# SHUBHAM SINGHAL

Apt 3666, SteelWorks, 1220, Mecaslin St Nw, Atlanta GA, +14045435015

shubham07iiit@gmail.com | https://shubham07iiit.github.io | https://in.linkedin.com/in/shubhamsinghal7

Passionate Software Developer and Machine Learning Engineer focused on building intelligent solutions which humans can interact intuitively.

## EDUCATION

| | | |
|---|---|---|
| MSCS, Machine Learning, **Georgia Institute of Technology, Atlanta, USA** | Graduating in May 2020 | 3.33/4 |
| B.tech. (I.T.), **Indian Institute of Information Technology, Allahabad, India** | June 2010 –July 2014 | 9.27/10 |

## PUBLICATION

D. Tomar, S. Singhal, and S. Agarwal, "Weighted Least Square Twin Support Vector Machine for Imbalanced Dataset," Int. J. Database Theory Appl., vol. 7, no. 2, pp. 25–36, 2014

## RESEARCH EXPERIENCE

**Indian Institute of Technology, Bombay**                                                    **Jan 2014-June 2014**
- *Eye Tracking for Natural Language Processing*.

To identify cognitive underpinnings in the text, an algorithm to generate consensus scanpath (eye movements) out of multiple scanpaths using **Bayesian Probability Reasoning** and **Hidden Markov Model** was proposed.

**Indian Institute of Science, Bangalore**                                                    **May 2012-June 2012**
- *Analysis of eye gaze scanpath data.*

To determine the dependency between different sentences in the text, an algorithm was proposed to convert scanpaths into an **undirected weighted graph** by combining saccades to form edges and fixations as nodes.

## WORK EXPERIENCE

**Booking.com, Amsterdam | Software Engineer**                                            **August 2018 to August 2019**
- *Deal of the Day (DOTD)*

Designed an algorithm to determine partners which would be eligible for the DOTD program. The program provides them better ranking in search results on the particular day. Wrote an **oozie Job** in **pyspark** on **Hadoop** clusters.
- *Campaign Microservice*

Implemented a microservice to create campaigns. **Alerting Monitoring** and **A/B testing** were integral part of the service.

**Microsoft India Development Center, India | Software Engineer**                         **March 2017 to August 2018**
- *Drive Vicinity*

Researched, designed and implemented the algorithm to fetch the drives data from **Amazon Redshift**, determining regions where people drive the most, pushing most visited locations to **S3** and exporting further down to **Postgres**.
- *Reporting Microservice*

Wrote a **microservice** to generate reports on users' drives. Microservice was deployed on **Kubernetes**, as **Docker** Containers. Service was written in **Python 3, Django**. Reports were generated asynchronously using **Redis** queue and **celery** workers. **Integration Testing**, code coverage with **Unit Testing** was maintained above 80%**.**
- *GDPR*

Wrote the service which will delete the users' data on request. Delete request could be withdrawn within 30 days. Users' delete requests were stored in **Azure Cosmos DB** (**NoSql**). Cron Job will run every day to delete 30 days older requests.

**Adobe Systems, India | Member of Technical Staff**                                      **July 2014 to March 2017**
- *2 way SSL in Adobe Experience Manager (AEM)*

Added the support for **2 way SSL** authentication using **Java** in AEM on web.
- *Adobe SignIn workflow*

Integrated Adobe Sign in the AEM workflow, asynchronously by **multithreading**. Threads kept on polling the Adobe Sign service to check for user's action, later they callback the workflow when user sign the document.

## ACADEMIC PROJECTS

**Generating sketches from photos and vice versa**                                         **Aug 2019 – Dec 2019**

Used semi supervised technique by generating pseudo image features - patch matching between input and output images. Later **GANs** implemented by **pytorch** and **numpy** were trained with pseudo features rather than ground truth features.

**Classification of Images using Artificial Neural Network**                               **Jan 2019 – May 2019**

Classified images using the connectionist model **ANN** implemented in **C**. Classes were further classified into subcategories using another ANN for each class. Experimented with **SVM's** using **Scikit Learn** to compare the efficiency.

**Improving the efficiency of the Information Retrieval system**                           **July 2013 – Dec 2013**

Disambiguated the sense of the ambiguous word in a query by looking at the context in which it is used to retrieve the best relevant documents in the **Information Retrieval System**.