Saransh Chopra

saransh-cpp.github.io | saransh0701@gmail.com | saransh-cpp | Saransh-cpp | Blogs

About

Saransh is an engineering junior who, by day, works on his academic and professional commitments, and by night, develops and maintains Open-Source Software, which he believes are the key to reproducible and collaborative research.
He is passionate about Software Development, DevOps, Research Software Engineering, Machine Learning, Scientific Machine Learning, and Open-Source Software.

Education

Cluster Innovation Centre, University of Delhi — New Delhi, India
Major: Information Technology and Mathematical Innovations — *2020 – 2024*
Minor: Computational Biology — *CGPA 9.40*

Research and Work Experience

FluxML, The Julia Programming Language — Remote
Technical Wroter and Open-Source Developer [Funded by The Julia Programming Language] — *May 2022 – November 2022*
Mr. Dhairya Gandhi (Julia Computing)

- Worked on FluxML, an ML ecosystem providing lightweight abstractions on top of Julia's native GPU and AD support.
- Fixed bugs and developed the infrastructure of numerous ML packages such as Flux.jl, NNlib.jl and Metalhead.jl.
- Wrote original Machine Learning/Deep Learning tutorials, documentation and API references for FluxML's ecosystem.

Institute for Research and Innovation in Software for High Energy Physics (IRIS-HEP), Princeton University — Remote
Research and Development Fellow [Funded by IRIS-HEP] — *June 2022 – October 2022*
Dr. Henry Schreiner (Princeton University / CERN) and Dr. Jim Pivarski (Princeton University)

- Prepared Vector for v0.9.0, v0.10.0, and v1.0.0 (first major release) by developing new public API, fixing bugs, and building new infrastructure.
- The releases are currently being used by researchers at CERN, Princeton University, and other research institutes to construct 4D jagged (awkward) vectors and perform Just-In-Time compiled vector operations in Python.
- Worked closely with CERN and Princeton researchers to build infrastructures for the HEP ecosystem.

AiView — New Delhi, India
Research and Development Engineer — *September 2021 – December 2021*
Mr. Geetansh Saxena and Mr. Chappa Sri Vinay

- Developed end-to-end OCR and Object Detection libraries by wrapping existing software for visually impaired people.
- Built a CI/CD pipeline and a REST API for ease of integration with any existing software using FastAPI and Heroku.
- Collected a dataset of human images and their distance from a point by building a stereovision camera.

PyBaMM (Python Battery Mathematical Modelling), NumFOCUS — Remote
Google Summer of Code Student Developer [Funded by Google] — *May 2021 – August 2021*
Prof. Ferran Brosa Planella (University of Warwick), Dr. Valentin Sulzer (Carnegie Melon University), Dr. Robert Timms (University of Oxford)

- Built an automated Twitter Bot capable of constructing Mathematical Simulations of Batteries, including but not limited to different battery models, parameter sets, chemistries, degradation modes, and experiments.
- Developed a replying functionality to run Mathematical Simulations on command using Twitter API.
- Created a novel CI - CD infrastructure and followed a micro-services-based architecture.

Open Source Research Software

PyBaMM — *NumFOCUS' Steering Committee Member, Maintainer, and Core Developer - 450+ stars*
PyBaMM (Python Battery Mathematical Modelling) solves physics-based electrochemical DAE models by using state-of-the-art automatic differentiation and numerical solvers.

Vector — *Collaborator and Core Developer - 45+ stars*
Vector is a Python 3.7+ library for 2D, 3D, and Lorentz vectors, especially arrays of vectors, to solve common physics problems in a NumPy-like way.

[SSP] **S. Chopra**, H. Schreiner, and J. Pivarski. *Vector: vector classes and utilities.* DOI: 10.5281/zenodo.5942082. URL: https://github.com/scikit-hep/vector.

Flux.jl *Member (FluxML) and Core Contributor - 3,700+ stars*

Flux is an elegant approach to machine learning. It's a 100% pure-Julia stack, and provides lightweight abstractions on top of Julia's native GPU and AD support. Flux makes the easy things easy while remaining fully hackable.

BattBot *Member (pybamm-team), Maintainer, and Core Developer - 10+ stars*

An automated Twitter Bot that Tweets random Battery Mathematical Modeling Simulations and replies to the requested Battery Simulations.

liionpack *Member (pybamm-team), Maintainer, and Core Developer - 30+ stars*

liionpack takes a 1D PyBaMM model and makes it into a pack. You can either specify the configuration e.g. 16 cells in parallel and 2 in series (16p2s) or load a netlist.

Other notable contributions

- **Zarr**: *An implementation of chunked, compressed, N-dimensional arrays for Python.* - 1,000+ stars
  Migrated Zarr's build and packaging infrastructure to a pyproject.toml based configuration.
- **DeepXDE**: *DeepXDE is a library for scientific machine learning* - 1,200+ stars
  Implemented utility functions and improved the existing examples on solving partial differential PDEs using PINNs.
- **Colour**: *Colour Science for Python* - 1,500+ stars
  Implemented the conversion between RGB and HCL colourspaces, along with tests and documentation.
- **Scikit-HEP ecosystem**: *High Energy Physics in Python*
  Fixed minor bugs in awkward and hist, added support for coverage in cookie, and wrote new developer pages.
- **Scikits.odes**: *Offers extra ODE/DAE solvers; an extension to ones available in SciPy - 100+ stars*
  Debugged and fixed their build (which was failing because of setuptools) and removed support for Python 2.7-3.6.

Posters and Presentations

[SSP22a] **S. Chopra**, H. Schreiner, and J. Pivarski. *Compiling Awkward Lorentz vectors with Numba.* 21st International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT). Poster (upcoming). Oct. 2022. URL: https://indi.to/45Kzq.

[SSP22b] **S. Chopra**, H. Schreiner, and J. Pivarski. *Constructing HEP vectors and analyzing HEP data using Vector.* 5th International Workshop on Python in High Energy Physics (PyHEP). Presentation. Oct. 2022. DOI: https://doi.org/10.5281/zenodo.7081003. URL: https://indi.to/bPmMc.

[S C22a] **S. Chopra**. *Code coverage through unit tests running in sub-processes/threads: Locally and automated on GitHub.* 10th Annual Conference on Python - Asia-Pacific (PyCon APAC). Presentation. Sept. 2022. URL: https://tw.pycon.org/2022/en-us/conference/talk/243.

[S C22b] **S. Chopra**. *Code coverage through unit tests running in sub-processes/threads: Locally and automated on GitHub.* 21st Annual Conference On Python - Europe (EuroPython). Presentation. July 2022. URL: https://ep2022.europython.eu/session/code-coverage-through-unit-tests-running-in-sub-processes-threads-locally-and-automated-on-github.

Grants, Prizes and Achievements

- **PyCon Acia-Pacific (APAC)**'s grant to lead a talk virtually (NT$3,600) *July 2022*
- **Shubhra Kar Linux Foundation Training Scholarship** (500 recipients worldwide) for contributions to open-source research software ($695) *June 2022*
- **EuroPython**'s travel grant to lead a talk in-person (£460) *June 2022*
- **PyCon Italia**'s travel grant to lead a workshop in-person (£400 + lodging support) *April 2022*
- **Faraday Institution's Collaboration Award** for contributions to the PyBaMM ecosystem *November 2021*
- Won the **Elastic hackathon** (out of 2500+ registrations) with a Flutter-Node-Google Cloud-ELK (Elasticsearch, Logstash, Kibana) based application *May 2021*
- **Discovered an asteroid** having a fixed orbit around Sun by analysing the data **Pan-STARRS** observatory *August 2016*

Projects

Assessing suicidal tendencies on Twitter using BERTicle [BERT, Twitter API, Interaction Circles] *October 2022*

- Worked under Dr. Anjani Verma to detect (94.7% prediction accuracy) and prevent suicidal tendencies on Twitter using a novel framework.
- BERTicle scrapes off a user's Twitter interaction to construct an interaction circle with 8 most interacted users.

- Users in this interaction circle are then warned of the primary user's suicidal tendencies detected by a BERT model.

releaseup [NLP, Text summarisation, spaCy, Scikit-learn, TF-IDF, Python library] *September 2022*

- Releaseup uses an extractive approach for generating release notes from comments and docstrings added b/w two git tags.
- The Python library has **400+ installs** on PyPI, **5+ stars** on GitHub, and follows best development practices.

OCRed [OCR, Computer Vision, Tesseract-OCR, NLTK, Python library] *August 2022*

- OCRed provides clever, simple, and intuitive wrapper functionalities for OCRing specific textual materials.
- The Python library has **2,900+ installs** on PyPI, **10+ stars** on GitHub, and follows best development practices.

ForMente [NLP, Flutter, Dart, FastAPI, Python, Firebase, Firestore, Heroku, GitHub Actions] *June 2022*

- Using Natural Language Processing, ForMente lets you diagnose your feelings in the form of a secure personal diary.
- The NLP model is deployed on Heroku using FastAPI, and the app uses Firebase and Firestore as its backend.

Covid-19 chest image Classification and Segmentation [Transfer Learning, Wavelet Transforms, UNet] *May 2022*

- Worked under Dr. Nirmal Yadav on a joint classification and segmentation system for Covid-19 chest CT scans.
- Achieved a classification accuracy of 94.88% and a segmentation accuracy of 97.29%.

ChaoticEncryption.jl [Image processing, Encryption algorithms, ODEs, PRNGs, Julia package] *February 2022*

- Vectorised image encryption and PRNG algorithms that runs ~**40X faster** than ordinary nested-for implementations.
- The Julia package has **10+ installs** on JuliaHub and **28+ stars** on GitHub.

PDEsWithPINNs [PDEs, PINNs, Python, DeepXDE] *January 2022*

- Worked under Prof. Shobha Bagai to solve 1, 2, 3, and 4D Partial Differential Equations using Physics-Informed Neural Networks.
- Worked with Neural tangent Kernels, Multi-scale Fourier feature networks, and Spatio-temporal Multi-scale Fourier feature networks to predict high-frequency details.

Position of Responsibilities, Volunteering, and Mentoring

- Member of **PyBaMM's NumFOCUS steering committee**. *Present*
- Member of **FluxML**'s GitHub organisation. *Present*
- **Vector**'s collaborator on GitHub. *Present*
- Mentored a hackashop conducted at the **5th International Workshop on Python in High Energy Physics (PyHEP)**. *September 2022*
- Mentored students under PyBaMM, NumFOCUS for **Google Summer of Code 2022.** *May 2022*
- Mentored 10+ students under Codepeak 2021 on open-source application development. *December 2021*

Relevant Skills

Languages:	Python (proficient) | Julia (proficient) | Dart (proficient) | C/C++ | Java | JavaScript
Frameworks/Libraries:	Tensorflow | PyTorch | FluxML-Ecosystem | PyData-Ecosystem | SciML-Ecosystem | Flutter FastAPI | NodeJS | ExpressJS | Android
Platforms and Tools	Linux | Heroku | Google Cloud | Elasticsearch | Kibana | Shell Scripting | Docker | CI/CD | VCS SQL and NoSQL Databases | GitHub Actions

Languages

Hindi (mother tongue), English (fluent)