

# Deborshi Goswami

*Practicing impactful data-science with empathy, integrity & honesty*

## Contact

- 📍 San Francisco Bay Area, CA
- 📞 +1 (217) 693-9898
- ✉️ deborshi.goswami95@gmail.com
- 🌐 [linkedin.com/in/deborshi-goswami](https://www.linkedin.com/in/deborshi-goswami)

## Education

- Master of Science, Industrial Engg.** 2019  
University of Illinois at Urbana-Champaign, USA
- Bachelor of Engineering, Mechanical Engg.** 2017  
Ramaiah Institute of Technology, India

## Skills

### Languages

Python - SQL - JavaScript - R - C++

### Data Science

Natural Language Processing - Classification - Clustering - Topic Modeling - Deep Learning - Dimensionality Reduction - Visualization - Graph Analysis - eCommerce

### Libraries & Frameworks

Pytorch - TensorFlow - SciKitLearn - Pandas - Numpy pySpark

### Data Engineering

Spark - Hive - Presto - Flink

### Deployment & Frontend

Docker - Flask - React - HTML - CSS

### Soft Skills

Data Storytelling - Communication - Design Thinking - Product Ideation

## Work Experience

### Autodesk

Senior Data Scientist, eCommerce Current

- Enabled personalized content recommendation for 6 Million site users annually, using predictive model that recommends industry banner tab and product use-cases on Autodesk [Home-Page](#) by examining user's prior click-stream behavior
- Spearheading the development of a recommender engine to match customer's use-case, expressed in natural language, to the appropriate products

Data Scientist 2019 - 2021

- Trained and deployed ML model that drives Autodesk Virtual Agent ([AVA](#)); the model uses distilBERT to classify natural language queries into relevant customer support functions
- Reduced error-rate of estimating product-support case resolution times from 3+ days to 1 hr, by building a distributed data pipeline with Spark that runs an estimation algorithm over teraBytes of data

### Inference Analytics

Data Scientist 2019

- Automated radiological report summary generation using Attention based Seq2Seq modelling
- Segmented inbound patients requiring urgent care vs further screening using clustering and topic modeling on radiological reports

### University of Illinois

Graduate Research Assistant 2017

Led team of researchers to develop deep learning/computer vision algorithm for automatic detection of bicyclists & pedestrians in traffic

## Pro-Bono Data Science

### Bridges to Prosperity

Lead Data Scientist 2021

Leading a team of data-professionals to model river flow and flooding data from underdeveloped regions to predict viability and specifications for bridge construction in flood-prone communities

### Nexleaf Analytics

Data Visualization Specialist 2021

Built data visualization tools to intuitively & effectively visualize the health of vaccine cold-chain networks

### Yuno

Data Scientist 2020

Developed an unsupervised machine learning model to pair blue-collar workers with jobs that match their profiles