

# Bidyarthi Paul

+880 1970307986 | bidyarthipaul01@gmail.com  
Github | LinkedIn | ORCID | ResearchGate | Scopus

## EDUCATION

<b>Bachelor of Science (B.Sc.) in Computer Science and Engineering</b>	<b>2020–2024</b>
Ahsanullah University of Science and Technology, Tejgaon, Dhaka	CGPA: 3.882/4.00
Bachelor Thesis: Bengali Math Word Problem Solving with Chain of Thought Reasoning	
<b>Higher Secondary School Certificate</b>	<b>2017–2019</b>
Dhaka Residential Model College, Mohammadpur, Dhaka	GPA: 5.00/5.00
<b>Secondary School Certificate</b>	<b>2015–2017</b>
Scholars' School and College, Dhanmondi, Dhaka	GPA: 5.00/5.00

## PROFESSIONAL EXPERIENCE

<b>Adjunct Lecturer</b>	<b>23<sup>rd</sup> February, 2025 – Present</b>
Southeast University, Tejgaon, Dhaka, Bangladesh	
Course: C, Algorithm, Digital Logic Design, Data Communication	
<b>Computer Science Teacher</b>	<b>4<sup>th</sup> January, 2025 – Present</b>
Mastermind English Medium School, Uttara, Dhaka, Bangladesh	
Course: AI Fundamentals, Python, Database	

## RESEARCH EXPERIENCE

<b>Research Mentor</b>	<b>1<sup>st</sup> January, 2025 – Present</b>
Articulated Research Institute for Scientific Excellence (A.R.i.S.E.), Dhaka, Bangladesh	
• Developed the Bengali Regional Hate Speech NLP dataset and conducted model evaluation of deep learning approaches, including BERT-based and large language models (LLMs).	
<b>Research Assistant</b>	<b>1<sup>st</sup> March, 2024 – 31<sup>st</sup> December, 2024</b>
Articulated Research Institute for Scientific Excellence (A.R.i.S.E.), Dhaka, Bangladesh	
Advisor: Tashreef Muhammad	
• Developed a Bengali Regional Named Entity Recognition (NER) dataset along with its evaluation pipeline, and proposed a large language model (LLM) prompting pipeline for regional language detection.	

## RESEARCH INTERESTS

Natural Language Processing (NLP), Bangla Mathematical Reasoning, Regional Dialects, Low Resource Language, Large Language Models (LLMs), VLMs, Computer Vision, Deep Learning, Machine Learning.

## PUBLICATIONS ( [GOOGLE SCHOLAR](#))

(\* denotes equal contribution)

### Published

- **B. Paul**, J.J. Era, M.R. Zim, T.S. Aothoi, F.M. Shah. *Leveraging Bengali Math Word Problem Solving with Chain of Thoughts*. **ACM Transactions on Asian and Low-Resource Language Information Processing**, 2025 (Journal). [DOI](#)
- **B. Paul**, F.F. Preotee, S. Sarker, T. Muhammad. *Improving Bangla Regional Dialect Detection Using BERT, LLMs, and XAI*. In: 2024 **IEEE International Conference on Computing, Applications and Systems (COMPAS)**. [DOI](#)
- JJ Era\*, **B. Paul\***, TS Aothoi, MR Zim, FM Shah. *Empowering Bengali Education with AI: Solving Bengali Math Word Problems through Transformer Models*. In: 2024 **International Conference on Computer and Information Technology (ICCIT)**. [DOI](#)
- MAA Khandaker, ZS Raha, **B. Paul**, T. Muhammad. *Bridging Dialects: Translating Standard Bangla to Regional Variants Using Neural Models*. In: 2024 **International Conference on Computer and Information Technology (ICCIT)**. [DOI](#)

• **Bidyarthi Paul**, Musfiquir Rahman, Dipta Biswas, Ziaul Hasan, MD. Zahid Hossain. *Analyzing Emotions in Bangla Social Media Comments Using Machine Learning and LIME*. In: 2024 **International Conference on Machine Intelligence and Emerging Technologies (MIET)**. DOI

Under Review

• **B. Paul**, F.F. Preotee, S. Sarker, S.R. Refat, S. Islam, T. Muhammad, M.A. Hoque, S. Manzoor. *ANCHOLIK-NER: A Benchmark Dataset for Bangla Regional Named Entity Recognition*. *PLOS One*, 2025 (Journal). Pre-Print DOI

• A.H. Fayaz, M.D. Uddin, R.U. Bhuiyan, Z. Sultana, M.S. Islam, **B. Paul**, T. Muhammad, S. Manzoor. *BIDWESH: A Bangla Regional Based Hate Speech Detection Dataset*. *Natural Language Engineering*, 2025 (Journal). Pre-Print DOI

DATASETS

ANCHOLIK-NER <a href="#">[Link]</a>	Bangla Regional Named Entity Recognition Dataset	2024
BIDWESH <a href="#">[Link]</a>	Bangla Regional Hate Speech Detection Dataset	2024
PatiGonit <a href="#">[Link]</a>	Bengali Equation Based MWPs Solving Dataset	2023
PatiGonit22K <a href="#">[Link]</a>	Bengali Complex Equation Based 22K MWPs Solving Dataset	2023
SOMADHAN <a href="#">[Link]</a>	Bengali Complex Reasoning Math Word Problem Solving Dataset	2025

TECHNICAL SKILLS

**Programming and Mathematical Foundations:** C/C++, Python, Git, Probability, Statistics, Linear Algebra,  
**Deep Learning Tools:** Keras, Tensorflow, PyTorch, scikit-learn  
**Developer Tools:** Google Colab, Codeblocks, Visual Studios Code, Netbeans 8.2, Overleaf (Online LaTeX Editor)  
**Databases:** MsSql, SQL Server 2014  
**MS Office:** Word, Excel, Powerpoint, Outlook

ACHIEVEMENTS

• Received a research fund from the Institute of Research & Training (IRT), Southeast University for ANCHOLIK-NER. 2025

• Dean’s List of Honor, 2nd Position 2024

• Consistently received merit based waivers in the 1st, 2nd, 3rd, 5th, 6th, and 7th semesters based on top academic ranking and merit within the CSE Department. 2020 - 2024

• Scholars’ School and College Science Fair 2nd Runners up 2016

• General Scholarship in JSC (Junior School Certificate) 2014

LANGUAGES

**English** (Overall Band Score: 7.00), **Bengali** (native speaker), **Hindi** (fluent)

INTERESTS

Singing, Guitar Playing

REFERENCES

<b>1) Mr. Faisal Muhammad Shah</b> Associate Professor Department of Computer Science and Engineering Ahsanullah University of Science and Technology Tejgaon, Dhaka faisal.cse@aust.edu +8801729594777	<b>2) Mr. Tashreef Muhammad</b> Lecturer & Coordinator Department of Computer Science and Engineering Southeast University Tejgaon, Dhaka tashreef.muhammad@seu.edu.bd +8801719888448
---	---