Anurag Roy

Website: ranarag.github.io
Email: anurag_roy@iitkgp.ac.in
LinkedIn: ranurag
GitHub: github.com/ranarag

EDUCATION

Indian Institute of Technology Kharagpur — Kharagpur, India
Ph.D. in Multimodal Machine Learning, Advisor: Prof. Saptarshi Ghosh — 2019–Current

Indian Institute of Engineering Science and Technology Shibpur — Howrah, India
B.E.(Hons), CGPA: 8.13/10 — 2013–2017

EXPERIENCE

India ML Team, Amazon Development Center India Pvt. Ltd. — Bangaluru, India
Applied Scientist Intern — May 2019 - Jul 2019

- Predicting Product Trust Score
- Developed machine learning models to Predict Trust Score

Polaris Networks(a Motorola Solutions Company) — Kolkata, India
Software Developer — Aug 2017 - Jul 2018

- Development of 5G network test tools
  - ∗ Developed 5g core network stack
  - ∗ Developed MCPTT tester
  - ∗ **Nominated for Star Performer Award for the year 2017**

CSE Department, IIT Kanpur — Kanpur, India
Research Intern — May 2017 - Aug 2017

- A Language Independent Morphological Clustering Algorithm
  - ∗ Analyzed different kinds of morphological variations existing over different languages.
  - ∗ Developed of an un-supervised clustering algorithm to capture the morphological variants of a word.

CST Department, IIEST Shibpur — Howrah, India
Research Intern — Dec 2016 - Jan 2017

- Rumour Identification in Twitter
- Developed using python and scikit-learn a program which identified rumor tweets in real-time by learning from previous data.

CSE Department, IIT Kharagpur — Kharagpur, India
Research Intern at CNeRG Lab — May 2016 - Jul 2016

- Understanding Relative Importance of Content & Responses on Community Survival in Reddit.
  - ∗ Developed an on-line survey application in flask and jinja.
  - ∗ Used machine learning models to evaluate accuracy

Journal Publications

[1] **A. Roy**, S. Ghosh, K. Ghosh, and S. Ghosh, "An unsupervised normalization algorithm for noisy text: A case study for information retrieval and stance detection", *Journal of Data and Information Quality*, 2021, To appear.

[2] R. Dey, **A. Roy**, T. Chakraborty, and S. Ghosh, "Sleeping beauties in computer science: Characterization and early identification", *Scientometrics*, vol. 113, no. 3, pp. 1645–1663, 2017.

Conference Publications

[1] Y. Jain and **A. Roy**, "Distributed representation of tags for active zero shot learning", in *8th ACM IKDD CODS and 26th COMAD*, ser. CODS COMAD 2021, Bangalore, India: Association for Computing Machinery, 2021, pp. 228–232, ISBN: 9781450388177.

[2] **A. Roy**, V. K. Verma, K. Ghosh, and S. Ghosh, "ZSCRGAN: A GAN-based expectation maximization model for zero-shot retrieval of images from textual descriptions", in *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*, ser. CIKM '20, New York, NY, USA: Association for Computing Machinery, 2020, pp. 1315–1324, ISBN: 9781450368599.

[3] **A. Roy**, K. Ghosh, M. Basu, P. Gupta, and S. Ghosh, "Retrieving information from multiple sources", in *Companion Proceedings of the The Web Conference 2018*, ser. WWW '18, Lyon, France: International World Wide Web Conferences Steering Committee, 2018, pp. 43–44, ISBN: 9781450356404.

[4] M. Basu, **A. Roy**, K. Ghosh, S. Bandyopadhyay, and S. Ghosh, "A novel word embedding based stemming approach for microblog retrieval during disasters", in *Advances in Information Retrieval*, Cham: Springer International Publishing, 2017, pp. 589–597, ISBN: 978-3-319-56608-5.

[5] M. Basu, **A. Roy**, K. Ghosh, S. Bandyopadhyay, and S. Ghosh, "Microblog retrieval in a disaster situation: A new test collection for evaluation", in *Proceedings of the First International Workshop on Exploitation of Social Media for Emergency Relief and Preparedness co-located with European Conference on Information Retrieval, SMERP@ECIR 2017, Aberdeen, UK, April 9, 2017*, ser. CEUR Workshop Proceedings, 2017, pp. 22–31.

[6] **A. Roy**, T. Ghorai, K. Ghosh, and S. Ghosh, "Combining local and global word embeddings for microblog stemming", in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, ser. CIKM '17, Singapore, Singapore: Association for Computing Machinery, 2017, pp. 2267–2270, ISBN: 9781450349185.

Teaching

- **Teaching Assistant** at IIT Kharagpur   Spring 2020
  Deep Learning (CS60010)
- **Teaching Assistant** at IIT Kharagpur   Autumn 2020
  Programming & Data Structures Theory (CS10003)
- **Teaching Assistant** at IIT Kharagpur   Spring 2019
  Machine Learning (CS60050)
- **Teaching Assistant** at IIT Kharagpur   Autumn 2019
  Programming & Data Structures Theory (CS10001)
- **Teaching Assistant** at IIT Kharagpur   Spring 2018
  Programming & Data Structures Lab (CS19101)

Technical Skills

- **Programming Languages:** C, C++ ,Python, Java
- **Deep Learning Libraries:** Pytorch, Tensorflow, theano
- **Database Management tools:** MySQL, SQLite
- **Big Data Management:** Pyspark

Languages

- **Bengali:** Native or bilingual proficiency
- **Enlgish:** Full professional proficiency
- **Hindi:** Limited working proficiency

Extracurricular Activities

- Member at CodeIIEST   2014 –2017
  CodeIIEST is the coding society of IIEST Shibpur
- Member at Robodarshan   2016 –2017
  Robodarshan is the robotics society of IIEST Shibpur
- Student volunteer at ACM CODS COMAD   Jan 2019
  ACM CODS COMAD is a conference on data science and management