SARANSH CHOPRA

🌐 saransh-cpp.github.io | ✉ saransh0701@gmail.com | 💼 saransh-cpp | ⛎ Saransh-cpp | Ⓜ WhiteViolin

About

Saransh is an engineering junior who, by day, works on his academic and professional commitments, and by night, develops and maintains open-source research software, which he believes are the key to collaborative and reproducible research.

He is passionate about Research Software Engineering, Machine Learning, Deep Learning, Scientific Machine Learning, and Open-Source Research 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
Open-source developer and technical writer [Funded by The Julia Programming Language]   *May 2022 – Present*
Mr. Dhairya Gandhi (Julia Computing)

- Fixing bugs and developing the infrastructure of prominent Julia ML and DL libraries such as Flux.jl (3,700+ stars), NNlib.jl (Neural Network primitives - 150+ stars), Metalhead.jl (Computer vision models - 270+ stars), and Functors.jl.
- Writing original Machine Learning/Deep Learning tutorials, documentation and API references for FluxML's ecosystem.
- Working closely with doctoral researchers and FluxML maintainers from around the world.

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

- Prepared Vector (85,000+ installs - 45+ stars) 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.

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, created a CI/CD pipeline, and followed a micro-services-based architecture.
- Worked closely with post-doctoral research fellows and PyBaMM maintainers from around the world.

Publications

[Tra+21]   T. G. Tranter, R. Timms, V. Sulzer, F. B. Planella, G. M. Wiggins, S. V. Karra, P. Agarwal, **S. Chopra**, S. Allu, P. R. Shearing, and D. J. L. Brett. *liionpack: A Python package for simulating packs of batteries with PyBaMM*. Journal of Open Source Software, 7(70), 4051. Oct. 2021. DOI: https://doi.org/10.21105/joss.04051.

Open Source Research Software

PyBaMM   *Member (pybamm-team), Maintainer, and Core Developer - 235,000+ installs - 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 - 85,000+ installs - 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 - 80,000+ installs - 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 - 130+ followers - 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 - 2,400+ installs - 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.* - 4,500,000+ installs - 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* - 350,000+ installs - 1,200+ stars
  Implemented utility functions and improved the existing examples on solving partial differential PDEs using PINNs.
- **Colour**: *Colour Science for Python* - 6,007,000+ installs - 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 - 160,000+ installs - 100+ stars*
  Debugged and fixed the 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**. *Python packaging: from stone age to the future.* PyDelhi workshop. Presentation. Sept. 2022.

[S C22b] **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 C22c] **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.

[S C21] **S. Chopra**. *BattBot: Mathematical Modeling of Batteries using an automated Twitter bot.* 1st International Training Workshop on Python Battery Mathematical Modeling (PyBaMM). Presentation. Oct. 2021. URL: https://www.pybamm.org/training.

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

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 **200+ 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 **1,900+ 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.

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.

PopItUp [Android, Kotlin, Firebase, Firestore, Google Sceneform SDK, Google ARCore SDK]   *May – August 2021*
- An Augmented Reality shooting game developed in native android by manually adding and editing the Sceneform SDK.
- Developed a live leaderboard using the FirestoreRecyclerView.

MemeTastic [Flutter, Dart, NodeJS, Elasticsearch, Kibana, CI/CD, Google Cloud, Reddit API, Ngram Analyser] *May 2021*
- Developed and deployed a backend service to ingest the latest meme data into Google Cloud hosted Elasticsearch.
- Created a frontend with a clean UI to fetch the data, providing an auto-complete feature to search any meme.

Volunteering / Mentoring

PyHEP (Python in High Energy Physics) Hackashop 2022   *September 2022*
- Collaborated with PyHEP organisers to mentor at a hackathon+workshop (hackashop) conducted for Python's High Energy Physics ecosystem.

Google Summer of Code   *May - September 2022*
- Mentored students under PyBaMM, NumFOCUS, on projects involving, but not limited to, documentation, DevOps, parameterisation, and visualization.

Relevant Skills

Languages:	Python (proficient) | Julia (proficient) | Dart (proficient) | C/C++ | Java | JavaScript
Frameworks/Libraries:	Tensorflow | PyTorch | spaCy | NLTK | Tesseract-OCR | FluxML-Ecosystem | DeepXDE PyData-Ecosystem | SciML-Ecosystem | Flutter | FastAPI | NodeJS
Platforms and Tools	Linux | Heroku | Google Cloud | Elasticsearch | Kibana | Shell Scripting | Docker | CI/CD | VCS
Research Interests:	Deep Learning | Machine Learning | Scientific Machine Learning | Open Source Research Software

Languages

Hindi (mother tongue), English (fluent), German (basic), Punjabi (conversational)