

# Sourin Chakrabarti

---

## PERSONAL INFORMATION

Phone	:	+91 8797964287	LinkedIn	:	linkedin.com/in/sourinc
Nationality	:	Indian	GitHub	:	github.com/imsourin
Mail	:	sourin.chakrabarti@gmail.com	Website	:	imsourin.github.io

---

## EDUCATION

### B.Tech. (Honors) in Information Technology

Indian Institute of Information Technology (IIIT), Allahabad, India

CGPA - 9.16/10

July 2016 – July 2020

**Courses:** Data Structures, Algorithm Analysis, Object Oriented Methodologies, Artificial Intelligence, Database Management, Computer Graphics, Computer Networking

---

## WORK EXPERIENCE

### Amazon Inc.

#### Software Development Engineer

Sep 2021 – Present

As a part of the Subscribe and Save(SnS) Team, my project is aimed at the convergence of Recurring Delivery(RD) for Amazon Business(AB) and SnS. This would lead to a reduction in management costs for RD and hence would put RD on track to increase its revenue by 26% YoY and reach 500k subscribers in the US by 2022. I am involved in the following projects.

- Enabling PayStation support in the Zero Click Ordering(ZeCO) Platform in order to support Decoupled Payment Options(DPO) such as Pay By Invoice, Pay Later, etc., for RD and SnS.
- Development of US AB manual selection management through SnS' selection tool. It deals with implementing an RD featured offer on Detail Page defaulting to our current selection based on the SnS replenishment score ML model.
- Tech Stack - Java, Spring, JUnit, Python, etc.

### MindTickle Inc.

#### Software Development Engineer

July 2020 – Aug 2021

- Contributed to the end-to-end re-architecture, development, bug fixing, testing, and deployments of the User Management(UM) platform of MT. Moved to an async pipeline from the legacy synchronous processing services, which led to a reduction in approximately 80% of event misses.
- Optimized Elastic search(ES) queries by removing redundant joins, migrating to bulk requests, and introducing in-memory ES caches, in the Content and Gamification Engines. This led to a decrease in 20% API timeouts and a mean latency reduction of about 110ms.
- Migrated the Content and Gamification engines(>10 million requests per day) running on EC2 instances to micro-services running inside Kubernetes clusters and on serverless lambdas. Responsible for implementing REST APIs, GraphQL clients, GRPC servers, setting up CI/CD pipelines, etc., for the migration. This led to increased scalability, which in turn directly impacted the reduction of prod issues by 25% as compared to the previous quarters.
- Tech Stack - Play, Kafka, CouchBase, PostgreSQL, etc.

#### Software Development Engineering Intern

Jan 2020 – July 2020

Worked on an auto-attendance mechanism for Virtual Instructor-Led Training (ILTs) which determines users attending a meeting and provides them with an attentiveness score based on their attendance time using MT's CallAI. I was the sole engineer from the Learning Applications team working on this project and hence was responsible for the full-stack development and monitoring of the system. Around 50%(~2.5k) of the ILTs had the auto-attendance mechanism enabled within the first month of its release.

#### Software Development Engineering Intern

May 2019 – Aug 2019

Implemented an audit framework for Coaching Module which helped track the actions(>100k events per day) the users take on the platform. Used Kafka connectors for fetching data from CDC and a custom processor for mapping and storing the parsed data into SQL tables. I was extended a full-time offer in recognition of my work.

---

## RESEARCH EXPERIENCE

### IIIT Allahabad

#### Indo-Norway collaborative project on CPSEC

Jan 2020 – July 2020

#### Supervisor: Dr. Om Prakash Vyas

- Our work was directed towards a resilient architecture for smart grids which anticipate high load and attacks in the grid prematurely and re-configures through deep learning combined with convex optimization and graph search.
- Proposed a meta heuristic-based graph search mechanism for distributed smart grid restoration strategy modeled on the linearized DistFlow Model, with tree pruning to decrease the complexity to  $n^3$ . The model was tested on IEEE-37 and IEEE-1069 bus test systems. It achieved up to 35% load reduction in a 3-fault 1069 bus system.
- Also worked on load prediction using Gated Transformer Networks and a load balancing mechanism which used graph-based CSP for a weighted Convex Bipartite Graphs.

## **Sequential Minimal Optimization for One-Class Slab Support Vector Machines**

Aug 2019 - Jun 2020

**Supervisor: Dr. Om Prakash Vyas**

- Proposed a fast training method for One Class Slab SVMs using a modified Sequential Minimal Optimization(SMO) algorithm. Reduced the Lagrangian to a 2-variable function and devised a suitable variable pair selection heuristic. The optimality criterion was determined through KKT analysis.
- The proposed algorithm was compared against other state-of-the-art One-Class Classifiers such as SVDD, KPCA, Markov Sampled OCSVM, etc., for training time and accuracy(MCC). Our approach was seen to have MCCs comparable to the original OCSSVM(and better than other SOTA classifiers) with approx. 75% reduction in its training time.

## **Network Intrusion detection using deep learning**

Jan 2019 - May 2019

**Supervisor: Dr. Vijay Kumar Chaurasia**

- Implemented a system that detects attacks on a network server in real-time by converting network packets into multi-channel images and detecting anomalies using a novel combination of an SVM with a new feature descriptor and a deep CNN.
- The network packets retrieved from the CICIDS2017 dataset were normalized and quantized before converting to images. The system also ran live on our college network(~10k nodes) to detect any anomalies in the traffic.

## **Text Writing In Air**

Aug 2018 - Dec 2018

**Supervisor: Dr. Ranjana Vyas**

- Implemented a text writing system for the disabled using finger movements in front of a camera. Fingertip detection in dynamic backgrounds was carried out through semantic segmentation, upsampling, and skipping connections.
- OCR for the handwritten text was performed using a dense CNN followed by bidirectional LSTM layers. CTC objective function was used to minimize the loss and Levenshtein distance was chosen to track the performance of the network.
- The model was trained on the IAM Handwriting Database. It showed an accuracy of about 0.80 on the training set.

---

## **INDEPENDENT PROJECTS**

### **Efficient Image Retrieval Using Neural Hash codes and Bloom filters**

Jun 2020 – Sep 2020

- Proposed a novel image retrieval approach using PCA-compressed neural activations and bloom filter for query elimination. Retrieval was carried out in a hierarchical coarse-to-fine approach.
- Engineered a novel approach of extracting neural codes in a single feed-forward cycle from an 8-layer dense CNN pre-trained on Image-Net classes. The model was trained on the Oxford 5k, Paris 6k, and the Inria Holidays datasets with distractor data from the Flickr 100 database. The size of the filter was kept to be  $2n$  and  $5n$ .
- The model was competitive in terms of mAP against the SOTA retrieval mechanisms which used Fisher vectors, VLAD, and triangulation embeddings. But, the meantime for a query significantly improved due to query reduction by bloom filters.

---

## **PUBLICATIONS**

### **Journal Articles**

- Bagesh Kumar, Ayush Sinha, **Sourin Chakrabarti**, and O.P. Vyas. A fast learning algorithm for one-class slab support vector machines. **Knowledge-Based Systems**, volume 228. Elsevier, 2021,(Impact Factor : **8.038**).

### **Conference Proceedings**

- Ayush Sinha, **Sourin Chakrabarti**, and O.P. Vyas. Distributed grid restoration based on graph theory. In *2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC)*.
- Sourin Chakrabarti**. Efficient image retrieval using multi neural hash codes and bloom filters. In *2020 IEEE International Conference for Innovation in Technology (INOCON)*.

---

## **POSITIONS OF RESPONSIBILITY**

- Teaching Assistant**, *Competitive Programming, Data Structures, and Algorithms*, Teaching, conducting lab sessions, and catering to the doubts of about 100 freshers and sophomores.
- Subject Matter Expert**, *Chegg Inc.*, Helped high school and college students from around the world with problems in various computer science domains such as machine learning, data structures, etc.
- Organizer**, *Alkharizim 2019*, An online-rated coding contest held on Codechef with contestants from all over the world.
- Problem Curator**, *CodeChef, HackerEarth and HackerRank*.

---

## **ACHIEVEMENTS**

- Ranked **1<sup>st</sup>** in **Prayogshala**(An online org-wide hackathon organized by MindTickle).
- Ranked **1<sup>st</sup>** in our Institute and **64/4000** in the **ACM International Collegiate Programming Contest** Online round. Ranked **27/121** in the Gwalior-Pune Regionals and **109/313** in the Amritapuri Regionals. *ICPC ID - <https://icpc.global/ICPCID/VCKKEYXAIR9V>*.
- Ranked **11<sup>th</sup>** in the **ACM India Chennai Provincial Programming Contest**.
- Ranked **2<sup>nd</sup>** in **Nucleathon**(A national-level coding competition organized by Nucleus Software Corp.).
- Appeared on the **Dean's Merit List**(A list of students with great academic performance in the first year).
- Ranked in the **top 0.7** percent among 12,00,000 examinees in **Joint Entrance Examination Mains**, Apr 2016.