Sunit Bhattacharya

LinkedIn| GitHub
ÚFAL mail | +91 700 123 8166 | Gmail

AREAS OF INTEREST

Neural Network Interpretability• Machine Translation• Natural Language Processing• Cognitive Linguistics• Psycholinguistics•

EDUCATION

CHARLES UNIVERSITY
PhD in Computer Science
Oct 2019-present | Prague, Czechia

CENTRAL UNIVERSITY
B.Sc+M.Sc in Computer Science
May 2013-June 2018 | Ajmer, Rajasthan
Department of Computer Science
Central University of Rajasthan
Cum. GPA: 4.52 / 6

D.A.V. PUBLIC SCHOOL
Sec IV| Bokaro, India

COURSEWORK

PHD
Formal Linguistics
Machine Translation
Distributed Computing
Unsupervised Learning
Natural Language Processing

UNDERGRADUATE+GRADUATE
Operating Systems
Compilers
Computer Graphics
Neural Networks
Combinatorial Optimization
Data Mining
Information Retrieval
Cloud Computing
Machine Learning
Game Theory
Non-parametric statistics

SKILLS

PROGRAMMING
Over 5000 lines:
Python • LaTeX
Over 1000 lines:
C • C++ • JAVA
Familiar:
pyLink (SR Research) • OpenCV • PyAudio • R • Matlab • MySQL

RESEARCH

CHARLES UNIVERSITY | Researcher
Oct 2019 – present | Prague, Czech Republic
Worked with the research group of Ondřej Bojar on interpretability aspects of multilingual/multimodal language models (under the project NEUREM-3 and other problems of Machine Translation under other projects like ELITR and LINDAT.

INDIAN INSTITUTE OF TECHNOLOGY, DELHI | Summer Intern
May 2017 – July 2017 | Delhi, India
Summer intern of Samar Husain. Worked on psycholinguistic experiments to study bilingual language understanding. Worked on various algorithms for final verb prediction in Hindi using LSTMs and dependency trees.

INDIAN INSTITUTE OF TECHNOLOGY, GUWAHATI | Summer Intern
May 2016 – July 2016 | Guwahati, India
Summer intern of Bidisha Som. Worked on Psycholinguistic experiments to understand processing of nouns by bilinguals in multimodal settings.

JADAVPUR UNIVERSITY, KOLKATA | Summer Intern
May 2015 – July 2015 | Kolkata, India
Summer intern of Dipankar Das. Worked on coreference resolution using "classical" ML methods.

MACHINE LEARNING SKILLS

Data Collection/ System Preparation/ Data Exploration:
Hadoop • Pyspark • Shell • SLURM • Makefile/Snakemake • Pandas
Model Building/Training:
Pytorch • Keras • Tensorflow • scikit-learn
Model Deployment/Monitoring:
Flask •Tensorboard

AWARDS

2020	Academic Fellowship	Charles University
2019	UGC-NET/JRF	Qualified for both UGC-NET and JRF

SELECTED PUBLICATIONS

[1] S. Bhattacharya and O. Bojar. Unveiling multilinguality in transformer models: Exploring language specificity in feed-forward networks. In *Proceedings of the 6th BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP*, pages 120–126, 2023.

[2] S. Bhattacharya and O. Bojar. Understanding the role of ffns in driving multilingual behaviour in llms. *arXiv preprint arXiv:2404.13855*, 2024.

[3] S. Bhattacharya, R. Kumar, and O. Bojar. Team úfal at cmcl 2022 shared task: Figuring out the correct recipe for predicting eye-tracking features using pretrained language models. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, pages 130–135, 2022.

[4] V. Zouhar, S. Bhattacharya, and O. Bojar. Multimodal shannon game with images. *arXiv preprint arXiv:2303.11192*, 2023.