

# Joel Joseph

joeljosephjin@gmail.com | (+91) 9497319129 | <https://joeljosephjin.github.io/>

## EXPERIENCE

### Bajaj Finserv, Pune — Senior Data Scientist

October 2022 - PRESENT

- Developed a web app in flask to aid in tagging text data in a semi-automated way.
- Used ChatGPT to develop an ATS System and deployed it in an internal server using Streamlit and Nginx.
- Developed a Voice Identification System for customer care calls utilizing a pre-trained language model (Titanet) coupled with a Siamese Network for enhanced accuracy and efficiency.

### Innoplexus Consulting Services, Pune — Associate Data Scientist

July 2021 - October 2022

- Developed novel solutions to drug development problems using GPT2, LSTM and other NLP models/techniques, which got 90% Laboratory Accuracy.
- Worked on Data Extraction/Processing from PDFs using Camelot-py, NLTK, Stanza (Stanford NLP) taggers, BERT, etc. Got 85% accuracy on Beta Validation.
- Used ElasticSearch, etc. for working with databases.

## INTERNSHIP

### AUV-IITBHU, Varanasi — Computer Vision Developer

August 2018 - September 2020

- Worked with non-deep learning based Object Detection techniques in OpenCV such as edge detection.
- Led the Computer Vision team in developing the code framework for Object Detection on Underwater Images using FastRCNN model. The results helped the group in getting more funding from the Inter-IIT Board. [Github](#)

## EDUCATION

### Indian Institute of Technology (BHU), Varanasi — BTech 2017 - 2021 | CGPA: 7.83

## PROJECTS

### Pixelate (TechFest) — Autonomous Bot Development

Jan 2018 - Mar 2018

- Developed code for object detection, localization and navigation in MATLAB without the use of Machine Learning techniques (as per the rules) for navigating an autonomous bot through a maze using input from an overhead camera.

## OTHER WORKS

- Worked with Dr. Jie Fu (MILA, Quebec AI Institute) on coding a novel Meta-Reinforcement Learning solution based on Learnable Neural Optimizers. Read research papers and implementations on Github for developing the code.
- Open Source Contributions:
  - mlpack: Wrote documentation, fixed bugs and developed some functions (in C++) towards the Reinforcement Learning part of the Library.
  - Learn2learn: Implemented Meta-Learning technique called MAML (Model Agnostic Meta Learning) on toy datasets. [Pull Request \(Github\)](#)
- Participated in IEEE Cog 2021 Competition, in which I worked on developing a Reinforcement Learning model that can generalize across multiple levels of the same grid based video game present in the General Video Game AI (GVG-AI) framework. I made use of the Stable-Baselines package (PyTorch) for my work.

## PATENTS

**US18060187:** A Method And System For Generating A Plurality Of Antibody Sequences Of A Target From One Or More Framework Regions

**US18060218:** A Method And System For Generating A Plurality Of Antibody Sequences

**US18148474:** Method And System For Predicting Binding Affinity Of Protein Structures Based On Deep Learning

**US18148794:** Method And System For Converting A Protein Data Bank File Into A Two-Dimensional Numerical Matrix

**US18148854:** Method And System For Converting A Protein Data Bank File Into A Grayscale Image Array

## SKILLS

### LANGUAGES

Python, C++, SQL, Javascript, MATLAB

### TECHNOLOGIES

PyTorch, Tensorflow 1.x/2.x, Keras, Huggingface Transformers, Sklearn, Matplotlib, Seaborn, Numpy, Pandas, AWS, Azure, PySpark

## AWARDS

### Best Performer of Data Science Team

Received Best Performer [Award](#), during my work on the Drug Sequence Generation Project.

### WandB Reproducibility Grant 2021

Received a grant amount of \$500 for successfully reproducing a NeurIPS paper and publishing a report of the analysis.

### JEE Advanced 2017

Top 1 percentile among 0.1 million candidates

## CERTIFICATIONS

[Deep Learning Specialization](#) - Coursera

[Fundamentals of Reinforcement Learning](#) - Coursera

## LANGUAGES

English, Hindi, Malayalam