

Bipasha Banerjee

+1 (240) 971-9803

bipashabanerjee@vt.edu

[linkedin.com/in/bipasha-banerjee](https://www.linkedin.com/in/bipasha-banerjee)

RESEARCH INTEREST

Natural language processing, Information Retrieval, Machine Learning, Working with Scholarly Big data.

EDUCATION

Virginia Polytechnic Institute and State University

Blacksburg, Virginia, 24061 USA

Ph.D. Candidate

Computer Science

August 2018 – Present

Expected grad: Spring 2024

GPA: 3.85/4.0

Virginia Polytechnic Institute and State University

Blacksburg, Virginia, 24061 USA

M.S. Computer Science

December 2022

GPA: 3.85/4.0

Institute: Institute of Engineering & Management (IEM)

University: West Bengal University of Technology

Kolkata, West Bengal India

Bachelor of Technology, Computer Science and Engineering

2015

GPA: 8.63/10

TECHNICAL SKILLS

Programming and Scripting Languages: C, C++, C#, Java, Python, Shell scripting, HTML, JavaScript, CSS

Databases: Basics of 11g and 12c Oracle, MySQL

Others: Salesforce, Android Studio, Selenium, Cucumber

OS basic knowledge: Linux and Windows

RELEVANT COURSEWORK

Science-guided Machine Learning, Statistics in Research, Human AI Interaction, Research Methods in CS, Digital Libraries, Software Engineering, (Big) Data-Management Systems, Operating Systems, Big Data Text Summarization, Probabilistic and Discrete Mathematics, Computer Organization and Architecture, Design and Analysis of Algorithms, Data Structures, Computer Networks, Database Management Systems, Internet Technology, Cloud Computing, E-Commerce

EMPLOYMENT HISTORY

Virginia Polytechnic Institute and State University

August 2019 - Present

Graduate Research Assistantship

- Responsible for bringing computational access to 500,000 electronic theses and dissertation (ETD) corpus using various machine learning and deep learning techniques
- Developing a language model fine-tuned on ETDs to improve the state of the art on downstream tasks like classification and summarization

ADP

Summer Internship

May 2019 - August 2019

- Worked on the time and labor module.
- Mainly involved in the automation of the application using Selenium

Virginia Polytechnic Institute and State University

January 2019 – May 2019

Graduate Teaching Assistant

- TA for the course CS3754 - Cloud Software Development - responsible for holding office hours and grading for a class of 70 students.

TATA Consultancy Services, Kolkata

April 2016 - November 2017

Developer

- Worked on the development and enhancement of an e-commerce Android application to help clients improve product sales. Used Java-based Android studio IDE as the frontend with Salesforce as its backend, reducing screen latency by 50 percent vs. the web application.
- Worked as a developer on Microsoft .Net platform and developed both web and windows applications using Basic as well as MVC architecture.

ACADEMIC PROJECTS / INTERNSHIPS

Virginia Polytechnic Institute and State University

Nov 2018 – Present

Research supervised by Dr. Edward A. Fox

- Working on computational extraction of digital objects from book-length scholarly documents using various machine and deep learning techniques.
- Focus on chapter classification and chapter-level summarization for ETDs using transformer-based language models

Virginia Polytechnic Institute and State University

Jan 2020 - May 2020

Course Project for Human-AI CS 6724

- Studying the effects of using expert labels in the training of a machine learning model

Virginia Polytechnic Institute and State University

Aug 2018 - Dec 2018

Course Project for CS5984

- Generating summaries of two shooting events from a large collection of webpages.
- Project URI: <http://hdl.handle.net/10919/86407>

Central Computer Science and Engineering Lab, IEM, Kolkata.

Jan 2013 – June 2015

Undergraduate Researcher

- Designed a Recommendation Based On-Demand Protocol for MANET.
- Proposed Self-Organized key management based on fidelity relationship list and dynamic path for security in MANET.
- Published a review of different routing protocols and their vulnerabilities and countermeasures.
- Published a review on attacks and secure routing protocols in MANET.

Indian Statistical Institute, Kolkata

May 2014 – July 2014

Cryptology Summer Internship

- Studied the various approaches to cryptology.
- Analyzed “intractable problems” with respect to cryptology.
- Had hands-on training on the practical aspects of cryptology.

Saha Institute of Nuclear Physics, Kolkata

January 2014 – March 2014

Internship

The purpose of the project was to learn about Web development tools, Linux operating system, the Basic commands as well as shell scripting, and then to move on to learn how a client-side web browser communicates with a remote MySQL server using Common Gateway Interface (CGI) over an Apache web server.

- Learnt about how to reduce the server-side load by using JavaScript on the client side, in jQuery.
- Learned about Asynchronous JavaScript and XML (AJAX) for client-server communication.

PRESENTATION/PUBLICATION

Title: “Application of text analysis on scholarly long documents

December 2022

Accepted at IEEE Big Data 2022 workshop Computational Archival Science, Osaka, Japan

Title: “Help Me Help You - A Mixed-Initiative Approach To Explore Book-length Documents” October 2022

Talk presented at CIKM 2022 Workshop on Human-in-the-loop Data Curation

Title: “Opening scholarly documents through text analytics”

June 2022

Presented at ACM/IEEE JCDL 2022 Doctoral Consortium

Citation: Bipasha Banerjee. 2022. Opening scholarly documents through text analytics. In Proceedings of the 22nd ACM/IEEE Joint Conference on Digital Libraries (JCDL '22). Association for Computing Machinery, New York, NY, USA, Article 47, 1–2. <https://doi.org/10.1145/3529372.3530948>

Title: “Applications of mining ETDs”

November 2021

Presented at ETD 2021 conference

In preparation journal article for J-ETD

Title: Building A Large Collection of Multi-domain Electronic Theses and Dissertations

December 2021

Citation: S. Uddin, B. Banerjee, J. Wu, W. A. Ingram and E. A. Fox, "Building A Large Collection of Multi-domain Electronic Theses and Dissertations," 2021 IEEE International Conference on Big Data (Big Data), 2021, pp. 6043-6045, doi: 10.1109/BigData52589.2021.9672058.

Title: “Extracting Information from Electronic Thesis and Dissertations”

March 26-27, 2021

Conference: ACM Capital Region Celebration of Women (CAPWIC 2021), Virtual

Title: “Summarizing ETDs with deep learning”

November 2019

Citation: Ingram, William A., Bipasha Banerjee, and Edward A. Fox. "Summarizing ETDs with deep learning." *Cadernos BAD 1* (2020): 46-52

Title: “Different Routing Protocols And Their Vulnerabilities And Their Measures”.

March 8-9, 2014

Conference: “The International Conference on Advances in Computer Science and Electronics Engineering-CSEE”, held in Kuala Lumpur, Malaysia.

AWARDS & ACHIEVEMENTS

- “Student of the Year” award by Institute of Engineering and Management for academics as well as contribution to the college for 2014
- Selected by Indian Statistical Institute, Kolkata, as one of the 60 candidates from all over the country as a Summer Cryptology Intern, May 2014.

WORKSHOPS/ TRAININGS/ CERTIFICATIONS

- Online course Data Science: R Basics offered by Harvard University through edX
- Online course titled “Introduction to Python for Data Science”, offered by Microsoft (DAT208x) completed with a final grade of 96% - June 2017
- Microsoft Certification (Microsoft Specialist): Programming in C#, March 2017
- Completed an online course Programming for Everybody (Python) offered by University of Michigan, with distinction. – May 2015
- Microsoft Certification (Microsoft Technological Associate) on Database Fundamentals, June 2014.
- Training on MySQL: Institute of Engineering & Management, Kolkata, April-May 2014
- Training on Microsoft .Net: Webel Informatics Limited, Kolkata, Summer 2013

EXTRA-CURRICULAR

- Member of IEM Computer Society and E-cell. 2012-2015
- Social Service rendered for the destitute and dying at Missionaries of Charity, Kalighat, Kolkata, June-July 2014
- Active participant in elocution competition- Participated in Golpark Ramakrishna Mission Elocution Competition, Kolkata
- ACM, and IEEE student member.