MOITREE BASU

🌐 Website   📞 +33 751 903 192   ✉ sfurti.basu@gmail.com   in m-basu   🎓 Google Scholar

PROFESSIONAL SUMMARY

Experienced ML Engineer and Researcher with a Ph.D. in Privacy-preserving AI, 6+ years of research experience in machine learning, deep learning and 5+ years of hands-on experience in developing and deploying machine learning models. Specialized in NLP applications, deep learning, and data science, with a strong background in privacy and security considerations for AI systems. Proven track record in implementing innovative solutions for complex ML problems, particularly in the areas of privacy-preserving machine learning and natural language processing. Motivated team player with analytical thinking, sharp communication skills, and a dedication to promoting secure and ethical data-driven solutions.

AREAS OF EXPERTISE

- Natural Language Processing (NLP)
- Machine Learning & Deep Learning
- Privacy-Preserving AI
- Large Language Models
- Data Science & Analytics
- Optimization Algorithms

EDUCATION

UNIVERSITY OF LILLE & INRIA | *Ph.D., Privacy-preserving AI*   2019 – 2024
- Thesis: "Privacy-preserving AI using declarative constraints"
- Focused on creating privacy measures in decentralized collaborative computation platforms
- Developed interpretable privacy measures in decentralized machine learning settings

SAARLAND UNIVERSITY | *M.Sc., Computer Science*   2015 – 2019
- Specialized in AI, Machine Learning, Deep Learning, NLP, and Speech
- Thesis: "Multi-speaker, Multilingual Speech Synthesis using Deep Learning"

WEST BENGAL UNIVERSITY OF TECHNOLOGY | *B.Tech., Information Technology* 2008 – 2012

WORK EXPERIENCE

INRIA, FRANCE   Lille, France
Doctoral Researcher (Magnet Lab)   Oct 2019 – Mar 2024
- Studied privacy requirements for machine learning systems, focusing on interpretability and high utility
- Developed privacy-preserving AI models in decentralized health-care setup
- Implemented a scalable ML pipeline for processing and analyzing large datasets
- Created a specification syntax for declaring privacy constraints in ML systems

External Collaborator (Magnet Lab)   Sep 2019
- Conducted groundwork for the PhD thesis on "*Integrated privacy preserving AI*"

DFKI, GERMANY   Saarbrücken, Germany
Research Assistant (MLT Lab - MaryTTS)   2016 – 2018
- Implemented a new unit-selection module for the *MaryTTS* system
- Developed a neural grapheme-to-phoneme conversion model for MaryTTS *Blizzard 2018*
- Contributed to data preparation for *Blizzard Challenge 2017* and *2018*

TATA CONSULTANCY SERVICES LIMITED   Kolkata, India
Software Engineer (Data Intelligence Team)   2012 – 2015
- Connected SAP-HANA database using python to backend software, executed SQL queries, loaded data into dataframe
- Ingested BODS files, combined HANA and BODS, trained advanced ML models like regression, decision-trees, random forest
- Wrote ML model predictions back to HANA database, bridging enterprise data platforms with analytics
- Led a team of SQL developers and junior SAP consultants
- Worked closely on the business requirements, with client *Australian Gas and Lighting*

Variable Energy Cyclotron Centre, India │ *Technical Internship*   Winter 2011
- Developed a "Sign-Language Animator" software for interaction among physically challenged people
- Translation of text into sign language animations

Saha Institute of Nuclear Physics, India │ *Research Internship*   Summer 2011
- Developed a new improved algorithm by hybridizing Particle Swarm Optimization and Simulated Annealing algorithm, aimed to improve strength and efficiency

──── TECHNICAL SKILLS ────

Machine Learning:
Supervised/Unsupervised Learning, Feature Engineering, Model Evaluation

Deep Learning:
Neural Networks, CNN, RNN, LSTM, Transformers

NLP:
Text Classification, Named Entity Recognition, (Large) Language Modeling

Privacy-Preserving ML:
Differential Privacy, Federated Learning

Tools & Frameworks:
Python, PyTorch, Scikit-learn, Pandas, NLTK

Big Data:
SAP HANA, SQL, MongoDB

DevOps:
Git, Docker, CI/CD pipelines

──── PUBLICATIONS ────

1. Ramon, J., Basu, M. **Interpretable privacy with optimizable utility**, ECML/PKDD workshop on eXplainable Knowledge Discovery in Data mining 2020, Ghent, Belgium.

2. Basu, M., Ramon, J. **A method to translate privacy requirements into a configuration for privacy-preserving machine learning applied to multi-centric studies**, AI4HEALTH Winter School 2021 Poster Presentation.

3. Maguer, S., Steiner, I., Tombini, F., Deb, P., Basu, M., Kröger, I. **Agile MaryTTS Architecture for the Blizzard Challenge 2018**, Blizzard Challenge Satellite Workshops of INTERSPEECH 2018, Microsoft, Hyderabad, India.

4. Basu, M., Deb, P., Garai, G. **Hybrid of Particle Swarm Optimization and Simulated Annealing for Multidimensional Function Optimization.** International Journal of Information Technology, Vol. 20 No-1 Year-2014, India.

5. Basu, M., Deb, P. **Tag SNP Selection Using Quine-Mccluskey Optimization Method** International Journal of Advanced Research in Engineering and Technology (IJARET), Volume:4, Issue: 5, Pages: 74-81, 2013, India.

6. Banerjee, C., Kundu, A., Basu, M., Deb, P., Nag, D., Datta gupta, R. **A Service Based Trust Management Classifier Approach for Cloud Security.** International Conference on Advanced Computing Technologies (ICACT), IEEE, vol., pp.1,5, 21-22 Sept. 2013, India.

7. Deb, P., Basu, M. **Next Generation Sequencing using Ant Colony Optimization Algorithm and Binary Particle Swarm Optimizers.** International Journal of Information Technology, Vol. 20 No-1 Year-2014, India.

8. Deb, P., Basu, M. **An Analogy of Algorithms for Tagging of Single Nucleotide Polymorphism and Evaluation through Linkage Disequilibrium** International Journal of Computer Engineering & Technology (IJCET), Volume: 4, Issue: 4, Pages: 514-523, 2013, India.

──── AWARDS ────

- **Best Conference Paper Award**, IEEE International Conference on Advanced Computing Technologies (ICACT-2013), 2013
- **Gold Medal for" Best Student Award"**, TATA Consultancy Services Ltd, 2012
- **Best Student of the Year Award**, Netaji Subhash Engineering College, West Bengal University of Technology, 2011 & 2012
- **Scholarship for Undergraduate Studies**, Govt. of India, 2008-2012

CERTIFICATIONS

- **Deeplearning.ai:**
  - AI Python for Beginners:
    * Basics of AI Python Coding!
    * Automating Tasks with Python!
    * Working with Your Own Data and Documents in Python!
    * Extending Python with Packages and APIs!
  - AI for Medical Diagnosis
  - AI for Medical Prognosis
  - Deep Learning Specialization:
    * Neural Networks and Deep Learning
    * Improving Deep Neural Networks
    * Structuring ML Projects
    * Convolutional Neural Networks
    * Sequence Models
- **Coursera:**
  - Machine Learning
  - Algorithms: Design and Analysis
- **edX:**
  - Introduction to Computer Science
  - Introduction to CS and Programming
  - Introduction to Descriptive Statistics
  - Introduction to Probability
  - Introduction to Inference

- **Imperial College London:**
  - Mathematics for ML Specialization:
    * Linear Algebra
    * Multivariate Calculus
    * PCA
- **DataCamp:**
  - R Courses:
    * Introduction to Data Course
    * Introduction to R
    * Intermediate R
    * Cleaning Data in R
    * Intro to Student's T-test with R
    * Correlation and Regression
    * Intro to Statistics with R
    * Analysis of Variance (ANOVA)
    * Data Visualization with ggplot2 I
  - Python Courses:
    * Intro to Python for Data Science
    * Intermediate Python for Data Science
    * Python Data Science Toolbox I
    * Python Data Science Toolbox II
    * Python Programming Track
    * Deep Learning in Python

LANGUAGE PROFICIENCY

Proficient

English, Hindi, Bengali

Intermediate

French

Beginner

German

PASSIONATE ABOUT

- Creative Writing & Art
- Cooking
- Travel Videography