, Walmart, Netapp, Cisco, GE, Nestle, Siemens, Accenture, TCS, DELL, Qualcomm, Mahindra & Mahindra, Maruti, Honda, Bajaj Automobile, MG Motor, Tata Motors, Nestle, Eaton, HUL, Ather Energy, Siemens, Honeywell, Cameron, Fluor Daniel, and list is endless, visiting us annually and offering great opportunities to our students.

The Placement Season for the batch 2019-20 started with the high note and more than 96 companies have visited the campus for about 316 Full Time Offers, 50 pre-placement Offers and more than 189 internship offers. The CTC of Rs.43.3 Lakhs from Microsoft IDC and internship package of Rs.1.5 Lakhs per month from Uber are the highest in each category. First time Google made its presence in the campus and specially organised the "Kickstart" Roadshow especially for IGDTUW students.
It is a matter of stupendous pride to share that compared to the last year, there has been a rise in average CTC from 10.3 Lakhs per annum to 13.33 Lakhs per annum and a rise in highest package from 41.6 Lakhs to 43.3 Lakhs.

The highest CTC of Rs. 59.45 Lakhs from Google and internship package of Rs.1.6 Lakhs per month from Uber was the highest in each category. It is a matter of stupendous pride to share that compared to the last year, there has been a rise in average CTC from 13.3 Lakhs per annum to 19.15 Lakhs per annum and a rise in highest package from 43.3 Lakhs to 59.45 Lakhs.

Awards and Recognition
The University students are encouraged to participate in various National and International events, competitions including Hackathons etc. Recently, Ms Monika, a student of MTech-IT (ISM) is awarded with a internship at FS-ISAC, USA with a scholarship of worth USD10,000.

Last year, 3 students were selected for Google I/O Global competition and one student of IGDTUW won the Facebook Global Prize Award of USD 10000 in Facebook Annual Conference. Every year, teams from IGDTUW girls have won several prizes including first prize in Smart India Hackathon.

IGDTUW-Anveshan Foundation

University has promoted and incorporated a Section 8 Company with name IGDTUW-Anveshan Foundationk to propagate entrepreneurial culture and ecosystem among women in 2016. Anveshan Foundation is funded by Directorate of Training and Technical Education (DTTE), Govt. of NCT of Delhi with a support grant of Rs. 2 Crores and is also recognized by Department of Science and Technology (DST), GOI as Technical Business Incubator (TBI). All kind of facilitation in four phases of complete incubation and venture development cycle i.e., i) Pre-incubation, ii) Incubation, iii) Acceleration and iv) Post-incubation are provided by IGDTUW-Anveshan Foundation. Currently 12 companies led by young women entrepreneurs are doing excellent work in various domains of expertise. Also, more than 100 students are working on their ideas in pre-incubation phase.

Recently IGDTUW Anveshan Foundation has signed a Contract with GiZ, a German organization for Strengthening Incubation & Start-up Ecosystem for Women Entrepreneurs in Delhi NCR under which, events like boot camps, training sessions, workshops and mentoring will be sponsored by GiZ. Total contract value is Rs. 25 Lakhs.

Anveshan Foundation is conducting regular internships, pragmatic workshops and short-term courses on a) Start-up Management, b) Innovation & Entrepreneurship, c) Digital Marketing, d) Python, e) Family Business & Entrepreneurship and in many more technical and non-technical areas for confidence building among aspiring women entrepreneurs.

Incubation Center has also signed ten MoUs with other colleges, institutes and universities to facilitate and establish incubation centers in their premises and to accelerate and diversify businesses of their students and Incubatees through various channels.