



sc3653@columbia.edu
linkedin.com/chandrasiddhartha
+91 9161121986 +1 3477662347
github.com/siddhartha-chandra

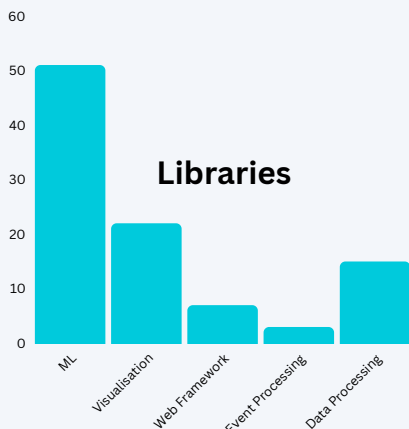


Siddhartha Chandra

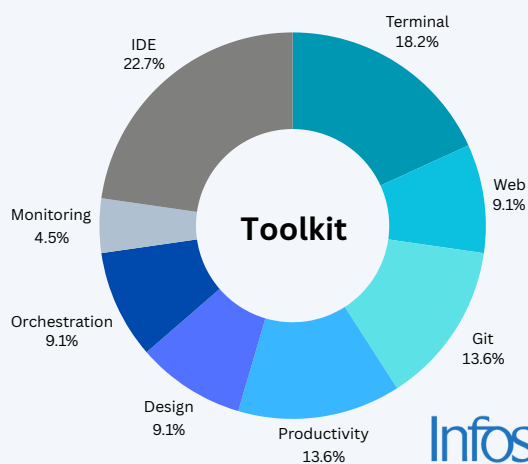
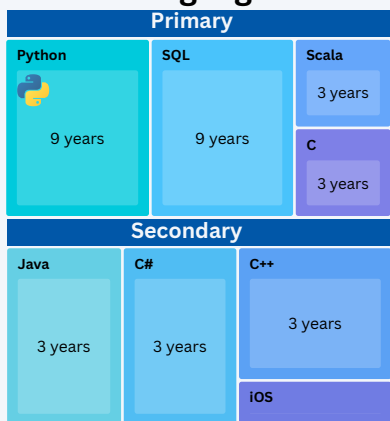
TECHNOLOGIST

education

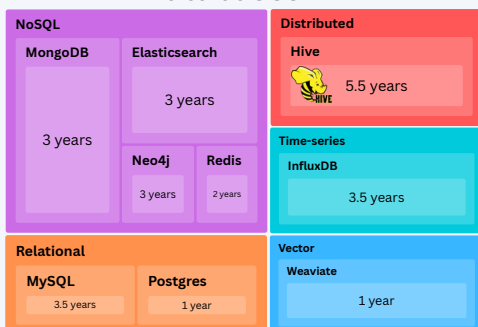
work



Languages



Databases



AUG 2023

Productivity
Stealth Startup

APR 2023
FEB 2023



NOV 2022
APR 2022



OCT 2016
SEP 2016



OCT 2014

MAY 2012

SEP 2008



MS in Computer Science

Columbia University,
The Fu Foundation
School of Engineering & Applied Science
New York, NY
[Sep 2013 - Dec 2014]

CGPA: 3.5/4

Machine Learning (ML), Computer Vision (CV),
Image Processing, Algorithms, AI, Security Architecture



Bachelor of Engineering in Information Technology

Mumbai University,
Sardar Patel College of Engineering (SPCE)
[Aug 2004 - Jul 2008]

MISSION

To creatively harness technology and visual storytelling for crafting compelling narratives that inspire, inform and drive positive change

NOTABLES

- SP-IT**: Pushed the performance median of 160 students from **50%** to **72%** for the course - Data Structures
- NNext**: Productionized outlier detection and dimensionality reduction for project NNifty in less than **3 months**
- Collective[i]**: Authored a data deduplication pipeline that ran **10x** faster & compressed **10%** more than existing solutions
- x.ai**: Built a crf-based email parser that outperformed **mailgun** by **~15%** on F1 score for signature detection

TEACHING

- Business Analytics with Python (2024)
- Data Structures, Advanced Algorithms (2023)

RECOGNITION

- Company-wide hackathon winner (2020)
- Tech talk @ SEAS, Columbia University (2016)
- SPOT award (quarterly) for task excellence (2011)
- PRIMA award (annual) for project execution (2010)

INTERESTS



EDUCATION





Sardar Patel Institute of Technology

Professor of Practice

[Aug 2023 - Present]

- **Projects:** Leading technology project in the Energy sector for the Indian Government and an alumni-incubated stealth startup on end-to-end software development
- **Mentorship:** Advisor for Final year computer engineering students, 5 ongoing project groups; mentored the winners of Govt. of India organised Smart India Hackathon 2023 in the category- Disaster Management
- **Alumni Relations:** Actively building a strong tripartite ecosystem of Alumni, Students and Faculty



NNext.ai

Lead ML Engineer

[Nov 2022 - Jan 2023]

- Productionized approach for transforming unstructured data into vector embeddings, followed by dimensionality reduction using techniques like PCA, UMAP & t-SNE. Coded & structured the entire functionality from scratch & surfaced it as API's, which could be tested from Postman. Used plotly to test visualisation.
- Built a module to detect outliers and misclassified points in vector space
- Prototyped a feature to empower sellers to query & view product reviews as points in a 3-D space
- Authored a library for database communications. Supported databases include Postgres, Weaviate & NNext DB



Collective[i], New York City

Data Science Engineer

[Oct 2016 - Apr 2022]

- **Incremental unification workflow**
Created a light workflow for unifying person & organisation entities; resulted in **10x** improvement in speed
- **Prediction Systems for deal pipeline insights(Python)**
Co-authored system for predicting deal win probability, close time & customer revenue based on deal pipeline. This included dashboard monitoring, Airflow orchestration, legacy code transition & technical-debt payoff
- **Automl architecture (Python)**
Prototyped a democratized ML architecture (automl) using h2o.ai for predicting deal pipeline insights
- **Quarterly business report generator (Python)**
Created a python script for generating quarterly business reports for the client success team. Reports would be generated from queries & sent to stakeholders via email; Simplified UI for inputting parameters using Jenkins
- **Active learning approach for Person profile unification (Python)**
Created pipeline for de-duplicating people records in the System, achieving compression of **10%**. Added visual time-based dashboards to capture metrics of processes
- **Parcel creation utility for Anaconda distribution (Python)**
Developed a utility that installs python packages and downloadable tars & then converts the setup into a Cloudera parcel for easy distribution across machines using Cloudera's parcel manager
- **Word2vec implementation for getting top news for sales reps (Python)**



x.ai, New York City

Data Scientist

[Oct 2014 - Sep 2016]

- **Person Entity Detection (Scala)**
Created data model & system for detecting Person entities using Semantic-role-labelling & anaphor resolution
- **Dashboard management (Scala)**
Authored & maintained dashboards for measuring annotation quality, model performance & company KPI's
- **Email-parser (Python & Scala)**
Created a parser for segmenting emails. Refined it into a linear chain conditional-random-fields (CRF) model. This boosted existing F1 score by **15%**



Infosys Ltd, Chandigarh, India

Sr. Systems Engineer

[Sep 2008 - May 2012]

- Administrator and maintainer of the website (ASP .net) designed for tracking team progress
- Certified security agent builds on App/Web Servers on various platforms (Windows, AIX, Linux, Solaris, z/OS)