

3.1.1 Promotion of Research and Facilities

Gayatri Vidya Parishad College of Engineering for Women deeply committed in promoting research and innovation to create vibrant academic environment among staff and students. The college regularly upgrades its laboratories and facilities, making sure that faculty and students get good support for their research and projects in their domain. A modern digital library with internet access is also available, providing resources from globally reputed publishers. This helps faculty and students get the latest tools, technologies, and scholarly content required for effective research outcomes.

A comprehensive Research Promotion Policy was introduced by the governing body of the college, including an incentive policy for faculty to support scholarly activities such as publications and patent filing. In the recent academic year 2024-2025, our faculty achieved good results in research: four books were published, four book chapters were contributed, five patents were filed, and research papers were published in Scopus (16) and Science Citation Indexed (5) journals. To appreciate the achievements, management provided cash incentives to the faculty. These outputs highlight the fruitful implementation of policy and faculty motivation. The college is committed to keep improving research facilities, establishing center of excellence labs and will continue these efforts in future also.

GAYATRI VIDYA PARISHAD COLLEGE OF ENGINEERING FOR WOMEN
Madhurawada :: Visakhapatnam — 530 048

No.GVPW/Internal Note/01

Date: 19/03/2024

INTERNAL NOTE

1. We are happy to inform the faculty of GVPCEW that as per the GVPCEW incentive policy for promotion of Publication and patents, the College has credited the cash incentives to the respective bank accounts of the faculty members on 19th March, 2023 as per the data available with the college.
2. The policy document is attached herewith for reference.


PRINCIPAL

Encl: Policy Document

To

All Faculty members

Copy to:

1. Secretaryj GVP for information please
2. Vice-principal, GVPW
3. All Heads of Departments

GAYATRI VIDYA PARISHAD COLLEGE OF ENGINEERING FOR WOMEN
Madhurawada :: Visakhapatnam —530 048

Incentive policy for Promotion of Publications and Patents

Gayatri Vidya Parishad College of Engineering for Women is committed to prioritizing research and innovation. This will serve as the basis for knowledge generation, which forms the bedrock of all academic fields. It plays a significant role in enhancing social welfare and fostering national development. Research activities thrive in an environment of academic freedom, which ensures that faculty and researchers have the liberty to inquire and share their findings. Nevertheless, it is essential that research conducted within the academic realm upholds principles of honesty, integrity, trust, and accountability. It is important to ensure that it meets the necessary scientific and ethical standards. The primary objective is to acquire knowledge that has a positive impact on society. The College aims to foster these values by implementing policies that facilitate the attainment of targeted research outputs, including patents.

The faculty of the College is committed to conducting both basic and applied research in order to enhance the institution's research capabilities. They are encouraged to pursue research goals through engagement in research, industry connections, and collaborations with reputable universities and scientific institutions, both in India and abroad. These efforts align with the College's core objectives.

An essential aspect of academia involves the publication of research in reputable journals. Publication serves as a means of communication among researchers in a specific field, preserving significant research endeavors, and acknowledging the contributions of both researchers and institutions.

As per the directives of the Secretary, Gayatri Vidya Parishad the guidelines for the sanctioning of cash incentives to the faculty of the college for publishing research papers in various international and national journals in every calendar year w.e.f. calendar year 2021-2022 as follows:

1. A maximum of four journal publications in a calendar year from a faculty member is considered for cash incentive with the following rule:
 - > A maximum of 2 SCI/SCIE indexed journals and 2 Scopus indexed journals.
 - In case of absence of publication in SCI/SCIE indexed journals then a maximum of 3 Scopus indexed journals are considered for incentives.

2. The cash incentive is as follows:
- For SCI/SCIE indexed journals : Rs 4000 (For 3rd/4th Positions of Authorship)
: Rs 2500 (For other positions of Authorship)
 - For Scopus indexed journals : Rs 2500 (For 1st / 2nd Positions of Authorship)
: Rs 2000 (For 3rd/4th Positions of Authorship)
: Rs 1500 (For other positions of Authorship)

:Rs 5000 (For 1st / 2nd Positions of Authorship)

- > For UGC Care journals : Rs 1000 (For Positions of Authorship)
: Rs 750 (For 3rd/4th Positions of Authorship)
• Rs 500 (For other positions of Authorship)
3. For Patents Granted
- Rs 10000 (For 1 (2nd Positions)
 - : Rs 7500 (For 3 / 4 Positions)
 - Rs 5000 (For others Positions)

4. Authoring print Books through reputed publishers:

- Full text book through reputed International publishers-Wiley, Elsevier, CRC Press, IEEE, ASME, Alp, OUP, AMS, SIAM, Taylor and Francis, ASCE, SPIE, IMA, Springer, Mc Graw Hill, Sage Rs.7500
- Full text book through other reputed publishers Rs.3000
- General Books (non-technical & non-scientific) Rs.1000
- Chapter of a book/Monograph (Not as a conference proceeding) Rs.2000
- Chapter of a book/Monograph (As a conference Proceedings) Rs.1000


Vice-Principal


Principal

**List of Incentives as per “GVPCEW incentive policy for
Promotion of publication and Patents” for Calendar Year 2024**

S.No	Faculty	Department	Research output	Amount	Total Amount
1.	Prof. Raj Kumar Goswami	ECE	4Patents (1/1,1/1, 1/4,1/6)	10000+10000 +10000+10000=40000	60000
	Prof. Raj Kumar Goswami	ECE	3Books (1/1,1/4,2/3) + 2Book Chapters (1/2,1/2)	9000+4000=13000	
	Prof. Raj Kumar Goswami	ECE	5Scopus (1,4,4,1,4,)	2500+2000+2000+2500+2000 =Limited(2500+2500+2000)= 7000	
2.	Dr.P.M.K.Prasad	ECE	2Scopus (3,6)	2000+1500	3500
3.	Dr.L.Ganesh	ECE	Book Chapter (2/2)	2000	5500
	Dr.L.Ganesh	ECE	2Scopus (6,3)	1500+2000	
4.	Dr.B.Vijaya Lakshmi	ECE	Book Chapter (2/2)	2000	2000
5.	Dr.R V S Lakshmi Kumari	EEE	Scopus(5)	1500	1500
6.	Botta Kusuma Kumari	EEE	Scopus(5)	1500	1500
7.	M.Krishna	EEE	IEEE(1)	2500	2500
8.	Prof.P V S Lakshmi Jagadamba	CSE	Patent (3/3)	7500	7500
9.	M.Aswini	CSE	SCOPUS (2)	2500	2500
10.	Dr.K.Rohini	CSE	Scopus(1)	2500	2500
11.	Dr. K. Purushotam Naidu	CSM	Books (4/4)	3000	4500
	Dr. K. Purushotam Naidu	CSM	Scopus (6)	1500	
12.	P.Sridevi	IT	UGC(1)	1000	1000
13.	Dr.K.L.Sai Prasad	BSH	Scopus (3,1)	2000+2500	4500

14.	Dr.Ch.S. Lakshmi	BSH	SCI(6,4,3)	2500+4000+4000	10500
15.	Dr.Vemakoti Siva Jahnvi,	BSH	SCI(1)	5000	5000
16.	B. Rajesh Babu,	BSH	SCI(2)	5000	5000
17.	Dr. T. Poorna Kantha,	BSH	Scopus(4)	2000	2000

1.Prof. Raj Kumar Goswami:

Patents Granted:

1)Title: "LI-FI based Data transfer device between mobile phones"

Inventors Name: Prof. Raj Kumar Goswami

Patent Office:Indian Patent, Government of India,

Month Year: August 2024.

2)Title: "High Performance Antenna for 5G Communication Networks"

Inventors Name: Prof. Raj Kumar Goswami

Registration Date:29-03-2024.

Date: 09-04-2024

Office: Intellectual Property office UK. Design Number:6355940,

3)Title: ACCIDENT NOTIFYING DEVICE FOR TWO- WHEELER VEHICLES

Inventors Names: 1.Dr. Raj Kumar Goswami 2.V. Porselvan 3.Jaydeep Parmar 4.G. Ramachandran

Design No: 422746-001

Date: 10-07-2024

Issue Data: 04-11-2024

4)Title:AI BASED HEALTH MONITORING CAR SEAT

Inventors Names:1.Dr. Raj Kumar Goswami 2. Dr Syed Haider Ali 3.Dr. Shan-E-Fatima 4.Suman Kr Mishra 5.Syed Asghar Hussain Rizvi 6.Digesh Pandey

Design No.411016-001

Date: 19-03-2024

Issue Date: 28-11-2024

Books:

1. Dr.Raj Kumar Goswami, "The Art of Design Thinking", SAN International Scientific Publications, ISBN:978-81-972500-9-5, Year: April 2024.

2.Dr.Raj Kumar Goswami, Mrs S.Nandhini Devi, Dr.ShwetaA.Bansal and Mrs. Pushpa Prasanna, “Data & Information Security,”ISBN: SAN International Scientific Publications, ISBN: 978-81-964397-4-3, Year:2024.

3.Gottapu Sasibhushan Rao, Dr.Raj Kumar Goswami and M.N.V.S.S.Kumar, “Mobile and Cellular Communication(Including 5G & beyond and Microstrip Antennas,” Paramount Publishers, ISBN: 978-81-967696-8-0, Year:2024.

Books Chapters:

1. Dr.Raj Kumar Goswami, Dr.GaneshLaveti, “Edge Computing in 5G and IoT Enhancing Data Processing Real-Time Analytics and Reducing Latency for Critical Applications”, Publisher - RADemics Research Institute, Chapter 10 , page 306-351.ISBN: 978-81-979336-3-9, Year:2024.

2.Dr.Raj Kumar Goswami and Kinjal Goswami, “Navigation Autonomy: Challenges and Opportunities in autonomous vehicle Communication published in Innovation Interconnected: Exploring the frontiers of computing and communication technologies, SAN International Scientific Publications, Year: Feb 2024.

Journal Publications:

1. Raj Kumar Goswami, “Trellis Coded Modulation having Rate 6/7, 128 State with 128 QAM modulation designed for fading Channel,” J. Electrical Systems 20-3 (2024): 6695-6701, ISSN: 1112-5209. SCOPUS.

2.Tammineni Shanmukha Prasanthi ,Swarajya Madhuri Rayavarapu, GottapuSasibhushanaRao,Raj Kumar Goswami, GottapuSanotosh Kumar, “RADIO FREQUENCY BASED INPAINTING FOR INDOOR LOCALIZATION USING MEMORYLESS TECHNIQUES AND WIRELESS TECHNOLOGY,” IAPGOS 4/2024, p-ISSN 2083-0157, e-ISSN 2391-6761,Dec.2024, SCOPUS.

3.Prasanthi S, Rayavarapu M, Rao S, Goswami R., “Investigation of Different Generative Adversarial Networks Techniques for Image Restoration,” International Journal of Computing and Digital Systems.,August 2024, ISSN:2210-142X. SCOPUS.

4.Raj Kumar Goswami , Swathi Nambari , Srinivasa Rao Kandula , P. Pavithra Roy and A.H. Sharief, “Rate 5/6 TCM Code Having 64 States with 64 QAM for Fading Channel,” International Journal of Electrical and Electronics Research (IJEER), e-ISSN: 2347-470X, April 2024. SCOPUS.

5.ShanmukhaprasanthiTammineni ,Swaraiya Madhuri Rayavarapu , Sasibhushana Rao Gottapu ,Raj Kumar Goswami, “DIGITAL IMAGE RESTORATION USING SURF ALGORITHM,” IAPGOS, p-ISSN 2083-0157, e-ISSN 2391-6761, March 2024. SCOPUS.

2.Dr.P.M.K.Prasad

Journal Publications:

1.D.Srinivasa Rao, Ch.Rajasekhar, **PMK Prasad** and GBSR Naidu, “Sustainability and Environmental impact on mining and maintaining cryptocurrencies: A review”, Nature Environment and Pollution Technology, Dec.2024, ISSN: 09726268(Print); ISSN: 2395.3454 (online). SCOPUS.

2.Chiranjeevi karri, João Santinha, Nikolaos Papanikolaou, Santosh Kumar Gottapu , Manohar Vuppula, **PMK Prasad**, “Pancreatic Cancer Detection Through Semantic Segmentation of CT Images: A Short Review”, Discover Artificial Intelligence, Springer Nature Journal (Scopus Indexed), Dec.2024, ISSN:2731-0809, SCOPUS.

3.Dr.L.Ganesh

Book Chapter:

1. Dr.Raj Kumar Goswami, Dr.GaneshLaveti, “Edge Computing in 5G and IoT Enhancing Data Processing Real-Time Analytics and Reducing Latency for Critical Applications”, Publisher - RADemics Research Institute, Chapter 10 , page 306-351.ISBN: 978-81-979336-3-9, Year:2024.

Journal Publications:

1.P. Sirish Kumar , A. Jayalaxmi, V. B. S. Srilatha Indira Dutt , P. Krishna Rao, P. Kameswara Rao and L. Ganesh, ““Advanced Algorithms for Pseudo-Range Estimation and Positioning Accuracy in Challenging Satellite Visibility Conditions”,Journal of Communications, Oct 2024, ISSN: 1796-2021 (Online); 2374-4367 (Print). SCOPUS.

2.M. N. V. S. S. Kumar, D. Eswara Chaitanya, L. Ganesh, and T. Sri Sudha, ““AN AUTOMATED BREAST CANCER DETECTION BY HEURISTIC-BASED ENSEMBLE DEEP CLASSIFIER USING MAMMOGRAM AND TOMOSYNTHESIS IMAGES,” YīxuéGōngchéng. Applications, Basis, Communications/Biomedical Engineering: Applications, Basis, and Communications, April 2024. ISSN:1016-2372 E-ISSN:1793-7132, SCOPUS.

Dr.B.Vijaya Lakshmi

Book Chapters:

1.Dr.V. Radhika and Dr.B.Vijayalakshmi , “NANOTECHNOLOGY APPLICATIONS IN VARIOUS FIELDS”, Futuristic Trends in Artificial Intelligence , IIP Series, Volume 3,Book 9,Part 5, Chapter 3, pp:97-103, Iterative International Publishers(IIP), e-ISBN: 978-93-6252-924-4, Year:2024.

Dr.R V S Lakshmi Kumari

Journal Publications:

1.Baddu Naik Bhukya,V. Venkataiah, S. Mani.Kuchibhatla, S. Koteswari, R V S Lakshmi Kumari, Yallapragada Ravi Raju, “Integrating the Internet of Things to Protect Electric Vehicle Control Systems from Cyber Attacks,” IAENG International Journal of Applied Mathematics, ISSN: 1992-9986 (online version); 1992-9978 (print version) Volume 54, Issue 3, March 2024, Pages 433-440, 2024. SCOPUS

Botta Kusuma Kumari

Journal Publications:

AjjannaMulakaledu, ChintamJagadeeswarReddy , Y Kamal Kishore, Vinod Bakka, Botta Kusuma Kumari and Busireddy Hemanth Kumar, “A New Quazi Z-Source Seven-Level inverter for Photovoltaic Applications,” International Journal of Electrical and Electronics Research (IJEER), e-ISSN: 2347-470X (online Version) 756-761pp, Volume 12 issue 3,2024. SCOPUS.

Krishna Molli, P. Ajay D Vimal Raj, “SST-Based 25-Level T-Type MLI For Integration of Micro-grids”, IEEE Journal of Emerging and Selected Topics in Industrial Electronics, 2024, Print ISSN: 2687-9735, Electronic ISSN: 2687-9743, IEEE Journal

Prof.P V S Lakshmi Jagadamba

Patent:

1.Title:“SYSTEM FOR DETECTION OF APPLICATION LAYER ATTACKS”,

Inventors Names:1.Pulugurta Krishna Subba Rao 2.Akasapu Ajaya Kumar 3.Paluri

V S Lakshmi Jagadamba

Date:14-03-2024Date (U/S 11A) 14/03/2024,

Patent Number : 526001,

Application number 201641037178 ,

M.Aswini:

Journal Publications:

1.Srinivasa Rao Illapu, Sankara, Aswini Mula, and Padmaja Malarowthu., “Multiobjective-energy centric honey badger optimization based routing for wireless body area network”, Journal of Intelligent & Fuzzy Systems, March 2024, ISSN 1064-1246. SCOPUS

Dr.K.Rohini:

Journal Publications:

Rohini Kancharapu, Sri Nagesh Ayyagari, “Suicidal ideation prediction based on social media posts using a GAN-infused deep learning framework with genetic optimization and word embedding fusion,” International Journal of Information Technology (Singapore), Feb.2024, 2511-2104 E-ISSN: 2511-2112. SCOPUS.

Dr. K. Purushotam Naidu

Book:

1)Milad Moheeni, Dr.Jagadish, Ravi kiranMahadasa, Dr. K. Purushotam Naidu, "Machine Learning", published by BR International, Jan 2024, ISBN: 978-81-965444-7-8

Journal Publications:

1.Dr. Hari Jyothula, Laxmana Rao Y, Dr. D NagaTej, Mr. Ch Srinivasa Reddy, Mr. Sagar Sathuluri, Dr. Purushotam Naidu K, "Implementation of Real And Accurate Level Set Formulation In Brain Mri/Cta Image Segmentation and Classification," Frontiers in Health Informatics 2024 Vol 13: Issue 3 (Scopus Index), ISSN: 2676-7104, SCOPUS

P.Sridevi

Journal Publications:

1)P. Sridevi, S.Himajasri, B.Nandini, M.Bhargavi and K.Jahnavi, ""Early Detection of Parkinson's Disease using Convolutional Neural Networks on Handwritten Patterns,"Mukt Shabd Journal, Vol. XIII, Issue IV, UGC Group-I Journal, April 2024, ISSN:2347-3150.UGC CARE 1.

Dr.K.L.Sai Prasad

Journal Publications:

1.I. V. Venkateswara Rao, S. Sunitha Devi and K. L. Sai Prasad, ""On ϕ -conharmonically flat Lorentzian para-Kenmotsu manifolds",Reliability: Theory & Applications, March 2024.ISSN:1932-2321, SCOPUS.

2.K. L. Sai Prasad, P. Naveen and S. Sunitha Devi,"On certain classes of conformally flat lorentzian para-kenmotsu manifolds", Reliability: Theory & Applications, June 2024, ISSN:1932-2321, SCOPUS.

Dr.Ch.S. Lakshmi

Journal Publications:

1.V. Lakshmi Savithri Vatsalya , G. Sunita Sundari , Ch.S.L.N. Sridhar, M. Durairaj , T.C. Sabari Girisun , Ch.S. Lakshmi , "Exploring nonlinear optical absorption in cobalt-doped Nickel-Zinc nanomaterials", Materials Chemistry and Physics, May 2024, ISSN:0254-0584, SCI.

2.P. B. Shettya, K. I. Maddani, K. S. Mahalaxmi, Ch. S. Lakshmi, "On the dielectric properties of nano-phased lanthanum doped nickel ferrites", Digest Journal of Nanomaterials and Biostructures, April 2024, ISSN:1842-3582, SCIE.

3.Vemakoti Siva Jahnavi, B. Rajesh Babu, Ch S. Lakshmi, “Synthesis, characterization of Al-doped SnO₂ nanoparticles for enhanced dielectric properties and antibacterial activity”, Physica B: Condensed Matter, Nov 2024, ISSN:0921-4526. SCI

Dr.Vemakoti Siva Jahnavi:

Journal Publications:

1.Vemakoti Siva Jahnavi, B. Rajesh Babu, Ch S. Lakshmi, “Synthesis, characterization of Al-doped SnO₂ nanoparticles for enhanced dielectric properties and antibacterial activity”, Physica B: Condensed Matter, Nov 2024, ISSN:0921-4526. SCI

Dr. B. Rajesh Babu

Journal Publications:

1.Vemakoti Siva Jahnavi, B. Rajesh Babu, Ch S. Lakshmi, “Synthesis, characterization of Al-doped SnO₂ nanoparticles for enhanced dielectric properties and antibacterial activity”, Physica B: Condensed Matter, Nov 2024, ISSN:0921-4526. SCI

Dr. T. Poorna Kantha

Journal Publications:

1.G Murali, G Deepa,Nirmala Kasturi V and Dr. T. Poorna Kantha,, “Joint effects of thermal diffusion and diffusion thermo on MHD three dimensional nano fluid flow towards a stretching sheet”, Mathematical Models in Engineering, Vol.9, Issue 04, 2023, 130-143. SCOPUS

पेटेंट कार्यालय
शासकीय जर्नल

**OFFICIAL JOURNAL
OF
THE PATENT OFFICE**

निर्गमन सं. 08/2025
ISSUE NO. 08/2025

शुक्रवार
FRIDAY

दिनांक: 21/02/2025
DATE: 21/02/2025

पेटेंट कार्यालय का एक प्रकाशन
PUBLICATION OF THE PATENT OFFICE

INTRODUCTION

In view of the recent amendment made in the Patents Act, 1970 by the Patents (Amendment) Act, 2005 effective from 01st January 2005, the Official Journal of The Patent Office is required to be published under the Statute. This Journal is being published on weekly basis on every Friday covering the various proceedings on Patents as required according to the provision of Section 145 of the Patents Act 1970. All the enquiries on this Official Journal and other information as required by the public should be addressed to the Controller General of Patents, Designs & Trade Marks. Suggestions and comments are requested from all quarters so that the content can be enriched.

(PROF. (DR) UNNAT P. PANDIT)
CONTROLLER GENERAL OF PATENTS, DESIGNS & TRADE MARKS

21th February, 2025

(12) PATENT APPLICATION PUBLICATION

(21) Application No.202541011712 A

(19) INDIA

(22) Date of filing of Application :12/02/2025

(43) Publication Date : 21/02/2025

(54) Title of the invention : SPEAKER DIARIZATION SYSTEM USING LSTM NETWORKS AND MFCC FEATURES FOR ENHANCED ACCURACY

(51) International classification :G10L0025240000, G10L0017000000, G06F0018232130, G06N0003080000, G10L0017040000

(86) International Application No :NA
Filing Date :NA

(87) International Publication No : NA

(61) Patent of Addition to Application Number :NA
Filing Date :NA

(62) Divisional to Application Number :NA
Filing Date :NA

(71)Name of Applicant :

1)Dasari Indu

Address of Applicant :W/o. Mr. A. Yaswanth, Flat No. 102, Happy Homes Apartment, Sri Valli Nagar, Kommadi, Visakhapatnam - 530048, Andra Pradesh, India. Visakhapatnam -----

2)Prof. Y. Srinivas

Name of Applicant : NA
Address of Applicant : NA

(72)Name of Inventor :

1)Dasari Indu

Address of Applicant :W/o. Mr. A. Yaswanth, Flat No. 102, Happy Homes Apartment, Sri Valli Nagar, Kommadi, Visakhapatnam - 530048, Andra Pradesh, India. Visakhapatnam -----

2)Prof. Y. Srinivas

Address of Applicant :Professor, Department of Computer Science Engineering, GST, GITAM University, Visakhapatnam - 530045, Andhra Pradesh, India
Visakhapatnam -----

(57) Abstract :

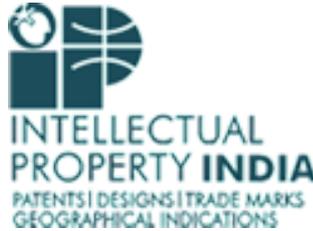
A method and system for speaker diarization combining Long Short-Term Memory (LSTM) networks and Mel-Frequency Cepstral Coefficients (MFCC) features to segment and cluster audio streams based on individual speakers. The system preprocesses audio data to extract MFCC features, segments the audio into overlapping chunks, and uses an LSTM-based neural network to capture temporal patterns and generate speaker embeddings. These embeddings are clustered using offline algorithms, such as k-means, to assign speaker labels. Contextual refinement ensures smooth transitions and improved accuracy, addressing challenges like overlapping speech and background noise. The invention enhances speaker recognition in multilingual and dynamic audio environments.

No. of Pages : 25 No. of Claims : 9



Office of the Controller General of Patents, Designs & Trade Marks
Department for Promotion of Industry and Internal Trade
Ministry of Commerce & Industry,
Government of India

(<http://ipindia.nic.in/index.htm>)



(<http://ipindia.nic.in/index.htm>)

Application Details

APPLICATION NUMBER	202541011712
APPLICATION TYPE	ORDINARY APPLICATION
DATE OF FILING	12/02/2025
APPLICANT NAME	1 . Dasari Indu 2 . Prof. Y. Srinivas
TITLE OF INVENTION	SPEAKER DIARIZATION SYSTEM USING LSTM NETWORKS AND MFCC FEATURES FOR ENHANCED ACCURACY
FIELD OF INVENTION	ELECTRONICS
E-MAIL (As Per Record)	patents@allinnov.org
ADDITIONAL-EMAIL (As Per Record)	allinnovrnd@gmail.com
E-MAIL (UPDATED Online)	
PRIORITY DATE	
REQUEST FOR EXAMINATION DATE	--
PUBLICATION DATE (U/S 11A)	21/02/2025

FORM 1
THE PATENTS ACT, 1970
(39 of 1970)
&
THE PATENTS RULES, 2003
APPLICATION FOR GRANT OF PATENT
[See sections 7,54 & 135 and rule 20(1)]

(FOR OFFICE USE ONLY)

Application No.:
Filing Date:
Amount of Fee Paid:
CBR No.:
Signature:

1. APPLICANT(S):

Sr.No.	Name	Nationality	Address	Country	State	Distict	City
1	Dasari Indu	India	W/o. Mr. A. Yaswanth, Flat No. 102, Happy Homes Apartment, Sri Valli Nagar, Kommadi, Visakhapatnam - 530048, Andra Pradesh, India.	India	Andhra Pradesh	Visakhapatnam	Visakhapatnam
2	Prof. Y. Srinivas	India	Professor, Department of Computer Science Engineering, GST, GITAM University, Visakhapatnam - 530045, Andhra Pradesh, India	India	Andhra Pradesh	Visakhapatnam	Visakhapatnam

2. INVENTOR(S):

Sr.No.	Name	Nationality	Address	Country	State	Distict	City
1	Dasari Indu	India	W/o. Mr. A. Yaswanth, Flat No. 102, Happy Homes Apartment, Sri Valli Nagar, Kommadi, Visakhapatnam - 530048, Andra Pradesh, India.	India	Andhra Pradesh	Visakhapatnam	Visakhapatnam
2	Prof. Y. Srinivas	India	Professor, Department of Computer Science Engineering, GST, GITAM	India	Andhra Pradesh	Visakhapatnam	Visakhapatnam

		University, Visakhapatnam - 530045, Andhra Pradesh, India				
--	--	--	--	--	--	--

3. TITLE OF THE INVENTION: SPEAKER DIARIZATION SYSTEM USING LSTM NETWORKS AND MFCC FEATURES FOR ENHANCED ACCURACY

**4. ADDRESS FOR CORRESPONDENCE OF APPLICANT /
AUTHORISED PATENT AGENT IN INDIA:**

Allinnov Research and Development Private Limited, #360E,
Senthur Murugan Kovil Street, Opp. SM Mahal, Oldpet,
Krishnagiri - 635001, Tamil Nadu, India.

Telephone No.:

Fax No.:

Mobile No: 9790586194

E-mail: patents@allinnov.org

5. PRIORITY PARTICULARS OF THE APPLICATION(S) FILED IN CONVENTION COUNTRY:

Sr.No.	Country	Application Number	Filing Date	Name of the Applicant	Title of the Invention
--------	---------	--------------------	-------------	-----------------------	------------------------

6. PARTICULARS FOR FILING PATENT COOPERATION TREATY (PCT) NATIONAL PHASE APPLICATION:

International Application Number	International Filing Date as Allotted by the Receiving Office
PCT//	

7. PARTICULARS FOR FILING DIVISIONAL APPLICATION

Original (first) Application Number	Date of Filing of Original (first) Application
-------------------------------------	--

8. PARTICULARS FOR FILING PATENT OF ADDITION:

Main Application / Patent Number:	Date of Filing of Main Application
-----------------------------------	------------------------------------

9. DECLARATIONS:

(i) Declaration by the inventor(s)

I/We ,Dasari Indu,Prof. Y. Srinivas, is/are the true & first inventor(s) for this invention and declare that the applicant(s) herein is/are my/our assignee or legal representative.

(a) Date: -----

(b) Signature(s) of the inventor(s):

(c) Name(s): Dasari Indu,Prof. Y. Srinivas

(ii) Declaration by the applicant(s) in the convention country

I/We, the applicant(s) in the convention country declare that the applicant(s) herein is/are my/our assignee or legal representative.

(a) Date: -----

(b) Signature(s) :

(c) Name(s) of the singnatory: Dasari Indu,Prof. Y. Srinivas

(iii) Declaration by the applicant(s)

- **The Complete specification relating to the invention is filed with this application.**
- **I am/We are, in the possession of the above mentioned invention.**
- **There is no lawful ground of objection to the grant of the Patent to me/us.**

10. FOLLOWING ARE THE ATTACHMENTS WITH THE APPLICATION:

Sr.	Document Description	FileName
------------	-----------------------------	-----------------

I/We hereby declare that to the best of my/our knowledge, information and belief the fact and matters stated hering are correct and I/We request that a patent may be granted to me/us for the said invention.

Dated this(Final Payment Date): -----

Signature:

Name: PREM CHARLES

To The Controller of Patents

The Patent office at CHENNAI

This form is electronically generated.

FORM 2

THE PATENTS ACT, 1970

(39 of 1970)

&

The Patent Rules, 2003

COMPLETE SPECIFICATION

(See sections 100 & rule 103)

1. TITLE OF THE INVENTION

**AN EFFICIENT SPEAKER DIARIZATION SYSTEM USING LSTM NETWORKS AND MFCC
FEATURES FOR ENHANCED ACCURACY**

2. APPLICANTS (S)

NAME(S)	NATION ALITY	ADDRESS
Dasari Indu	Indian	W/o. Mr. A. Yaswanth, Flat No. 102, Happy Homes Apartment, Sri Valli Nagar, Kommadi, Visakhapatnam - 530048, Andra Pradesh, India.
Prof. Y. Srinivas	Indian	Professor, Department of Computer Science Engineering, GST, GITAM University, Visakhapatnam - 530045, Andhra Pradesh, India

3. PREAMBLE TO THE DESCRIPTION

COMPLETE SPECIFICATION

The following description particularly describes the invention and the method in which it has to be performed.

AN EFFICIENT SPEAKER DIARIZATION SYSTEM USING LSTM NETWORKS AND MFCC FEATURES FOR ENHANCED ACCURACY

TECHNICAL FIELD

The invention relates to the field of speaker diarization and speech processing and methods and specifically to speaker diarization system using LSTM networks and MFCC features for enhanced accuracy.

BACKGROUND

[1]. Speaker diarization, the process of identifying "who spoke when" in an audio stream, is a critical component of various applications, including automated transcription, teleconferencing, voice authentication, and multimedia indexing. This process involves segmenting and clustering speech signals from multiple speakers in an audio stream and assigning unique labels to each speaker. Despite significant advancements, speaker diarization remains a challenging problem due to complexities such as overlapping speech, dynamic audio environments, and variations in speaker accents, dialects, and speaking styles.

[2]. Traditional speaker diarization methods have relied on statistical modeling techniques like Gaussian Mixture Models (GMMs) and Hidden Markov Models (HMMs) to analyze audio features. These methods utilize acoustic features such as pitch, energy, and spectral coefficients for speaker

segmentation and clustering. While effective in controlled environments, these approaches often struggle to maintain accuracy in real-world scenarios involving noise, non-uniform speaker contributions, and multilingual speech.

[3].The emergence of machine learning, particularly deep learning, has revolutionized speaker diarization systems. Techniques like Deep Neural Networks (DNNs) and Convolutional Neural Networks (CNNs) have shown significant promise in extracting high-level features from speech signals and improving segmentation accuracy. However, these methods often rely on large labeled datasets and are computationally expensive, limiting their applicability in resource-constrained environments.

[4].Long Short-Term Memory (LSTM) networks, a type of Recurrent Neural Network (RNN), have emerged as a powerful tool for capturing temporal dependencies in sequential data. LSTMs have proven effective in speech recognition, language modeling, and speaker verification tasks due to their ability to model long-range patterns in audio signals. When combined with robust feature extraction techniques like Mel-Frequency Cepstral Coefficients (MFCC), LSTMs provide a highly accurate and efficient framework for speaker diarization.

[5].MFCC features are widely used in speech and audio processing due to their ability to represent the short-term power spectrum of audio signals in a manner that closely resembles human auditory perception. By combining MFCC

features with the temporal modeling capabilities of LSTMs, it is possible to achieve enhanced speaker segmentation and clustering accuracy, even in challenging audio conditions.

[6].Despite these advancements, significant challenges remain in implementing practical speaker diarization systems. These include handling overlapping speech segments, refining speaker transitions, and maintaining accuracy in noisy and dynamic environments. Furthermore, real-world applications often require the system to operate in an offline mode to process the entire audio context, allowing for more accurate clustering and speaker label assignment.

[7].This invention addresses these challenges by presenting a novel speaker diarization system that leverages LSTM networks and MFCC features to achieve superior segmentation and clustering performance. By incorporating offline clustering techniques and contextual refinement, the system ensures smooth speaker transitions and accurate labeling, making it suitable for various practical applications such as call centers, voice-activated devices, and multimedia analysis.

[8].The proposed system represents a significant improvement over existing approaches by combining advanced deep learning techniques with efficient feature extraction and clustering methodologies, providing a robust solution for speaker diarization in multilingual and dynamic environments.

SUMMARY

[9]. In one aspect of the present disclosure, A method for speaker diarization using deep learning and feature extraction, comprising: Preprocessing audio data to extract Mel-Frequency Cepstral Coefficients (MFCC) features from audio streams; Segmenting the audio stream into overlapping audio segments; Feeding the MFCC features of each segment into a Long Short-Term Memory (LSTM) network to extract temporal patterns and generate speaker embeddings; Clustering the speaker embeddings using an offline clustering algorithm to group audio segments associated with the same speaker; Refining speaker labels based on contextual data to improve segmentation accuracy; Outputting a diarized audio stream with labeled speaker segments.

[10]. In some aspects of the present disclosure, the preprocessing step further comprises using a library, such as Librosa, to extract MFCC features from the audio stream.

[11]. In some aspects of the present disclosure, the LSTM network includes multiple layers configured to capture long-term dependencies in the MFCC sequences.

[12]. In some aspects of the present disclosure, the clustering step employs a k-means algorithm to group the speaker embeddings into clusters.

[13]. In some aspects of the present disclosure, the audio stream includes non-speech segments, and the method further comprises filtering out the non-speech segments during preprocessing.

[14]. In some aspects of the present disclosure, the refinement step involves
5 handling background noise and overlapping speech segments.

[15]. In one aspect of the present disclosure, A speaker diarization system, comprising: a preprocessing module configured to extract MFCC features from audio streams; a segmentation module configured to divide audio streams into overlapping segments; an LSTM-based neural network configured to process
10 the MFCC features and generate speaker embeddings; a clustering module employing an offline clustering algorithm to classify the speaker embeddings into groups corresponding to different speakers; a post-processing module configured to refine speaker labels and ensure smooth transitions between segments.

[16]. In some aspects of the present disclosure, the clustering module
15 performs offline clustering to leverage the entire audio context for improved speaker label assignment.

[17]. In some aspects of the present disclosure, the post-processing module
20 applies smoothing algorithms to adjust speaker labels for enhanced contextual consistency.

BRIEF DESCRIPTION OF DRAWINGS

[18]. The above and still further features and advantages of aspects of the present disclosure become apparent upon consideration of the following detailed description of aspects thereof, especially when taken in conjunction with the accompanying drawings, and wherein:

[19]. Figure 1 is a architecture, in accordance with an aspect of the present disclosure;

[20]. Figure 2 is a diagrammatic representation of exploratory data analysis of audio, in accordance with an aspect of the present disclosure;

[21]. Figure 3 is a diagrammatic representation of the raw audio to MFCC feature extraction, in accordance with an aspect of the present disclosure;

[22]. Figure 4 is a diagrammatic representation of the test accuracy and loss for the hold LSTM model, in accordance with an aspect of the present disclosure;

[23]. Figure 5 is a diagrammatic representation of confusion matrix for the model, in accordance with an aspect of the present disclosure;

DETAILED DESCRIPTION

[24]. The following description provides specific details of certain aspects of the disclosure illustrated in the drawings to provide a thorough understanding of those aspects. It should be recognized, however, that the present disclosure

can be reflected in additional aspects and the disclosure may be practiced without some of the details in the following description.

5 [25]. The various aspects including the example aspects are now described more fully with reference to the accompanying drawings, in which the various aspects of the disclosure are shown. The disclosure may, however, be embodied in different forms and should not be construed as limited to the aspects set forth herein. Rather, these aspects are provided so that this disclosure is thorough and complete, and fully conveys the scope of the disclosure to those skilled in the art. In the drawings, the sizes of components may be exaggerated for clarity.

10 [26]. It is understood that when an element or layer is referred to as being “on,” “connected to,” or “coupled to” another element or layer, it can be directly on, connected to, or coupled to the other element or layer or intervening elements or layers that may be present. As used herein, the term “and/or” includes any and all combinations of one or more of the associated listed items.

15 [27]. The subject matter of example aspects, as disclosed herein, is described with specificity to meet statutory requirements. However, the description itself is not intended to limit the scope of this patent. Rather, the inventor/inventors have contemplated that the claimed subject matter might also be embodied in other ways, to include different features or combinations of features similar to
20 the ones described in this document, in conjunction with other technologies.

[28]. The present invention provides an advanced method and system for speaker diarization, which integrates Long Short-Term Memory (LSTM) networks with Mel-Frequency Cepstral Coefficients (MFCC) features to segment and cluster audio streams, accurately identifying and labeling individual speakers. This approach addresses significant limitations in traditional methods, such as difficulties in handling overlapping speech, variations in speaker accents, and environmental noise. The invention offers robust performance, scalability, and suitability for real-world applications like automated transcription, teleconferencing, and voice authentication.

10 [29]. The process begins with preprocessing the audio data to extract MFCC features, which capture the spectral characteristics of the speech signal. MFCC features are widely recognized for their ability to represent audio in a way that closely mimics human auditory perception. Libraries such as Librosa can be employed to perform this step, ensuring high-quality feature extraction. During preprocessing, non-speech segments and background noise are filtered out, enhancing the overall efficiency and accuracy of the system.

15 [30]. The audio stream is then segmented into overlapping chunks, typically 1 to 2 seconds in duration. This segmentation ensures that temporal features spanning across segment boundaries are captured, which is crucial for downstream processing. Each segment's MFCC features are fed into an LSTM network, which is specially designed to handle sequential data and capture

long-term temporal dependencies. The LSTM network generates speaker embeddings, which are high-dimensional vectors that uniquely represent the voice characteristics of each speaker.

[31]. These speaker embeddings are grouped using an offline clustering algorithm, such as k-means, which classifies similar embeddings into clusters. Each cluster corresponds to a distinct speaker. Offline clustering is particularly effective as it processes the entire audio stream at once, leveraging the full context of the conversation. This approach outperforms online clustering methods, which process data incrementally and may lack global context.

10 [32]. To refine the speaker labels and ensure contextual consistency, a post-processing module is included. This module addresses challenges such as overlapping speech, abrupt transitions between speakers, and background noise. By analyzing contextual information from adjacent segments, the system ensures smooth transitions and coherent labeling of speakers throughout the audio stream. Additionally, overlapping speech is handled by examining the distribution of embeddings and accurately assigning labels to segments with multiple active speakers.

[33]. The invention employs a modular design, with distinct preprocessing, LSTM processing, clustering, and refinement components. This modularity allows for flexibility in system configuration and future upgrades. For example, the LSTM architecture can be adjusted to include additional layers or features

to improve accuracy, while the clustering algorithm can be replaced with more advanced techniques as needed.

[34]. The system is trained using datasets such as the Speech Accent Archive, which includes a diverse collection of voice samples from speakers of various accents and languages. This ensures the system's robustness in handling multilingual and dynamic audio environments. During training, the LSTM network is optimized using a loss function designed for sequence labeling tasks, such as categorical cross-entropy, and hyperparameters like batch size and learning rate are tuned for optimal performance.

10 [35]. One of the key innovations of the invention is its ability to process audio streams offline, which allows for more accurate clustering and speaker label assignment. By processing the entire audio stream as a single batch, the system gains access to global context, enabling it to handle complex scenarios such as dynamic speaker interactions and overlapping speech. This offline approach makes the system particularly well-suited for applications like call center analytics, automated transcription, and voice-based authentication.

15 [36]. The invention's performance is evaluated using standard metrics such as Diarization Error Rate (DER), accuracy, precision, recall, and F1 score. These metrics demonstrate the system's effectiveness in accurately segmenting and labeling speakers, even in challenging environments with background noise or non-uniform speaker contributions. In testing, the system achieved high

accuracy and robust performance, with significant improvements over traditional models.

[37]. Furthermore, the invention is scalable and resource-efficient, making it deployable on both cloud-based platforms and edge devices with limited computational power. The preprocessing and clustering steps are optimized to minimize latency, enabling efficient operation in real-world applications. The system can also be integrated with other speech processing frameworks, such as automatic speech recognition (ASR) and natural language processing (NLP) systems, to provide comprehensive audio analysis solutions.

[38]. The invention addresses key challenges in speaker diarization, including handling overlapping speech, refining transitions between speakers, and maintaining accuracy in noisy environments. It also supports multilingual audio streams, making it suitable for global applications. By combining LSTM networks with MFCC features and incorporating offline clustering and contextual refinement, the invention offers a robust and accurate solution for speaker diarization in a wide range of use cases.

[39]. In summary, this invention represents a significant advancement in speaker diarization technology. Its innovative combination of LSTM networks, MFCC features, and offline clustering ensures accurate speaker segmentation and labeling in complex audio environments. The system's modular design,

scalability, and flexibility make it a valuable tool for applications in transcription, teleconferencing, voice authentication, and multimedia analysis.

[40]. The invention further enhances speaker diarization by incorporating mechanisms to address specific challenges encountered in real-world scenarios.

5 One significant challenge is overlapping speech, where multiple speakers are active simultaneously. The system analyzes the distribution of speaker embeddings for overlapping segments and uses contextual information to assign appropriate labels to each active speaker. This is particularly beneficial in applications like teleconferencing, where overlapping conversations are
10 common.

[41]. Another critical aspect of the invention is noise resilience. The preprocessing module filters out non-speech segments and background noise using advanced algorithms, ensuring that only meaningful speech data is processed. This noise reduction capability improves the accuracy of the
15 diarization system, especially in dynamic environments like public spaces or call centers, where background noise can vary significantly.

[42]. The use of offline clustering is a major innovation in this system. Unlike online clustering, which assigns speaker labels incrementally as data becomes available, offline clustering processes the entire audio stream as a single batch.
20 This enables the system to leverage global context, resulting in more precise speaker segmentation and labeling. Offline clustering is particularly

advantageous in applications where post-event analysis, such as meeting transcription or forensic audio analysis, is required.

[43]. The invention's architecture is highly adaptable to different audio environments and application requirements. For example, the system can handle audio streams with varying numbers of speakers by dynamically adjusting the clustering parameters. Additionally, the system supports multilingual audio streams, making it suitable for global applications. This capability is achieved by training the LSTM network on diverse datasets that include speakers of different languages and accents.

[44]. The modular design of the system also allows for easy integration with other technologies. For instance, it can be combined with Automatic Speech Recognition (ASR) systems to provide both speaker identification and transcription in a single workflow. This integration is particularly useful for applications like customer service analytics, where understanding both "who spoke" and "what was said" is critical.

[45]. Moreover, the invention is designed to be resource-efficient, enabling deployment on edge devices with limited computational power. This is achieved through optimized preprocessing, lightweight LSTM architectures, and clustering algorithms. The system's efficiency makes it suitable for applications in mobile devices, smart home assistants, and IoT platforms, where computational resources are often constrained.

[46]. To ensure robustness across diverse use cases, the system employs advanced training techniques. The LSTM network is trained on a large, labeled dataset containing speaker annotations, allowing it to generalize well to unseen audio streams. Transfer learning can also be applied, where the network is pre-trained on a broad dataset and fine-tuned for specific applications, further enhancing its adaptability.

[47]. The system incorporates measures to ensure smooth transitions between speaker segments. Abrupt transitions, which can occur due to rapid speaker changes, are refined during post-processing. This involves analyzing contextual information from adjacent segments and adjusting speaker labels to align with the natural flow of the conversation. This feature is particularly valuable in scenarios like meeting transcription, where speaker transitions can occur frequently.

[48]. Additionally, the invention supports fine-grained analysis of audio streams. Beyond identifying speakers, the system can extract additional metadata, such as speaker duration, speaking rate, and turn-taking patterns. This metadata can be used for advanced analytics, such as identifying dominant speakers in a meeting or analyzing conversational dynamics in a customer service interaction.

[49]. The invention's ability to handle multilingual audio streams is further enhanced by its reliance on MFCC features, which are language-agnostic and

focus on the acoustic properties of speech. This ensures that the system performs consistently across languages, making it a versatile solution for applications in global markets.

[50]. Testing and evaluation of the system have demonstrated its effectiveness in handling complex audio conditions. In scenarios with overlapping speech, background noise, and multiple speakers, the system consistently achieved high accuracy in speaker segmentation and labeling. This performance highlights its potential for deployment in challenging environments, such as courtrooms, public events, and emergency response systems.

[51]. The system's offline processing capability also allows for advanced analysis and refinement of speaker labels. For example, in legal applications, the system can provide detailed, time-aligned speaker labels that facilitate evidence review and transcription. Similarly, in media production, it can help streamline the editing process by accurately identifying and separating contributions from multiple speakers.

[52]. Real-time adaptation is another potential enhancement for this system. While the current implementation focuses on offline processing, the underlying architecture can be modified to support real-time operation. By employing incremental clustering and online inference with the LSTM network, the system

could provide speaker diarization in live scenarios, such as broadcasting or live event transcription.

[53]. Finally, the invention's versatility and performance make it a valuable tool for various industries. In healthcare, for instance, it can be used to analyze doctor-patient interactions, improving communication and care delivery. In education, it can facilitate the analysis of classroom discussions, providing insights into student participation and engagement. The system's ability to adapt to diverse environments and requirements ensures its relevance across multiple domains.

[54]. In conclusion, the present invention represents a groundbreaking advancement in speaker diarization technology. By combining LSTM networks with MFCC features and incorporating offline clustering and contextual refinement, the system achieves unmatched accuracy and robustness. Its modular design, scalability, and adaptability make it a powerful tool for a wide range of applications, from transcription and teleconferencing to analytics and multimedia indexing. This invention sets a new standard for speaker diarization, addressing longstanding challenges and unlocking new possibilities for speech processing and analysis.

#####DIGITALLY SIGNED#####
PREM CHARLES I
REGISTERED PATENT AGENT INPA-3311
On Behalf of the Applicants.

Claims:

Claim:

1. A method for speaker diarization using deep learning and feature extraction, comprising:

- 5
- Preprocessing audio data to extract Mel-Frequency Cepstral Coefficients (MFCC) features from audio streams;
 - Segmenting the audio stream into overlapping audio segments;
 - Feeding the MFCC features of each segment into a Long Short-Term Memory (LSTM) network to extract temporal patterns and generate
- 10
- speaker embeddings;
 - Clustering the speaker embeddings using an offline clustering algorithm to group audio segments associated with the same speaker;
 - Refining speaker labels based on contextual data to improve segmentation accuracy;
- 15
- Outputting a diarized audio stream with labeled speaker segments.

2. The method as claimed in claim 1, wherein the preprocessing step further comprises using a library, such as Librosa, to extract MFCC features from the audio stream.

3. The method as claimed in claim 1, wherein the LSTM network includes multiple layers configured to capture long-term dependencies in the MFCC sequences.
4. The method as claimed in claim 1, wherein the clustering step employs a k-means algorithm to group the speaker embeddings into clusters.
5. The method as claimed in claim 1, wherein the audio stream includes non-speech segments, and the method further comprises filtering out the non-speech segments during preprocessing.
6. The method as claimed in claim 1, wherein the refinement step involves handling background noise and overlapping speech segments.
7. A speaker diarization system, comprising:
 - A preprocessing module configured to extract MFCC features from audio streams;
 - A segmentation module configured to divide audio streams into overlapping segments;
 - An LSTM-based neural network configured to process the MFCC features and generate speaker embeddings;
 - A clustering module employing an offline clustering algorithm to classify the speaker embeddings into groups corresponding to different speakers;

- A post-processing module configured to refine speaker labels and ensure smooth transitions between segments.

8. The system as claimed in claim 7, wherein the clustering module performs offline clustering to leverage the entire audio context for improved speaker label assign.

5

9. The system as claimed in claim 7, wherein the post-processing module applies smoothing algorithms to adjust speaker labels for enhanced contextual consistency.

#####DIGITALLY SIGNED#####
PREM CHARLES I
REGISTERED PATENT AGENT INPA-3311
On Behalf of the Applicants

10

15

20

ABSTRACT

**SPEAKER DIARIZATION SYSTEM USING LSTM NETWORKS AND
MFCC FEATURES FOR ENHANCED ACCURACY**

A method and system for speaker diarization combining Long Short-Term Memory
5 (LSTM) networks and Mel-Frequency Cepstral Coefficients (MFCC) features to
segment and cluster audio streams based on individual speakers. The system
preprocesses audio data to extract MFCC features, segments the audio into overlapping
chunks, and uses an LSTM-based neural network to capture temporal patterns and
generate speaker embeddings. These embeddings are clustered using offline
10 algorithms, such as k-means, to assign speaker labels. Contextual refinement ensures
smooth transitions and improved accuracy, addressing challenges like overlapping
speech and background noise. The invention enhances speaker recognition in
multilingual and dynamic audio environments.

15

**#####DIGITALLY SIGNED#####
PREM CHARLES I
REGISTERED PATENT AGENT INPA-3311
On Behalf of the Applicants**

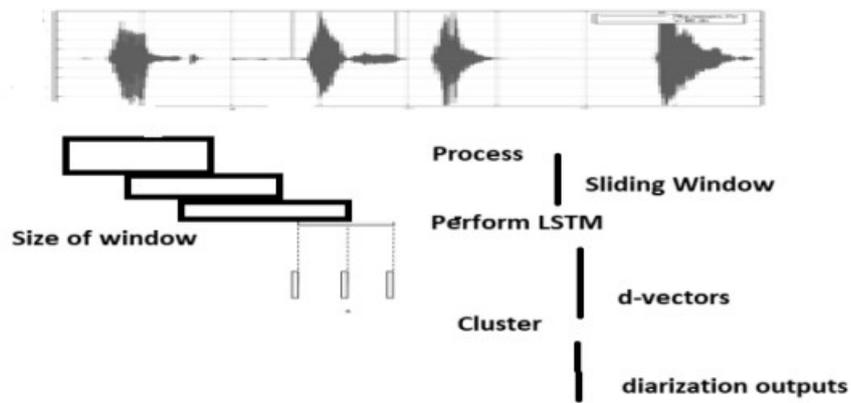


Figure 1

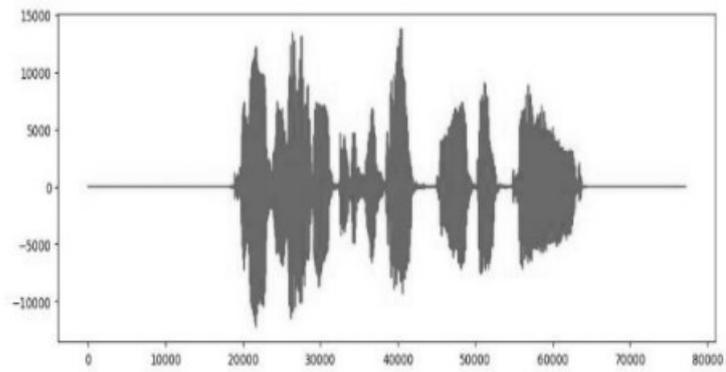


Figure 2

#####DIGITALLY SIGNED#####
PREM CHARLES I
REGISTERED PATENT AGENT INPA-3311
On Behalf of the Applicants

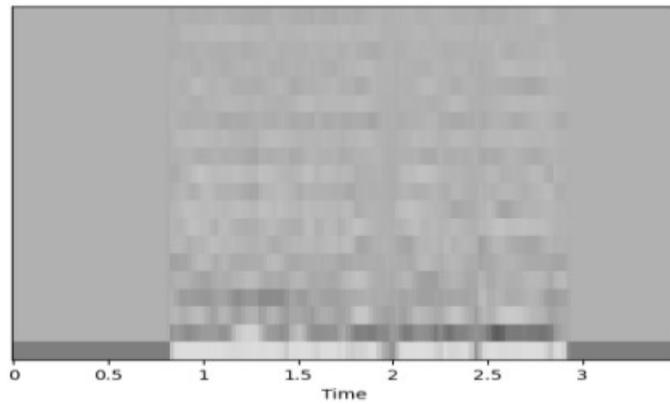


Figure 3

#####DIGITALLY SIGNED#####
PREM CHARLES I
REGISTERED PATENT AGENT INPA-3311
On Behalf of the Applicants

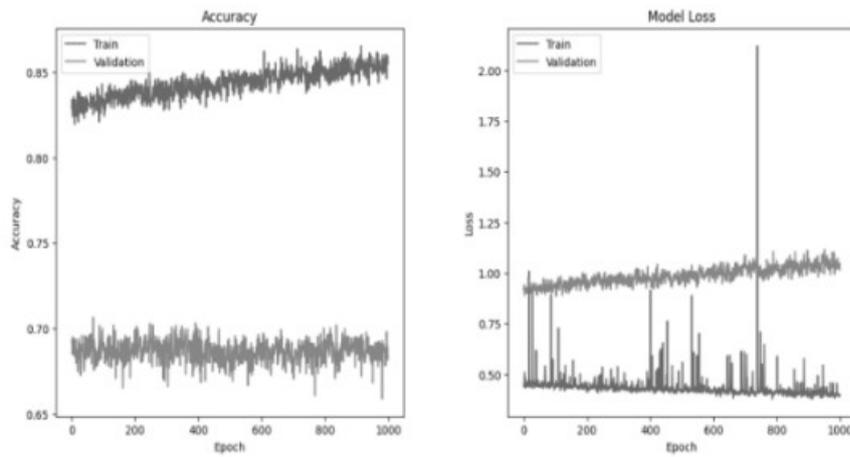


Figure 4

#####DIGITALLY SIGNED#####
PREM CHARLES I
REGISTERED PATENT AGENT INPA-3311
On Behalf of the Applicants

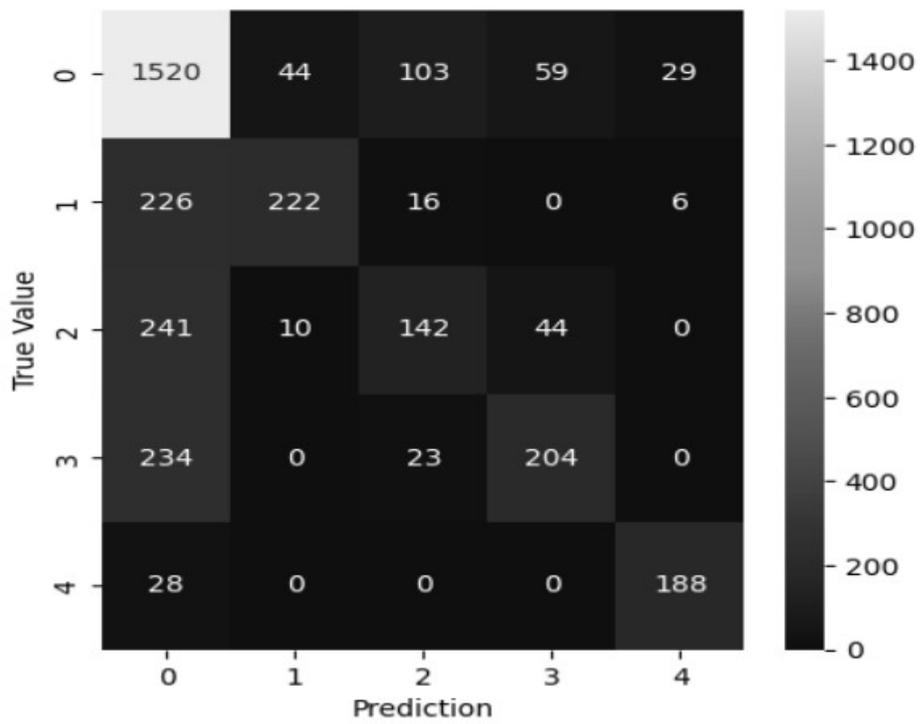


Figure 5

#####DIGITALLY SIGNED#####
PREM CHARLES I
REGISTERED PATENT AGENT INPA-3311
On Behalf of the Applicants

FORM 3

THE PATENT ACT, 1970
(39 of 1970)

AND

THE PATENTS RULES, 2003

STATEMENT AND UNDERTAKING UNDER SECTION 8

[(See section 8; Rule 12)]

1. Name of the applicant(s).	I/We Dasari Indu, W/o. Mr. A. Yaswanth, Flat No. 102, Happy Homes Apartment, Sri Valli Nagar, Kommadi, Visakhapatnam - 530048, Andhra Pradesh, India. , Applicant Type - NP , hereby declare:				
2. Name, address and nationality of the joint applicant.	(i) that I/We who have made the application for patent number 202541011712 in India, dated 12/02/2025 13:12:56. , alone / jointly with Prof. Y. Srinivas. (ii) that I/We have not made any application for the same/substantially the same invention outside India Or (iii) that I/We have made for the same/ substantially same invention, application(s) for patent in the other countries, the particulars of which are given below:				
Name of the country	Date of application	Application No.	Status of the application	Date of publication	Date of disposal
NA	NA	NA	NA	NA	NA
3. Name and address of the assignee	(i) that the rights in the application(s) filed in India has/have been assigned to None. All rights remain with the applicants only. . (ii) that I/We undertake that upto the date of grant of the patent by the Controller, I/We would keep him informed in writing regarding the details of corresponding applications for patents filed outside India in accordance with the provisions contained in section 8 and rule 12. Dated this 12th day of February 2025				
4. To be signed by the applicant or his authorized registered patent agent.	Signature(s)				
5. Name of the natural person who has signed.				
	To The Controller of Patents, The Patent Office, at New Delhi				
Note: ” Strike out whichever is not applicable; ”					

This form is electronically generated.

FORM 5

THE PATENT ACT, 1970
(39 of 1970)
&
THE PATENTS RULES, 2003

DECLARATION AS TO INVENTORSHIP

[See section 10(6) and rule 13(6)]

1. NAME OF APPLICANT(S) Dasari Indu, Prof. Y. Srinivas,

hereby declare that the true and first inventor(s) of the invention disclosed in the complete specification filed in pursuance of my/our application numbered **202541011712** dated **12/02/2025** is/are

2. INVENTOR(S)

Name	Country	Nationality	Address
Dasari Indu	India	India	W/o. Mr. A. Yaswanth, Flat No. 102, Happy Homes Apartment, Sri Valli Nagar, Kommadi, Visakhapatnam - 530048, Andra Pradesh, India.
Prof. Y. Srinivas	India	India	Professor, Department of Computer Science Engineering, GST, GITAM University, Visakhapatnam - 530045, Andhra Pradesh, India

Dated this. **12/02/2025** Day of **2025**

Signature

Name of the signatory

3. DECLARATION TO BE GIVEN WHEN THE APPLICATION IN INDIA IS FILED BY THE APPLICANT(S) IN THE CONVENTION COUNTRY:--

We the applicant(s) in the convention country hereby declare that our right to apply for a patent in India is by way of assignment from the true and first inventor(s).

Dated this. **12/02/2025**. Day of **2025**

Signature

Name of the signatory

4. STATEMENT (to be signed by the additional inventor(s) not mentioned in the application form)

I/We assent to the invention referred to in the above declaration, being included in the complete specification filed in pursuance of the stated application.

Dated this(Final Payment Date):-----

Signature

Name of the signatory

FORM 9

THE PATENT ACT, 1970
(39 of 1970)
&
THE PATENTS RULES, 2003

REQUEST FOR PUBLICATION

[See section 11A (2) rule 24A]

I/We **Dasari Indu, Prof. Y. Srinivas** hereby request for early publication of my/our [Patent Application No.]
TEMP/E-1/13448/2025-CHE

Dated **12/02/2025 00:00:00** under section 11A(2) of the Act.

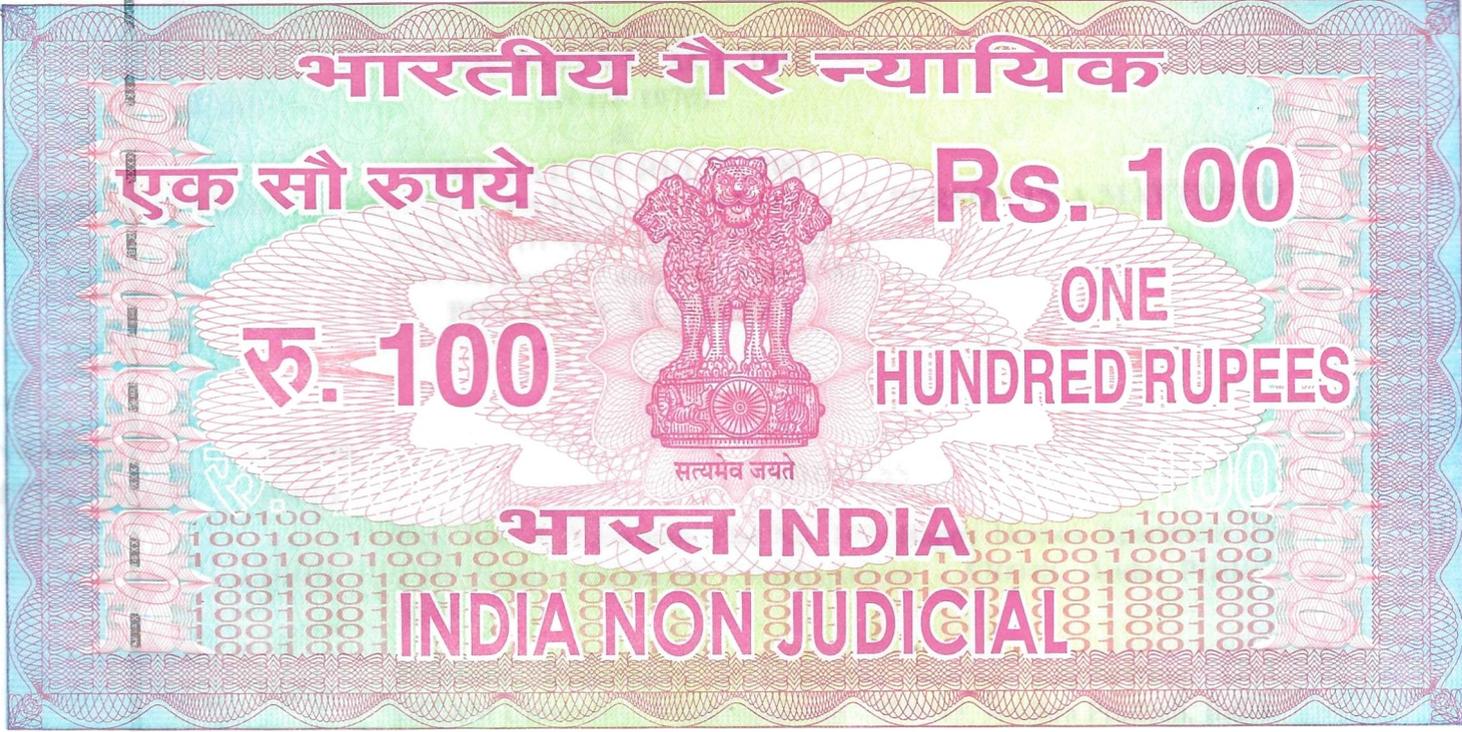
Dated this(Final Payment Date):-----

Signature

Name of the signatory

To,
The Controller of Patents,
The Patent Office,
At Chennai

This form is electronically generated.



தமிழ்நாடு தமிழ்நாடு TAMILNADU

31.1.2025

*V. Pream Charless
Krishnagiri.*

DS 085689

V. Radha

V. RADHA
S.V.L. No: 3930/B1/2000
163-A, Salem Road
KRISHNAGIRI-635 001

STAMP DUTY FOR
APPLICATION NO: 20254011712

FORM 26
THE PATENTS ACT, 1970
(39 OF 1970)

&

THE PATENT RULES, 2003
FORM FOR AUTHORIZATION OF A PATENT AGENT/OR ANY PERSON IN A MATTER OF
PROCEEDING UNDER THE ACT
(See Section 127 and 132; rule 135)
POWER OF ATTORNEY

I we,

NAME(S) OF APPLICANT(S)	NATION ALITY	ADDRESS
Dasari Indu	Indian	W/o. Mr. A. Yaswanth, Flat No. 102, Happy Homes Apartment, Sri Valli Nagar, Kommadi, Visakhapatnam - 530048, Andra Pradesh, India.
Prof. Y. Srinivas	Indian	Professor, Department of Computer Science Engineering, GST, GITAM University, Visakhapatnam - 530045, Andhra Pradesh, India

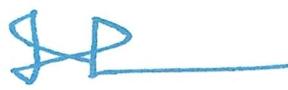
hereby authorize, Prem Charles I, Registered Patent Agent INPA 3311, at#360E, First Floor, SenthurMuruganKovil Street, Opp. S.M. Mahal, Oldpet, Krishnagiri - 635001, Tamil Nadu, India, to act on our behalf in connection with Patent filing and further prosecution, filing of assignments and any document related thereto, with reference to our patent application/ reference no. 202541011712 dated 12-02-25 and all further patent applications filed by them in future, on our behalf and request that all notices, requisitions and communications relating thereto may be sent to such person at the above address unless otherwise specified. They are also authorized to substitute another attorney agent to attend hearings (if any) in relation to the patent. We hereby revoke all previous authorizations, if any, in respect of same matter or proceeding. We hereby assent to the action already taken by the said persons in the above matter.

Dated - 13-02-25


Dasari Indu


Prof. Y. Srinivas

To,
The Controller of Patents
The Indian Patent Office
At Chennai, Kolkata, Delhi, Mumbai


PREM CHARLES I
REGISTERED INDIAN PATENT AGENT
REGN. ID.: INPA 3311

Controller General of Patents, Designs & Trade
Marks



सत्यमेव जयते

G.A.R.6
[See Rule 22(1)]
RECEIPT



Docket No 13872

Date/Time 2025/02/12 13:12:56

To
PREM CHARLES

UserId: prem1987

360E, SENTHUR MURUGAN KOVIL
STREET, OPPOSITE SM MAHAL,
OLDPET

CBR Detail:

Sr. No.	App. Number	Ref. No./Application No.	Amount Paid	C.B.R. No.	Form Name	Remarks
1	202541011712	TEMP/E-1/13448/2025-CHE	1600	7988	FORM 1	SPEAKER DIARIZATION SYSTEM USING LSTM NETWORKS AND MFCC FEATURES FOR ENHANCED ACCURACY
2	E-12/2686/2025/CHE	202541011712	2500	7988	FORM 9	----

TransactionID	Payment Mode	Challan Identification Number	Amount Paid	Head of A/C No
N-0001600319	Online Bank Transfer	1202250019905	4100.00	1475001020000001

Total Amount : ₹ 4100.00

Amount in Words: Rupees Four Thousand One Hundred Only

Received from PREM CHARLES the sum of ₹ 4100.00 on account of Payment of fee for above mentioned Application/Forms.

* This is a computer generated receipt, hence no signature required.

[Print](#)

[Home](#)

[About Us](#)

[Contact Us](#)

Welcome PREM CHARLES [Sign out](#)

Controller General of Patents, Designs & Trade
Marks
G.S.T. Road, Guindy, Chennai-600032
Tel No. (091)(044) 22502081-84 Fax No. 044 22502066
E-mail: chennai-patent@nic.in
Web Site: www.ipindia.gov.in



सत्यमेव जयते



Docket No 13881

Date/Time 12/02/2025

To
PREM CHARLES

User Id: prem1987

360E, SENTHUR MURUGAN KOVIL
STREET, OPPOSITE SM MAHAL,
OLDPET

Sr. No.	App. Number	Ref. No./Application No.	Amount Paid	C.B.R. No.	Form Name	Remarks
1	202541011712	E-5/1241/2025/CHE	0	----	FORM 5	
2	202541011712	E-3/2587/2025/CHE	0	----	FORM 3	

Total Amount : ₹ 0

Amount in Words: Rupees Only

[Print](#)

[Home](#)

[About Us](#)

[Contact Us](#)

Welcome PREM CHARLES [Sign out](#)

Controller General of Patents, Designs & Trade
Marks
G.S.T. Road, Guindy, Chennai-600032
Tel No. (091)(044) 22502081-84 Fax No. 044 22502066
E-mail: chennai-patent@nic.in
Web Site: www.ipindia.gov.in



सत्यमेव जयते



Docket No 14637

Date/Time 14/02/2025

To
PREM CHARLES

User Id: prem1987

360E, SENTHUR MURUGAN
KOVIL STREET, OPPOSITE SM
MAHAL, OLDPET

Sr. No.	App. Number	Ref. No./Application No.	Amount Paid	C.B.R. No.	Form Name	Remarks
1	202541011712	E-45/2014/2025/CHE	0	----	FORM 26	

Total Amount : ₹ 0

Amount in Words: Rupees Only

[Print](#)

[Home](#)

[About Us](#)

[Contact Us](#)