

Dr. Sandeep Albert Mathias

Ph.D., IIT Bombay, Mumbai

✉ sandeepalbert@presidencyuniversity.in | mathiassandeep@gmail.com
☎ +91-8433780647
🏠 lwsam.github.io | www.cse.iitb.ac.in/~sam
🔗 lwsam | in Sandeep Mathias | 📁 Sandeep Mathias
📧 Sandeep Mathias | 📁 Sandeep-Mathias

RESEARCH INTERESTS

My primary area of interest is the field of Natural Language Processing (NLP). Currently, I am working on problems combining the usage of gaze behaviour with NLP, in particular, in the domain of education, like automatic essay grading, text simplification and automatic short-answer scoring.

EDUCATION

Qualification	Board/University	Month & Year	CPI (out of) / %
Ph.D. (Computer Science)	IIT Bombay, Mumbai	July, 2020	8.00 (10)
M.Tech. (Information Technology)	IIT Bangalore, Bengaluru	July, 2012	3.22 (4)
B.Tech. (Information Technology)	NITK, Surathkal	June, 2010	7.44 (10)
Higher Secondary Examination (10+2)	Karnataka Pre-University College Board in: St. Joseph's Pre-University College	April, 2006	82.33%
Secondary Examination (10)	Indian Certificate of Secondary Education in: St. Joseph's Boys' High School	April, 2004	86.00%

WORK EXPERIENCE/INTERNSHIP

- **Assistant Professor**, CSE Department, Presidency University, Bengaluru *October 2021 onwards*
Responsibilities: I am currently a faculty member of the Department of Computer Science and Engineering at Presidency University, Bengaluru. The list of courses that I teach are listed later on in the resume. As a faculty member, I was also involved in some administration work.
- **Research Associate**, MALL Lab, CDS Dept., IISc, Bengaluru *November 2020 to September 2021*
Responsibilities: I am currently collaborating with Ashutosh Kumar, a PhD student of the lab. We are working on personalized cross-lingual sentence simplification.
- **Research Intern**, Kaybus India Pvt. Ltd., Bengaluru *May 2014 to August 2014*
Responsibilities: During the internship, I worked on automatically generating questions from text, in particular user manuals.
- **Software Engineer**, Nokia India Private Ltd., Bengaluru *July 2012 to June 2013*
Responsibilities: During my time in Nokia, I worked on user authentication and sign-in for the Nokia account.

PROFESSIONAL SERVICE

- **Journal Reviewer** for:
 1. ACM Transactions on Asian and Low-Resource Language Information Processing
TALLIP 2020, 2021, & 2023
 2. Pattern Recognition Letters
PRL 2022
 3. International Journal of Artificial Intelligence in Education
IJAIED 2023
- **Conference Meta-Reviewer** for:
 1. International Joint Conference on Natural Language Processing and Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (IJCNLP-AAACL)
IJCNLP-AAACL 2023 System Demo Track.
- **Program Committee member / Conference Reviewer** for:
 1. International Conference on Computational Linguistics (COLING)
COLING 2024

2. European Chapter of the Association for Computational Linguistics (EACL)
EACL 2023, 2024
 3. International Conference on Natural Language Processing (ICON)
ICON 2022
 4. European Conference on Artificial Intelligence (ECAI)
ECAI 2022
 5. Association of Computational Linguistics Rolling Review (ARR)
ARR October 2021 onwards
 6. Conference on Empirical Methods in Natural Language Processing (EMNLP)
EMNLP 2021, 2022 & 2023
 7. International Joint Conference on Artificial Intelligence (IJCAI)
IJCAI 2021, 2022, 2023 & 2024
 8. Annual Meeting of the Association for Computational Linguistics (ACL)
ACL 2021, 2022 & 2023
 9. International Joint Conference on Natural Language Processing
IJCNLP 2021, 2023
 10. Conference of the North-American Chapter of Association for Computational Linguistics (NAACL)
NAACL 2021, 2022 & 2024
 11. AAAI Conference on Artificial Intelligence (AAAI)
AAAI 2021, 2022 & 2023
 12. Workshop on Innovative Use of NLP for Building Educational Applications (BEA)
BEA 2020, 2021, 2022 & 2023
 13. Language Resources and Evaluation Conference (LREC)
LREC 2020, 2022 & 2024
- **Mentor** for the following conferences / workshop:
 1. Workshop on Innovative Use of NLP for Building Educational Applications (BEA)
BEA 2022
 - **Teaching Assistant** for the following courses:

1. Topics in Natural Language Processing @ IIT Bombay	<i>Spring 2015</i>
2. Speech, Natural Language Processing and the Web @ IIT Bombay	<i>Fall 2014</i>
3. Automata Theory and Logic @ IIT Bombay	<i>Spring 2014</i>
4. Design and Implementation of Functional Programming Languages @ IIT Bombay	<i>Fall 2013</i>
5. Mathematical Models of Computation @ IIIT Bangalore	<i>Spring 2012</i>

COURSES TAUGHT

As a faculty member of Presidency University, I have taught the following courses:

1. **CSE 228** Principles of Artificial Intelligence
Fall 2021
2. **CSE 234** Social Media Analytics
Fall 2021
3. **CSE 1001** Problem Solving Using Java
Winter 2022
4. **CSE 5006** Knowledge Engineering and Expert Systems
Winter 2022, Winter 2023
5. **BCA 217** Artificial Intelligence for Gaming
Spring 2022

6. **CSE 6002** Natural Language Processing Techniques
Spring 2022, Spring 2023
7. **CSE 243** Natural Language Processing
Fall 2022
8. **CSE 3014** Fundamentals of Natural Language Processing
Spring 2023, Fall 2023, Spring 2024
9. **CSA 3020** Artificial Intelligence for Game Development
Spring 2023
10. **CSE 3001** Artificial Intelligence and Machine Learning
Fall 2023
11. **CSA 3005** Applied Artificial Intelligence
Fall 2023, Spring 2024

ADMINISTRATIVE WORK

As a faculty member in the Department of Computer Science and Engineering, I was involved in the following administrative work:

- Research Cluster Head - Research cluster head for the Information Retrieval Research Cluster at Presidency University.
- Talent acquisition - Interviewing prospective faculty members for recruitment at Presidency University.
- Department Academic Council - I was a part of the Department Academic Council in Computer Science and Engineering at Presidency University, responsible for auditing course syllabi, lecture plans, etc.
- Time-table Co-Ordinator - I was one of the members of the Time-table committee for the Department of Computer Science and Engineering at Presidency University.

THESIS AND PROJECTS¹

- **Ph.D. Thesis: Cognitively Aided Automatic Essay Grading** *(Defended: July 24, 2020)*
Under the guidance of Prof. Pushpak Bhattacharyya.
Description: An essay is a piece of text, written in response to a topic, called a prompt. In the thesis, cognitive information, which was recorded / learnt using a reader's gaze behaviour, was used to score the essays. As part of the thesis, we first showed that gaze behaviour can be used for predicting how a reader would score a text qualitatively. Next, we showed how we can learn the gaze behaviour of a reader to score the text. In addition to the above works, we also investigated ways of scoring essay traits - different aspects of the essay, like content, organization, style, etc. - and using the trait scores to score the essay.
- **Automatic Text Simplification** *Industry Collaboration with PracTutor | July 2014 to May 2016*
Description: In this project, I worked on simplifying text by splitting sentences and substituting words and phrases. One of the outcomes of the project was coming up with a measure of text complexity, called the E-Score.
- **Automatic Question Generation from Text** *Internship Project | May 2014 to August 2014*
Description: This project was done as my internship with Kaybus. As part of the project, I was to create and extract questions from texts. This was done by first splitting the sentences using a number of rules. Next, important question words (like names, places, dates, etc.) would be identified for each sub-sentence, and appropriate questions would be generated.
- **M.Tech Thesis: Automatic Classification of Folktales** *January 2012 to June 2012*
Under the guidance of Prof. V.N. Muralidhara
Description: The aim of the thesis was, given an input text and a set of classes, we have to find out which class the text belongs to. Tagging was done using Stanford Core NLP, after which I found out which was the most similar class that a text belonged to, using similarity measures. The results were compared using the Arne-Thompson Tale Type Index for classifying folktales.

¹Only important ones that are completed are listed here.

- **Academic Data Analytics** *Project @ IIITB | January 2011 to April 2011*
Description: Implementation of data mining algorithms to extract potentially useful information from a data set. The algorithms implemented were k-means clustering (for clustering), Dynamic Itemset Counting and Frequency Pattern Tree (for Association Rule Mining) and ID3 (for Classification). I was responsible for implementing the k-means clustering, as well as making an offline version using Java.
- **Column Data Store in C** *Project @ IIITB | August 2010 to November 2010*
Description: We had to implement a simple data base application using C as a programming language. We implemented simple database operations, such as insertion, deletion and searching of records, based on certain parameters. The biggest issue we faced was in generalizing our application to handle any type of database query. In this project, I was responsible for implementing insertion and unloading functions, as well as reducing errors.

PUBLICATIONS

Conference Publications

1. Rahul Kumar, **Sandeep Mathias**, Sriparna Saha, and Pushpak Bhattacharyya. Many Hands Make Light Work: Using Essay Traits to Automatically Score Essays. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1485 – 1495. **NAACL 2022**. Seattle, WA, USA (Hybrid). July 10 – 15, 2022. [CORE rating: A] [Acceptance Rate: 22.0%] [Oral Presentation].
2. **Sandeep Mathias**, Diptesh Kanojia, Abhijit Mishra, and Pushpak Bhattacharyya. A Survey on Using Gaze Behaviour for Natural Language Processing. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence*, pages 4907 – 4913. **IJCAI-PRICAI 2020**. Yokohama, Japan (Online). January 11 – 15, 2021. [CORE rating: A*] [Acceptance Rate: 25.2%]. [Recognized as a **Premier Paper** by IKDD²]
3. **Sandeep Mathias**, Rudra Murthy, Diptesh Kanojia, and Pushpak Bhattacharyya. Cognitively Aided Zero-Shot Automatic Essay Grading. In *Proceedings of the 17th International Conference on Natural Language Processing*, pages 175 – 180. **ICON 2020**. Patna, India (Online). December 19 – 20, 2020. [CORE rating: Unrated] [Acceptance Rate: 48.5%].
4. **Sandeep Mathias**, Rudra Murthy, Diptesh Kanojia, Abhijit Mishra, and Pushpak Bhattacharyya. Happy Are Those Who Grade without Seeing: A Multi-Task Learning Approach to Grade Essays Using Gaze Behaviour. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, pages 858–872. **AACL-IJCNLP 2020**. Suzhou, China (Online). December 5 – 6, 2020. [CORE Rating: B] [Acceptance Rate: 27.0%].
5. **Sandeep Mathias**, Diptesh Kanojia, Kevin Patel, Samarth Agrawal, Abhijit Mishra, and Pushpak Bhattacharyya. Eyes are the Windows to the Soul: Predicting the Rating of Text Quality Using Gaze Behaviour. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2352 – 2362. **ACL 2018**. Melbourne, Australia. July 16 – 18, 2018. [CORE rating: A*] [Acceptance Rate: 24.9%]
6. **Sandeep Mathias**, and Pushpak Bhattacharyya. ASAP++: Enriching the ASAP Automated Essay Grading Dataset with Essay Attribute Scores. In *Proceedings of the 11th International Conference on Language Resources and Evaluation*, pages 1169 – 1173. **LREC 2018**. Miyazaki, Japan. May 8 – 10, 2018. [CORE rating: C] [Acceptance Rate: 65.2%].

Workshop Publications

1. Peniel John Whistely, **Sandeep Mathias**, and Galiveeti Poornima. 2022. PresiUniv at TSAR-2022 Shared Task: Generation and Ranking of Simplification Substitutes of Complex Words in Multiple Languages. In *Proceedings of the 2022 Workshop on Text Simplification, Accessibility and Readability*, pages 213 – 217. **TSAR 2022**. Abu Dhabi, UAE. December 8, 2022. [Acceptance Rate: 63.0%]

²Link: <https://ikdd.acm.org/premier-papers.php/>

2. **Sandeep Mathias**, and Pushpak Bhattacharyya. Can Neural Networks Automatically Score Essay Traits? In *Proceedings of the 15th Workshop on Innovative Use of NLP for Building Educational Applications*, pages 85 – 91. **BEA 2020**. Seattle, WA, USA (Online). July 10, 2020. [Acceptance Rate: 42.9%]
3. **Sandeep Mathias**, and Pushpak Bhattacharyya. Thank “Goodness”! A Way to Measure Style in Student Essays. In *Proceedings of the 5th Workshop on Natural Language Processing Techniques for Educational Applications*, pages 35 – 41. **NLP-TEA 2018**. Melbourne, Australia. July 19, 2018. [Oral Presentation] [Acceptance Rate: 25% (for oral presentation)]
4. Nikhil Wani, **Sandeep Mathias**, Jayashree Aanand Gajjam, and Pushpak Bhattacharyya. The Whole is Greater than the Sum of its Parts: Towards the Effectiveness of Voting Ensemble Classifiers for Complex Word Identification. In *Proceedings of the 13th Workshop on Innovative Use of NLP for Building Educational Applications*, pages 200 – 205. **BEA 2018**. June 5, 2018. [Acceptance Rate: 63.4%]
5. **Sandeep Mathias**, and Pushpak Bhattacharyya. How Hard Can it Be? The E-Score - A Scoring Metric to Assess the Complexity of Text. In *Proceedings of the 1st Workshop on Quality Assessment for Text Simplification*, pages 10 – 14. **QATS 2016**. Portorož, Slovenia. May 28, 2016. [Oral Presentation].
6. **Sandeep Mathias**, and Pushpak Bhattacharyya. Using Machine Translation Evaluation Techniques to Evaluate Text Simplification Systems. In *Proceedings of the 1st Workshop on Quality Assessment for Text Simplification*, pages 38 – 41. **QATS 2016**. Portorož, Slovenia. May 28, 2016. [Oral Presentation].

CERTIFICATIONS

NPTEL Certifications

- Natural Language Processing - **Certified** (October 2022)
- Affective Computing - **Certified with Elite Grade** (April 2023)
- Python for Data Science - **Certified with Elite + Silver Grade (TOP 1% TOPPER)** (September 2023)
- Data Science for Engineers - **Certified with Elite + Silver Grade** (September 2023)
- Deep Learning - **Certified with Elite Grade** (October 2023)
- Learning Analytics Tools - **Certified with Elite + Silver Grade** (October 2023)
- Programming in Java - **Certified with Elite Grade** (October 2023)
- Data Structure and Algorithms Using Java - **Certified with Elite + Silver Grade** (October 2023)

STUDENTS MENTORED

List of students mentored for internships, undergraduate project, masters project, and PhD research scholars³.

- Peniel John Whistely as his **guide** for his M.Tech. Project @ Presidency University, Bangalore (**Duration:** September 2022 to June 2023)
Topic: Lexical Simplification
- Hameed Moqbel as his **guide** for his Ph.D. @ Presidency University, Bangalore (**Duration:** August 2022 to present)
Topic: NLP for Educational Applications in Arabic
- Sunil Kumar Sahoo as his **guide** for his Ph.D. @ Presidency University, Bangalore (**Duration:** February 2022 to present)
Topic: Multidomain, Multilingual and Multimodal Automatic Essay Grading
- Rahul Kumar during his M.Tech @ IIT Kanpur (**Duration:** August 2022 to present)
Topic: Zero-Shot and Few-Shot Automatic Essay Grading
- Ashutosh Kumar, as part of his PhD @ IISc Bangalore (**Duration:** March 2021 to March 2022)
Topic: Personalized Cross-Lingual Sentence Simplification

³Excluding B.Tech. Project Groups of more than 1 person

- Rahul Kumar for his internship @ IIT Bombay (**Duration:** *June to July 2019*) & his B.Tech. Project @ IIT Patna (**Duration:** *August 2019 to June 2020*)
Topic: Trait-Specific Automatic Essay Grading
- Yashasvi Sriram for his R&D project @ IIT Bombay (**Duration:** *July to December 2018*)
Topic: Automatic Essay Grading
- Nikhil Wani for his internship @ IIT Bombay (**Duration:** *January to June 2018*)
Topic: Complex Word Identification

OTHER ACHIEVEMENTS AND AWARDS

- Was selected as one of the speakers at the AICTE Training and Learning (ATAL) Academy from January 8 to January 12, 2024 for the Faculty Development Program on Natural Language Processing.
- Achieved a **TOP 1 % TOPPER** in the NPTEL Course “Python for Data Science” in September 2023.
- One of my papers (see above) that was written and published during the time of COVID-19 was recognized as a **premier paper** by IKDD (ACM India Special Interest Group on Knowledge Discovery and Data Mining).
- Was awarded Rs. 4,75,000 as for the Presidency University Seed Grant **ACC/19/11/2021-2**, dated November 19, 2021, for the project “**Natural Language Processing (NLP) for Educational Applications**”, for 2 years, starting from November 2021. Till date, the project has resulted in 2 publications, with 2 more in the pipeline.
- Was awarded the **Cultural Roll of Honour** at IIT Bombay. This is the highest award given by the IIT Bombay Gymkhana to graduating students for excellence in cultural activities during their stay at IIT Bombay.
- Was awarded the **ACM India IARCS Travel Grant** to attend ACL 2018 in Melbourne, Australia.
- Was awarded the **Cultural Person of the Year** (2018) and the **Cultural Freshman of the Year** (2014) for excellence in cultural activities during the Academic Years 2017 – ’18, and 2013 – ’14 respectively.

Last Updated: April 18, 2024
Place: Bengaluru

Sandeep Albert Mathias