UTKARSH MATHUR (Data Scientist)

I'm a Data Scientist with experience in Machine Learning, programming with Python, C++, and Perl. I'm pursuing B.Tech. in Polymer Science and Engineering at Indian Institute of Technology, Roorkee and I aim to do Masters in Data Science.

FORMAL EDUCATION

YEAR	DEGREE / BOARD	INSTITUTE / BOARD	Score
2022	B.Tech. Polymer Science and Engineering	Indian Institute of Technology, Roorkee	6.15 / 10
2018	Twelfth	Central Board of Secondary Education (CBSE)	89.6%
2016	Tenth	Central Board of Secondary Education (CBSE)	9.6 / 10

ACADEMIC & PROFESSIONAL COURSES

Deep Learning Specialisation (Coursera), **Data Structures and Algorithms in Java** (NPTEL), **IEE-03 Artififcial Neural Networks** (IIT Roorkee), **Machine Learning A-Z** (Udemy), **Data Science using Python** (EICT IIT Roorkee), **PEN-103 Computer Programming and Numerical Methods** (IIT Roorkee), **CHN-323 Computer Application in Chemical Engineering** (IIT Roorkee)

SKILLS

- Object Oriented Programming (**OOP**), Data Structures and Algorithms (**DSA**), Machine Learning (**ML**), Artificial Neural Network (**ANN**), Deep Learning (**DL**), Computer Vision (**CV**), Natural Language Processing (**NLP**), Big Data
- **Python**, **C++**, Java, Perl, **MATLAB**, **Scala**, **SQL**
- **NumPy**, **Pandas**, **Matplotlib**, Seaborn, Scikit-Learn, SciPy, **TensorFlow**, **Keras**, PyTorch, PySpark, OpenCV, **PIL** (Python Imaging Library), OpenCV, Git, Material Studio

EXPERIENCE

Research Intern | IIT ROORKEE | 6 Months (Present)

- Working as a research intern on a **Deep Learning** project under Dr. Mayank Goswami, Department of Physics, IIT Roorkee.
- Learning valuable Data Science skills like Data annotation, Data preprocessing, and Hyperparameter tuning.

Research Intern | IIT ROORKEE | 1 Month

- Worked as a research intern on a **Material Simulation** project under Dr. Gaurav Manik, Department of Polymer and Process Engineering, IIT Roorkee.
- Scripted an extended library in **Perl** for the Forcite module of Material Studio, that helped in calculating the contact angle and motion of a liquid droplet on an inclined surface coated with super-hydrophobic polymers.
- Learned how to work with a team of researchers and how to tailor a software code according to requirements.

PROJECTS

Breast Cancer Classification | Course Project

- Project Title - *Comparison of SVM and Neural Networks (with and without Backpropagation) on performance over Classification of Breast Cancer.*
- This was a course project for **IEE 03 Artificial Neural Networks** done with my classmate Aman Arora under the guidance of **Dr. G.N. Pillai**.
- The project aims to develop a classifier for the Classification of Breast Tumor into Malignant (cancer tumor) and Benign (noncancer tumor) using features obtained from several cell images. The dataset we used for this purpose was Breast Cancer Wisconsin dataset from Scikit Learn The classification process was carried out by three models, **Support Vector Machine** (**SVM**), **Neural Network** (with **Particle Swarm Optimizer**), and **Neural Network** (with **Gradient Descent**). The main objective is to compare these three models and find the most suitable model.
- The main conclusion was - Support Vector Machine (SVM) models perform a better classification task than Neural Networks models when the number of training examples is small (in our case only 512).

Facial Expression Recognition | Group Project

- Collaborated with a colleague over a comparative study of various Machine Learning and Computer Vision models over the FER2013 dataset.
- I worked on Deep Learning Models to build a Facial Expression Recognizer while my colleague worked over conventional Machine Learning model.

Melbourne Housing Prices | Personal Project

- Trained a housing price prediction model on Melbourne Housing Snapshot dataset on Kaggle.
- R-squares score of prediction from Polynomial Regression, Lasso Regression, and Decision Tree Regressor were in decreasing order while being above **0.999**.

| Website: https://007mathur.github.io/ | Email: umathur@ch.iitr.ac.in | Github: https://github.com/007mathur |

CIFAR -10 | Personal Project

- Comparative study of Neural Networks, Convolution Neural Networks, and ResNet models over CIFAR-10 dataset.
- Accuracies achieved were **48.42%, 72.899%, 77.759%** Neural Networks, Convolutional Neural Networks, and Residual Neural Network Models respectively.

Digit Recognizer | Personal Project

- Built a Digit Recognizer to recognize handwritten digits, based on the concepts of Convolution Neural Networks.
- The Digit Recognizer was trained on the MNSIT dataset, and it was able to achieve an accuracy of **99.21%**.

Titanic | Personal Project

- Compared Logistic Regression, Support Vector Machines, and Neural Networks on classification over the famous Titanic dataset.
- For the above models, accuracies obtained were **81%**, **78.77%**, and **74.44%** in the order mentioned.

POSITIONS OF RESPONSIBILITIES & EXTRA-CURRICULARS

- Manager, **TEDx IITROORKEE** (September 2019 – Present)
- Company Coordinator, **Training and Placement Office (TPO), IIT Roorkee** (January 2020 – August 2020)
- Core Team Member, **Cognizance – Technical Festival of IIT Roorkee** (November 2018 – January 2020)
- Volunteer, **Prahari Kaksh**, **NSS (3 UK CTR), IIT Roorkee** (September 2018 – Present)