MUSKAN GOYAL

720-313-5502 | **Email:** muskan.goyal@colorado.edu | **Portfolio:** muskan-goyal6.github.io | **Linkedin:** linkedin.com/in/Muskan-goyal6/ | Google Scholar | **Github:** github.com/Muskan-goyal6

EDUCATION

University of Colorado Boulder, USA | Master of Science (Computer Science) **Aug 2022 – May 2024**

- Design and Analysis of Algorithms, Natural Language Processing, Software Engineering, Machine learning

Maharaja Agrasen Institute of Technology, Delhi, India | B. Tech (Computer Science and Engineering) **Aug 2017 – June 2021**

- Machine Learning, Applied Mathematics, Artificial Intelligence, Theory of Computation| **CGPA: 8.31/10**

TECHNICAL SKILLS

Languages: Java, Python, HTML, CSS, Javascript, SQL

Tools & Frameworks: Numpy, Pandas, Matplotlib, SciKit, Spacy, NLTK, TensorFlow, Keras, PyTorch, Hugging Face transformers & datasets, Git, Colab, PostgreSQL, MySQL

Technical: Data Structures, Computer Vision, Deep Learning, Machine Learning, Data Analysis, Time Series Forecasting, Data Visualization.

WORK EXPERIENCE

Research Assistant| Indraprastha Institute of Technology, Delhi, India **Sept 2021 – July 2022**

- Collaborated with a team of 5 members to conduct a detailed research analysis on types of privacy attacks on machine learning and deep learning models (model extraction, model inversion, membership and property inference).
- Reviewed literature on various defensive measures to protect the privacy and confidentiality of models against different attacks.
- Reproduced black-box model attacks on ML Systems using techniques like ActiveTheif and Knockoff Nets in **PyTorch**.
- Proposed a technique to build a shadow model that helps in black-box model explanations to explore and understand the behavior of any black-box model in different feature spaces using **PyTorch**.

Research Engineer Intern | University of Fortaleza, Brazil **Mar 2020 – May 2020**

- Collected and pre-processed the chest x-ray data for training and evaluation pipeline. Applied techniques such as resizing, normalization, and image hashing algorithm for removing the duplicates.
- Constructed a CNN architecture by adding custom layers to VGG16 model. Fined tuned the CNN architecture for the new Covid dataset.
- Designed and executed an Auxiliary Classifier Generative Adversarial Network (AC-GAN) based model called CovidGAN that generated synthetic chest x-ray images using **Keras and TensorFlow**.
- **Result:** Visualized the results with Principal Component Analysis (PCA) and confusion matrix. The addition of synthetic images produced by CovidGAN increased the accuracy of CNN for Covid-19 detection from 85% (F1-score 0.83) to 95% (F1 score 0.95). Published in IEEE Access.

Research Engineer Intern | Center Edge Hill University, Lancashire, England **Sept 2019 – Dec 2019**

- Trained and evaluated 4 established CNN architectures for corn leaf disease classification: VGGNet, XceptionNet, EfficientNet, NASNet.
- Performed Model Compression and developed an optimized DenseNet model for corn leaf disease identification with **Keras and TensorFlow.**
- Proved that its performance was close to the established CNN architectures with significantly fewer parameters and computation time.
- Used Grid Search method to find the optimal hyperparameter values and analyzed the models' performance through rigorous simulations.
- **Result:** The proposed DenseNet was computationally cost-effective with 98.06% accuracy, 0.07 million parameters and it took 3 minutes per epoch. Paper published in Computers and Electronics in Agriculture, Elsevier.

PUBLICATIONS

- **"Domain-Controlled Title Generation with Human Evaluation".** In Proceedings of The International Conference on Innovative Computing and Communications (pp. 461-474), Springer, Singapore.
- **"CovidGAN: Data Augmentation using Auxiliary Classifier GAN for improved Covid-19 Detection".** Published in IEEE Access, volume 8.
- **"An optimized dense convolutional neural network model for disease recognition and classification in corn leaf".** Published in Computers and Electronics in Agriculture, Elsevier.

SELECTED PROJECTS

- **Excel Clone:** This web application clone of MS-Excel is created using **HTML, CSS, and JavaScript** entirely. It has a completely responsive design with features like text formatting, address bar, formula evaluation, multiple sheet handling functionality, and two-way binding for cell properties. It also uses cycle detection algorithm and color tracking for formula cycle validation.
- **Code-mixed VQA Dataset:** Implemented code-mix generation to make a novel dataset that helps multilingual speakers and medical practitioners. The dataset composed of medical images that posed clinical queries in Hindi, English, or Code-mixed (Hinglish: Hindi – English) language. Used **Stanza for POS tagging**, and **GIZA++** tool for creating alignments between English and Hindi questions. Lastly, calculated the **Switch-Point-Factor and Code-Mixed-Index** for all the code-mixed transliterations.
- **Sentiment Analyzer:** Conducted **Exploratory Data Analysis** and built a Sentiment Analyzer for Amazon reviews and IMDB dataset (total 64000 reviews). Analyzed the stopwords, frequency of words, distribution of rating scores, and the results. Cleaned the dataset and experimented with various models (Naive Bayes, XGBoost, MLP, RNN, BERT) for comparative insights. BERT gave the highest accuracy with 89% on IMDB and 95% on Amazon dataset.

ACHIEVEMENTS and SERVICES

- Grand Finalist in Smart India Hackathon 2019.
- Volunteer at the International Conference on Innovative Computing and Communication 2020 (ICICC 2020).
- Collaboration story of CovidGAN appeared in WTT Ventures' study "From Pelotas to Boa Vista: scientific partnerships to face the pandemic".