Divyansh Singh

Email: divyanshucs127@gmail.com
Mobile: (352) 740 6854

Linkedin: linkedin.com/in/divyansh-singh-604
Github: github.com/divyanshrm

EDUCATION

- **University Of Florida - Herbert Wertheim College of Engineering** — Gainesville, USA
  Master of Science - Computer Science, GPA 3.94/4 — *August 2022 - May 2024*
  Courses: *Distributed Operating System Principles, Advanced Data Structures, Natural Language Processing*

- **The LNM Institute of Information Technology** — Jaipur, India
  Bachelor of Technology - Computer Science and Engineering — *August 2018 - June 2022*
  Courses: *Operating Systems, Data Structures, Analysis Of Algorithms, Artificial Intelligence, Machine Learning, Networking, Database Management*

SKILLS SUMMARY

- **Languages**: Python, JAVA, C/C++, SQL, Bash, Javascript, CSS
- **Frameworks**: Scikit, NLTK, TensorFlow, Keras, Torch, Spark, Teradata, React, Django, MongoDB, Flask
- **Specialization**: Software Development, Data Science, Machine Learning, Business Analytics, Web Development

EXPERIENCE

- **The University of Florida** — Gainesville, FL, USA
  Research Assistant — *April 2023 - Present*
  - **Working in Prof. Bonnie Dorr's Natural Language Processing lab**: Leading the Topic Modelling Portion in the ARLIS Information Competition Simulation project, using clustered BERT Embeddings to detect topics in over 10,000 un-annotated documents and achieving a cosine similarity of 0.72 compared to ground truth.
  - **ARLIS Information Competition Simulation**: This project aims to develop a simulated population that replicates the dissemination of information among various actors and individuals.
  - **Deploying open source large language models (LLMs)**: Researching and deploying LLMs including LLAMA 2, Falcon 40B etc. And using them for NLP tasks like Machine Translation and Topic Detection.

- **ZS** — Pune, India
  Decison Analytics Associate (Internship) — *January 2022 - June 2022*
  - **Raw Patient Level and Physician level Claims data**: Prepared analytics reports related to market share, tracking patient's treatment journey using their claims data. Compiled data from 6 retail claims datasets from Hive databases.
  - **Physician Deciling for Promotions**: Identified 500 high decile physicians and territories using diagnosis statistics across the United States and recommended them for promotions.
  - **Weekly Report of New Patients**: Prepared an automated pipeline to identify new patients each week (AML) and their physicians for sending out biopsy reminders using their Diagnosis Statistics. Increased Biopsies count by 23 percent.

- **The LNM Institute of Information Technology, Jaipur** — Remote
  Research Assistant — *July 2021 - December 2021*
  - **Computer Vision based Exam Proctoring System**: Worked under Prof. Miller and developed an AI powered proctoring system for online exams using Computer Vision for over 2,000 Students.
  - **Inference System for Efficient Proctoring from Multiple Data Feeds**: Developed the Proctoring system to monitor the camera feed and audio which performed Face Authentication, Phone Detection, Spoofing, Speech to text using **Opencv, Dlib, YOLOv3, Darknet**. Also deployed multi-threading for optimized usage of cloud computing.

- **Indian Institute of Technology, Kanpur** — Remote
  Research Intern — *May 2021 - July 2021*
  - **Aggregated a dataset of about 1,250,000 samples containing hate speech.**: Scraped tweets and comments containing hate speech from twitter, Reddit data. Collected a labeled dataset, having data points from over 10,000 web pages.
  - **Building a Classifier to Identify Hate Speech in Online Text Streams**: Worked on a multiclass classifier using BERT, ELECTRA encoding architectures and ensembling the baselines. Used transformers based models with a testing accuracy of 98 percent.

PROJECTS

- **JobPly:(Full Stack Development)** : Engineered JobPly, a full-stack Django app simplifying job posting and application processes. Enabling employers to list jobs seamlessly and users to apply with ease. Tech: Python, Django, Javascript, HTML, CSS

- **Reddit Chatbot: A chatbot modelled on Reddit corpus.(Natural Language Processing, Deep Learning)**: An innovative question answering chatbot trained on reddit posts and their replies (Over 1,000,000 samples) using attention based seq2seq neural networks. Tech: Python, Nltk, Tensorflow, MySql, SQLite

PUBLICATIONS

- **Polyth-Net: Classification of Polythene Bags for Garbage Segregation using Deep Learning**: Research project published at 2021 International Conference on Sustainable Energy and Future Electric Transportation - IEEE · 2021

- **Detection of Emotions in Hindi-English Code Mixed Text Data**: Research Project accepted to be Published at IEEE International Conference on Artificial Intelligence and Smart Systems - IEEE . 2022

VOLUNTEER EXPERIENCE

- **Microsoft Student Partner** — Noida, India
  Conducted official technical workshops and hackathons impacting over 3000 students. — *April 2020 - December 2021*