

Field of Interest

Computer Vision,
Deep Learning,
Machine Learning.

Kanchan Sarkar

Sr. Data Scientist (Computer Vision), Shopee

Tool for Data Science

PyTorch ★★★★★

keras ★★★★★

Tensorflow ★★★★★

Python ★★★★★

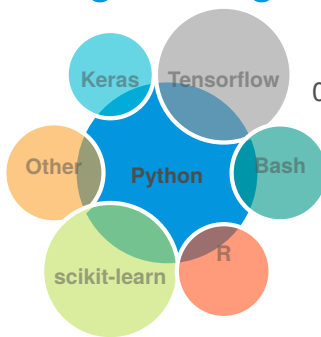
Key Courses

Computer
Vision ★★★★★

Deep
Learning ★★★★★

Machine
Learning ★★★★★
Algorithm ★★★★★

Programming



Personal Skills



Full stack Data Scientist with a history of working in Computer Vision and NLP. Skilled in Computer Vision, NLP, Machine Learning, and Deep Learning.

Experience in computer vision and image processing algorithms, such as image composition, segmentation, image matting, object detection, tracking, and classification.

Experience

03/19 - Now **Sr. Data Scientist**

• [Shopee \(Singapore\)](#)

- Building tools for AI-driven personalize (recommendation) auto Poster/Campaign/Banner generation to reduce the manual effort of the design team.

- Building detection and segmentation model for few thousands of product types to incorporate in the image search (product discovery) on Shopee platform (Apps).

- Experience in computer vision and image processing algorithms, such as image composition, segmentation, image matting, object detection, tracking, and classification.

02/17 - 12/18 **Data Engineer (Data Semantics - Data Science)**

• [DataWeave \(Bangalore\)](#)

- **Product Matching in E-Commerce:** Developed text and image based deep learning model for finding similar products across thousands of e-commerce stores containing millions of products. Text based mapping includes classification of products based on their normalized attribute features. Image based mapping includes image preprocessing (background removal, object detection, upper and lower body detection) and image classification (using CNN, imagenet and inception v3). This is the most revenue generating product at Dataweave and I am leading the semantics part of this product from end-to-end.

- **Sentiment Analysis and Product Review Summarization:** Have wide knowledge of sentiment analysis and summarization of product reviews. Developed deep learning based model for multi-document product review summarization.

- **Description Parsing:** Working on Deep Learning model for product description parsing. In e-commerce platform description of a product is written in plain text which is the rich source of product metadata. This description is error-prone, unstructured and noisy. Building deep learning model for parsing product description to extract meaning full information and product specifications.

- **Attribute Tagging:** Working on Deep Learning based Name Entity Recognition model for tagging different attributes of e-commerce products. Using combination of LSTM and Knowledge Graph for this purpose.

07/16 - 02/17 **Software Engineer**

• Riverbed India Pvt. Ltd (R&D Lab Bangalore)

- Designation: Member of Technical Staff.
- Experience in: C++, Python, Hyper-v, Microsoft Azure.
- Job profile: Developer at SteelFusion product team.
- Worked on design of virtual granite core (SteelFusion for virtual platform) for Hyper-V.

11/13 - 07/14 **DataStage Specialist**

• IBM India Pvt. Ltd (IBM Global Business Service)

- Designation: Associate System Engineer.
- Experience in: InfoSphere DataStage v8.5, IBM DB2.
- Job profile: Build Datastage jobs. Cleanse and transform the data as per business logic. Designed and build shell script utility to automate the tasks for running DataStage batches.

Key Project

Jun, 2015

A Facebook Profile-Based TV Recommender System

IIT , Bombay

- Objective : Building a recommendation system for TV shows based on data collected from Facebook profiles of several users.
- Implemented and evaluated several algorithms in the context of developing a recommender system.
- Used Matrix Factorization technique (SVD), a Clustering algorithm (K-Means), Collaborative Filtering algorithms, Latent Semantic Analysis (LSA), Link Prediction, and Naïve Bayes.

Jan, 2016

Data Analysis and Regression Modeling on a Bike Share System

IIT , Bombay

- Objective : Build a multiple linear regression model that will be used to predict the number of bikes that will be required at any given hour of the day relative to the weather conditions present at that particular time.
- Tool Used : R

Mar, 2016

Predicting Airline Delays

IIT , Bombay

- Objective : Predict flight departure delays for a subset of commercial flights in the United States.
- Used a classifier (SVM) to predict if there will be a delay. To estimate the magnitude of delays, we use a non-parametric quadratic regression algorithm.
- Tool Used : Python

07/15-04/16

Design and Analysis of Multi-Hop Network for TV White Spaces

Master Thesis, IIT , Bombay

- Objective : Central Govt. supported project with the aim of providing broadband connectivity in rural and semi-urban areas.
- Designed multi-hop network for TV white space communication and evaluated performance metric in NS3.
- Performed real experiment of designed topology on Rice WARP board inside IIT Campus.

- 05/12-05/13 **Designing a Network Simulator** [BE Project, BESU, Shibpur](#)
- Objective : To design a network simulator for network protocol such as FTP, HTTP, DNS, DHCP, IP, TCP, ARP and Ethernet with interactive interface.
 - Design : The simulator includes a GUI which allows users to quickly build simulation scenarios with a point-and-click interface.

Hands-on Projects in

1. Sentiment analysis/Classification (Deep Learning).
2. Text/Document Summarization (Multi-Document) (Deep Learning).
3. Named Entity Recognition (Both CRF and Deep Learning).
4. Document Categorization (Topic Modeling - LSA).
5. Finding Similar Documents (TF-IDF, Deep Learning).

Technical Articles

Recent Trends in Natural Language Processing Using Deep Learning. (click me)
ReLU : Not a Differentiable Function: Why used in Gradient Based Optimization? (click me)

Selected Book Chapters

- Dec, 2019 **Deep Learning for Medical Image Segmentation** [IGI Global](#)
- K. Sarkar., and Li, Bohang, "Deep Learning for Medical Image Segmentation", Chapter 8 in Handbook of Research in "Deep Learning Applications in Medical Imaging", Dec, 2019, D. Guide, IGI Global.

Education

- 2014 - 2016 **M. Tech. in Computer Science & Engineering** [Indian Institute of Technology, Bombay](#)
Main subjects: Machine Learning, Algorithm, Artificial Intelligence, Probabilistic Model, Computer Network.
CPI: 8.14
- 2010 - 2013 **BE in Information Technology** [Indian Institute of Engineering Science and Technology, Kolkata \(formally BESU, Shibpur\)](#)
Percentile: 82.71
Title of the Thesis: "Designing a Network Simulator".
- 2007 - 2010 **Computer Science Diploma** [WBSCTE](#)
Central Calcutta Polytechnic
Main subjects: Mathematics, Computer Science. Percentile: 84.5%

Certifications

OS Preference

GNU/Linux ★★★★★

Unix ★★★★★

MacOS ★★★★★

Windows ★★★★★

Technical Skills

Languages

C, C++, Python, Java, R

Cloud Platform

Microsoft Azure

Tools

Eclipse, Latex,

Gnuplot, Weka

API Tools

Stanford NLP Parser,

Wordnet,

Sentiwordnet, JSON

Libraries

Pandas, Numpy,

Sckit-learn, NLTK

Streaming

kafka

Current Location

Singapore

Tel & Skype

+65 82865671

kanchan8308

Mail

kanchan.besu@

gmail.com

03/2015

Specialization in Data Mining

[Coursera](#)

• Courses: Pattern Discovery in Data Mining, Text Retrieval and Search Engines, Cluster Analysis in Data Mining, Text Mining and Analytics.

01/2016

Specialization in Data Science

[Coursera](#)

• Courses: R Programming, The Data Scientist's Toolbox, Getting and Cleaning Data, Exploratory Data Analysis, Reproducible Research.

02/2016

Specialization in Big Data

[Coursera](#)

• Courses: Introduction to Big Data, Hadoop Platform and Application Framework, Introduction to Big Data Analytics

03/2016

Data Visualization and Communication with Tableau

[Coursera](#)

Position and Responsibility

Teaching Assistantship, IIT Bombay:

• Database

– Performed checking and evaluation of 100 students examination copy.

• Computer Network

– Conducted lab sessions, prepared solutions of lab assignments and exam paper.

– Evaluated assignments of 120 students.

Achievements and Extra-Curriculum Activities

• Participated in a National Level **Ethical Hacking** Workshop organized by IIT Bombay Tech-fest in association with TechDefence Pvt. Ltd.

• Participated in **Digital Forensic** workshop Organized by IIT KGP in association with Chakravayuh.

• Successfully cleared the assessment of **Cognizant Certified Student Program** on "IT FOUNDATION SKILLS".

• Participated in **Oracle Database 11g** an oracle university training class in association with IBM and earned "CERTIFICATE OF COMPLETION".

• Got **GAABESU scholarship** for academic achievement in 2010, 2011 and 2012 at undergraduate university.

Nov 30th, 2019

Kanchan Sarkar

Singapore