Rishit Toteja

Roll No: 2K20/EE/217
B.Tech - Electrical Engineering
Minor in Computer Science and Engineering
Delhi Technological University

+91-8588970916
totejarishit262@gmail.com
GitHub
LinkedIn
Website

Education

Degree/Certificate	Institute/Board	CGPA/Percentage	Year
B.Tech. Major	Delhi Technological University, India	9.53 (Department Rank 1)	2020-2024
B.Tech. Minor	Delhi Technological University, India	10.00	2020-2024
AISSCE (Class XII)	CBSE Board	93.00%	2020
AISSE (Class X)	CBSE Board	87.80%	2018

Experience

- **Microsoft, USA** *Jul. 2024 - Present*
  Research Fellow Redmond
  - **Resident** at the **Microsoft Program Synthesis (PROSE Team) Research Fellowship Program**
  - Part of the core team developing **Co-Pilot for Excel**. Working on **smart copy paste** problem for **east data migration** across **spreadsheets**
  - **Used Tools/Frameworks: Python | C# | SQL | Pandas | Microsoft Azure | Azure DevOps**

- **Amazon, India** *Jan. 2024 - Jun. 2024*
  Applied Scientist Intern Bangalore
  - Interned under the **Central Machine Learning Team**
  - Worked on developing a **Text-to-SQL system** using **Large Language Models (LLMs)** on internal**Amazon Databases**
  - Designed and implemented novel **schema linking techniques** to improve **Text-to-SQL generation accuracy**
  - Developed a **Retrieval-Augmented Generation (RAG)** pipeline for **retrieving relevant tables** and improving the **SQL query generation process**
  - Conducted extensive **experiments** for evaluating **accuracy** and **efficiency** of **Text-to-SQL models** across diverse **Amazon databases**
  - Proposed a **novel approach** that **outperformed state-of-the-art models** in **schema retrieval** by **15-20%**, and achieved **up to 2% higher execution accuracy** in **SQL generation tasks**
  - **Used Tools/Frameworks: Python | PyTorch | SQL | Pandas | Amazon SageMaker | Amazon Web Services**

- **Amazon, India** *Jun. 2023 - Aug.2023*
  Applied Scientist Intern Bangalore
  - Interned under the **International Machine Learning Team**
  - Worked on **designing** and **implementing** new **architectures** for **Parametric Efficient Fine-Tuning Large Language Models** (LLMs) on **Amazon Datasets**
  - New **fine-tuning methods** for LLMs **outperformed traditional ML Models** and **Deep Neural Networks** by a **large margin** over various experiments
  - New fine-tuning techniques **reduced training time** by **95.8%. Accelerated** the **training process** by **distributed training** over **8 NVIDIA GPUs**
  - **Performed extensive experiments** for **comparison** between existing **supervised/unsupervised techniques** and using **LLMs** and **LLM based embeddings** for **ranking** and **outlier prediction tasks**. LLM based techniques **significantly outperformed existing models** for **ranking** and **outlier prediction tasks**
  - **Used Tools/Frameworks: Python | PyTorch | PySpark | SQL | Pandas | Amazon SageMaker | Amazon Web Services**

- **Bosch Global Software Technologies (BGSW)** *May. 2022 - Jul. 2022*
  Software Development Engineering Intern (Metaverse) Bangalore
  - Interned under the Core Metaverse Project Team
  - Worked on **deploying metaverse platforms** on immersive devices like **Oculus Quest** and **HoloLens**
  - Helped in creating **intelligent avatars** for metaverse platforms using **Unity3D** and **C#**
  - Built **spatial audio** and **voice chat functionalities** for integration in Metaverse using Normcore and Photon SDK
  - **Used Tools/Frameworks: C# | Unity3D | Microsoft Mixed Reality Toolkit (MRTK) | Photon PUN | Agora**

Research Work

- **Research Paper - Skin Disease Detection Using Saliency Maps and Segmentation Techniques** *May. 2023*
  Published in Springer Series on Communications in Computer and Information Science (CCIS) Paper Link
  - **First authored** and presented the paper at the **7th International Conference on Computer Vision and Image Processing, (CVIP, 2022)** in VNIT Nagpur
  - Worked on devising a **Deep Learning Model** for **early detection of Skin Lesion images** using the **PH2** and **HAM10000** datasets

– Performed **image-preprocessing** using **Gaussian Filters** and trained **U-Net** model for **segmentation** and **MobileNetV2** model for inference
– Achieved **State-of-the-Art** results by using **data-prepossessing** and **segmentation** techniques

- **Research Project - ParkinSIGHT** *Jan. 2023*
  Computer Vision-based Early Detection of Parkinson's Disease using Brain SPECT Scans View Project
  – Developed an **Active Contour (Spline Method) algorithm** for **segmentation** of **axial slides** of **brain SPECT scans**
  – Performed image preprocessing steps, including **Gaussian filtering**, to **reduce noise of SPECT scan**. Did **feature selection** by **uniformly sampling coordinates** from **segmented curve**
  – Wrote a **custom script** to sort the left and right curves of the **substantia nigra coordinates** in a **clockwise** direction to ensure consistent and **standardized representation** of the **segmented regions** for training
  – **Demonstrated** that the utilization of **autoencoder-based dimensionality** reduction yielded **improved results** compared to the approach without dimensionality reduction
  – **Results obtained** with using **autoencoder** as dimensionality reduction for **PD and Normal Classification** was **Accuracy** as **0.95**, **F1-score** as **0.97**, **AUC-ROC** as **0.88** and for the **PD and SWEDD classification** was **Accuracy** as **0.93**, **F1-score** as **0.96**, **AUC-ROC** as **0.83**

Projects

- **MelodyAI** *GitHub Demo Samples*
  Innovative Music Generation with Artificial Intelligence
  – Built an **AI model** that **generates various melodies** was trained on famous Irish Folks Songs and Christmas Carols using **Long Short Term Memory Network** (LSTM)
  – Using the **NMD2ABC** program, dataset was converted to **ABC Notation.**
  – Used **mitdeeplearning module** to convert LSTM output from ABC notation to **.wav** format
  – **Used Tools/Frameworks: Python | TensorFlow | Keras | NumPy | LSTM**

- **ProctorLess** *GitHub Demo Video*
  AI Based Autonomous Proctoring Solution
  – Developed an **autonomous remote exam proctoring** solution using **Machine Learning** and **Deep Learning** techniques. The project won **1st prize** in Innovathon Hackathon
  – Built various AI Proctoring models which Included: **Eye-Detection**, **Mouth open Detection**, **Malicious Object Detection**, **Plagiarism Checker**, **Headphones Detection** and **Speech Recognition Surveillance**
  – Built a **demo website** for **deployment** of AI solutions (Website)
  – **Used Tools/Frameworks: Python | OpenCV | TensorFlow | Keras | YoloV5 | ReactJS | CSS**

- **ChromaVision** *GitHub*
  Image Colorization using Autoencoder
  – Built and trained an **autoencoder neural network** in **Keras** using **TensorFlow backend** for converting **old/grayscale images** into **colored RGB images**
  – For **encoding** the input image, used pre-trained **VGG-16** model trained on **ImageNet** Dataset
  – For **decoding** the encoded **latent representation**, used **UpSampling Layers** from **TensorFlow Layers API**
  – **Used Tools/Frameworks: Python | TensorFlow | Keras | OpenCV | VGG-16 | Matplotlib**

Achievements

- **Codechef – Max Rating: 1975 (User ID - iceonfire26)**: Global Rank – 46 in Starters 44 (1800+ participants), Global Rank – 155 in Starters – 33 (10000+ participants), Global Rank – 239 in Starters 43 (2500+ participants)
- **Leetcode (500+ Problems Solved) (User ID – RishitToteja)**: Global Rank – 1184 in Bi-Weekly 87 (23000+ participants)
- **Google Kickstart 2022**: Secured Global Rank 1761 & AIR (1329) in Round E **(Certificate)**
- **Selected** for **Amazon ML Summer School**, 2022, conducted by **Amazon Scientists** over a span of **4 weeks (Certificate)**
- Won the **2nd Prize** in **Bosch Technical Case Matrix Hackathon** Organized by **E-Cell DTU (Certificate)**
- Won the **3rd Prize** in **Innerve Hacks'21 Hackathon** Organized by **IGDTUW (Certificate)**
- **Departmental Rank 3** in my university with an **Aggregate CGPA** of **9.53**
- **Department Topper** in **6th Semester** with **CGPA** of **9.87**

Other Projects

- **CureIT Hackathon - AI Medical Chatbot** *View Project*
- **FUNANCO - ML based Financial Management System** *View Project Presentation Demo*
- **Gesture Controlled Hill Climb Racing** *View Project Demo Video*
- **ForecastNet - Stock Price Prediction with stacked LSTM** *View Project*
- **VIBE - Intelligent NLP-Based Profile Matching** *View Project Demo Video*

Technical Skills

- **Programming**: C++, C#, C, Python, Java, JavaScript, MATLAB, SQL, Assembly Language (x86)
- **Tools/Frameworks**: PyTorch, TensorFlow, OpenCV, Deepstream, MySQL, Git/GitHub, Operating Systems, Database Management Systems, Object Oriented Programming Systems
- **Operating Systems**: Windows, Linux

Key Courses Taken

- **Computer Engineering**: Data Structures and Algorithms, Programming Fundamentals, Operating Systems Design, Database Management Systems, Object Oriented Programming Systems
- **Mathematics**: Applied Mathematics I/II, Numerical Engineering Optimization Methods, Machine Learning, Control Systems, Deep Learning and ANN
- **Electrical Engineering**: Microprocessors and Microcontroller Applications, Electronic Devices and Circuits, Digital Circuits and System, Network Analysis and Synthesis

Open-Source Work / Volunteering

- **Google Explore ML with Crowdsource Facilitator (Profile)**: **Conducted** workshops in various universities covering introduction to **Machine Learning** and understanding **Neural Networks**
- **Guest Speaker at AMITY University**: Invited by the AMITY-AI Club to give **lecture** on **Transformers and Large Language Models**
- **Chief Guest at PSIT University**: Invited by IEEE Society to give **lecture** on **Basics of ML and Deep Learning**