

Praveen Kumar Sridhar

Website: PraveenKumarSridhar.com

Github: github.com/PraveenKumarSridhar

Email: sridharpraveenkumar@gmail.com

Linkedin: [in/praveenks23/](https://www.linkedin.com/in/praveenks23/)

Mobile: +1 857 424 8758

EXPERIENCE

Khoury college - NULabs

Research Assistant

Boston,MA

Aug 2022 - Present

- ◇ Conducted research on the impact of influential actors on the #MeToo online social movement, utilizing machine learning techniques.
- ◇ Developed and optimized a Natural Language Processing (NLP) pipeline, incorporating pre-processing, topic modeling, and author **demographic inference**.
- ◇ Implemented an **emotion detection** module using **DistilBERT** and a pagination module to increase overall efficiency and improve functionality.

Meta

Data Scientist Intern

Seattle,WA

May 2022 - Aug 2022

- ◇ Spearheaded a project to redefine the organization's approach to **advertiser churn**, resulting in the creation of a robust framework for predicting and reducing churn.
- ◇ **Devised new features** to identify and isolate controllable factors that contribute to advertiser churn, utilizing highly optimized SQL queries.
- ◇ Applied advanced statistical techniques such as **causal inference models (X learners, Causal trees)** to measure the impact of various features on churn.
- ◇ Utilized **SHAP values** to pinpoint critical features that led to advertisers leaving the platform, and developed accurate machine learning models (89% ACC) to forecast churn.
- ◇ Leveraged **calibration curves** to refine the models and ensure optimal accuracy.

Intellect Design Arena

Data Scientist

Chennai,India

Jun 2018 - Aug 2021

- ◇ Designed, developed, and successfully deployed deep learning models such as Convolutional Neural Networks (CNNs), Long Short-Term Memory networks (LSTMs), and Bidirectional LSTMs with attention, achieving **accuracy above 90%** in a production environment.
- ◇ Developed and integrated a user feedback module to gather user feedback and improve model performance.
- ◇ Designed, built, and deployed a complex ensemble classifier using **BERT and ROBERTA**.
- ◇ Experimented with various Optical Character Recognition (OCR) tools, including **Tesseract, easyOCR, paddleOCR, and aOCR**, and combined CRAFT with Tesseract to achieve a **5% increase** in accuracy.
- ◇ Utilized image processing techniques in combination with **Tesseract and CRAFT** to extract data from the Machine Readable Zone in passports.
- ◇ Developed a complete Natural Language Processing (NLP) pipeline using RabbitMQ, from tokenization to spell-checking, that runs on multiple customizable servers with adjustable worker/consumer numbers and flow.

EDUCATION

Northeastern University, Khoury College

Masters, Data Science; GPA: 4.0/4.0

Boston,MA

Aug 2021 - May 2023

VIT University

B.Tech, Computer Science;

Chennai,India

May 2014 - Apr 2018

TECHNICAL SKILLS

- ◇ **Languages:** Python, R, Scala, C/C++, Java
- ◇ **ML & DL Packages:** TensorFlow, Keras, PyTorch, sklearn, Plotly, Matplotlib, NEAT, OpenCV, tesseract, EasyOCR.
- ◇ **Databases:** MongoDB, Redis, SQL Server, PostgreSQL.
- ◇ **Technical Skills:** NLP, Image Processing, Deep Learning, Machine learning, Tableau, Data Cleaning & Interpretation, A/B testing, Causal inference techniques, git.
- ◇ **Coursework:** NLP, Supervised ML, Algorithms, Deep Learning, Information Retrieval, Unsupervised ML.

PROJECTS

- ◇ **Poetry Generator:** Developed a multilingual poetry generator using advanced **Bidirectional LSTM** neural networks. The model was trained to generate beautiful and engaging poems in English, and Tamil, with the latter being regional South Indian language.
- ◇ **Question Answering model:** Designed and developed a powerful question-answering model leveraging state-of-the-art transformer-based models such as **BERT, DistilBERT, and ALBERT**. The model achieved an impressive F1 score of **81%**, showcasing its ability to accurately answer a wide range of questions.
- ◇ **AI Flappy bird:** Created a captivating version of the classic Flappy Bird game using pygame and incorporated cutting-edge AI technology powered by **NEAT** (NeuroEvolution of Augmenting Topologies) to train the AI player. The AI learned quickly and achieved a **high score of 5000**, demonstrating the effectiveness of the training algorithm.
- ◇ **Art generation** Produced stunning and visually appealing art using pre-trained **VGG-19** deep neural network. The model generated images that combined the content of one image with the style of another image, creating unique and breathtaking results.
- ◇ **The Song Search**The Song Search: Built an innovative Information Retrieval system to retrieve songs from a curated music dataset. The system achieved a MAP of 80% by using **Tensorflow MAGENTA's MT3** model that uses a **T5 architecture** for music transcription. With this system, music enthusiasts can easily find and enjoy their favorite tunes with ease.