

# Sayan Dasgupta

## *Curriculum Vitae*

### WORK EXPERIENCE

---

MAR 2019 – PRESENT

Embibe

*Principal Data Scientist* Managing 3

Teams. Consumer Intelligence, Search Relevance and Data Science Platform with a focus to drive learning outcomes for millions of students.

15+ members across teams, comprising of Data Scientists, Data Engineers and ML Engineers. Consumer Intelligence team built and delivered state of the art recommendation platform, by assessing current academic and behavioral state of student leveraging different knowledge tracing models, and then taking the student through a learning journey to help her achieve her goal.

Search relevance team was responsible for launch of lightning fast personalised search, with improved relevance algorithms.

OCT 2017 – MAR 2019

Careem

*Member of technical Staff in Data Science* Responsible for detection and pre-

vention of fraud on the driver side leveraging ML.

OCT 2016 – SEPT 2017

C1X Inc

*Senior Manager of Data Science* Team

Architected data science and analytics platform using AWS EMR(ETL), MySQL, Apache Superset (BI) and reduced time to insights to less than hour from previously being multiple days.

Implemented RTB enrichment based on extensive A/B testing on ad impression bid prices.

MAY 2013 – SEPT 2016

Linkedin

*Senior Data Scientist*

*Anti-abuse engineering Team*

Designed and implemented machine learning models to identify fake accounts on a social network in order to prevent fraud, spam, and

abuse on the site. The work involved designing and implementing an extensible framework to enhance the model by applying different modeling technique, including new features, and benchmarking them.

Trained and implemented classifiers to detect spam in short text like name and headline.

Designed and implemented a framework that takes outcome of classification models for abuse detection available on hdfs and takes appropriate action against them online.

Worked on detection and prevention of scraping on LinkedIn.

Designed and implemented intrusion detection system on LinkedIn production system.

JAN 2012 – APR 2013

Symphony (Shopzilla)

*Senior Data Scientist*




Forecasting Revenue Per Click of each keyword that is being marketed to acquire SEM traffic leveraging big data technologies and statistical learning models.

AUG 2010 - JAN 2012

Innovation Labs, 24/7 Inc

*Senior Analytics Consultant*

Built and deployed machine learning models utilizing click stream data to score visitor propensity to attrite or purchase.

 Bangalore, India  
 +919591374000  
 dasgupta.sayan@gmail.com

## EDUCATION

---

- 2007 – 2009 **Masters in Statistics**  
*Indian Statistical Institute,  
Delhi*
- 2004 – 2007 **Statistics Honours, Bachelor of Science**  
*St Xavier's College, Kolkata*

## INVITED TALKS

---

- 2016 **(Smart Data summit Dubai)**  
*Presented our work on leveraging data science to prevent abuse on LinkedIn at Smart Data Summit, Dubai*
- 2015 **(Facebook Spam at Scale conference)**  
*Presented the work on detecting spam in short text at Spam at Scale Conference hosted by Facebook.*

## MISCELLANEOUS

---

- IIT JAM (Statistics) 2007 All India Rank 15
- Granted patent for Classification of members in a social networking service. Patent Number: US10204307B1

## SOFTWARE SKILLS

---

PROGRAMMING AND SCRIPTING	R, Python,
BIG DATA TECHNOLOGIES	Bash, SQL
	Hadoop,
	Spark,
	Hive,
	Splunk,
	Pig
MISCELLANEOUS	AWS
	Stack,
	Docker,
	Git,
	Python-
	Flask,
	GNU
	Emacs

## REFERENCE

---

Available on request.