# Divya Prakash Manivannan

(+1) 917-388-6994 | dippuprakash@gmail.com | LinkedIn | Github

## Technical Skills

**Languages**: Python, C/C++, SQL, PL/SQL, NoSQL, HTML, Shell Scripting, Javascript
**Tools & Technologies**: Github, Jira, CI/CD, Linux, Snowflake, OracleDB, AWS, Azure, Tableau, Hadoop, Docker, dbt, Spark
**Python Libraries**: Pytorch, TensorFlow, Numpy, Pandas, Matplotlib, Scikit-Learn, PySpark, Scipy
**Role Related Skills**: Data Structures & Algorithms, ETL, ELT, Data Analysis, Predictive Modelling, Dimensional Modelling
**ML/DL Domains**: Computer Vision (CV), Natural Language Processing (NLP), Large Language Models (LLMs), GenAI
**Industry Skills**: Agile Practices, Test Driven Development, Design Thinking & Analysis, Cloud Computing

## Experience

**Research Scientist** Jul 2023 - Sep 2023
*New York University, Tandon School of Engineering* *Brooklyn , New York*
- Enhanced Out-Of-Distribution detection techniques for detecting backdoor samples and achieved a maximum of **95% accuracy**.
- Executed established methodologies and managed the data pipeline for generating results for over 5000 configurations.
- Achieved a 50% reduction in run times by optimizing GPU cluster usage, implemented on NYU's Greene HPC cluster.

**Graduate Student Researcher & Developer** Jan 2022 - Dec 2022
*New York University, Tandon School of Engineering* *Brooklyn , New York*
- Revamped the Augmented Library APP by developing a new backend system for reserving 40 study spaces.
- Proposed a cost-effective **AWS Cloud Hosting** solution for migrating the backend PostgreSQL database, for scalability.

**Software Engineer** Sept 2018 - Jul 2021
*IBM India Pvt. Ltd.,* *Bengaluru, India*
- Developed high-performance PL/SQL APIs using **TDD**, managing 400K daily transactions of **AT&T Enterprise** accounts.
- Led the implementation of an automated mailing system using **Shell Scripting** and **Dynamic SQL**, enhancing database insights and averting outages, reducing manual workload from 2 hours/week to zero.
- Utilized **SQL** for data analysis and managed over **50** critical APIs, reflecting proficiency in **EDA** tasks and **ELT** processes.
- Optimized PL/SQL APIs, cutting Production Database execution time by 10%.
- Facilitated team meetings and Knowledge Transfer sessions for new database team members while serving as Subject Matter Expert **(SME)** for approximately 10 sub-applications.

## Education

**New York University, Tandon School of Engineering** Brooklyn , New York
*Master of Science, Computer Engineering* *Sept 2021 - May 2023*

**PDPM Indian Institute of Information Technology, Design & Manufacturing** Jabalpur, India
*Bachelor of Technology, Electronics and Communication Engineering* *Aug. 2014 – May 2018*

## Projects

**Security Vulnerabilities of Self-supervised Multi-Modal Models** Jan 2023 - May 2023
- Evaluated the security vulnerabilities of Self-supervised Multi-Modal models, highlighting the need for robust backdoor defenses.
- Devised **DKAC** (Dual Key Attack on CLIP), a novel backdoor attack, leveraging CLIP's characteristics in Visual Question Answering (VQA) setups.

**Deep Dive into Google Play Store Apps & Reviews** Sep 2022 - Dec 2022
- Performed comprehensive analysis on a Google Play Store dataset comprising 2.3 million entries using **Python** and **PySpark**, uncovering insights into app popularity, advertising efficacy, and user engagement dynamics.
- Executed extensive data pre-processing, statistical analysis, causal inference, and classification tasks, achieving a 80% **test accuracy** for predicting 'Maximum Installs' and uncovering key factors influencing app ratings and installs.

**Data-Driven Movie Analysis: Unveiling Patterns in Viewer Preferences** Sep 2022 - Dec 2022
- Applied hypothesis testing on movie data set having ratings of over 1000 users to assess popularity and viewer preferences.
- Maximized movie rating predictions through advanced regression techniques against performance metrics such as **COD**, **RMSE** and **AUC** values, achieving up to 95% accuracy in predicting average user enjoyment for individual films.

**Stealthy Syntactical Backdoor Attack on Language Models [Link]** Sep 2022 - Dec 2022
- Proposed a novel backdoor attack on Language Models, extending the Hidden Killer attack by integrating poison data generation with the Pre-trained T5 model, resulting in a maximum **ASR** increase of 23%.
- Implemented a robust backdoor defense using GPT3, reducing ASR by 50% in most tested model and data set configurations.

**Cross-Architectural Self-Supervision for Multi-Modal Learning [Link]** Sep 2022 - Dec 2022
- Explored Unimodal and Multi-modal Self-supervised approaches for classifying memes in Facebook AI's hateful meme dataset, evaluating performance with AUROCs and accuracies.
- Introduced **CASS-MM (CASS-Multi-modal)**, a novel technique, outperforming CLIP-trained models in accuracy by 10%.

## Publications

**NeurIPS 2023 Workshop BUGS:** On the Limitation of Backdoor Detection Methods **[Link]**